Taller 8

Métodos Computacionales para Políticas Públicas - URosario

Entrega: viernes 12-abr-2019 11:59 PM

Juan Sebastián Valbuena Silva

juans.valbuena@urosario.edu.co

Instrucciones:

- Guarde una copia de este *Jupyter Notebook* en su computador, idealmente en una carpeta destinada al material del curso.
- Modifique el nombre del archivo del *notebook*, agregando al final un guión inferior y su nombre y apellido, separados estos últimos por otro guión inferior. Por ejemplo, mi *notebook* se llamaría: mcpp_taller8_santiago_matallana
- Marque el *notebook* con su nombre y e-mail en el bloque verde arriba. Reemplace el texto "[Su nombre acá]" con su nombre y apellido. Similar para su e-mail.
- Desarrolle la totalidad del taller sobre este *notebook*, insertando las celdas que sea necesario debajo de cada pregunta. Haga buen uso de las celdas para código y de las celdas tipo *markdown* según el caso.
- Recuerde salvar periódicamente sus avances.
- Cuando termine el taller:
  1. Descárguelo en PDF. Si tiene algún problema con la conversión, descárguelo en HTML.
  2. Suba todos los archivos a su repositorio en GitHub, en una carpeta destinada exclusivamente para este taller, antes de la fecha y hora límites.

---

1. [1 punto]

Usando expresiones regulares extraiga en una lista todos los números presentes en el siguiente objeto de Python:

ob1 = "JEFF BEZOS, the founder of Amazon, has reached a divorce settlement with his wife, MacKenzie. Mr Bezos will keep all the shares in the Washington Post and Blue Origin, a space-exploration firm, as well as 75% of the couple's Amazon stock. Mrs Bezos will retain a 4% stake in the tech giant, worth nearly $36bn, which is likely to make her the third-richest woman alive when the divorce is finalised."

In [8]:

```
import requests
import re
from bs4 import BeautifulSoup
```

In [9]:

```
ob1 = "JEFF BEZOS, the founder of Amazon, has reached a divorce settlement with his wife,
MacKenzie. Mr Bezos will keep all the shares in the Washington Post and Blue Origin, a space-explo
ration firm, as well as 75% of the couple's Amazon stock. Mrs Bezos will retain a 4% stake in the
tech giant, worth nearly $36bn, which is likely to make her the third-richest woman alive when the
divorce is finalised."
pun1 = re.findall('[0-9]+', ob1)
punt1 = re.findall('\d+', ob1)
print(pun1)
print(punt1)
```

```
['75', '4', '36']
['75', '4', '36']
```

2. [1 punto]

2. [1 punto]

Usando expresiones regulares ahora extraiga de *ob1* sólo los números que correspondan a porcentajes.

In [10]:

```
ob1 = "JEFF BEZOS, the founder of Amazon, has reached a divorce settlement with his wife,
MacKenzie. Mr Bezos will keep all the shares in the Washington Post and Blue Origin, a space-explo
ration firm, as well as 75% of the couple's Amazon stock. Mrs Bezos will retain a 4% stake in the
tech giant, worth nearly $36bn, which is likely to make her the third-richest woman alive when the
divorce is finalised."
pun2 = re.findall('\d+%', ob1)
print(pun2)
```

```
['75%', '4%']
```

3. [2 puntos]

Usando expresiones regulares, escriba una función de Python que reciba una fecha en formato **Marzo 7, 2019** y retorne la fecha en formato **2019-07-03**

In [11]:

```

  File "<ipython-input-11-cb8e39560293>", line 4
    from datetime import datetime:
                                 ^
SyntaxError: invalid syntax
```

4. [3 puntos]

ob2 es un string que reune una lista de clases en una universidad. Use expresiones regulares para extraer los códigos de cada una de las clases. Ejemplo: El código de la clase **COMPSCI 143 (Spring 2012): Machine Learning** es 143.

ob2 = "COMPSCI 270 (Spring 2019): Introduction to Artificial Intelligence. COMPSCI 590.2 (Fall 2018): Computational Microeconomics: Game Theory, Social Choice, and Mechanism Design. COMPSCI 223 (Spring 2018): Computational Microeconomics. COMPSCI 570 (Fall 2017): Artificial Intelligence. COMPSCI 590.3 (Fall 2017): Ethics and AI. COMPSCI 590.2 (Spring 2017): Computation, Information, and Learning in Market Design. COMPSCI 590.4 (Spring 2016): Computational Microeconomics: Game Theory, Social Choice, and Mechanism Design. COMPSCI 290.4/590.4 (Spring 2015): Crowdsourcing Societal Tradeoffs. COMPSCI 570 (Fall 2014): Artificial Intelligence. COMPSCI 590.4 (Spring 2014): Computational Microeconomics: Game Theory, Social Choice, and Mechanism Design. COMPSCI 590.1 (Fall 2012): Linear and Integer Programming. COMPSCI 173 (Spring 2012): Computational Microeconomics. COMPSCI 296.1 (Fall 2011): Computational Microeconomics: Game Theory, Social Choice, and Mechanism Design. COMPSCI 296.1 (Fall 2010): Linear and Integer Programming. COMPSCI 173 (Spring 2010): Computational Microeconomics. COMPSCI 196.1/296.1 (Fall 2009): Computational Microeconomics: Game Theory, Social Choice, and Mechanism Design. COMPSCI 170 (Spring 2009): Introduction to Artificial Intelligence. COMPSCI 270 (Fall 2008): Artificial Intelligence. COMPSCI 196/296.2 (Spring 2008): Linear and Integer Programming. COMPSCI 196.2 (Fall 2007): Introduction to Computational Economics. COMPSCI 296.3 (Spring 2007): Topics in Computational Economics. COMPSCI 296.2 (Fall 2006): Computational Game Theory and Mechanism Design."

In [12]:

```
import re
ob2 = "COMPSCI 270 (Spring 2019): Introduction to Artificial Intelligence. COMPSCI 590.2 (Fall
2018): Computational Microeconomics: Game Theory, Social Choice, and Mechanism Design. COMPSCI 223
(Spring 2018): Computational Microeconomics. COMPSCI 570 (Fall 2017): Artificial Intelligence. COM
PSCI 590.3 (Fall 2017): Ethics and AI. COMPSCI 590.2 (Spring 2017): Computation, Information, and
Learning in Market Design. COMPSCI 590.4 (Spring 2016): Computational Microeconomics: Game Theory,
Social Choice, and Mechanism Design. COMPSCI 290.4/590.4 (Spring 2015): Crowdsourcing Societal Tra
deoffs. COMPSCI 570 (Fall 2014): Artificial Intelligence. COMPSCI 590.4 (Spring 2014):
Computational Microeconomics: Game Theory, Social Choice, and Mechanism Design. COMPSCI 590.1 (Fal
l 2012): Linear and Integer Programming. COMPSCI 173 (Spring 2012): Computational Microeconomics.
COMPSCI 296.1 (Fall 2011): Computational Microeconomics: Game Theory, Social Choice, and Mechanism
Design. COMPSCI 296.1 (Fall 2010): Linear and Integer Programming. COMPSCI 173 (Spring 2010): Comp
utational Microeconomics. COMPSCI 196.1/296.1 (Fall 2009): Computational Microeconomics: Game Theo
ry, Social Choice, and Mechanism Design. COMPSCI 170 (Spring 2009): Introduction to Artificial Int
elligence. COMPSCI 270 (Fall 2008): Artificial Intelligence. COMPSCI 196/296.2 (Spring 2008): Line
ar and Integer Programming. COMPSCI 196.2 (Fall 2007): Introduction to Computational Economics. CO
MPSCI 296.3 (Spring 2007): Topics in Computational Economics. COMPSCI 296.2 (Fall 2006):
```

```
Computational Game Theory and Mechanism Design."
punt4 = re.findall(('COMPSCI (\d\S+\d)+ \('), ob2)
print(punt4)
```

```
['270', '590.2', '223', '570', '590.3', '590.2', '590.4', '290.4/590.4', '570', '590.4', '590.1',
'173', '296.1', '296.1', '173', '196.1/296.1', '170', '270', '196/296.2', '196.2', '296.3', '296.2
']
```

5. [5 puntos]

ob3 es un string que reune una lista de publicaciones. Use expresiones regulares para extraer todos los *Journals* en los cuales el autor ha publicado. Ejemplo: El paper **Bail, CA. "The configuration of symbolic boundaries against immigrants in Europe." American Sociological Review 73.1 (January 1, 2008): 37-59. Full Text** fue publicado en el Journal *American Sociological Review*

ob3 = "Bail, CA, Argyle, LP, Brown, TW, Bumpus, JP, Chen, H, Hunzaker, MBF, Lee, J, Mann, M, Merhout, F, and Volfovsky, A. "Exposure to opposing views on social media can increase political polarization." Proceedings of the National Academy of Sciences of the United States of America 115.37 (September 2018): 9216-9221. Full Text Open Access Copy. Bail, CA, Merhout, F, and Ding, P. "Using Internet search data to examine the relationship between anti-Muslim and pro-ISIS sentiment in U.S. counties." Science Advances 4.6 (June 6, 2018): eaao5948-null. Full Text Open Access Copy. Bail, CA, Brown, TW, and Mann, M. "Channeling Hearts and Minds: Advocacy Organizations, Cognitive-Emotional Currents, and Public Conversation." American Sociological Review 82.6 (December 1, 2017): 1188-1213. Full Text. Bail, CA. "Taming Big Data: Using App Technology to Study Organizational Behavior on Social Media." Sociological Methods and Research 46.2 (March 1, 2017): 189-217. Full Text. McDonnell, TE, Bail, CA, and Tavory, I. "A Theory of Resonance." Sociological Theory 35.1 (March 1, 2017): 1-14. Full Text. Bail, CA. "Combining natural language processing and network analysis to examine how advocacy organizations stimulate conversation on social media." Proceedings of the National Academy of Sciences of the United States of America 113.42 (October 2016): 11823-11828. Full Text. Bail, CA. "Emotional Feedback and the Viral Spread of Social Media Messages About Autism Spectrum Disorders." American journal of public health 106.7 (July 2016): 1173-1180. Full Text. Bail, CA. "The public life of secrets: Deception, disclosure, and discursive framing in the policy process." Sociological Theory 33.2 (January 1, 2015): 97-124. Full Text. Bail, CA. "The cultural environment: Measuring culture with big data." Theory and Society 43.3 (January 1, 2014): 465-524. Full Text."

In [15]:

```
ob3 = '''Bail, CA, Argyle, LP, Brown, TW, Bumpus, JP, Chen, H, Hunzaker, MBF, Lee, J, Mann, M, Mer
hout, F, and Volfovsky, A. "Exposure to opposing views on social media can increase political pola
rization." Proceedings of the National Academy of Sciences of the United States of America 115.37
(September 2018): 9216-9221. Full Text Open Access Copy.\n", "Bail, CA, Merhout, F, and Ding, P. "
Using Internet search data to examine the relationship between anti-Muslim and pro-ISIS sentiment
in U.S. counties." Science Advances 4.6 (June 6, 2018): eaao5948-null. Full Text Open Access Copy.
\n", "Bail, CA, Brown, TW, and Mann, M. "Channeling Hearts and Minds: Advocacy Organizations, Cogn
itive-Emotional Currents, and Public Conversation." American Sociological Review 82.6 (December 1,
2017): 1188-1213. Full Text.\n", "Bail, CA. "Taming Big Data: Using App Technology to Study Organi
zational Behavior on Social Media." Sociological Methods and Research 46.2 (March 1, 2017): 189-21
7. Full Text.\n", "McDonnell, TE, Bail, CA, and Tavory, I. "A Theory of Resonance." Sociological T
heory 35.1 (March 1, 2017): 1-14. Full Text.\n", "Bail, CA. "Combining natural language processing
and network analysis to examine how advocacy organizations stimulate conversation on social media.
" Proceedings of the National Academy of Sciences of the United States of America 113.42 (October
2016): 11823-11828. Full Text.\n", "Bail, CA. "Emotional Feedback and the Viral Spread of Social M
edia Messages About Autism Spectrum Disorders." American journal of public health 106.7 (July 2016
): 1173-1180. Full Text.\n", "Bail, CA. "The public life of secrets: Deception, disclosure, and di
scursive framing in the policy process." Sociological Theory 33.2 (January 1, 2015): 97-124. Full
Text.\n", "Bail, CA. "The cultural environment: Measuring culture with big data." Theory and Socie
ty 43.3 (January 1, 2014): 465-524. Full Text." '''
journals = re.findall("\" (.+?) \d", ob3)
print(journals)
```

```
['Proceedings of the National Academy of Sciences of the United States of America', 'Science Advan
ces', 'American Sociological Review', 'Sociological Methods and Research', 'Sociological Theory',
'Proceedings of the National Academy of Sciences of the United States of America', 'American journ
al of public health', 'Sociological Theory', 'Theory and Society']
```

6. [10 puntos]

Vamos a hacer "scraping" a esta página: https://archive.ics.uci.edu/ml/datasets.php, que contiene un listado de 468 bases de datos que hacen parte del repositorio de la Universidad de California, Irvine.

Su tarea consiste en crear un "Pandas dataframe" que contenga 468 filas (una por base de datos) y las siguientes columnas:

- Nombre de la base de datos

- Nombre de la base de datos
- Link a la base de datos
- Tipo de datos
- Tipo de tarea a resolver (default task)
- Tipo de las variables
- Número de observaciones
- Número de variables
- Año
- Descripción de la base (Pista: Utilice la opción list view: https://archive.ics.uci.edu/ml/datasets.php?format=&task=&att=&area=&numAtt=&numIns=&type=&sort=nameUp&view=list)

In [6]:

```
import requests
from bs4 import BeautifulSoup
```

In [7]:

```
url = "https://archive.ics.uci.edu/ml/datasets.php"
```

In [8]:

```
html = requests.get(url).text
```

In [9]:

```
soup = BeautifulSoup(html, "lxml")
```

In [10]:

```
print(soup.prettify())
```

```
<!DOCTYPE HTML>
<html>
 <body>
  <p>
   "-//W3C//DTD HTML 4.01 Transitional//EN\"&gt;
  </p>
  <title>
   UCI Machine Learning Repository: Data Sets
  </title>
  <!-- Stylesheet link -->
  <link href="assets/ml.css" rel="stylesheet" type="text/css"/>
  <script language="JavaScript" type="text/javascript">
   <!--
function checkform ( form )
{
  // see http://www.thesitewizard.com/archive/validation.shtml
  // for an explanation of this script and how to use it on your
  // own website

  // ** START **
  if (form.q.value == "")
  {
    alert( "Please enter search terms." );
    form.q.focus();
    return false ;
  }

  if (getCheckedValue(form.sitesearch) == "ics.uci.edu" &&
form.q.value.indexOf("site:archive.ics.uci.edu/ml") == -1)
  {
    form.q.value = form.q.value + " site:archive.ics.uci.edu/ml";
  }

  // ** END **
  return true ;
}

// return the value of the radio button that is checked
```

```
// return an empty string if none are checked, or
// there are no radio buttons
function getCheckedValue(radioObj) {
 if(!radioObj)
  return "";
 var radioLength = radioObj.length;
 if(radioLength == undefined)
  if(radioObj.checked)
   return radioObj.value;
  else
   return "";
 for(var i = 0; i < radioLength; i++) {
  if(radioObj[i].checked) {
   return radioObj[i].value;
  }
 }
 return "";
}




//-->
  </script>
  <!-- SITE HEADER (INCLUDES LOGO AND SEARCH BOX) -->
  <!-- SITE HEADER (INCLUDES LOGO AND SEARCH BOX) -->
  <table bgcolor="#003366" width="100%">
   <tr>
    <td>
     <span class="normal">
      <a alt="Home" href="index.html">
       <img border="0" src="assets/logo.gif"/>
      </a>
      <br/>
      <a href="http://cml.ics.uci.edu">
       <font color="FFDD33">
        Center for Machine Learning and Intelligent Systems
       </font>
      </a>
     </span>
    </td>
    <td align="right" valign="top" width="100%">
     <span class="whitetext">
      <a href="about.html">
       About
      </a>
      <a href="citation_policy.html">
       Citation Policy
      </a>
      <a href="donation_policy.html">
       Donate a Data Set
      </a>
      <a href="contact.html">
       Contact
      </a>
     </span>
     <br/>
     <br/>
     <!-- Search Google -->
     <form action="http://www.google.com/custom" method="GET" onsubmit="return checkform(this);">
      <input maxlength="255" name="q" size="30" type="text" value=""/>
      <input name="sa" type="submit" value="Search"/>
      <input name="cof" type="hidden"
value="AH:center;LH:130;L:http://archive.ics.uci.edu/assets/logo.gif;LW:384;AWFID:869c0b2eaa8d518e;

      <input name="domains" type="hidden" value="ics.uci.edu"/>
      <br/>
      <input checked="" name="sitesearch" type="radio" value="ics.uci.edu"/>
      <span class="whitetext">
       <font size="1">
        Repository
       </font>
      </span>
      <input name="sitesearch" type="radio" value=""/>
      <span class="whitetext">
       <font size="1">
```

```
        Web
       </font>
      </span>
      <a href="http://www.google.com/search">
       <img align="middle" alt="Google" border="0" height="27"
src="http://www.google.com/logos/Logo_25blk.gif"/>
      </a>
      <br/>
     </form>
     <!-- Search Google -->
     <span class="whitetext">
      <a href="datasets.php">
       <font color="#FFDD33" size="3">
        <b>
         View ALL Data Sets
        </b>
       </font>
      </a>
     </span>
     <br/>
    </td>
   </tr>
  </table>
  <br/>
  <table cellpadding="3">
   <tr>
    <td valign="top">
     <table>
      <tr>
       <td>
        <p align="center" class="big">
         Browse Through:
        </p>
       </td>
      </tr>
     </table>
     <table border="2" cellpadding="2" width="150">
      <tr>
       <td bgcolor="#003366">
        <p class="whitetext">
         <b>
          Default Task
         </b>
        </p>
       </td>
      </tr>
      <tr>
       <td valign="top">
        <p class="normal">
         <a href="datasets.php?
format=&amp;task=cla&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view=
e">
          Classification
         </a>
         <font color="red">
          (350)
         </font>
         <br/>
         <a href="datasets.php?
format=&amp;task=reg&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view=
e">
          Regression
         </a>
         <font color="red">
          (96)
         </font>
         <br/>
         <a href="datasets.php?
format=&amp;task=clu&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view=
e">
          Clustering
         </a>
         <font color="red">
          (84)
         </font>
         <br/>
         <a href="datasets.php?
```

format=&task=other&att=&area=&numAtt=&numIns=&type=&sort=nameUp&vie
ble">
          Other
         </a>
         <font color="red">
          (55)
         </font>

       </td>
      </tr>
      <tr>
       <td bgcolor="#003366">

          Attribute Type

       </td>
      </tr>
      <tr>
       <td valign="top">

         <a href="datasets.php?
format=&task=&att=cat&area=&numAtt=&numIns=&type=&sort=nameUp&view=
e">
          Categorical
         </a>
         <font color="red">
          (38)
         </font>

         <a href="datasets.php?
format=&task=&att=num&area=&numAtt=&numIns=&type=&sort=nameUp&view=
e">
          Numerical
         </a>
         <font color="red">
          (307)
         </font>

         <a href="datasets.php?
format=&task=&att=mix&area=&numAtt=&numIns=&type=&sort=nameUp&view=
e">
          Mixed
         </a>
         <font color="red">
          (55)
         </font>

       </td>
      </tr>
      <tr>
       <td bgcolor="#003366">

          Data Type

       </td>
      </tr>
      <tr>
       <td valign="top">

         <a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=mvar&sort=nameUp&view
le">
          Multivariate
         </a>
         <font color="red">
          (357)
         </font>

         <a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=uvar&sort=nameUp&view
le">
          Univariate
         </a>

```
          <font color="red">
           (23)
          </font>
          <br/>
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=seq&amp;sort=nameUp&amp;view=
e">
           Sequential
          </a>
          <font color="red">
           (47)
          </font>
          <br/>
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=ts&amp;sort=nameUp&amp;view=t
">
           Time-Series
          </a>
          <font color="red">
           (91)
          </font>
          <br/>
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=text&amp;sort=nameUp&amp;view
le">
           Text
          </a>
          <font color="red">
           (53)
          </font>
          <br/>
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=dt&amp;sort=nameUp&amp;view=t
">
           Domain-Theory
          </a>
          <font color="red">
           (23)
          </font>
          <br/>
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=other&amp;sort=nameUp&amp;vie
ble">
           Other
          </a>
          <font color="red">
           (21)
          </font>
          <br/>
         </p>
        </td>
       </tr>
       <tr>
        <td bgcolor="#003366">
         <p class="whitetext">
          <b>
           Area
          </b>
         </p>
        </td>
       </tr>
       <tr>
        <td valign="top">
         <p class="normal">
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=life&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view
le">
           Life Sciences
          </a>
          <font color="red">
           (107)
          </font>
          <br/>
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=phys&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view
le">
           Physical Sciences
```

```
          </a>
          <font color="red">
           (49)
          </font>
          <br/>
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=comp&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view
le">
           CS / Engineering
          </a>
          <font color="red">
           (170)
          </font>
          <br/>
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=soc&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view=
e">
           Social Sciences
          </a>
          <font color="red">
           (26)
          </font>
          <br/>
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=bus&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view=
e">
           Business
          </a>
          <font color="red">
           (29)
          </font>
          <br/>
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=game&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view
le">
           Game
          </a>
          <font color="red">
           (10)
          </font>
          <br/>
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=other&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;vie
ble">
           Other
          </a>
          <font color="red">
           (73)
          </font>
         </p>
        </td>
       </tr>
       <tr>
        <td bgcolor="#003366">
         <p class="whitetext">
          <b>
           # Attributes
          </b>
         </p>
        </td>
       </tr>
       <tr>
        <td valign="top">
         <p class="normal">
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=less10&amp;numIns=&amp;type=&amp;sort=nameUp&amp;vi
able">
           Less than 10
          </a>
          <font color="red">
           (113)
          </font>
          <br/>
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=10to100&amp;numIns=&amp;type=&amp;sort=nameUp&amp;v
table">
           10 to 100
```

```
          </a>
          <font color="red">
           (210)
          </font>
          <br/>
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=greater100&amp;numIns=&amp;type=&amp;sort=nameUp&am
ew=table">
           Greater than 100
          </a>
          <font color="red">
           (84)
          </font>
         </p>
        </td>
       </tr>
       <tr>
        <td bgcolor="#003366">
         <p class="whitetext">
          <b>
           # Instances
          </b>
         </p>
        </td>
       </tr>
       <tr>
        <td valign="top">
         <p class="normal">
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=less100&amp;type=&amp;sort=nameUp&amp;v
table">
           Less than 100
          </a>
          <font color="red">
           (27)
          </font>
          <br/>
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=100to1000&amp;type=&amp;sort=nameUp&amp
w=table">
           100 to 1000
          </a>
          <font color="red">
           (162)
          </font>
          <br/>
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=greater1000&amp;type=&amp;sort=nameUp&a
iew=table">
           Greater than 1000
          </a>
          <font color="red">
           (246)
          </font>
         </p>
        </td>
       </tr>
       <tr>
        <td bgcolor="#003366">
         <p class="whitetext">
          <b>
           Format Type
          </b>
         </p>
        </td>
       </tr>
       <tr>
        <td valign="top">
         <p class="normal">
          <a href="datasets.php?
format=mat&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view=
e">
           Matrix
          </a>
          <font color="red">
           (324)
          </font>
```

```
         <br/>
         <a href="datasets.php?
format=nonmat&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;vi
able">
          Non-Matrix
         </a>
         <font color="red">
          (145)
         </font>
        </p>
       </td>
      </tr>
     </table>
    </td>
    <td valign="top">
     <table width="100%">
      <tr>
       <td>
        <p class="big">
         <b>
          469
         </b>
         Data Sets
        </p>
       </td>
       <td align="right">
        <p class="normal">
         <font color="gray">
          Table View
         </font>
         <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view=lis

          List View
         </a>
        </p>
       </td>
      </tr>
     </table>
     <table border="1" cellpadding="5">
      <tr bgcolor="#003366">
       <td class="normal, whitetext">
        <p class="normal, whitetext">
         <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameDown&amp;view=t
">
          <b>
           Name
          </b>
         </a>
        </p>
       </td>
       <!-- <td><p class="normal, whitetext"><b>Abstract</b></p></td> -->
       <td>
        <p class="normal, whitetext">
         <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=typeUp&amp;view=tab

          <b>
           Data Types
          </b>
         </a>
        </p>
       </td>
       <td>
        <p class="normal, whitetext">
         <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=taskUp&amp;view=tab

          <b>
           Default Task
          </b>
         </a>
        </p>
       </td>
       <td>
        <p class="normal, whitetext">
```

```
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=attTypeUp&amp;view=
e">
           <b>
            Attribute Types
           </b>
          </a>
         </p>
        </td>
        <td>
         <p class="normal, whitetext">
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=instUp&amp;view=tab

           <b>
            # Instances
           </b>
          </a>
         </p>
        </td>
        <td>
         <p class="normal, whitetext">
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=attUp&amp;view=tabl

           <b>
            # Attributes
           </b>
          </a>
         </p>
        </td>
        <td>
         <p class="normal, whitetext">
          <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=dateUp&amp;view=tab

           <b>
            Year
           </b>
          </a>
         </p>
        </td>
        <!-- <td><p class="normal, whitetext"><b>Area</b></p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Abalone">
             <img border="1" src="assets/MLimages/SmallLarge1.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Abalone">
               Abalone
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Predict the age of abalone from physical measurements </p><
/td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
```

```
        <td>
         <p class="normal">
          Categorical, Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          4177
         </p>
        </td>
        <td>
         <p class="normal">
          8
         </p>
        </td>
        <td>
         <p class="normal">
          1995
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Adult">
             <img border="1" src="assets/MLimages/SmallLarge2.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Adult">
               Adult
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Predict whether income exceeds $50K/yr based on census data.  Als
o known as "Census Income" dataset. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical, Integer
         </p>
        </td>
        <td>
         <p class="normal">
          48842
         </p>
        </td>
        <td>
         <p class="normal">
          14
         </p>
        </td>
        <td>
         <p class="normal">
          1996
         </p>
        </td>
        <!-- <td><p class="normal">Social </p></td> -->
       </tr>
```

```
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Annealing">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Annealing">
              Annealing
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Steel annealing data </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         798
        </p>
       </td>
       <td>
        <p class="normal">
         38
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <!-- <td><p class="normal">Physical </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Anonymous+Microsoft+Web+Data">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Anonymous+Microsoft+Web+Data">
              Anonymous Microsoft Web Data
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Log of anonymous users of www.microsoft.com; predict areas of the
web site a user visited based on data on other areas the user visited. </p></td> -->
       <td>
```

```
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         Recommender-Systems
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical
        </p>
       </td>
       <td>
        <p class="normal">
         37711
        </p>
       </td>
       <td>
        <p class="normal">
         294
        </p>
       </td>
       <td>
        <p class="normal">
         1998
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Arrhythmia">
            <img border="1" src="assets/MLimages/SmallLarge5.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Arrhythmia">
              Arrhythmia
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Distinguish between the presence and absence of cardiac
arrhythmia and classify it in one of the 16 groups. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         452
        </p>
       </td>
       <td>
        <p class="normal">
         279
        </p>
```

```
      </td>
      <td>
       <p class="normal">
        1998
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Artificial+Characters">
           <img border="1" src="assets/MLimages/SmallLarge6.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Artificial+Characters">
             Artificial Characters
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Dataset artificially generated by using first order theory which
describes structure of ten capital letters of English alphabet </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical, Integer, Real
       </p>
      </td>
      <td>
       <p class="normal">
        6000
       </p>
      </td>
      <td>
       <p class="normal">
        7
       </p>
      </td>
      <td>
       <p class="normal">
        1992
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Audiology+%28Original%29">
           <img border="1" src="assets/MLimages/SmallLarge7.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Audiology+%28Original%29">
             Audiology (Original)
```

```
            Audiology (Original)
           </a>
          </b>
         </p>
        </td>
       </tr>
      </table>
     </td>
     <!-- <td><p class="normal">Nominal audiology dataset from Baylor </p></td> -->
     <td>
      <p class="normal">
       Multivariate
      </p>
     </td>
     <td>
      <p class="normal">
       Classification
      </p>
     </td>
     <td>
      <p class="normal">
       Categorical
      </p>
     </td>
     <td>
      <p class="normal">
       226
      </p>
     </td>
     <td>
      <p class="normal">
      </p>
     </td>
     <td>
      <p class="normal">
       1987
      </p>
     </td>
     <!-- <td><p class="normal">Life </p></td> -->
    </tr>
    <tr bgcolor="DDEEFF">
     <td>
      <table>
       <tr>
        <td>
         <a href="datasets/Audiology+%28Standardized%29">
          <img border="1" src="assets/MLimages/SmallLarge7.jpg"/>
         </a>
        </td>
        <td>
         <p class="normal">
          <b>
           <a href="datasets/Audiology+%28Standardized%29">
            Audiology (Standardized)
           </a>
          </b>
         </p>
        </td>
       </tr>
      </table>
     </td>
     <!-- <td><p class="normal">Standardized version of the original audiology database 
</p></td> -->
     <td>
      <p class="normal">
       Multivariate
      </p>
     </td>
     <td>
      <p class="normal">
       Classification
      </p>
     </td>
     <td>
      <p class="normal">
       Categorical
      </p>
     </td>
```

```
      <td>
       <p class="normal">
        226
       </p>
      </td>
      <td>
       <p class="normal">
        69
       </p>
      </td>
      <td>
       <p class="normal">
        1992
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Auto+MPG">
           <img border="1" src="assets/MLimages/SmallLarge9.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Auto+MPG">
             Auto MPG
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Revised from CMU StatLib library, data concerns city-cycle fuel
consumption </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical, Real
       </p>
      </td>
      <td>
       <p class="normal">
        398
       </p>
      </td>
      <td>
       <p class="normal">
        8
       </p>
      </td>
      <td>
       <p class="normal">
        1993
       </p>
      </td>
      <!-- <td><p class="normal">Other </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
```

```
          <td>
           <a href="datasets/Automobile">
            <img border="1" src="assets/MLimages/SmallLarge10.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Automobile">
              Automobile
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">From 1985 Ward's Automotive Yearbook </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         205
        </p>
       </td>
       <td>
        <p class="normal">
         26
        </p>
       </td>
       <td>
        <p class="normal">
         1987
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Badges">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Badges">
              Badges
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Badges labeled with a "+" or "-" as a function of a person's
name </p></td> -->
       <td>
        <p class="normal">
         Univariate, Text
        </p>
       </td>
```

```
</td>
<td>
 <p class="normal">
  Classification
 </p>
</td>
<td>
 <p class="normal">
 </p>
</td>
<td>
 <p class="normal">
  294
 </p>
</td>
<td>
 <p class="normal">
  1
 </p>
</td>
<td>
 <p class="normal">
  1994
 </p>
</td>
<!-- <td><p class="normal">Other </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Balance+Scale">
      <img border="1" src="assets/MLimages/SmallLarge12.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Balance+Scale">
        Balance Scale
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">Balance scale weight & distance database </p></td> -->
 <td>
  <p class="normal">
   Multivariate
  </p>
 </td>
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Categorical
  </p>
 </td>
 <td>
  <p class="normal">
   625
  </p>
 </td>
 <td>
  <p class="normal">
   4
  </p>
 </td>
 <td>
  <p class="normal">
   1994
  </p>
```

```
        </p>
       </td>
       <!-- <td><p class="normal">Social </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Balloons">
            <img border="1" src="assets/MLimages/SmallLarge13.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Balloons">
              Balloons
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Data previously used in cognitive psychology experiment; 4 data
sets represent different conditions of an experiment </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical
        </p>
       </td>
       <td>
        <p class="normal">
         16
        </p>
       </td>
       <td>
        <p class="normal">
         4
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <!-- <td><p class="normal">Social </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Breast+Cancer">
            <img border="1" src="assets/MLimages/SmallLarge14.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Breast+Cancer">
              Breast Cancer
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
```

```
 </table>
</td>
<!-- <td><p class="normal">Breast Cancer Data (Restricted Access) </p></td> -->
<td>
 <p class="normal">
  Multivariate
 </p>
</td>
<td>
 <p class="normal">
  Classification
 </p>
</td>
<td>
 <p class="normal">
  Categorical
 </p>
</td>
<td>
 <p class="normal">
  286
 </p>
</td>
<td>
 <p class="normal">
  9
 </p>
</td>
<td>
 <p class="normal">
  1988
 </p>
</td>
<!-- <td><p class="normal">Life </p></td> -->
</tr>
<tr>
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Breast+Cancer+Wisconsin+%28Original%29">
      <img border="1" src="assets/MLimages/SmallLarge14.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Breast+Cancer+Wisconsin+%28Original%29">
        Breast Cancer Wisconsin (Original)
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">Original Wisconsin Breast Cancer Database </p></td> -->
 <td>
  <p class="normal">
   Multivariate
  </p>
 </td>
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Integer
  </p>
 </td>
 <td>
  <p class="normal">
   699
  </p>
 </td>
```

```
 <td>
  <p class="normal">
   10
  </p>
 </td>
 <td>
  <p class="normal">
   1992
  </p>
 </td>
 <!-- <td><p class="normal">Life </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Breast+Cancer+Wisconsin+%28Prognostic%29">
      <img border="1" src="assets/MLimages/SmallLarge14.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Breast+Cancer+Wisconsin+%28Prognostic%29">
        Breast Cancer Wisconsin (Prognostic)
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">Prognostic Wisconsin Breast Cancer Database </p></td> -->
 <td>
  <p class="normal">
   Multivariate
  </p>
 </td>
 <td>
  <p class="normal">
   Classification, Regression
  </p>
 </td>
 <td>
  <p class="normal">
   Real
  </p>
 </td>
 <td>
  <p class="normal">
   198
  </p>
 </td>
 <td>
  <p class="normal">
   34
  </p>
 </td>
 <td>
  <p class="normal">
   1995
  </p>
 </td>
 <!-- <td><p class="normal">Life </p></td> -->
</tr>
<tr>
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29">
      <img border="1" src="assets/MLimages/SmallLarge14.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
```

```
            <b>
             <a href="datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29">
              Breast Cancer Wisconsin (Diagnostic)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Diagnostic Wisconsin Breast Cancer Database </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         569
        </p>
       </td>
       <td>
        <p class="normal">
         32
        </p>
       </td>
       <td>
        <p class="normal">
         1995
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Pittsburgh+Bridges">
            <img border="1" src="assets/MLimages/SmallLarge18.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Pittsburgh+Bridges">
              Pittsburgh Bridges
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Bridges database that has original and numeric-discretized
datasets </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
```

```
        <p class="normal">
         Categorical, Integer
        </p>
       </td>
       <td>
        <p class="normal">
         108
        </p>
       </td>
       <td>
        <p class="normal">
         13
        </p>
       </td>
       <td>
        <p class="normal">
         1990
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Car+Evaluation">
            <img border="1" src="assets/MLimages/SmallLarge19.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Car+Evaluation">
              Car Evaluation
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Derived from simple hierarchical decision model, this database m
ay be useful for testing constructive induction and structure discovery methods. </p></td> --
>
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical
        </p>
       </td>
       <td>
        <p class="normal">
         1728
        </p>
       </td>
       <td>
        <p class="normal">
         6
        </p>
       </td>
       <td>
        <p class="normal">
         1997
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
```

```
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Census+Income">
            <img border="1" src="assets/MLimages/SmallLarge2.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Census+Income">
              Census Income
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Predict whether income exceeds $50K/yr based on census data.  Als
o known as "Adult" dataset. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer
        </p>
       </td>
       <td>
        <p class="normal">
         48842
        </p>
       </td>
       <td>
        <p class="normal">
         14
        </p>
       </td>
       <td>
        <p class="normal">
         1996
        </p>
       </td>
       <!-- <td><p class="normal">Social </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Chess+%28King-Rook+vs.+King-Knight%29">
            <img border="1" src="assets/MLimages/SmallLarge24.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Chess+%28King-Rook+vs.+King-Knight%29">
              Chess (King-Rook vs. King-Knight)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Knight Pin Chess End-Game Database Creator </p></td> -->
```

```
        <td>
         <p class="normal">
          Multivariate, Data-Generator
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical, Integer
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          22
         </p>
        </td>
        <td>
         <p class="normal">
          1988
         </p>
        </td>
        <!-- <td><p class="normal">Game </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Chess+%28King-Rook+vs.+King-Pawn%29">
             <img border="1" src="assets/MLimages/SmallLarge24.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Chess+%28King-Rook+vs.+King-Pawn%29">
               Chess (King-Rook vs. King-Pawn)
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">King+Rook versus King+Pawn on a7 (usually abbreviated
KRKPA7). </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical
         </p>
        </td>
        <td>
         <p class="normal">
          3196
         </p>
        </td>
        <td>
         <p class="normal">
          36
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         1989
        </p>
       </td>
       <!-- <td><p class="normal">Game </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Chess+%28King-Rook+vs.+King%29">
            <img border="1" src="assets/MLimages/SmallLarge24.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Chess+%28King-Rook+vs.+King%29">
              Chess (King-Rook vs. King)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Chess Endgame Database for White King and Rook against Black
King (KRK). </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer
        </p>
       </td>
       <td>
        <p class="normal">
         28056
        </p>
       </td>
       <td>
        <p class="normal">
         6
        </p>
       </td>
       <td>
        <p class="normal">
         1994
        </p>
       </td>
       <!-- <td><p class="normal">Game </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Chess+%28Domain+Theories%29">
            <img border="1" src="assets/MLimages/SmallLarge24.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Chess+%28Domain+Theories%29">
```

```
              Chess (Domain Theories)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">6 different domain theories for generating legal moves of
chess </p></td> -->
       <td>
        <p class="normal">
         Domain-Theory
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <!-- <td><p class="normal">Game </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Bach+Chorales">
            <img border="1" src="assets/MLimages/SmallLarge25.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Bach+Chorales">
              Bach Chorales
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Time-series data based on chorales; challenge is to learn generat
ive grammar; data in Lisp </p></td> -->
       <td>
        <p class="normal">
         Univariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer
        </p>
       </td>
       <td>
        <p class="normal">
         100
```

```
  </p>
 </td>
 <td>
  <p class="normal">
   6
  </p>
 </td>
 <td>
  <p class="normal">
  </p>
 </td>
 <!-- <td><p class="normal">Other </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Connect-4">
      <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Connect-4">
        Connect-4
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">Contains connect-4 positions </p></td> -->
 <td>
  <p class="normal">
   Multivariate, Spatial
  </p>
 </td>
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Categorical
  </p>
 </td>
 <td>
  <p class="normal">
   67557
  </p>
 </td>
 <td>
  <p class="normal">
   42
  </p>
 </td>
 <td>
  <p class="normal">
   1995
  </p>
 </td>
 <!-- <td><p class="normal">Game </p></td> -->
</tr>
<tr>
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Credit+Approval">
      <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
     </a>
    </td>
    <td>
```

```
            <p class="normal">
             <b>
              <a href="datasets/Credit+Approval">
               Credit Approval
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This data concerns credit card applications; good mix of
attributes </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical, Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          690
         </p>
        </td>
        <td>
         <p class="normal">
          15
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <!-- <td><p class="normal">Financial </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Japanese+Credit+Screening">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Japanese+Credit+Screening">
               Japanese Credit Screening
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Includes domain theory (generated by talking to Japanese domain
experts); data in Lisp </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Domain-Theory
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
```

```
        <td>
         <p class="normal">
          Categorical, Real, Integer
         </p>
        </td>
        <td>
         <p class="normal">
          125
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          1992
         </p>
        </td>
        <!-- <td><p class="normal">Financial </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Computer+Hardware">
             <img border="1" src="assets/MLimages/SmallLarge29.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Computer+Hardware">
               Computer Hardware
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Relative CPU Performance Data, described in terms of its cycle ti
me, memory size, etc. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          209
         </p>
        </td>
        <td>
         <p class="normal">
          9
         </p>
        </td>
        <td>
         <p class="normal">
          1987
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
```

```
<td>
 <table>
  <tr>
   <td>
    <a href="datasets/Contraceptive+Method+Choice">
     <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
    </a>
   </td>
   <td>
    <p class="normal">
     <b>
      <a href="datasets/Contraceptive+Method+Choice">
       Contraceptive Method Choice
      </a>
     </b>
    </p>
   </td>
  </tr>
 </table>
</td>
<!-- <td><p class="normal">Dataset is a subset of the 1987 National Indonesia Contraceptive
Prevalence Survey. </p></td> -->
<td>
 <p class="normal">
  Multivariate
 </p>
</td>
<td>
 <p class="normal">
  Classification
 </p>
</td>
<td>
 <p class="normal">
  Categorical, Integer
 </p>
</td>
<td>
 <p class="normal">
  1473
 </p>
</td>
<td>
 <p class="normal">
  9
 </p>
</td>
<td>
 <p class="normal">
  1997
 </p>
</td>
<!-- <td><p class="normal">Life </p></td> -->
</tr>
<tr>
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Covertype">
      <img border="1" src="assets/MLimages/SmallLarge31.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Covertype">
        Covertype
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">Forest CoverType dataset </p></td> -->
 <td>
```

```
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical, Integer
         </p>
        </td>
        <td>
         <p class="normal">
          581012
         </p>
        </td>
        <td>
         <p class="normal">
          54
         </p>
        </td>
        <td>
         <p class="normal">
          1998
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Cylinder+Bands">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Cylinder+Bands">
               Cylinder Bands
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Used in decision tree induction for mitigating process delays
known as "cylinder bands" in rotogravure printing </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical, Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          512
         </p>
        </td>
        <td>
         <p class="normal">
          39
```

```
</p>
</td>
<td>
<p class="normal">
1995
</p>
</td>
<!-- <td><p class="normal">Physical </p></td> -->
</tr>
<tr>
<td>
<table>
<tr>
<td>
<a href="datasets/Dermatology">
<img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>
</td>
<td>
<p class="normal">
<b>
<a href="datasets/Dermatology">
Dermatology
</a>
</b>
</p>
</td>
</tr>
</table>
</td>
<!-- <td><p class="normal">Aim for this dataset is to determine the type of Eryhemato-Squamo
us Disease. </p></td> -->
<td>
<p class="normal">
Multivariate
</p>
</td>
<td>
<p class="normal">
Classification
</p>
</td>
<td>
<p class="normal">
Categorical, Integer
</p>
</td>
<td>
<p class="normal">
366
</p>
</td>
<td>
<p class="normal">
33
</p>
</td>
<td>
<p class="normal">
1998
</p>
</td>
<!-- <td><p class="normal">Life </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
<td>
<table>
<tr>
<td>
<a href="datasets/Diabetes">
<img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>
</td>
<td>
<p class="normal">
<b>
<a href="datasets/Diabetes">
```

```
            <a href="datasets/Diabetes">
             Diabetes
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This diabetes dataset is from AIM '94 </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical, Integer
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        20
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/DGP2+-+The+Second+Data+Generation+Program">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/DGP2+-+The+Second+Data+Generation+Program">
             DGP2 - The Second Data Generation Program
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Generates application domains based on specific parameters,
number of features, and proportion of positive to negative examples </p></td> -->
      <td>
       <p class="normal">
        Data-Generator
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
```

```
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Document+Understanding">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Document+Understanding">
              Document Understanding
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Five concepts, expressed as predicates, to be learned </p>
</td> -->
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         1994
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/EBL+Domain+Theories">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/EBL+Domain+Theories">
              EBL Domain Theories
             </a>
```

```
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Assorted small-scale domain theories </p></td> -->
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Echocardiogram">
           <img border="1" src="assets/MLimages/SmallLarge38.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Echocardiogram">
             Echocardiogram
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Data for classifying if patients will survive for at least one y
ear after a heart attack </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical, Integer, Real
       </p>
      </td>
      <td>
       <p class="normal">
        132
       </p>
      </td>
```

```
 <td>
  <p class="normal">
   12
  </p>
 </td>
 <td>
  <p class="normal">
   1989
  </p>
 </td>
 <!-- <td><p class="normal">Life </p></td> -->
</tr>
<tr>
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Ecoli">
      <img border="1" src="assets/MLimages/SmallLarge120.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Ecoli">
        Ecoli
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">This data contains protein localization sites </p></td> -->
 <td>
  <p class="normal">
   Multivariate
  </p>
 </td>
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Real
  </p>
 </td>
 <td>
  <p class="normal">
   336
  </p>
 </td>
 <td>
  <p class="normal">
   8
  </p>
 </td>
 <td>
  <p class="normal">
   1996
  </p>
 </td>
 <!-- <td><p class="normal">Life </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Flags">
      <img border="1" src="assets/MLimages/SmallLarge40.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
```

```
            <p>
             <a href="datasets/Flags">
              Flags
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">From Collins Gem Guide to Flags, 1986 </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer
        </p>
       </td>
       <td>
        <p class="normal">
         194
        </p>
       </td>
       <td>
        <p class="normal">
         30
        </p>
       </td>
       <td>
        <p class="normal">
         1990
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Function+Finding">
            <img border="1" src="assets/MLimages/SmallLarge41.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Function+Finding">
              Function Finding
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Cases collected mostly from investigations in physical science;
intention is to evaluate function-finding algorithms </p></td> -->
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         Function-Learning
        </p>
       </td>
       <td>
        <p class="normal">
```

```
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          352
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          1990
         </p>
        </td>
        <!-- <td><p class="normal">Physical </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Glass+Identification">
             <img border="1" src="assets/MLimages/SmallLarge42.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Glass+Identification">
               Glass Identification
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">From USA Forensic Science Service; 6 types of glass; defined in
terms of their oxide content (i.e. Na, Fe, K, etc) </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          214
         </p>
        </td>
        <td>
         <p class="normal">
          10
         </p>
        </td>
        <td>
         <p class="normal">
          1987
         </p>
        </td>
        <!-- <td><p class="normal">Physical </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
```

```
          <tr>
           <td>
            <a href="datasets/Haberman%27s+Survival">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Haberman%27s+Survival">
               Haberman's Survival
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Dataset contains cases from study conducted on the survival of
patients who had undergone surgery for breast cancer </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          306
         </p>
        </td>
        <td>
         <p class="normal">
          3
         </p>
        </td>
        <td>
         <p class="normal">
          1999
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Hayes-Roth">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Hayes-Roth">
               Hayes-Roth
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Topic: human subjects study </p></td> -->
        <td>
         <p class="normal">
          Multivariate
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical
        </p>
       </td>
       <td>
        <p class="normal">
         160
        </p>
       </td>
       <td>
        <p class="normal">
         5
        </p>
       </td>
       <td>
        <p class="normal">
         1989
        </p>
       </td>
       <!-- <td><p class="normal">Social </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Heart+Disease">
            <img border="1" src="assets/MLimages/SmallLarge45.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Heart+Disease">
              Heart Disease
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">4 databases: Cleveland, Hungary, Switzerland, and the VA Long
Beach </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         303
        </p>
       </td>
       <td>
        <p class="normal">
         75
        </p>
       </td>
```

```
        <td>
         <p class="normal">
          1988
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Hepatitis">
             <img border="1" src="assets/MLimages/SmallLarge46.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Hepatitis">
               Hepatitis
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">From G.Gong: CMU; Mostly Boolean or numeric-valued attribute
types; Includes cost data (donated by Peter Turney) </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical, Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          155
         </p>
        </td>
        <td>
         <p class="normal">
          19
         </p>
        </td>
        <td>
         <p class="normal">
          1988
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Horse+Colic">
             <img border="1" src="assets/MLimages/SmallLarge47.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Horse+Colic">
               Horse Colic
              </a>
```

```
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Well documented attributes; 368 instances with 28 attributes
(continuous, discrete, and nominal); 30% missing values </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         368
        </p>
       </td>
       <td>
        <p class="normal">
         27
        </p>
       </td>
       <td>
        <p class="normal">
         1989
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/ICU">
            <img border="1" src="assets/MLimages/SmallLarge49.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/ICU">
              ICU
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Data set prepared for the use of participants for the 1994 AAAI
Spring Symposium on Artificial Intelligence in Medicine. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
```

```
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Image+Segmentation">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Image+Segmentation">
               Image Segmentation
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Image data described by high-level numeric-valued attributes, 7
classes </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          2310
         </p>
        </td>
        <td>
         <p class="normal">
          19
         </p>
        </td>
        <td>
         <p class="normal">
          1990
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Internet+Advertisements">
             <img border="1" src="assets/MLimages/SmallLarge51.jpg"/>
            </a>
```

```
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Internet+Advertisements">
              Internet Advertisements
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset represents a set of possible advertisements on
Internet pages. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         3279
        </p>
       </td>
       <td>
        <p class="normal">
         1558
        </p>
       </td>
       <td>
        <p class="normal">
         1998
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Ionosphere">
            <img border="1" src="assets/MLimages/SmallLarge52.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Ionosphere">
              Ionosphere
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Classification of radar returns from the ionosphere </p>
</td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
```

```
    -
    Classification
   </p>
  </td>
  <td>
   <p class="normal">
    Integer, Real
   </p>
  </td>
  <td>
   <p class="normal">
    351
   </p>
  </td>
  <td>
   <p class="normal">
    34
   </p>
  </td>
  <td>
   <p class="normal">
    1989
   </p>
  </td>
  <!-- <td><p class="normal">Physical </p></td> -->
 </tr>
 <tr bgcolor="DDEEFF">
  <td>
   <table>
    <tr>
     <td>
      <a href="datasets/Iris">
       <img border="1" src="assets/MLimages/SmallLarge53.jpg"/>
      </a>
     </td>
     <td>
      <p class="normal">
       <b>
        <a href="datasets/Iris">
         Iris
        </a>
       </b>
      </p>
     </td>
    </tr>
   </table>
  </td>
  <!-- <td><p class="normal">Famous database; from Fisher, 1936 </p></td> -->
  <td>
   <p class="normal">
    Multivariate
   </p>
  </td>
  <td>
   <p class="normal">
    Classification
   </p>
  </td>
  <td>
   <p class="normal">
    Real
   </p>
  </td>
  <td>
   <p class="normal">
    150
   </p>
  </td>
  <td>
   <p class="normal">
    4
   </p>
  </td>
  <td>
   <p class="normal">
    1988
   </p>
  </td>
```

```
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/ISOLET">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/ISOLET">
             ISOLET
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal"> Goal: Predict which letter-name was spoken--a simple classificat
ion task. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        7797
       </p>
      </td>
      <td>
       <p class="normal">
        617
       </p>
      </td>
      <td>
       <p class="normal">
        1994
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Kinship">
           <img border="1" src="assets/MLimages/SmallLarge55.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Kinship">
             Kinship
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
```

```
 </td>
 <!-- <td><p class="normal">Relational dataset </p></td> -->
 <td>
  <p class="normal">
   Relational
  </p>
 </td>
 <td>
  <p class="normal">
   Relational-Learning
  </p>
 </td>
 <td>
  <p class="normal">
   Categorical
  </p>
 </td>
 <td>
  <p class="normal">
   104
  </p>
 </td>
 <td>
  <p class="normal">
   12
  </p>
 </td>
 <td>
  <p class="normal">
   1990
  </p>
 </td>
 <!-- <td><p class="normal">Social </p></td> -->
</tr>
<tr>
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Labor+Relations">
      <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Labor+Relations">
        Labor Relations
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">From Collective Bargaining Review </p></td> -->
 <td>
  <p class="normal">
   Multivariate
  </p>
 </td>
 <td>
  <p class="normal">
  </p>
 </td>
 <td>
  <p class="normal">
   Categorical, Integer, Real
  </p>
 </td>
 <td>
  <p class="normal">
   57
  </p>
 </td>
 <td>
  <p class="normal">
```

```
        <p class="normal">
         16
        </p>
       </td>
       <td>
        <p class="normal">
         1988
        </p>
       </td>
       <!-- <td><p class="normal">Social </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/LED+Display+Domain">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/LED+Display+Domain">
              LED Display Domain
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">From Classification and Regression Trees book; We provide here 2
C programs for generating sample databases </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Data-Generator
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         7
        </p>
       </td>
       <td>
        <p class="normal">
         1988
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Lenses">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Lenses">
```

```
            <a href="datasets/Lenses">
             Lenses
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Database for fitting contact lenses </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical
       </p>
      </td>
      <td>
       <p class="normal">
        24
       </p>
      </td>
      <td>
       <p class="normal">
        4
       </p>
      </td>
      <td>
       <p class="normal">
        1990
       </p>
      </td>
      <!-- <td><p class="normal">Other </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Letter+Recognition">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Letter+Recognition">
             Letter Recognition
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Database of character image features; try to identify the
letter </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
```

```
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          20000
         </p>
        </td>
        <td>
         <p class="normal">
          16
         </p>
        </td>
        <td>
         <p class="normal">
          1991
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Liver+Disorders">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Liver+Disorders">
               Liver Disorders
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">BUPA Medical Research Ltd. database donated by Richard S.
Forsyth </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical, Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          345
         </p>
        </td>
        <td>
         <p class="normal">
          7
         </p>
        </td>
        <td>
         <p class="normal">
          1990
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
```

```
    <tr>
     <td>
      <a href="datasets/Logic+Theorist">
       <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
      </a>
     </td>
     <td>
      <p class="normal">
       <b>
        <a href="datasets/Logic+Theorist">
         Logic Theorist
        </a>
       </b>
      </p>
     </td>
    </tr>
   </table>
  </td>
  <!-- <td><p class="normal">All code for Logic Theorist </p></td> -->
  <td>
   <p class="normal">
    Domain-Theory
   </p>
  </td>
  <td>
   <p class="normal">
   </p>
  </td>
  <td>
   <p class="normal">
   </p>
  </td>
  <td>
   <p class="normal">
   </p>
  </td>
  <td>
   <p class="normal">
   </p>
  </td>
  <td>
   <p class="normal">
   </p>
  </td>
  <!-- <td><p class="normal">Computer </p></td> -->
 </tr>
 <tr>
  <td>
   <table>
    <tr>
     <td>
      <a href="datasets/Lung+Cancer">
       <img border="1" src="assets/MLimages/SmallLarge62.jpg"/>
      </a>
     </td>
     <td>
      <p class="normal">
       <b>
        <a href="datasets/Lung+Cancer">
         Lung Cancer
        </a>
       </b>
      </p>
     </td>
    </tr>
   </table>
  </td>
  <!-- <td><p class="normal">Lung cancer data; no attribute definitions </p></td> -->
  <td>
   <p class="normal">
    Multivariate
   </p>
  </td>
  <td>
   <p class="normal">
    Classification
   </p>
  </td>
```

```
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
       <p class="normal">
        32
       </p>
      </td>
      <td>
       <p class="normal">
        56
       </p>
      </td>
      <td>
       <p class="normal">
        1992
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Lymphography">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Lymphography">
             Lymphography
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This lymphography domain was obtained from the University
Medical Centre, Institute of Oncology, Ljubljana, Yugoslavia.  (Restricted access) </p></td>
-->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical
       </p>
      </td>
      <td>
       <p class="normal">
        148
       </p>
      </td>
      <td>
       <p class="normal">
        18
       </p>
      </td>
      <td>
       <p class="normal">
        1988
       </p>
      </td>
```

```
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Mechanical+Analysis">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Mechanical+Analysis">
             Mechanical Analysis
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Fault diagnosis problem of electromechanical devices; also PUMPS
DATA SET is newer version with domain theory and results </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical, Integer, Real
       </p>
      </td>
      <td>
       <p class="normal">
        209
       </p>
      </td>
      <td>
       <p class="normal">
        8
       </p>
      </td>
      <td>
       <p class="normal">
        1990
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Meta-data">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Meta-data">
             Meta-data
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
```

```
        </td>
        <!-- <td><p class="normal">Meta-Data was used in order to give advice about which classifica
tion method is appropriate for a particular dataset (taken from results of Statlog project). 
</p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical, Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          528
         </p>
        </td>
        <td>
         <p class="normal">
          22
         </p>
        </td>
        <td>
         <p class="normal">
          1996
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Mobile+Robots">
             <img border="1" src="assets/MLimages/SmallLarge66.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Mobile+Robots">
               Mobile Robots
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Learning concepts from sensor data of a mobile robot; set of
data sets </p></td> -->
        <td>
         <p class="normal">
          Domain-Theory
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical, Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
```

```
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          1995
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Molecular+Biology+%28Promoter+Gene+Sequences%29">
             <img border="1" src="assets/MLimages/SmallLarge67.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Molecular+Biology+%28Promoter+Gene+Sequences%29">
               Molecular Biology (Promoter Gene Sequences)
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">E. Coli promoter gene sequences (DNA) with partial domain
theory </p></td> -->
        <td>
         <p class="normal">
          Sequential, Domain-Theory
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical
         </p>
        </td>
        <td>
         <p class="normal">
          106
         </p>
        </td>
        <td>
         <p class="normal">
          58
         </p>
        </td>
        <td>
         <p class="normal">
          1990
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Molecular+Biology+%28Protein+Secondary+Structure%29">
             <img border="1" src="assets/MLimages/SmallLarge67.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
```

```
            <b>
             <a href="datasets/Molecular+Biology+%28Protein+Secondary+Structure%29">
              Molecular Biology (Protein Secondary Structure)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">From CMU connectionist bench repository; Classifies secondary
structure of certain globular proteins </p></td> -->
       <td>
        <p class="normal">
         Sequential
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical
        </p>
       </td>
       <td>
        <p class="normal">
         128
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Molecular+Biology+%28Splice-junction+Gene+Sequences%29">
            <img border="1" src="assets/MLimages/SmallLarge67.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Molecular+Biology+%28Splice-junction+Gene+Sequences%29">
              Molecular Biology (Splice-junction Gene Sequences)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Primate splice-junction gene sequences (DNA) with associated impe
rfect domain theory </p></td> -->
       <td>
        <p class="normal">
         Sequential, Domain-Theory
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
```

```
          Categorical
         </p>
        </td>
        <td>
         <p class="normal">
          3190
         </p>
        </td>
        <td>
         <p class="normal">
          61
         </p>
        </td>
        <td>
         <p class="normal">
          1992
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/MONK%27s+Problems">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/MONK%27s+Problems">
               MONK's Problems
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">A set of three artificial domains over the same attribute space;
Used to test a wide range of induction algorithms </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical
         </p>
        </td>
        <td>
         <p class="normal">
          432
         </p>
        </td>
        <td>
         <p class="normal">
          7
         </p>
        </td>
        <td>
         <p class="normal">
          1992
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
```

```
        <table>
         <tr>
          <td>
           <a href="datasets/Moral+Reasoner">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Moral+Reasoner">
              Moral Reasoner
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Horn-clause model that qualitatively simulates moral reasoning; T
heory includes negated literals </p></td> -->
       <td>
        <p class="normal">
         Domain-Theory
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         202
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         1994
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Multiple+Features">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Multiple+Features">
              Multiple Features
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset consists of features of handwritten numerals (`0'--
`9') extracted from a collection of Dutch utility maps </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
```

```
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          2000
         </p>
        </td>
        <td>
         <p class="normal">
          649
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Mushroom">
             <img border="1" src="assets/MLimages/SmallLarge73.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Mushroom">
               Mushroom
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">From Audobon Society Field Guide; mushrooms described in terms of
physical characteristics; classification: poisonous or edible </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical
         </p>
        </td>
        <td>
         <p class="normal">
          8124
         </p>
        </td>
        <td>
         <p class="normal">
          22
         </p>
        </td>
        <td>
         <p class="normal">
```

```
          1987
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Musk+%28Version+1%29">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Musk+%28Version+1%29">
               Musk (Version 1)
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The goal is to learn to predict whether new molecules will be mus
ks or non-musks </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          476
         </p>
        </td>
        <td>
         <p class="normal">
          168
         </p>
        </td>
        <td>
         <p class="normal">
          1994
         </p>
        </td>
        <!-- <td><p class="normal">Physical </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Musk+%28Version+2%29">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Musk+%28Version+2%29">
               Musk (Version 2)
              </a>
             </b>
            </p>
```

```
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The goal is to learn to predict whether new molecules will be mus
ks or non-musks </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         6598
        </p>
       </td>
       <td>
        <p class="normal">
         168
        </p>
       </td>
       <td>
        <p class="normal">
         1994
        </p>
       </td>
       <!-- <td><p class="normal">Physical </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Nursery">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Nursery">
              Nursery
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal"> Nursery Database was derived from a hierarchical decision model
originally developed to rank applications for nursery schools. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical
        </p>
       </td>
       <td>
```

```
        <p class="normal">
         12960
        </p>
       </td>
       <td>
        <p class="normal">
         8
        </p>
       </td>
       <td>
        <p class="normal">
         1997
        </p>
       </td>
       <!-- <td><p class="normal">Social </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Othello+Domain+Theory">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Othello+Domain+Theory">
              Othello Domain Theory
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Used in research to generate features for an inductive learning
system </p></td> -->
       <td>
        <p class="normal">
         Domain-Theory
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         1991
        </p>
       </td>
       <!-- <td><p class="normal">Game </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Page+Blocks+Classification">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
```

```
          <p class="normal">
           <b>
            <a href="datasets/Page+Blocks+Classification">
             Page Blocks Classification
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The problem consists of classifying all the blocks of the page l
ayout of a document that has been detected by a segmentation process. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer, Real
       </p>
      </td>
      <td>
       <p class="normal">
        5473
       </p>
      </td>
      <td>
       <p class="normal">
        10
       </p>
      </td>
      <td>
       <p class="normal">
        1995
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Optical+Recognition+of+Handwritten+Digits">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Optical+Recognition+of+Handwritten+Digits">
             Optical Recognition of Handwritten Digits
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Two versions of this database available; see folder </p>
</td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
```

```
  </td>
  <td>
   <p class="normal">
    Integer
   </p>
  </td>
  <td>
   <p class="normal">
    5620
   </p>
  </td>
  <td>
   <p class="normal">
    64
   </p>
  </td>
  <td>
   <p class="normal">
    1998
   </p>
  </td>
  <!-- <td><p class="normal">Computer </p></td> -->
 </tr>
 <tr>
  <td>
   <table>
    <tr>
     <td>
      <a href="datasets/Pen-Based+Recognition+of+Handwritten+Digits">
       <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
      </a>
     </td>
     <td>
      <p class="normal">
       <b>
        <a href="datasets/Pen-Based+Recognition+of+Handwritten+Digits">
         Pen-Based Recognition of Handwritten Digits
        </a>
       </b>
      </p>
     </td>
    </tr>
   </table>
  </td>
  <!-- <td><p class="normal">Digit database of 250 samples from 44 writers </p></td> -->
  <td>
   <p class="normal">
    Multivariate
   </p>
  </td>
  <td>
   <p class="normal">
    Classification
   </p>
  </td>
  <td>
   <p class="normal">
    Integer
   </p>
  </td>
  <td>
   <p class="normal">
    10992
   </p>
  </td>
  <td>
   <p class="normal">
    16
   </p>
  </td>
  <td>
   <p class="normal">
    1998
   </p>
  </td>
  <!-- <td><p class="normal">Computer </p></td> -->
 </tr>
```

```
<tr bgcolor="DDEEFF">
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Post-Operative+Patient">
      <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Post-Operative+Patient">
        Post-Operative Patient
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">Dataset of patient features </p></td> -->
 <td>
  <p class="normal">
   Multivariate
  </p>
 </td>
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Categorical, Integer
  </p>
 </td>
 <td>
  <p class="normal">
   90
  </p>
 </td>
 <td>
  <p class="normal">
   8
  </p>
 </td>
 <td>
  <p class="normal">
   1993
  </p>
 </td>
 <!-- <td><p class="normal">Life </p></td> -->
</tr>
<tr>
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Primary+Tumor">
      <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Primary+Tumor">
        Primary Tumor
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">From Ljubljana Oncology Institute </p></td> -->
 <td>
```

```
<td>
 <p class="normal">
  Multivariate
 </p>
</td>
<td>
 <p class="normal">
  Classification
 </p>
</td>
<td>
 <p class="normal">
  Categorical
 </p>
</td>
<td>
 <p class="normal">
  339
 </p>
</td>
<td>
 <p class="normal">
  17
 </p>
</td>
<td>
 <p class="normal">
  1988
 </p>
</td>
<!-- <td><p class="normal">Life </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
<td>
 <table>
  <tr>
   <td>
    <a href="datasets/Prodigy">
     <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
    </a>
   </td>
   <td>
    <p class="normal">
     <b>
      <a href="datasets/Prodigy">
       Prodigy
      </a>
     </b>
    </p>
   </td>
  </tr>
 </table>
</td>
<!-- <td><p class="normal">Assorted domains like blocksworld, eightpuzzle, and schedworld.&n
bsp;</p></td> -->
<td>
 <p class="normal">
  Domain-Theory
 </p>
</td>
<td>
 <p class="normal">
 </p>
</td>
<td>
 <p class="normal">
 </p>
</td>
<td>
 <p class="normal">
 </p>
</td>
<td>
 <p class="normal">
 </p>
</td>
<td>
 <p class="normal">
```

```
        <p class="normal">
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Qualitative+Structure+Activity+Relationships">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Qualitative+Structure+Activity+Relationships">
              Qualitative Structure Activity Relationships
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Two sets of datasets are given: pyrimidines and triazines <
/p></td> -->
       <td>
        <p class="normal">
         Domain-Theory
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <!-- <td><p class="normal">Physical </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Quadruped+Mammals">
            <img border="1" src="assets/MLimages/SmallLarge86.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Quadruped+Mammals">
              Quadruped Mammals
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal"> The file animals.c is a data generator of structured instances
representing quadruped animals </p></td> -->
```

representing quadruped animals </td> -->
<td>

Multivariate, Data-Generator

</td>
<td>

Classification

</td>
<td>

Real

</td>
<td>

</td>
<td>

72

</td>
<td>

1992

</td>
<!-- <td> Life </td> -->
</tr>
<tr>
<td>
<table>
<tr>
<td>
<a href="datasets/Servo">
<img border="1" src="assets/MLimages/SmallLarge87.jpg"/>
</a>
</td>
<td>

<a href="datasets/Servo">
Servo
</a>

</td>
</tr>
</table>
</td>
<!-- <td> Data was from a simulation of a servo system </td> -->
<td>

Multivariate

</td>
<td>

Regression

</td>
<td>

Categorical, Integer

</td>
<td>

167

</td>
<td>

4

</td>
</tr>
</table>
</td>
</tr>
</table>
</td>
</tr>

```
  </p>
 </td>
 <td>
  <p class="normal">
   1993
  </p>
 </td>
 <!-- <td><p class="normal">Computer </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Shuttle+Landing+Control">
      <img border="1" src="assets/MLimages/SmallLarge92.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Shuttle+Landing+Control">
        Shuttle Landing Control
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">Tiny database; all nominal values </p></td> -->
 <td>
  <p class="normal">
   Multivariate
  </p>
 </td>
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Categorical
  </p>
 </td>
 <td>
  <p class="normal">
   15
  </p>
 </td>
 <td>
  <p class="normal">
   6
  </p>
 </td>
 <td>
  <p class="normal">
   1988
  </p>
 </td>
 <!-- <td><p class="normal">Physical </p></td> -->
</tr>
<tr>
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Solar+Flare">
      <img border="1" src="assets/MLimages/SmallLarge89.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Solar+Flare">
        Solar Flare
```

```
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Each class attribute counts the number of solar flares of a
certain class that occur in a 24 hour period </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical
       </p>
      </td>
      <td>
       <p class="normal">
        1389
       </p>
      </td>
      <td>
       <p class="normal">
        10
       </p>
      </td>
      <td>
       <p class="normal">
        1989
       </p>
      </td>
      <!-- <td><p class="normal">Physical </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Soybean+%28Large%29">
           <img border="1" src="assets/MLimages/SmallLarge90.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Soybean+%28Large%29">
             Soybean (Large)
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Michalski's famous soybean disease database </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical
       </p>
```

```
 </td>
 <td>
  <p class="normal">
   307
  </p>
 </td>
 <td>
  <p class="normal">
   35
  </p>
 </td>
 <td>
  <p class="normal">
   1988
  </p>
 </td>
 <!-- <td><p class="normal">Life </p></td> -->
</tr>
<tr>
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Soybean+%28Small%29">
      <img border="1" src="assets/MLimages/SmallLarge90.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Soybean+%28Small%29">
        Soybean (Small)
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">Michalski's famous soybean disease database </p></td> -->
 <td>
  <p class="normal">
   Multivariate
  </p>
 </td>
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Categorical
  </p>
 </td>
 <td>
  <p class="normal">
   47
  </p>
 </td>
 <td>
  <p class="normal">
   35
  </p>
 </td>
 <td>
  <p class="normal">
   1987
  </p>
 </td>
 <!-- <td><p class="normal">Life </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
 <td>
  <table>
   <tr>
    <td>
```

```
            <a href="datasets/Challenger+USA+Space+Shuttle+O-Ring">
             <img border="1" src="assets/MLimages/SmallLarge92.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Challenger+USA+Space+Shuttle+O-Ring">
               Challenger USA Space Shuttle O-Ring
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Task: predict the number of O-rings that experience thermal distr
ess on a flight at 31 degrees F given data on the previous 23 shuttle flights </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          23
         </p>
        </td>
        <td>
         <p class="normal">
          4
         </p>
        </td>
        <td>
         <p class="normal">
          1993
         </p>
        </td>
        <!-- <td><p class="normal">Physical </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Low+Resolution+Spectrometer">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Low+Resolution+Spectrometer">
               Low Resolution Spectrometer
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">From IRAS data -- NASA Ames Research Center </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
```

```
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Integer, Real
  </p>
 </td>
 <td>
  <p class="normal">
   531
  </p>
 </td>
 <td>
  <p class="normal">
   102
  </p>
 </td>
 <td>
  <p class="normal">
   1988
  </p>
 </td>
 <!-- <td><p class="normal">Physical </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Spambase">
      <img border="1" src="assets/MLimages/SmallLarge94.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Spambase">
        Spambase
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">Classifying Email as Spam or Non-Spam </p></td> -->
 <td>
  <p class="normal">
   Multivariate
  </p>
 </td>
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Integer, Real
  </p>
 </td>
 <td>
  <p class="normal">
   4601
  </p>
 </td>
 <td>
  <p class="normal">
   57
  </p>
 </td>
 <td>
  <p class="normal">
   1999
```

```
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/SPECT+Heart">
            <img border="1" src="assets/MLimages/SmallLarge45.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/SPECT+Heart">
              SPECT Heart
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Data on cardiac Single Proton Emission Computed Tomography
(SPECT) images. Each patient classified into two categories: normal and abnormal. </p></td> -
->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical
        </p>
       </td>
       <td>
        <p class="normal">
         267
        </p>
       </td>
       <td>
        <p class="normal">
         22
        </p>
       </td>
       <td>
        <p class="normal">
         2001
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/SPECTF+Heart">
            <img border="1" src="assets/MLimages/SmallLarge45.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/SPECTF+Heart">
              SPECTF Heart
             </a>
            </b>
           </p>
```

```
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Data on cardiac Single Proton Emission Computed Tomography
(SPECT) images. Each patient classified into two categories: normal and abnormal. </p></td> -
->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         267
        </p>
       </td>
       <td>
        <p class="normal">
         44
        </p>
       </td>
       <td>
        <p class="normal">
         2001
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Sponge">
            <img border="1" src="assets/MLimages/SmallLarge97.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Sponge">
              Sponge
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Data on sponges; Attributes in spanish </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer
        </p>
       </td>
       <td>
```

```
        <p class="normal">
         76
        </p>
       </td>
       <td>
        <p class="normal">
         45
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Statlog+Project">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Statlog+Project">
              Statlog Project
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Various Databases: Vehicle silhouttes, Landsat Sattelite, Shuttle
, Australian Credit Approval, Heart Disease, Image Segmentation, German Credit </p></td> -->
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         1992
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Student+Loan+Relational">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
```

```
              <a href="datasets/Student+Loan+Relational">
               Student Loan Relational
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Student Loan Relational Domain </p></td> -->
        <td>
         <p class="normal">
          Domain-Theory
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          1000
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          1993
         </p>
        </td>
        <!-- <td><p class="normal">Social </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Teaching+Assistant+Evaluation">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Teaching+Assistant+Evaluation">
               Teaching Assistant Evaluation
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The data consist of evaluations of teaching performance; scores
are "low", "medium", or "high" </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical, Integer
         </p>
        </td>
```

```
        <td>
         <p class="normal">
          151
         </p>
        </td>
        <td>
         <p class="normal">
          5
         </p>
        </td>
        <td>
         <p class="normal">
          1997
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Tic-Tac-Toe+Endgame">
             <img border="1" src="assets/MLimages/SmallLarge101.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Tic-Tac-Toe+Endgame">
               Tic-Tac-Toe Endgame
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Binary classification task on possible configurations of tic-
tac-toe game </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical
         </p>
        </td>
        <td>
         <p class="normal">
          958
         </p>
        </td>
        <td>
         <p class="normal">
          9
         </p>
        </td>
        <td>
         <p class="normal">
          1991
         </p>
        </td>
        <!-- <td><p class="normal">Game </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
```

```
          <a href="datasets/Thyroid+Disease">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Thyroid+Disease">
             Thyroid Disease
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">10 separate databases from Garavan Institute </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Domain-Theory
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical, Real
       </p>
      </td>
      <td>
       <p class="normal">
        7200
       </p>
      </td>
      <td>
       <p class="normal">
        21
       </p>
      </td>
      <td>
       <p class="normal">
        1987
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Trains">
           <img border="1" src="assets/MLimages/SmallLarge103.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Trains">
             Trains
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">2 data formats (structured, one-instance-per-line) </p>
</td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
```

```
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Categorical
  </p>
 </td>
 <td>
  <p class="normal">
   10
  </p>
 </td>
 <td>
  <p class="normal">
   32
  </p>
 </td>
 <td>
  <p class="normal">
   1994
  </p>
 </td>
 <!-- <td><p class="normal">Other </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/University">
      <img border="1" src="assets/MLimages/SmallLarge104.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/University">
        University
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">Data in original (LISP-readable) form </p></td> -->
 <td>
  <p class="normal">
   Multivariate
  </p>
 </td>
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Categorical, Integer
  </p>
 </td>
 <td>
  <p class="normal">
   285
  </p>
 </td>
 <td>
  <p class="normal">
   17
  </p>
 </td>
 <td>
  <p class="normal">
   1988
```

```
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Congressional+Voting+Records">
            <img border="1" src="assets/MLimages/SmallLarge105.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Congressional+Voting+Records">
              Congressional Voting Records
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">1984 United Stated Congressional Voting Records; Classify as
Republican or Democrat </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical
        </p>
       </td>
       <td>
        <p class="normal">
         435
        </p>
       </td>
       <td>
        <p class="normal">
         16
        </p>
       </td>
       <td>
        <p class="normal">
         1987
        </p>
       </td>
       <!-- <td><p class="normal">Social </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Water+Treatment+Plant">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Water+Treatment+Plant">
              Water Treatment Plant
             </a>
            </b>
           </p>
          </td>
```

```
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">Multiple classes predict plant state </p></td> -->
 <td>
  <p class="normal">
   Multivariate
  </p>
 </td>
 <td>
  <p class="normal">
   Clustering
  </p>
 </td>
 <td>
  <p class="normal">
   Integer, Real
  </p>
 </td>
 <td>
  <p class="normal">
   527
  </p>
 </td>
 <td>
  <p class="normal">
   38
  </p>
 </td>
 <td>
  <p class="normal">
   1993
  </p>
 </td>
 <!-- <td><p class="normal">Physical </p></td> -->
</tr>
<tr>
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Waveform+Database+Generator+%28Version+1%29">
      <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Waveform+Database+Generator+%28Version+1%29">
        Waveform Database Generator (Version 1)
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">CART book's waveform domains </p></td> -->
 <td>
  <p class="normal">
   Multivariate, Data-Generator
  </p>
 </td>
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Real
  </p>
 </td>
 <td>
  <p class="normal">
   5000
  </p>
```

```
  </p>
 </td>
 <td>
  <p class="normal">
   21
  </p>
 </td>
 <td>
  <p class="normal">
   1988
  </p>
 </td>
 <!-- <td><p class="normal">Physical </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Waveform+Database+Generator+%28Version+2%29">
      <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Waveform+Database+Generator+%28Version+2%29">
        Waveform Database Generator (Version 2)
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">CART book's waveform domains </p></td> -->
 <td>
  <p class="normal">
   Multivariate, Data-Generator
  </p>
 </td>
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Real
  </p>
 </td>
 <td>
  <p class="normal">
   5000
  </p>
 </td>
 <td>
  <p class="normal">
   40
  </p>
 </td>
 <td>
  <p class="normal">
   1988
  </p>
 </td>
 <!-- <td><p class="normal">Physical </p></td> -->
</tr>
<tr>
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Wine">
      <img border="1" src="assets/MLimages/SmallLarge109.jpg"/>
     </a>
    </td>
    <td>
```

```
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Wine">
              Wine
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Using chemical analysis determine the origin of wines </p><
/td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         178
        </p>
       </td>
       <td>
        <p class="normal">
         13
        </p>
       </td>
       <td>
        <p class="normal">
         1991
        </p>
       </td>
       <!-- <td><p class="normal">Physical </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Yeast">
            <img border="1" src="assets/MLimages/SmallLarge110.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Yeast">
              Yeast
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Predicting the Cellular Localization Sites of Proteins </p>
</td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
```

```
  </p>
 </td>
 <td>
  <p class="normal">
   Real
  </p>
 </td>
 <td>
  <p class="normal">
   1484
  </p>
 </td>
 <td>
  <p class="normal">
   8
  </p>
 </td>
 <td>
  <p class="normal">
   1996
  </p>
 </td>
 <!-- <td><p class="normal">Life </p></td> -->
</tr>
<tr>
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Zoo">
      <img border="1" src="assets/MLimages/SmallLarge111.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Zoo">
        Zoo
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">Artificial, 7 classes of animals </p></td> -->
 <td>
  <p class="normal">
   Multivariate
  </p>
 </td>
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Categorical, Integer
  </p>
 </td>
 <td>
  <p class="normal">
   101
  </p>
 </td>
 <td>
  <p class="normal">
   17
  </p>
 </td>
 <td>
  <p class="normal">
   1990
  </p>
 </td>
 <!-- <td><p class="normal">Life </p></td> -->
</tr>
```

```
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Undocumented">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Undocumented">
              Undocumented
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Various datasets without documentation (feel free to explore!)&nb
sp;</p></td> -->
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Twenty+Newsgroups">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Twenty+Newsgroups">
              Twenty Newsgroups
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data set consists of 20000 messages taken from 20
newsgroups. </p></td> -->
       <td>
        <p class="normal">
         Text
        </p>
       </td>
```

```
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        20000
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        1999
       </p>
      </td>
      <!-- <td><p class="normal">Other </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Australian+Sign+Language+signs">
           <img border="1" src="assets/MLimages/SmallLarge114.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Australian+Sign+Language+signs">
             Australian Sign Language signs
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This data consists of sample of Auslan (Australian Sign
Language) signs. Examples of 95 signs were collected from five signers with a total of 6650 sign s
amples. </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical, Real
       </p>
      </td>
      <td>
       <p class="normal">
        6650
       </p>
      </td>
      <td>
       <p class="normal">
        15
       </p>
      </td>
      <td>
       <p class="normal">
        1999
```

```
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Australian+Sign+Language+signs+%28High+Quality%29">
            <img border="1" src="assets/MLimages/SmallLarge114.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Australian+Sign+Language+signs+%28High+Quality%29">
              Australian Sign Language signs (High Quality)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data consists of sample of Auslan (Australian Sign
Language) signs. 27 examples of each of 95 Auslan signs were captured from a native signer using h
igh-quality position trackers </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         2565
        </p>
       </td>
       <td>
        <p class="normal">
         22
        </p>
       </td>
       <td>
        <p class="normal">
         2002
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/US+Census+Data+%281990%29">
            <img border="1" src="assets/MLimages/SmallLarge2.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/US+Census+Data+%281990%29">
              US Census Data (1990)
             </a>
            </b>
           </p>
```

```
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The USCensus1990raw data set contains a one percent sample of
the Public Use Microdata Samples (PUMS) person records drawn from the full 1990 census
sample. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical
        </p>
       </td>
       <td>
        <p class="normal">
         2458285
        </p>
       </td>
       <td>
        <p class="normal">
         68
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <!-- <td><p class="normal">Social </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Census-Income+%28KDD%29">
            <img border="1" src="assets/MLimages/SmallLarge2.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Census-Income+%28KDD%29">
              Census-Income (KDD)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data set contains weighted census data extracted from the 1
994 and 1995 current population surveys conducted by the U.S. Census Bureau. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer
        </p>
       </td>
       <td>
```

```
        <p class="normal">
         299285
        </p>
       </td>
       <td>
        <p class="normal">
         40
        </p>
       </td>
       <td>
        <p class="normal">
         2000
        </p>
       </td>
       <!-- <td><p class="normal">Social </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Coil+1999+Competition+Data">
            <img border="1" src="assets/MLimages/SmallLarge118.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Coil+1999+Competition+Data">
              Coil 1999 Competition Data
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data set is from the 1999 Computational Intelligence and
Learning (COIL) competition. The data contains measurements of river chemical concentrations and a
lgae densities. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Real
        </p>
       </td>
       <td>
        <p class="normal">
         340
        </p>
       </td>
       <td>
        <p class="normal">
         17
        </p>
       </td>
       <td>
        <p class="normal">
         1999
        </p>
       </td>
       <!-- <td><p class="normal">Physical </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Corel+Image+Features">
```

```
            <img border="1" src="assets/MLimages/SmallLarge119.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Corel+Image+Features">
              Corel Image Features
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset contains image features extracted from a Corel
image collection. Four sets of features are available based on the color histogram, color
histogram layout, color moments, and co-occurence </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         68040
        </p>
       </td>
       <td>
        <p class="normal">
         89
        </p>
       </td>
       <td>
        <p class="normal">
         1999
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/E.+Coli+Genes">
            <img border="1" src="assets/MLimages/SmallLarge120.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/E.+Coli+Genes">
              E. Coli Genes
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Data giving characteristics of each ORF (potential gene) in the
E. coli genome. Sequence, homology (similarity to other genes) and structural information, and fun
ction (if known) are provided. </p></td> -->
       <td>
        <p class="normal">
         Relational
        </p>
```

```
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          2001
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/EEG+Database">
             <img border="1" src="assets/MLimages/SmallLarge121.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/EEG+Database">
               EEG Database
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This data arises from a large study to examine EEG correlates of genetic predisposition to alcoholism. It contains measurements from 64 electrodes placed on the scalp sampled at 256 Hz </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical, Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          122
         </p>
        </td>
        <td>
         <p class="normal">
          4
         </p>
        </td>
        <td>
         <p class="normal">
          1999
         </p>
        </td>
```

```
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/El+Nino">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/El+Nino">
               El Nino
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The data set contains oceanographic and surface meteorological
readings taken from a series of buoys positioned throughout the equatorial Pacific. </p></td>
-->
        <td>
         <p class="normal">
          Spatio-temporal
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          178080
         </p>
        </td>
        <td>
         <p class="normal">
          12
         </p>
        </td>
        <td>
         <p class="normal">
          1999
         </p>
        </td>
        <!-- <td><p class="normal">Physical </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Entree+Chicago+Recommendation+Data">
             <img border="1" src="assets/MLimages/SmallLarge123.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Entree+Chicago+Recommendation+Data">
               Entree Chicago Recommendation Data
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
```

```
      </td>
      <!-- <td><p class="normal">This data contains a record of user interactions with the Entree
Chicago restaurant recommendation system. </p></td> -->
      <td>
       <p class="normal">
        Transactional, Sequential
       </p>
      </td>
      <td>
       <p class="normal">
        Recommender-Systems
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical
       </p>
      </td>
      <td>
       <p class="normal">
        50672
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        2000
       </p>
      </td>
      <!-- <td><p class="normal">Other </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/CMU+Face+Images">
           <img border="1" src="assets/MLimages/SmallLarge124.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/CMU+Face+Images">
             CMU Face Images
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This data consists of 640 black and white face images of people
taken with varying pose (straight, left, right, up), expression (neutral, happy, sad, angry), eyes
(wearing sunglasses or not), and size </p></td> -->
      <td>
       <p class="normal">
        Image
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
       <p class="normal">
        640
       </p>
```

```
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          1999
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Insurance+Company+Benchmark+%28COIL+2000%29">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Insurance+Company+Benchmark+%28COIL+2000%29">
               Insurance Company Benchmark (COIL 2000)
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This data set used in the CoIL 2000 Challenge contains
information on customers of an insurance company. The data consists of 86 variables and includes p
roduct usage data and socio-demographic data </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Regression, Description
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical, Integer
         </p>
        </td>
        <td>
         <p class="normal">
          9000
         </p>
        </td>
        <td>
         <p class="normal">
          86
         </p>
        </td>
        <td>
         <p class="normal">
          2000
         </p>
        </td>
        <!-- <td><p class="normal">Social </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Internet+Usage+Data">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
```

```
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Internet+Usage+Data">
              Internet Usage Data
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data contains general demographic information on internet u
sers in 1997. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer
        </p>
       </td>
       <td>
        <p class="normal">
         10104
        </p>
       </td>
       <td>
        <p class="normal">
         72
        </p>
       </td>
       <td>
        <p class="normal">
         1999
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/IPUMS+Census+Database">
            <img border="1" src="assets/MLimages/SmallLarge2.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/IPUMS+Census+Database">
              IPUMS Census Database
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data set contains unweighted PUMS census data from the Los
Angeles and Long Beach areas for the years 1970, 1980, and 1990. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
```

```
        <td>
         <p class="normal">
          Categorical, Integer
         </p>
        </td>
        <td>
         <p class="normal">
          256932
         </p>
        </td>
        <td>
         <p class="normal">
          61
         </p>
        </td>
        <td>
         <p class="normal">
          1999
         </p>
        </td>
        <!-- <td><p class="normal">Social </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Japanese+Vowels">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Japanese+Vowels">
               Japanese Vowels
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This dataset records 640 time series of 12 LPC cepstrum
coefficients taken from nine male speakers. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          640
         </p>
        </td>
        <td>
         <p class="normal">
          12
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr>
```

```
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/KDD+Cup+1998+Data">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/KDD+Cup+1998+Data">
             KDD Cup 1998 Data
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This is the data set used for The Second International Knowledge
Discovery and Data Mining Tools Competition, which was held in conjunction with KDD-98 </p></
td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical, Integer
       </p>
      </td>
      <td>
       <p class="normal">
        191779
       </p>
      </td>
      <td>
       <p class="normal">
        481
       </p>
      </td>
      <td>
       <p class="normal">
        1998
       </p>
      </td>
      <!-- <td><p class="normal">Other </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/KDD+Cup+1999+Data">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/KDD+Cup+1999+Data">
             KDD Cup 1999 Data
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This is the data set used for The Third International Knowledge
```

```
<td><p class="normal">This is the data set used for the third International Knowledge
Discovery and Data Mining Tools Competition, which was held in conjunction with KDD-99 </p></
td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical, Integer
       </p>
      </td>
      <td>
       <p class="normal">
        4000000
       </p>
      </td>
      <td>
       <p class="normal">
        42
       </p>
      </td>
      <td>
       <p class="normal">
        1999
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/M.+Tuberculosis+Genes">
           <img border="1" src="assets/MLimages/SmallLarge131.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/M.+Tuberculosis+Genes">
             M. Tuberculosis Genes
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal"> Data giving characteristics of each ORF (potential gene) in the
M. tuberculosis bacterium. Sequence, homology (similarity to other genes) and structural
information, and function (if known) are provided </p></td> -->
      <td>
       <p class="normal">
        Relational
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
```

```
        <p class="normal">
         </p>
       </td>
       <td>
        <p class="normal">
         2001
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Movie">
            <img border="1" src="assets/MLimages/SmallLarge132.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Movie">
              Movie
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data set contains a list of over 10000 films including many
older, odd, and cult films. There is information on actors, casts, directors, producers, studios,
etc. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Relational
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         10000
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         1999
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/MSNBC.com+Anonymous+Web+Data">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/MSNBC.com+Anonymous+Web+Data">
              MSNBC.com Anonymous Web Data
             </a>
```

```
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This data describes the page visits of users who visited
msnbc.com on September 28, 1999. Visits are recorded at the level of URL category (see
description) and are recorded in time order. </p></td> -->
      <td>
       <p class="normal">
        Sequential
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical
       </p>
      </td>
      <td>
       <p class="normal">
        989818
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/NSF+Research+Award+Abstracts+1990-2003">
           <img border="1" src="assets/MLimages/SmallLarge134.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/NSF+Research+Award+Abstracts+1990-2003">
             NSF Research Award Abstracts 1990-2003
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This data set consists of (a) 129,000 abstracts describing NSF a
wards for basic research, (b) bag-of-word data files extracted from the abstracts, (c) a list of w
ords used for indexing the bag-of-word </p></td> -->
      <td>
       <p class="normal">
        Text
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
```

```
      <p class="normal">
       129000
      </p>
     </td>
     <td>
      <p class="normal">
      </p>
     </td>
     <td>
      <p class="normal">
       2003
      </p>
     </td>
     <!-- <td><p class="normal">Other </p></td> -->
    </tr>
    <tr>
     <td>
      <table>
       <tr>
        <td>
         <a href="datasets/Pioneer-1+Mobile+Robot+Data">
          <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
         </a>
        </td>
        <td>
         <p class="normal">
          <b>
           <a href="datasets/Pioneer-1+Mobile+Robot+Data">
            Pioneer-1 Mobile Robot Data
           </a>
          </b>
         </p>
        </td>
       </tr>
      </table>
     </td>
     <!-- <td><p class="normal">This dataset contains time series sensor readings of the
Pioneer-1 mobile robot. The data is broken into "experiences" in which the robot takes action for
some period of time and experiences a control </p></td> -->
     <td>
      <p class="normal">
       Multivariate, Time-Series
      </p>
     </td>
     <td>
      <p class="normal">
      </p>
     </td>
     <td>
      <p class="normal">
       Categorical, Real
      </p>
     </td>
     <td>
      <p class="normal">
      </p>
     </td>
     <td>
      <p class="normal">
      </p>
     </td>
     <td>
      <p class="normal">
       1999
      </p>
     </td>
     <!-- <td><p class="normal">Computer </p></td> -->
    </tr>
    <tr bgcolor="DDEEFF">
     <td>
      <table>
       <tr>
        <td>
         <a href="datasets/Pseudo+Periodic+Synthetic+Time+Series">
          <img border="1" src="assets/MLimages/SmallLarge136.jpg"/>
         </a>
        </td>
        <td>
```

```
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Pseudo+Periodic+Synthetic+Time+Series">
              Pseudo Periodic Synthetic Time Series
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data set is designed for testing indexing schemes in time s
eries databases. The data appears highly periodic, but never exactly repeats itself. </p></td
> -->
       <td>
        <p class="normal">
         Univariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         100000
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         1999
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Reuters-21578+Text+Categorization+Collection">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Reuters-21578+Text+Categorization+Collection">
              Reuters-21578 Text Categorization Collection
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This is a collection of documents that appeared on Reuters
newswire in 1987. The documents were assembled and indexed with categories. </p></td> -->
       <td>
        <p class="normal">
         Text
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
```

```
        <td>
         <p class="normal">
          Categorical
         </p>
        </td>
        <td>
         <p class="normal">
          21578
         </p>
        </td>
        <td>
         <p class="normal">
          5
         </p>
        </td>
        <td>
         <p class="normal">
          1997
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Robot+Execution+Failures">
             <img border="1" src="assets/MLimages/SmallLarge138.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Robot+Execution+Failures">
               Robot Execution Failures
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This dataset contains force and torque measurements on a robot
after failure detection. Each failure is characterized by 15 force/torque samples collected at
regular time intervals </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          463
         </p>
        </td>
        <td>
         <p class="normal">
          90
         </p>
        </td>
        <td>
         <p class="normal">
          1999
         </p>
        </td>
        <!-- <td><p class="normal">Physical </p></td> -->
```

```
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Synthetic+Control+Chart+Time+Series">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Synthetic+Control+Chart+Time+Series">
              Synthetic Control Chart Time Series
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data consists of synthetically generated control
charts. </p></td> -->
       <td>
        <p class="normal">
         Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         600
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         1999
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Syskill+and+Webert+Web+Page+Ratings">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Syskill+and+Webert+Web+Page+Ratings">
              Syskill and Webert Web Page Ratings
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This database contains HTML source of web pages plus the ratings
```

of a single user on these web pages. Web pages are on four seperate subjects (Bands- recording art
ists; Goats; Sheep; and BioMedical) </td> -->

```
       <td>
        <p class="normal">
         Multivariate, Text
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical
        </p>
       </td>
       <td>
        <p class="normal">
         332
        </p>
       </td>
       <td>
        <p class="normal">
         5
        </p>
       </td>
       <td>
        <p class="normal">
         1998
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/UNIX+User+Data">
            <img border="1" src="assets/MLimages/SmallLarge141.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/UNIX+User+Data">
              UNIX User Data
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
```
       <!-- <td> This file contains 9 sets of sanitized user data drawn from the
command histories of 8 UNIX computer users at Purdue over the course of up to 2 years. </
td> -->

```
       <td>
        <p class="normal">
         Text, Sequential
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
```

```
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Volcanoes+on+Venus+-+JARtool+experiment">
            <img border="1" src="assets/MLimages/SmallLarge142.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Volcanoes+on+Venus+-+JARtool+experiment">
              Volcanoes on Venus - JARtool experiment
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The JARtool project was a pioneering effort to develop an automat
ic system for cataloging small volcanoes in the large set of Venus images returned by the Magellan
spacecraft. </p></td> -->
       <td>
        <p class="normal">
         Image
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <!-- <td><p class="normal">Physical </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Statlog+%28Australian+Credit+Approval%29">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Statlog+%28Australian+Credit+Approval%29">
              Statlog (Australian Credit Approval)
             </a>
            </b>
           </p>
```

```
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This file concerns credit card applications. This database
exists elsewhere in the repository (Credit Screening Database) in a slightly different form <
/p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         690
        </p>
       </td>
       <td>
        <p class="normal">
         14
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <!-- <td><p class="normal">Financial </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Statlog+%28German+Credit+Data%29">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Statlog+%28German+Credit+Data%29">
              Statlog (German Credit Data)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset classifies people described by a set of attributes
as good or bad credit risks. Comes in two formats (one all numeric). Also comes with a cost matrix
 </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer
        </p>
       </td>
```

```
        <td>
         <p class="normal">
          1000
         </p>
        </td>
        <td>
         <p class="normal">
          20
         </p>
        </td>
        <td>
         <p class="normal">
          1994
         </p>
        </td>
        <!-- <td><p class="normal">Financial </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Statlog+%28Heart%29">
             <img border="1" src="assets/MLimages/SmallLarge45.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Statlog+%28Heart%29">
               Statlog (Heart)
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This dataset is a heart disease database similar to a database
already present in the repository (Heart Disease databases) but in a slightly different form 
</p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical, Real
         </p>
        </td>
        <td>
         <p class="normal">
          270
         </p>
        </td>
        <td>
         <p class="normal">
          13
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
```

```
            <a href="datasets/Statlog+%28Landsat+Satellite%29">
             <img border="1" src="assets/MLimages/SmallLarge146.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Statlog+%28Landsat+Satellite%29">
               Statlog (Landsat Satellite)
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Multi-spectral values of pixels in 3x3 neighbourhoods in a
satellite image, and the classification associated with the central pixel in each
neighbourhood </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          6435
         </p>
        </td>
        <td>
         <p class="normal">
          36
         </p>
        </td>
        <td>
         <p class="normal">
          1993
         </p>
        </td>
        <!-- <td><p class="normal">Physical </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Statlog+%28Image+Segmentation%29">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Statlog+%28Image+Segmentation%29">
               Statlog (Image Segmentation)
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This dataset is an image segmentation database similar to a datab
ase already present in the repository (Image segmentation database) but in a slightly different fo
rm. </p></td> -->
        <td>
         <p class="normal">
```

```
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          2310
         </p>
        </td>
        <td>
         <p class="normal">
          19
         </p>
        </td>
        <td>
         <p class="normal">
          1990
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Statlog+%28Shuttle%29">
             <img border="1" src="assets/MLimages/SmallLarge92.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Statlog+%28Shuttle%29">
               Statlog (Shuttle)
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The shuttle dataset contains 9 attributes all of which are
numerical. Approximately 80% of the data belongs to class 1 </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          58000
         </p>
        </td>
        <td>
         <p class="normal">
          9
         </p>
```

```
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <!-- <td><p class="normal">Physical </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Statlog+%28Vehicle+Silhouettes%29">
             <img border="1" src="assets/MLimages/SmallLarge149.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Statlog+%28Vehicle+Silhouettes%29">
               Statlog (Vehicle Silhouettes)
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">3D objects within a 2D image by application of an ensemble of
shape feature extractors to the 2D silhouettes of the objects. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          946
         </p>
        </td>
        <td>
         <p class="normal">
          18
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Connectionist+Bench+%28Nettalk+Corpus%29">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Connectionist+Bench+%28Nettalk+Corpus%29">
               Connectionist Bench (Nettalk Corpus)
              </a>
             </b>
```

```
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The file "nettalk.data" contains a list of 20,008 English words,
along with a phonetic transcription for each word. The task is to train a network to produce the p
roper phonemes </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical
       </p>
      </td>
      <td>
       <p class="normal">
        20008
       </p>
      </td>
      <td>
       <p class="normal">
        4
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <!-- <td><p class="normal">Other </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Connectionist+Bench+%28Sonar%2C+Mines+vs.+Rocks%29">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Connectionist+Bench+%28Sonar%2C+Mines+vs.+Rocks%29">
             Connectionist Bench (Sonar, Mines vs. Rocks)
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The task is to train a network to discriminate between sonar sigr
als bounced off a metal cylinder and those bounced off a roughly cylindrical rock. </p></td>
-->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
```

```
      <td>
       <p class="normal">
        208
       </p>
      </td>
      <td>
       <p class="normal">
        60
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <!-- <td><p class="normal">Physical </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Connectionist+Bench+%28Vowel+Recognition+-+Deterding+Data%29">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Connectionist+Bench+%28Vowel+Recognition+-+Deterding+Data%29">
             Connectionist Bench (Vowel Recognition - Deterding Data)
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Speaker independent recognition of the eleven steady state
vowels of British English using a specified training set of lpc derived log area ratios. </p>
</td> -->
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        528
       </p>
      </td>
      <td>
       <p class="normal">
        10
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <!-- <td><p class="normal">Other </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Economic+Sanctions">
           <img border="1" src="assets/MLimages/SmallLarge153.jpg"/>
```

```
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Economic+Sanctions">
             Economic Sanctions
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Domain Theory on Economic Sanctions; Undocumented </p></td>
-->
      <td>
       <p class="normal">
        Domain-Theory
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <!-- <td><p class="normal">Financial </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Protein+Data">
           <img border="1" src="assets/MLimages/SmallLarge154.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Protein+Data">
             Protein Data
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Undocumented </p></td> -->
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
```

```
 <td>
  <p class="normal">
  </p>
 </td>
 <td>
  <p class="normal">
  </p>
 </td>
 <td>
  <p class="normal">
  </p>
 </td>
 <!-- <td><p class="normal">Life </p></td> -->
</tr>
<tr>
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Cloud">
      <img border="1" src="assets/MLimages/SmallLarge155.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Cloud">
        Cloud
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">Little Documentation </p></td> -->
 <td>
  <p class="normal">
   Multivariate
  </p>
 </td>
 <td>
  <p class="normal">
  </p>
 </td>
 <td>
  <p class="normal">
   Real
  </p>
 </td>
 <td>
  <p class="normal">
   1024
  </p>
 </td>
 <td>
  <p class="normal">
   10
  </p>
 </td>
 <td>
  <p class="normal">
   1989
  </p>
 </td>
 <!-- <td><p class="normal">Physical </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/CalIt2+Building+People+Counts">
      <img border="1" src="assets/MLimages/SmallLarge156.jpg"/>
     </a>
    </td>
    <td>
```

```
          <td>
           <p class="normal">
            <b>
             <a href="datasets/CalIt2+Building+People+Counts">
              CalIt2 Building People Counts
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data comes from the main door of the CalIt2 building at
UCI. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer
        </p>
       </td>
       <td>
        <p class="normal">
         10080
        </p>
       </td>
       <td>
        <p class="normal">
         4
        </p>
       </td>
       <td>
        <p class="normal">
         2006
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Dodgers+Loop+Sensor">
            <img border="1" src="assets/MLimages/SmallLarge157.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Dodgers+Loop+Sensor">
              Dodgers Loop Sensor
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Loop sensor data was collected for the Glendale on ramp for the
101 North freeway in Los Angeles </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
```

```
 <td>
  <p class="normal">
   Categorical, Integer
  </p>
 </td>
 <td>
  <p class="normal">
   50400
  </p>
 </td>
 <td>
  <p class="normal">
   3
  </p>
 </td>
 <td>
  <p class="normal">
   2006
  </p>
 </td>
 <!-- <td><p class="normal">Other </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Poker+Hand">
      <img border="1" src="assets/MLimages/SmallLarge158.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Poker+Hand">
        Poker Hand
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">Purpose is to predict poker hands </p></td> -->
 <td>
  <p class="normal">
   Multivariate
  </p>
 </td>
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Categorical, Integer
  </p>
 </td>
 <td>
  <p class="normal">
   1025010
  </p>
 </td>
 <td>
  <p class="normal">
   11
  </p>
 </td>
 <td>
  <p class="normal">
   2007
  </p>
 </td>
 <!-- <td><p class="normal">Game </p></td> -->
</tr>
<tr>
 <td>
```

```
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/MAGIC+Gamma+Telescope">
           <img border="1" src="assets/MLimages/SmallLarge159.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/MAGIC+Gamma+Telescope">
             MAGIC Gamma Telescope
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Data are MC generated to simulate registration of high energy
gamma particles in an atmospheric Cherenkov telescope </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        19020
       </p>
      </td>
      <td>
       <p class="normal">
        11
       </p>
      </td>
      <td>
       <p class="normal">
        2007
       </p>
      </td>
      <!-- <td><p class="normal">Physical </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/UJI+Pen+Characters">
           <img border="1" src="assets/MLimages/SmallLarge160.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/UJI+Pen+Characters">
             UJI Pen Characters
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Data consists of written characters in a UNIPEN-like
format </p></td> -->
      <td>
```

```
      <td>
       <p class="normal">
        Multivariate, Sequential
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
       <p class="normal">
        1364
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        2007
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Mammographic+Mass">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Mammographic+Mass">
             Mammographic Mass
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Discrimination of benign and malignant mammographic masses based
on BI-RADS attributes and the patient's age. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
       <p class="normal">
        961
       </p>
      </td>
      <td>
       <p class="normal">
        6
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         2007
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Forest+Fires">
            <img border="1" src="assets/MLimages/SmallLarge162.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Forest+Fires">
              Forest Fires
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This is a difficult regression task, where the aim is to predict
the burned area of forest fires, in the northeast region of Portugal, by using meteorological and
other data (see details at: http://www.dsi.uminho.pt/~pcortez/forestfires). </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         517
        </p>
       </td>
       <td>
        <p class="normal">
         13
        </p>
       </td>
       <td>
        <p class="normal">
         2008
        </p>
       </td>
       <!-- <td><p class="normal">Physical </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Reuters+Transcribed+Subset">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
```

```
            <a href="datasets/Reuters+Transcribed+Subset">
             Reuters Transcribed Subset
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This dataset is created by reading out 200 files from the 10
largest Reuters
classes and using an Automatic Speech Recognition system to create
corresponding transcriptions. </p></td> -->
      <td>
       <p class="normal">
        Text
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        200
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        2008
       </p>
      </td>
      <!-- <td><p class="normal">Business </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Bag+of+Words">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Bag+of+Words">
             Bag of Words
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This data set contains five text collections in the form of
bags-of-words. </p></td> -->
      <td>
       <p class="normal">
        Text
       </p>
      </td>
      <td>
       <p class="normal">
        Clustering
       </p>
      </td>
      <td>
```

```
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         8000000
        </p>
       </td>
       <td>
        <p class="normal">
         100000
        </p>
       </td>
       <td>
        <p class="normal">
         2008
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Concrete+Compressive+Strength">
            <img border="1" src="assets/MLimages/SmallLarge165.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Concrete+Compressive+Strength">
              Concrete Compressive Strength
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Concrete is the most important material in civil engineering.
The concrete compressive strength is a highly nonlinear function of age and ingredients.  </p
></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         1030
        </p>
       </td>
       <td>
        <p class="normal">
         9
        </p>
       </td>
       <td>
        <p class="normal">
         2007
        </p>
       </td>
       <!-- <td><p class="normal">Physical </p></td> -->
      </tr>
```

```
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Hill-Valley">
            <img border="1" src="assets/MLimages/SmallLarge166.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Hill-Valley">
              Hill-Valley
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Each record represents 100 points on a two-dimensional graph.
When plotted in order (from 1 through 100) as the Y co-ordinate, the points will create either a H
ill (a bump in the terrain) or a Valley (a dip in the terrain). </p></td> -->
       <td>
        <p class="normal">
         Sequential
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         606
        </p>
       </td>
       <td>
        <p class="normal">
         101
        </p>
       </td>
       <td>
        <p class="normal">
         2008
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Arcene">
            <img border="1" src="assets/MLimages/SmallLarge167.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Arcene">
              Arcene
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
```

```
      <!-- <td><p class="normal">ARCENE's task is to distinguish cancer versus normal patterns
from mass-spectrometric data. This is a two-class classification problem with continuous input
variables. This dataset is one of 5 datasets of the NIPS 2003 feature selection challenge. </
p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        900
       </p>
      </td>
      <td>
       <p class="normal">
        10000
       </p>
      </td>
      <td>
       <p class="normal">
        2008
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Dexter">
           <img border="1" src="assets/MLimages/SmallLarge168.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Dexter">
             Dexter
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">DEXTER is a text classification problem in a bag-of-word
representation. This is a two-class classification problem with sparse continuous input variables.
This dataset is one of five datasets of the NIPS 2003 feature selection challenge.
 </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
```

```
        <p class="normal">
         2600
        </p>
       </td>
       <td>
        <p class="normal">
         20000
        </p>
       </td>
       <td>
        <p class="normal">
         2008
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Dorothea">
            <img border="1" src="assets/MLimages/SmallLarge169.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Dorothea">
              Dorothea
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">DOROTHEA is a drug discovery dataset. Chemical compounds
represented by structural molecular features must be classified as active (binding to thrombin) or
inactive. This is one of 5 datasets of the NIPS 2003 feature selection challenge. </p></td> -
->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         1950
        </p>
       </td>
       <td>
        <p class="normal">
         100000
        </p>
       </td>
       <td>
        <p class="normal">
         2008
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
```

```
          <td>
           <a href="datasets/Gisette">
            <img border="1" src="assets/MLimages/SmallLarge170.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Gisette">
              Gisette
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">GISETTE is a handwritten digit recognition problem. The problem
is to separate the highly confusible digits '4' and '9'. This dataset is one of five datasets of t
he NIPS 2003 feature selection challenge.
 </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         13500
        </p>
       </td>
       <td>
        <p class="normal">
         5000
        </p>
       </td>
       <td>
        <p class="normal">
         2008
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Madelon">
            <img border="1" src="assets/MLimages/SmallLarge171.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Madelon">
              Madelon
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">MADELON is an artificial dataset, which was part of the NIPS
2003 feature selection challenge. This is a two-class classification problem with continuous input
variables. The difficulty is that the problem is multivariate and highly non-linear.  </p></t
```

```
d> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         4400
        </p>
       </td>
       <td>
        <p class="normal">
         500
        </p>
       </td>
       <td>
        <p class="normal">
         2008
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Ozone+Level+Detection">
            <img border="1" src="assets/MLimages/SmallLarge172.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Ozone+Level+Detection">
              Ozone Level Detection
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Two ground ozone level data sets are included in this
collection. One is the eight hour peak set (eighthr.data), the other is the one hour peak set (one
hr.data). Those data were collected from 1998 to 2004 at the Houston, Galveston and Brazoria area.
 </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Sequential, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         2536
        </p>
```

```
        </td>
        <td>
         <p class="normal">
          73
         </p>
        </td>
        <td>
         <p class="normal">
          2008
         </p>
        </td>
        <!-- <td><p class="normal">Physical </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Abscisic+Acid+Signaling+Network">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Abscisic+Acid+Signaling+Network">
               Abscisic Acid Signaling Network
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The objective is to determine the set of boolean rules that descr
ibe the interactions of the nodes within this plant signaling network.  The dataset includes 300 s
eparate boolean pseudodynamic simulations using an asynchronous update scheme.  </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Causal-Discovery
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          300
         </p>
        </td>
        <td>
         <p class="normal">
          43
         </p>
        </td>
        <td>
         <p class="normal">
          2008
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Parkinsons">
             <img border="1" src="assets/MLimages/SmallLarge174.jpg"/>
            </a>
```

```
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Parkinsons">
              Parkinsons
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Oxford Parkinson's Disease Detection Dataset </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         197
        </p>
       </td>
       <td>
        <p class="normal">
         23
        </p>
       </td>
       <td>
        <p class="normal">
         2008
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Character+Trajectories">
            <img border="1" src="assets/MLimages/SmallLarge175.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Character+Trajectories">
              Character Trajectories
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Multiple, labelled samples of pen tip trajectories recorded
whilst writing individual characters. All samples are from the same writer, for the purposes of pr
imitive extraction. Only characters with a single pen-down segment were considered. </p></td>
-->
       <td>
        <p class="normal">
         Time-Series
        </p>
       </td>
       <td>
```

```
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         2858
        </p>
       </td>
       <td>
        <p class="normal">
         3
        </p>
       </td>
       <td>
        <p class="normal">
         2008
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Blood+Transfusion+Service+Center">
            <img border="1" src="assets/MLimages/SmallLarge176.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Blood+Transfusion+Service+Center">
              Blood Transfusion Service Center
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Data taken from the Blood Transfusion Service Center in Hsin-Chu
City in Taiwan -- this is a classification problem.  </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         748
        </p>
       </td>
       <td>
        <p class="normal">
         5
        </p>
       </td>
       <td>
        <p class="normal">
         2008
```

```
        </p>
       </td>
       <!-- <td><p class="normal">Business </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/UJI+Pen+Characters+%28Version+2%29">
            <img border="1" src="assets/MLimages/SmallLarge160.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/UJI+Pen+Characters+%28Version+2%29">
              UJI Pen Characters (Version 2)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">A pen-based database with more than 11k isolated handwritten cha
racters </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Sequential
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         11640
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         2009
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Semeion+Handwritten+Digit">
            <img border="1" src="assets/MLimages/SmallLarge178.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Semeion+Handwritten+Digit">
              Semeion Handwritten Digit
             </a>
            </b>
           </p>
          </td>
         </tr>
```

```
        </table>
       </td>
       <!-- <td><p class="normal">1593 handwritten digits from around 80 persons were scanned,
stretched in a rectangular box 16x16 in a gray scale of 256 values. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         1593
        </p>
       </td>
       <td>
        <p class="normal">
         256
        </p>
       </td>
       <td>
        <p class="normal">
         2008
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/SECOM">
            <img border="1" src="assets/MLimages/SmallLarge179.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/SECOM">
              SECOM
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Data from a semi-conductor manufacturing process </p></td>
-->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Causal-Discovery
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         1567
```

```
        1007
        </p>
       </td>
       <td>
        <p class="normal">
         591
        </p>
       </td>
       <td>
        <p class="normal">
         2008
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Plants">
            <img border="1" src="assets/MLimages/SmallLarge180.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Plants">
              Plants
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Data has been extracted from the USDA plants database. It
contains all plants (species and genera) in the database and the states of USA and Canada where th
ey occur. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical
        </p>
       </td>
       <td>
        <p class="normal">
         22632
        </p>
       </td>
       <td>
        <p class="normal">
         70
        </p>
       </td>
       <td>
        <p class="normal">
         2008
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Libras+Movement">
            <img border="1" src="assets/MLimages/SmallLarge181.jpg"/>
```

```
          <img border="1" src="assets/MLimages/SmallLarge101.jpg"/>
         </a>
        </td>
        <td>
         <p class="normal">
          <b>
           <a href="datasets/Libras+Movement">
            Libras Movement
           </a>
          </b>
         </p>
        </td>
       </tr>
      </table>
     </td>
     <!-- <td><p class="normal">The data set contains 15 classes of 24 instances each. Each class
references to a hand movement type in LIBRAS (Portuguese
name 'L⍰ngua BRAsileira de Sinais', oficial brazilian signal language). </p></td> -->
     <td>
      <p class="normal">
       Multivariate, Sequential
      </p>
     </td>
     <td>
      <p class="normal">
       Classification, Clustering
      </p>
     </td>
     <td>
      <p class="normal">
       Real
      </p>
     </td>
     <td>
      <p class="normal">
       360
      </p>
     </td>
     <td>
      <p class="normal">
       91
      </p>
     </td>
     <td>
      <p class="normal">
       2009
      </p>
     </td>
     <!-- <td><p class="normal">Other </p></td> -->
    </tr>
    <tr bgcolor="DDEEFF">
     <td>
      <table>
       <tr>
        <td>
         <a href="datasets/Concrete+Slump+Test">
          <img border="1" src="assets/MLimages/SmallLarge165.jpg"/>
         </a>
        </td>
        <td>
         <p class="normal">
          <b>
           <a href="datasets/Concrete+Slump+Test">
            Concrete Slump Test
           </a>
          </b>
         </p>
        </td>
       </tr>
      </table>
     </td>
     <!-- <td><p class="normal">Concrete is a highly complex material. The slump flow of
concrete is not only determined by the water content, but that is also influenced by other
concrete ingredients. </p></td> -->
     <td>
      <p class="normal">
       Multivariate
      </p>
```

```
      </p>
     </td>
     <td>
      <p class="normal">
       Regression
      </p>
     </td>
     <td>
      <p class="normal">
       Real
      </p>
     </td>
     <td>
      <p class="normal">
       103
      </p>
     </td>
     <td>
      <p class="normal">
       10
      </p>
     </td>
     <td>
      <p class="normal">
       2009
      </p>
     </td>
     <!-- <td><p class="normal">Computer </p></td> -->
    </tr>
    <tr>
     <td>
      <table>
       <tr>
        <td>
         <a href="datasets/Communities+and+Crime">
          <img border="1" src="assets/MLimages/SmallLarge183.jpg"/>
         </a>
        </td>
        <td>
         <p class="normal">
          <b>
           <a href="datasets/Communities+and+Crime">
            Communities and Crime
           </a>
          </b>
         </p>
        </td>
       </tr>
      </table>
     </td>
     <!-- <td><p class="normal">Communities within the United States. The data combines socio-ec
onomic data from the 1990 US Census, law enforcement data from the 1990 US LEMAS survey, and crime
data from the 1995 FBI UCR. </p></td> -->
     <td>
      <p class="normal">
       Multivariate
      </p>
     </td>
     <td>
      <p class="normal">
       Regression
      </p>
     </td>
     <td>
      <p class="normal">
       Real
      </p>
     </td>
     <td>
      <p class="normal">
       1994
      </p>
     </td>
     <td>
      <p class="normal">
       128
      </p>
     </td>
```

```
        </td>
        <td>
         <p class="normal">
          2009
         </p>
        </td>
        <!-- <td><p class="normal">Social </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Acute+Inflammations">
             <img border="1" src="assets/MLimages/SmallLarge184.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Acute+Inflammations">
               Acute Inflammations
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The data was created by a medical expert as a data set to test
the expert system,
which will perform the presumptive diagnosis of two diseases of the urinary system.
 </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Categorical, Integer
         </p>
        </td>
        <td>
         <p class="normal">
          120
         </p>
        </td>
        <td>
         <p class="normal">
          6
         </p>
        </td>
        <td>
         <p class="normal">
          2009
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Wine+Quality">
             <img border="1" src="assets/MLimages/SmallLarge186.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Wine+Quality">
```

```
            <a href="datasets/Wine+Quality">
             Wine Quality
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Two datasets are included, related to red and white vinho verde
wine samples, from the north of Portugal. The goal is to model wine quality based on
physicochemical tests (see [Cortez et al., 2009], http://www3.dsi.uminho.pt/pcortez/wine/). <
/p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        4898
       </p>
      </td>
      <td>
       <p class="normal">
        12
       </p>
      </td>
      <td>
       <p class="normal">
        2009
       </p>
      </td>
      <!-- <td><p class="normal">Business </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/URL+Reputation">
           <img border="1" src="assets/MLimages/SmallLarge187.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/URL+Reputation">
             URL Reputation
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Anonymized 120-day subset of the ICML-09 URL data containing 2.4
million examples and 3.2 million features. </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
```

```
        </td>
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          2396130
         </p>
        </td>
        <td>
         <p class="normal">
          3231961
         </p>
        </td>
        <td>
         <p class="normal">
          2009
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/p53+Mutants">
             <img border="1" src="assets/MLimages/SmallLarge188.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/p53+Mutants">
               p53 Mutants
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The goal is to model mutant p53 transcriptional activity (active
vs inactive) based on data extracted from biophysical simulations.
 </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          16772
         </p>
        </td>
        <td>
         <p class="normal">
          5409
         </p>
        </td>
        <td>
         <p class="normal">
          2010
         </p>
        </td>
```

```
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Parkinsons+Telemonitoring">
             <img border="1" src="assets/MLimages/SmallLarge174.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Parkinsons+Telemonitoring">
               Parkinsons Telemonitoring
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Oxford Parkinson's Disease Telemonitoring Dataset </p></td>
-->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          5875
         </p>
        </td>
        <td>
         <p class="normal">
          26
         </p>
        </td>
        <td>
         <p class="normal">
          2009
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Demospongiae">
             <img border="1" src="assets/MLimages/SmallLarge190.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Demospongiae">
               Demospongiae
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
```

```
      </td>
      <!-- <td><p class="normal">Marine sponges of the Demospongiae class classification
domain. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
       <p class="normal">
        503
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        2010
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Opinosis+Opinion+%26frasl%3B+Review">
           <img border="1" src="assets/MLimages/SmallLarge191.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Opinosis+Opinion+%26frasl%3B+Review">
             Opinosis Opinion ⁄ Review
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This dataset contains sentences extracted from user reviews on a
given topic. Example topics are “performance of Toyota Camry” and “sound quality of ipod nano”. &n
bsp;</p></td> -->
      <td>
       <p class="normal">
        Text
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        51
       </p>
      </td>
      <td>
```

```
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         2010
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Breast+Tissue">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Breast+Tissue">
              Breast Tissue
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Dataset with electrical impedance measurements of freshly excised
tissue samples from the breast. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         106
        </p>
       </td>
       <td>
        <p class="normal">
         10
        </p>
       </td>
       <td>
        <p class="normal">
         2010
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Cardiotocography">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
```

```
            <a href="datasets/Cardiotocography">
             Cardiotocography
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The dataset consists of measurements of fetal heart rate (FHR) ar
d uterine contraction (UC) features on cardiotocograms classified by expert obstetricians. </
p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        2126
       </p>
      </td>
      <td>
       <p class="normal">
        23
       </p>
      </td>
      <td>
       <p class="normal">
        2010
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Wall-Following+Robot+Navigation+Data">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Wall-Following+Robot+Navigation+Data">
             Wall-Following Robot Navigation Data
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The data were collected as the SCITOS G5 robot navigates through
the room following the wall in a clockwise direction, for 4 rounds, using 24 ultrasound sensors ar
ranged circularly around its 'waist'. </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Sequential
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
```

```
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          5456
         </p>
        </td>
        <td>
         <p class="normal">
          24
         </p>
        </td>
        <td>
         <p class="normal">
          2010
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Spoken+Arabic+Digit">
             <img border="1" src="assets/MLimages/SmallLarge195.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Spoken+Arabic+Digit">
               Spoken Arabic Digit
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This dataset contains timeseries of mel-frequency cepstrum
coefficients (MFCCs) corresponding to spoken Arabic digits.  Includes data from 44 male and 44 fem
ale native Arabic speakers. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          8800
         </p>
        </td>
        <td>
         <p class="normal">
          13
         </p>
        </td>
        <td>
         <p class="normal">
          2010
         </p>
        </td>
```

```
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Localization+Data+for+Person+Activity">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Localization+Data+for+Person+Activity">
               Localization Data for Person Activity
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Data contains recordings of five people performing different
activities. Each person wore four sensors (tags) while performing the same scenario five times. &n
bsp;</p></td> -->
        <td>
         <p class="normal">
          Univariate, Sequential, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          164860
         </p>
        </td>
        <td>
         <p class="normal">
          8
         </p>
        </td>
        <td>
         <p class="normal">
          2010
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/AutoUniv">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/AutoUniv">
               AutoUniv
              </a>
             </b>
            </p>
           </td>
          </tr>
```

```
        </table>
       </td>
       <!-- <td><p class="normal">AutoUniv is an advanced data generator for classifications
tasks. The aim is to reflect the nuances and heterogeneity of real data.  Data can be generated in
.csv, ARFF or C4.5 formats. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Categorical, Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         2010
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Steel+Plates+Faults">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Steel+Plates+Faults">
              Steel Plates Faults
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">A dataset of steel plates' faults, classified into 7 different ty
pes.
The goal was to train machine learning for automatic pattern recognition.
 </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
```

```
-
1941
</p>
</td>
<td>
<p class="normal">
27
</p>
</td>
<td>
<p class="normal">
2010
</p>
</td>
<!-- <td><p class="normal">Physical </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
<td>
<table>
<tr>
<td>
<a href="datasets/MiniBooNE+particle+identification">
<img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>
</td>
<td>
<p class="normal">
<b>
<a href="datasets/MiniBooNE+particle+identification">
MiniBooNE particle identification
</a>
</b>
</p>
</td>
</tr>
</table>
</td>
<!-- <td><p class="normal">This dataset is taken from the MiniBooNE experiment and is used
to distinguish electron neutrinos (signal) from muon neutrinos (background). </p></td> -->
<td>
<p class="normal">
Multivariate
</p>
</td>
<td>
<p class="normal">
Classification
</p>
</td>
<td>
<p class="normal">
Real
</p>
</td>
<td>
<p class="normal">
130065
</p>
</td>
<td>
<p class="normal">
50
</p>
</td>
<td>
<p class="normal">
2010
</p>
</td>
<!-- <td><p class="normal">Physical </p></td> -->
</tr>
<tr>
<td>
<table>
<tr>
<td>
<a href="datasets/YearPredictionMSD">
<img border="1" src="assets/MLimages/SmallLarge203.jpg"/>
```

```
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/YearPredictionMSD">
             YearPredictionMSD
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Prediction of the release year of a song from audio features.
Songs are mostly western, commercial tracks ranging from 1922 to 2011, with a peak in the year 200
0s. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        515345
       </p>
      </td>
      <td>
       <p class="normal">
        90
       </p>
      </td>
      <td>
       <p class="normal">
        2011
       </p>
      </td>
      <!-- <td><p class="normal">Other </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/PEMS-SF">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/PEMS-SF">
             PEMS-SF
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">15 months worth of daily data (440 daily records) that describes
the occupancy rate, between 0 and 1, of different car lanes of the San Francisco bay area freeways
across time. </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Time-Series
       </p>
```

```
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        440
       </p>
      </td>
      <td>
       <p class="normal">
        138672
       </p>
      </td>
      <td>
       <p class="normal">
        2011
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/OpinRank+Review+Dataset">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/OpinRank+Review+Dataset">
             OpinRank Review Dataset
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This data set contains user reviews of cars and and hotels
collected from Tripadvisor (~259,000
reviews) and Edmunds (~42,230 reviews).    </p></td> -->
      <td>
       <p class="normal">
        Text
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        2011
       </p>
```

```
</p>
</td>
<!-- <td><p class="normal">Computer </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
<td>
<table>
<tr>
<td>
<a href="datasets/Relative+location+of+CT+slices+on+axial+axis">
<img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>
</td>
<td>
<p class="normal">
<b>
<a href="datasets/Relative+location+of+CT+slices+on+axial+axis">
Relative location of CT slices on axial axis
</a>
</b>
</p>
</td>
</tr>
</table>
</td>
<!-- <td><p class="normal">The dataset consists of 384 features extracted from CT images.
The class variable is numeric and denotes the relative location of the CT slice on the axial axis
of the human body. </p></td> -->
<td>
<p class="normal">
Domain-Theory
</p>
</td>
<td>
<p class="normal">
Regression
</p>
</td>
<td>
<p class="normal">
Real
</p>
</td>
<td>
<p class="normal">
53500
</p>
</td>
<td>
<p class="normal">
386
</p>
</td>
<td>
<p class="normal">
2011
</p>
</td>
<!-- <td><p class="normal">Computer </p></td> -->
</tr>
<tr>
<td>
<table>
<tr>
<td>
<a href="datasets/Online+Handwritten+Assamese+Characters+Dataset">
<img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>
</td>
<td>
<p class="normal">
<b>
<a href="datasets/Online+Handwritten+Assamese+Characters+Dataset">
Online Handwritten Assamese Characters Dataset
</a>
</b>
</p>
</td>
```

```
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This is a dataset of 8235 online handwritten assamese
characters. The “online” process involves capturing of data as text is written on a digitizing tab
let with an electronic pen. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Sequential
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         8235
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         2011
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/PubChem+Bioassay+Data">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/PubChem+Bioassay+Data">
              PubChem Bioassay Data
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">These highly imbalanced bioassay datasets are from the differing
types of screening that can be performed using HTS technology. 21 datasets were created from 12 bi
oassays. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
```

```
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        2011
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Record+Linkage+Comparison+Patterns">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Record+Linkage+Comparison+Patterns">
             Record Linkage Comparison Patterns
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Element-wise comparison of records with personal data from a
record linkage setting. The task is to decide from a comparison pattern whether the underlying rec
ords belong to one person. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        5749132
       </p>
      </td>
      <td>
       <p class="normal">
        12
       </p>
      </td>
      <td>
       <p class="normal">
        2011
       </p>
      </td>
      <!-- <td><p class="normal">Other </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Communities+and+Crime+Unnormalized">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
```

```
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Communities+and+Crime+Unnormalized">
              Communities and Crime Unnormalized
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Communities in the US. Data combines socio-economic data from th
e '90 Census, law enforcement data from the 1990 Law Enforcement Management and Admin Stats survey
, and crime data from the 1995 FBI UCR </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         2215
        </p>
       </td>
       <td>
        <p class="normal">
         147
        </p>
       </td>
       <td>
        <p class="normal">
         2011
        </p>
       </td>
       <!-- <td><p class="normal">Social </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Vertebral+Column">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Vertebral+Column">
              Vertebral Column
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Data set containing values for six biomechanical features used t
o classify orthopaedic patients into 3 classes (normal, disk hernia or spondilolysthesis) or 2 cla
sses (normal or abnormal). </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
```

```
      </p>
     </td>
     <td>
      <p class="normal">
       Classification
      </p>
     </td>
     <td>
      <p class="normal">
       Real
      </p>
     </td>
     <td>
      <p class="normal">
       310
      </p>
     </td>
     <td>
      <p class="normal">
       6
      </p>
     </td>
     <td>
      <p class="normal">
       2011
      </p>
     </td>
     <!-- <td><p class="normal"> </p></td> -->
    </tr>
    <tr bgcolor="DDEEFF">
     <td>
      <table>
       <tr>
        <td>
         <a href="datasets/EMG+Physical+Action+Data+Set">
          <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
         </a>
        </td>
        <td>
         <p class="normal">
          <b>
           <a href="datasets/EMG+Physical+Action+Data+Set">
            EMG Physical Action Data Set
           </a>
          </b>
         </p>
        </td>
       </tr>
      </table>
     </td>
     <!-- <td><p class="normal">The Physical Action Data Set includes 10 normal and 10
aggressive physical actions that measure the human activity. The data have been collected by 4 sub
jects using the Delsys EMG wireless apparatus. </p></td> -->
     <td>
      <p class="normal">
       Time-Series
      </p>
     </td>
     <td>
      <p class="normal">
       Classification
      </p>
     </td>
     <td>
      <p class="normal">
       Real
      </p>
     </td>
     <td>
      <p class="normal">
       10000
      </p>
     </td>
     <td>
      <p class="normal">
       8
      </p>
     </td>
```

```
        </td>
        <td>
         <p class="normal">
          2011
         </p>
        </td>
        <!-- <td><p class="normal">Physical </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Vicon+Physical+Action+Data+Set">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Vicon+Physical+Action+Data+Set">
               Vicon Physical Action Data Set
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The Physical Action Data Set includes 10 normal and 10
aggressive physical actions that measure the human activity. The data have been collected by 10 su
bjects using the Vicon 3D tracker. </p></td> -->
        <td>
         <p class="normal">
          Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          3000
         </p>
        </td>
        <td>
         <p class="normal">
          27
         </p>
        </td>
        <td>
         <p class="normal">
          2011
         </p>
        </td>
        <!-- <td><p class="normal">Physical </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Amazon+Commerce+reviews+set">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Amazon+Commerce+reviews+set">
```

```
            Amazon Commerce reviews set
           </a>
          </b>
         </p>
        </td>
       </tr>
      </table>
     </td>
     <!-- <td><p class="normal">The dataset is used for authorship identification in online Write
print which is a new research field of pattern recognition.  </p></td> -->
     <td>
      <p class="normal">
       Multivariate, Text, Domain-Theory
      </p>
     </td>
     <td>
      <p class="normal">
       Classification
      </p>
     </td>
     <td>
      <p class="normal">
       Real
      </p>
     </td>
     <td>
      <p class="normal">
       1500
      </p>
     </td>
     <td>
      <p class="normal">
       10000
      </p>
     </td>
     <td>
      <p class="normal">
       2011
      </p>
     </td>
     <!-- <td><p class="normal">Physical </p></td> -->
    </tr>
    <tr>
     <td>
      <table>
       <tr>
        <td>
         <a href="datasets/Amazon+Access+Samples">
          <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
         </a>
        </td>
        <td>
         <p class="normal">
          <b>
           <a href="datasets/Amazon+Access+Samples">
            Amazon Access Samples
           </a>
          </b>
         </p>
        </td>
       </tr>
      </table>
     </td>
     <!-- <td><p class="normal">Amazon's InfoSec is getting smarter about the way Access data is
leveraged. This is an anonymized sample of access provisioned within the company. </p></td> -
->
     <td>
      <p class="normal">
       Time-Series, Domain-Theory
      </p>
     </td>
     <td>
      <p class="normal">
       Regression, Clustering, Causal-Discovery
      </p>
     </td>
     <td>
```

```
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         30000
        </p>
       </td>
       <td>
        <p class="normal">
         20000
        </p>
       </td>
       <td>
        <p class="normal">
         2011
        </p>
       </td>
       <!-- <td><p class="normal">Business </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Reuter_50_50">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Reuter_50_50">
              Reuter_50_50
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The dataset is used for authorship identification in online Write
print which is a new research field of pattern recognition.  </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Text, Domain-Theory
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         2500
        </p>
       </td>
       <td>
        <p class="normal">
         10000
        </p>
       </td>
       <td>
        <p class="normal">
         2011
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
```

```
        <table>
         <tr>
          <td>
           <a href="datasets/Farm+Ads">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Farm+Ads">
              Farm Ads
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data was collected from text ads found on twelve websites t
hat deal with various farm animal related topics.  The binary labels are based on whether or not t
he content owner approves of the ad. </p></td> -->
       <td>
        <p class="normal">
         Text
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         4143
        </p>
       </td>
       <td>
        <p class="normal">
         54877
        </p>
       </td>
       <td>
        <p class="normal">
         2011
        </p>
       </td>
       <!-- <td><p class="normal">Business </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/DBWorld+e-mails">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/DBWorld+e-mails">
              DBWorld e-mails
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">It contains 64 e-mails which I have manually collected from
DBWorld mailing list. They are classified in: 'announces of conferences' and 'everything
else'. </p></td> -->
```

```
        <td>
         <p class="normal">
          Text
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          64
         </p>
        </td>
        <td>
         <p class="normal">
          4702
         </p>
        </td>
        <td>
         <p class="normal">
          2011
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/KEGG+Metabolic+Relation+Network+%28Directed%29">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/KEGG+Metabolic+Relation+Network+%28Directed%29">
               KEGG Metabolic Relation Network (Directed)
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">KEGG Metabolic pathways modeled as directed relation network. Va
riety of graphical features presented. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Univariate, Text
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Regression, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          53414
         </p>
        </td>
        <td>
         <p class="normal">
          24
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         2011
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/KEGG+Metabolic+Reaction+Network+%28Undirected%29">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/KEGG+Metabolic+Reaction+Network+%28Undirected%29">
              KEGG Metabolic Reaction Network (Undirected)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">KEGG Metabolic pathways modeled as un-directed reaction network.
Variety of graphical features presented. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Univariate, Text
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         65554
        </p>
       </td>
       <td>
        <p class="normal">
         29
        </p>
       </td>
       <td>
        <p class="normal">
         2011
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Bank+Marketing">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Bank+Marketing">
```

```
              Bank Marketing
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The data is related with direct marketing campaigns (phone
calls) of a Portuguese banking institution. The classification goal is to predict if the client wi
ll subscribe a term deposit (variable y). </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         45211
        </p>
       </td>
       <td>
        <p class="normal">
         17
        </p>
       </td>
       <td>
        <p class="normal">
         2012
        </p>
       </td>
       <!-- <td><p class="normal">Business </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/YouTube+Comedy+Slam+Preference+Data">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/YouTube+Comedy+Slam+Preference+Data">
              YouTube Comedy Slam Preference Data
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset provides user vote data on which video from a pair
of videos is funnier collected on YouTube Comedy Slam. The task is to automatically predict this p
reference based on video metadata. </p></td> -->
       <td>
        <p class="normal">
         Text
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
```

```
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          1138562
         </p>
        </td>
        <td>
         <p class="normal">
          3
         </p>
        </td>
        <td>
         <p class="normal">
          2012
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Gas+Sensor+Array+Drift+Dataset">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Gas+Sensor+Array+Drift+Dataset">
               Gas Sensor Array Drift Dataset
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This archive contains 13910 measurements from 16 chemical
sensors utilized in simulations for drift compensation in a discrimination task of 6 gases at vari
ous levels of concentrations. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          13910
         </p>
        </td>
        <td>
         <p class="normal">
          128
         </p>
        </td>
        <td>
         <p class="normal">
          2012
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
```

```
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/ILPD+%28Indian+Liver+Patient+Dataset%29">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/ILPD+%28Indian+Liver+Patient+Dataset%29">
              ILPD (Indian Liver Patient Dataset)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data set contains 10 variables that are age, gender, total
Bilirubin, direct Bilirubin, total proteins, albumin, A/G ratio, SGPT, SGOT and Alkphos. </p>
</td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         583
        </p>
       </td>
       <td>
        <p class="normal">
         10
        </p>
       </td>
       <td>
        <p class="normal">
         2012
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/OPPORTUNITY+Activity+Recognition">
            <img border="1" src="assets/MLimages/SmallLarge226.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/OPPORTUNITY+Activity+Recognition">
              OPPORTUNITY Activity Recognition
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
```

```
      <!-- <td><p class="normal">The OPPORTUNITY Dataset for Human Activity Recognition from
Wearable, Object, and Ambient Sensors is a dataset devised to benchmark human activity recognition
algorithms (classification, automatic data segmentation, sensor fusion, feature extraction,
etc). </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        2551
       </p>
      </td>
      <td>
       <p class="normal">
        242
       </p>
      </td>
      <td>
       <p class="normal">
        2012
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Nomao">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Nomao">
             Nomao
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Nomao collects data about places (name, phone, localization...)
from many sources.
Deduplication consists in detecting what data refer to the same place.
Instances in the dataset compare 2 spots. </p></td> -->
      <td>
       <p class="normal">
        Univariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
```

```
        <p class="normal">
         34465
        </p>
       </td>
       <td>
        <p class="normal">
         120
        </p>
       </td>
       <td>
        <p class="normal">
         2012
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/SMS+Spam+Collection">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/SMS+Spam+Collection">
              SMS Spam Collection
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The SMS Spam Collection is a public set of SMS labeled messages
that have been collected for mobile phone spam research. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Text, Domain-Theory
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         5574
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         2012
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Skin+Segmentation">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
```

```
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Skin+Segmentation">
             Skin Segmentation
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The Skin Segmentation dataset is constructed over B, G, R color s
pace. Skin and Nonskin dataset is generated using skin textures from face images of diversity of a
ge, gender, and race people. </p></td> -->
      <td>
       <p class="normal">
        Univariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        245057
       </p>
      </td>
      <td>
       <p class="normal">
        4
       </p>
      </td>
      <td>
       <p class="normal">
        2012
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Planning+Relax">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Planning+Relax">
             Planning Relax
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The dataset concerns with the classification of two mental
stages from recorded EEG signals: Planning (during imagination of motor act) and Relax state. &nbs
p;</p></td> -->
      <td>
       <p class="normal">
        Univariate
       </p>
```

```
</p>
</td>
<td>
 <p class="normal">
  Classification
 </p>
</td>
<td>
 <p class="normal">
  Real
 </p>
</td>
<td>
 <p class="normal">
  182
 </p>
</td>
<td>
 <p class="normal">
  13
 </p>
</td>
<td>
 <p class="normal">
  2012
 </p>
</td>
<!-- <td><p class="normal">Computer </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
<td>
 <table>
  <tr>
   <td>
    <a href="datasets/PAMAP2+Physical+Activity+Monitoring">
     <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
    </a>
   </td>
   <td>
    <p class="normal">
     <b>
      <a href="datasets/PAMAP2+Physical+Activity+Monitoring">
       PAMAP2 Physical Activity Monitoring
      </a>
     </b>
    </p>
   </td>
  </tr>
 </table>
</td>
<!-- <td><p class="normal">The PAMAP2 Physical Activity Monitoring dataset contains data of 18 different physical activities, performed by 9 subjects wearing 3 inertial measurement units and a heart rate monitor. </p></td> -->
<td>
 <p class="normal">
  Multivariate, Time-Series
 </p>
</td>
<td>
 <p class="normal">
  Classification
 </p>
</td>
<td>
 <p class="normal">
  Real
 </p>
</td>
<td>
 <p class="normal">
  3850505
 </p>
</td>
<td>
 <p class="normal">
  52
 </p>
</td>
```

</td>

```
      <td>
       <p class="normal">
        2012
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Restaurant+%26+consumer+data">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Restaurant+%26+consumer+data">
             Restaurant &amp; consumer data
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The dataset was obtained from a recommender system prototype.
The task was to generate a top-n list of restaurants according to the consumer preferences.  
</p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        138
       </p>
      </td>
      <td>
       <p class="normal">
        47
       </p>
      </td>
      <td>
       <p class="normal">
        2012
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/CNAE-9">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/CNAE-9">
             CNAE-9
            </a>
           </b>
```

```
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This is a data set containing 1080 documents of free text
business descriptions of Brazilian companies categorized into a
subset of 9 categories </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Text
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         1080
        </p>
       </td>
       <td>
        <p class="normal">
         857
        </p>
       </td>
       <td>
        <p class="normal">
         2012
        </p>
       </td>
       <!-- <td><p class="normal">Business </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Individual+household+electric+power+consumption">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Individual+household+electric+power+consumption">
              Individual household electric power consumption
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Measurements of electric power consumption in one household with
a one-minute sampling rate over a period of almost 4 years. Different electrical quantities and so
me sub-metering values are available. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Regression, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
```

```
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        2075259
       </p>
      </td>
      <td>
       <p class="normal">
        9
       </p>
      </td>
      <td>
       <p class="normal">
        2012
       </p>
      </td>
      <!-- <td><p class="normal">Physical </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/seeds">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/seeds">
             seeds
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Measurements of geometrical properties of kernels belonging to th
ree different varieties of wheat. A soft X-ray technique and GRAINS package were used to construct
all seven, real-valued attributes. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Clustering
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        210
       </p>
      </td>
      <td>
       <p class="normal">
        7
       </p>
      </td>
      <td>
       <p class="normal">
        2012
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr>
      <td>
```

```
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Northix">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Northix">
             Northix
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Northix is designed to be a schema matching benchmark problem
for data integration of two entity relationship databases.  </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Univariate, Text
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer, Real
       </p>
      </td>
      <td>
       <p class="normal">
        115
       </p>
      </td>
      <td>
       <p class="normal">
        200
       </p>
      </td>
      <td>
       <p class="normal">
        2012
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/QtyT40I10D100K">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/QtyT40I10D100K">
             QtyT40I10D100K
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Since there is no numerical sequential data stream available in s
tandard data sets, this data set is generated from the original T40I10D100K data set </p></td
```

```
> -->
        <td>
         <p class="normal">
          Sequential
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          3960456
         </p>
        </td>
        <td>
         <p class="normal">
          4
         </p>
        </td>
        <td>
         <p class="normal">
          2012
         </p>
        </td>
        <!-- <td><p class="normal"> </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Legal+Case+Reports">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Legal+Case+Reports">
               Legal Case Reports
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">A textual corpus of 4000 legal cases for automatic summarization
and citation analysis. For each document we collect catchphrases, citations sentences, citation ca
tchphrases and citation classes. </p></td> -->
        <td>
         <p class="normal">
          Text
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
```

```
        </td>
        <td>
         <p class="normal">
          2012
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Human+Activity+Recognition+Using+Smartphones">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Human+Activity+Recognition+Using+Smartphones">
               Human Activity Recognition Using Smartphones
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Human Activity Recognition database built from the recordings of
30 subjects performing activities of daily living (ADL) while carrying a waist-mounted smartphone
with embedded inertial sensors. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          10299
         </p>
        </td>
        <td>
         <p class="normal">
          561
         </p>
        </td>
        <td>
         <p class="normal">
          2012
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/One-hundred+plant+species+leaves+data+set">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/One-hundred+plant+species+leaves+data+set">
               One-hundred plant species leaves data set
```

```
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Sixteen samples of leaf each of one-hundred plant species. For
each sample, a shape descriptor, fine scale margin and texture histogram are given. </p></td>
-->
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        1600
       </p>
      </td>
      <td>
       <p class="normal">
        64
       </p>
      </td>
      <td>
       <p class="normal">
        2012
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Energy+efficiency">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Energy+efficiency">
             Energy efficiency
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This study looked into assessing the heating load and cooling
load requirements of buildings (that is, energy efficiency) as a function of building
parameters. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Regression
       </p>
      </td>
      <td>
       <p class="normal">
```

```
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          768
         </p>
        </td>
        <td>
         <p class="normal">
          8
         </p>
        </td>
        <td>
         <p class="normal">
          2012
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Yacht+Hydrodynamics">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Yacht+Hydrodynamics">
               Yacht Hydrodynamics
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Delft data set, used to predict the hydodynamic performance of
sailing yachts from dimensions and velocity. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          308
         </p>
        </td>
        <td>
         <p class="normal">
          7
         </p>
        </td>
        <td>
         <p class="normal">
          2013
         </p>
        </td>
        <!-- <td><p class="normal">Physical </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
```

```
        <table>
         <tr>
          <td>
           <a href="datasets/Fertility">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Fertility">
              Fertility
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">100 volunteers provide a semen sample analyzed according to the
WHO 2010 criteria. Sperm concentration are related to socio-demographic data, environmental
factors, health status, and life habits </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         100
        </p>
       </td>
       <td>
        <p class="normal">
         10
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Daphnet+Freezing+of+Gait">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Daphnet+Freezing+of+Gait">
              Daphnet Freezing of Gait
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset contains the annotated readings of 3 acceleration s
ensors at the hip and leg of Parkinson's disease patients that experience freezing of gait (FoG) d
```

uring walking tasks.
  </td> -->
<td>

Multivariate, Time-Series

</td>
<td>

Classification

</td>
<td>

Real

</td>
<td>

237

</td>
<td>

9

</td>
<td>

2013

</td>
<!-- <td> Life </td> -->
</tr>
<tr bgcolor="DDEEFF">
<td>
<table>
<tr>
<td>
<a href="datasets/3D+Road+Network+%28North+Jutland%2C+Denmark%29">
<img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>
</td>
<td>

<a href="datasets/3D+Road+Network+%28North+Jutland%2C+Denmark%29">
3D Road Network (North Jutland, Denmark)
</a>

</td>
</tr>
</table>
</td>
<!-- <td> 3D road network with highly accurate elevation information (+-20 cm) from Denmark used in eco-routing and fuel/Co2-estimation routing algorithms. </td> -->
<td>

Sequential, Text

</td>
<td>

Regression, Clustering

</td>
<td>

Real

</td>
<td>

434874

```
        </td>
        <td>
         <p class="normal">
          4
         </p>
        </td>
        <td>
         <p class="normal">
          2013
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/ISTANBUL+STOCK+EXCHANGE">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/ISTANBUL+STOCK+EXCHANGE">
               ISTANBUL STOCK EXCHANGE
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Data sets includes returns of Istanbul Stock Exchange with seven
other international index; SP, DAX, FTSE, NIKKEI, BOVESPA, MSCE_EU, MSCI_EM from Jun 5, 2009 to Fe
b 22, 2011. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Univariate, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          536
         </p>
        </td>
        <td>
         <p class="normal">
          8
         </p>
        </td>
        <td>
         <p class="normal">
          2013
         </p>
        </td>
        <!-- <td><p class="normal">Business </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Buzz+in+social+media+">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
```

```
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Buzz+in+social+media+">
              Buzz in social media
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data-set contains examples of buzz events from two
different social networks: Twitter, and Tom's Hardware, a forum network focusing on new technology
with more conservative dynamics. </p></td> -->
       <td>
        <p class="normal">
         Time-Series, Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Regression, Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         140000
        </p>
       </td>
       <td>
        <p class="normal">
         77
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/First-order+theorem+proving">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/First-order+theorem+proving">
              First-order theorem proving
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Given a theorem, predict which of five heuristics will give the
fastest proof when used by a first-order prover. A sixth prediction declines to attempt a proof, s
hould the theorem be too difficult. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
```

```
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          6118
         </p>
        </td>
        <td>
         <p class="normal">
          51
         </p>
        </td>
        <td>
         <p class="normal">
          2013
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a
href="datasets/Wearable+Computing%3A+Classification+of+Body+Postures+and+Movements+%28PUC-Rio%29">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a
href="datasets/Wearable+Computing%3A+Classification+of+Body+Postures+and+Movements+%28PUC-Rio%29">
               Wearable Computing: Classification of Body Postures and Movements (PUC-Rio)
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">A dataset with 5 classes (sitting-down, standing-up, standing,
walking, and sitting) collected on 8 hours of activities of 4 healthy subjects. We also
established a baseline performance index. </p></td> -->
        <td>
         <p class="normal">
          Sequential
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          165632
         </p>
        </td>
        <td>
         <p class="normal">
          18
         </p>
```

```
      </td>
      <td>
       <p class="normal">
        2013
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Gas+sensor+arrays+in+open+sampling+settings">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Gas+sensor+arrays+in+open+sampling+settings">
             Gas sensor arrays in open sampling settings
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The dataset contains 18000 time-series recordings from a chemical
detection platform at six different locations in a wind tunnel facility in response to ten high-pr
iority chemical gaseous substances </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        18000
       </p>
      </td>
      <td>
       <p class="normal">
        1950000
       </p>
      </td>
      <td>
       <p class="normal">
        2013
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Climate+Model+Simulation+Crashes">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Climate+Model+Simulation+Crashes">
```

```
              Climate Model Simulation Crashes
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Given Latin hypercube samples of 18 climate model input parameter
values, predict climate model simulation crashes and determine the parameter value combinations th
at cause the failures. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         540
        </p>
       </td>
       <td>
        <p class="normal">
         18
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Physical </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/MicroMass">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/MicroMass">
              MicroMass
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">A dataset to explore machine learning approaches for the
identification of microorganisms from mass-spectrometry data. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
```

```
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         931
        </p>
       </td>
       <td>
        <p class="normal">
         1300
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/QSAR+biodegradation">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/QSAR+biodegradation">
              QSAR biodegradation
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Data set containing values for 41 attributes (molecular
descriptors) used to classify 1055 chemicals into 2 classes (ready and not ready
biodegradable). </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         1055
        </p>
       </td>
       <td>
        <p class="normal">
         41
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
```

```
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/BLOGGER">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/BLOGGER">
              BLOGGER
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">In this paper, we look for to recognize the causes of users tend
to cyber space in Kohkiloye and Boyer Ahmad Province in
Iran </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         100
        </p>
       </td>
       <td>
        <p class="normal">
         6
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Daily+and+Sports+Activities">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Daily+and+Sports+Activities">
              Daily and Sports Activities
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The dataset comprises motion sensor data of 19 daily and sports
```

```
<!-- <td><p class="normal">The dataset comprises motion sensor data of 19 daily and sports
activities each performed by 8 subjects in their own style for 5 minutes. Five Xsens MTx units are
used on the torso, arms, and legs.
 </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Clustering
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        9120
       </p>
      </td>
      <td>
       <p class="normal">
        5625
       </p>
      </td>
      <td>
       <p class="normal">
        2013
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/User+Knowledge+Modeling">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/User+Knowledge+Modeling">
             User Knowledge Modeling
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">It is the real dataset about the students' knowledge status
about the subject of Electrical DC Machines. The dataset had been obtained from Ph.D.
Thesis. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Clustering
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
       <p class="normal">
        403
```

```
        403
        </p>
       </td>
       <td>
        <p class="normal">
         5
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a
href="datasets/Reuters+RCV1+RCV2+Multilingual%2C+Multiview+Text+Categorization+Test+collection">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a
href="datasets/Reuters+RCV1+RCV2+Multilingual%2C+Multiview+Text+Categorization+Test+collection">
              Reuters RCV1 RCV2 Multilingual, Multiview Text Categorization Test collection
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This test collection contains feature characteristics of
documents originally written in five different languages and their translations, over a common set
of 6 categories.  </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         111740
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/NYSK">
```

```
          <a href="datasets/NYSK">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/NYSK">
             NYSK
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">NYSK (New York v. Strauss-Kahn) is a collection of English news
articles about the case relating to allegations of sexual assault against the former IMF director
Dominique Strauss-Kahn (May 2011). </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Sequential, Text
       </p>
      </td>
      <td>
       <p class="normal">
        Clustering
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        10421
       </p>
      </td>
      <td>
       <p class="normal">
        7
       </p>
      </td>
      <td>
       <p class="normal">
        2013
       </p>
      </td>
      <!-- <td><p class="normal">Social </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Turkiye+Student+Evaluation">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Turkiye+Student+Evaluation">
             Turkiye Student Evaluation
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This data set contains a total 5820 evaluation scores provided b
y students from Gazi University in Ankara (Turkey). There is a total of 28 course specific
questions and additional 5 attributes. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         5820
        </p>
       </td>
       <td>
        <p class="normal">
         33
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a
href="datasets/ser+Knowledge+Modeling+Data+%28Students%27+Knowledge+Levels+on+DC+Electrical+Machine
">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a
href="datasets/ser+Knowledge+Modeling+Data+%28Students%27+Knowledge+Levels+on+DC+Electrical+Machine
">
              ser Knowledge Modeling Data (Students' Knowledge Levels on DC Electrical Machines)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The dataset is about the users' learning activities and
knowledge levels on subjects of DC  Electrical Machines. The dataset had been obtained from online
web-courses  and reported in my Ph.D. Thesis. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         403
        </p>
       </td>
       <td>
```

```
        <p class="normal">
         5
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/EEG+Eye+State">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/EEG+Eye+State">
              EEG Eye State
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The data set consists of 14 EEG values and a value indicating
the eye state. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Sequential, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         14980
        </p>
       </td>
       <td>
        <p class="normal">
         15
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Physicochemical+Properties+of+Protein+Tertiary+Structure">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
```

```
            <p>
             <a href="datasets/Physicochemical+Properties+of+Protein+Tertiary+Structure">
              Physicochemical Properties of Protein Tertiary Structure
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This is a data set of Physicochemical Properties of Protein Terti
ary Structure. The data set is taken from CASP 5-9. There are 45730 decoys and size varying from 0
to 21 armstrong. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         45730
        </p>
       </td>
       <td>
        <p class="normal">
         9
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/seismic-bumps">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/seismic-bumps">
              seismic-bumps
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The data describe the problem of high energy (higher than 10^4
J) seismic bumps forecasting in a coal
mine. Data come from two of longwalls located in a Polish coal mine. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         2584
        </p>
       </td>
       <td>
        <p class="normal">
         19
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/banknote+authentication">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/banknote+authentication">
              banknote authentication
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Data were extracted from images that were taken for the
evaluation of an authentication procedure for banknotes. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         1372
        </p>
       </td>
       <td>
        <p class="normal">
         5
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
```

```
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/USPTO+Algorithm+Challenge%2C+run+by+NASA-
Harvard+Tournament+Lab+and+TopCoder++++Problem%3A+Pat">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/USPTO+Algorithm+Challenge%2C+run+by+NASA-
Harvard+Tournament+Lab+and+TopCoder++++Problem%3A+Pat">
              USPTO Algorithm Challenge, run by NASA-Harvard Tournament Lab and TopCoder    Problem
: Pat
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Data used for USPTO Algorithm Competition. Contains drawing
pages from US patents with manually labeled figure and part labels. </p></td> -->
       <td>
        <p class="normal">
         Domain-Theory
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         306
        </p>
       </td>
       <td>
        <p class="normal">
         5
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/YouTube+Multiview+Video+Games+Dataset">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/YouTube+Multiview+Video+Games+Dataset">
              YouTube Multiview Video Games Dataset
             </a>
            </b>
           </p>
```

```
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset contains about 120k instances, each described by 13
feature types, with class information, specially useful for exploring multiview topics
(cotraining, ensembles, clustering,..). </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Text
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         120000
        </p>
       </td>
       <td>
        <p class="normal">
         1000000
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Gas+Sensor+Array+Drift+Dataset+at+Different+Concentrations">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Gas+Sensor+Array+Drift+Dataset+at+Different+Concentrations">
              Gas Sensor Array Drift Dataset at Different Concentrations
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This archive contains 13910 measurements from 16 chemical
sensors exposed to 6 different gases at various concentration levels. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression, Clustering, Causa
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
```

```
        <td>
         <p class="normal">
          13910
         </p>
        </td>
        <td>
         <p class="normal">
          129
         </p>
        </td>
        <td>
         <p class="normal">
          2013
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a
href="datasets/Activities+of+Daily+Living+%28ADLs%29+Recognition+Using+Binary+Sensors">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a
href="datasets/Activities+of+Daily+Living+%28ADLs%29+Recognition+Using+Binary+Sensors">
               Activities of Daily Living (ADLs) Recognition Using Binary Sensors
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This dataset comprises information regarding the ADLs performed
by two users on a daily basis in their
own homes.   </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Sequential, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          2747
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          2013
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
```

```
          <td>
           <a href="datasets/SkillCraft1+Master+Table+Dataset">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/SkillCraft1+Master+Table+Dataset">
              SkillCraft1 Master Table Dataset
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data was used in Thompson et al. (2013). A list of possible
game actions is discussed in Thompson, Blair, Chen, & Henrey (2013). </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         3395
        </p>
       </td>
       <td>
        <p class="normal">
         20
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Game </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Weight+Lifting+Exercises+monitored+with+Inertial+Measurement+Units">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Weight+Lifting+Exercises+monitored+with+Inertial+Measurement+Units">
              Weight Lifting Exercises monitored with Inertial Measurement Units
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Six young health subjects were asked to perform 5 variations of t
he biceps curl weight lifting exercise. One of the variations is the one predicted by the health p
rofessional. </p></td> -->
       <td>
        <p class="normal">
```

```
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          39242
         </p>
        </td>
        <td>
         <p class="normal">
          152
         </p>
        </td>
        <td>
         <p class="normal">
          2013
         </p>
        </td>
        <!-- <td><p class="normal">Physical </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/SML2010">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/SML2010">
               SML2010
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This dataset is collected from a monitor system mounted in a
domotic house. It corresponds to approximately 40 days of monitoring data. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Sequential, Time-Series, Text
         </p>
        </td>
        <td>
         <p class="normal">
          Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          4137
         </p>
        </td>
        <td>
         <p class="normal">
          24
         </p>
```

```
        </td>
        <td>
         <p class="normal">
          2014
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Bike+Sharing+Dataset">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Bike+Sharing+Dataset">
               Bike Sharing Dataset
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This dataset contains the hourly and daily count of rental bikes
between years 2011 and 2012 in Capital bikeshare system with the corresponding weather and
seasonal information. </p></td> -->
        <td>
         <p class="normal">
          Univariate
         </p>
        </td>
        <td>
         <p class="normal">
          Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          17389
         </p>
        </td>
        <td>
         <p class="normal">
          16
         </p>
        </td>
        <td>
         <p class="normal">
          2013
         </p>
        </td>
        <!-- <td><p class="normal">Social </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Predict+keywords+activities+in+a+online+social+media">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Predict+keywords+activities+in+a+online+social+media">
```

```
              Predict keywords activities in a online social media
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The data from Twitter was collected during 360 consecutive days.
It was done by querying 1497 English keywords sampled from Wikipedia. This dataset is proposed in
a Learning to rank setting. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Sequential, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         51
        </p>
       </td>
       <td>
        <p class="normal">
         35
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Thoracic+Surgery+Data">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Thoracic+Surgery+Data">
              Thoracic Surgery Data
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The data is dedicated to classification problem related to the po
st-operative life expectancy in the lung cancer patients: class 1 - death within one year after su
rgery, class 2 - survival. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
```

```
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         470
        </p>
       </td>
       <td>
        <p class="normal">
         17
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/EMG+dataset+in+Lower+Limb">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/EMG+dataset+in+Lower+Limb">
              EMG dataset in Lower Limb
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">3 different exercises: sitting, standing and walking in the
muscles: biceps femoris, vastus medialis, rectus femoris and semitendinosus addition to goniometry
in the exercises. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         132
        </p>
       </td>
       <td>
        <p class="normal">
         5
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
```

```
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/SUSY">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/SUSY">
             SUSY
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This is a classification problem to distinguish between a signal
process which produces supersymmetric particles and a background process which does not. </p>
</td> -->
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        5000000
       </p>
      </td>
      <td>
       <p class="normal">
        18
       </p>
      </td>
      <td>
       <p class="normal">
        2014
       </p>
      </td>
      <!-- <td><p class="normal">Physical </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/HIGGS">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/HIGGS">
             HIGGS
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This is a classification problem to distinguish between a signal
process which produces Higgs bosons and a background process which does not. </p></td> -->
```

process which produces huge bursts and a background process which does not.   </td>
<td>

</td>
<td>

Classification

</td>
<td>

Real

</td>
<td>

11000000

</td>
<td>

28

</td>
<td>

2014

</td>
<!-- <td> Physical </td> -->
</tr>
<tr>
<td>
<table>
<tr>
<td>
<a href="datasets/Qualitative_Bankruptcy">
<img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>
</td>
<td>

<a href="datasets/Qualitative_Bankruptcy">
Qualitative_Bankruptcy
</a>

</td>
</tr>
</table>
</td>
<!-- <td> Predict the Bankruptcy from Qualitative parameters from experts. </td> -->
<td>

Multivariate

</td>
<td>

Classification

</td>
<td>

</td>
<td>

250

</td>
<td>

7

```
      </p>
     </td>
     <td>
      <p class="normal">
       2014
      </p>
     </td>
     <!-- <td><p class="normal">Computer </p></td> -->
    </tr>
    <tr bgcolor="DDEEFF">
     <td>
      <table>
       <tr>
        <td>
         <a href="datasets/LSVT+Voice+Rehabilitation">
          <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
         </a>
        </td>
        <td>
         <p class="normal">
          <b>
           <a href="datasets/LSVT+Voice+Rehabilitation">
            LSVT Voice Rehabilitation
           </a>
          </b>
         </p>
        </td>
       </tr>
      </table>
     </td>
     <!-- <td><p class="normal">126 samples from 14 participants, 309 features. Aim: assess wheth
er voice rehabilitation treatment lead to phonations considered 'acceptable' or 'unacceptable'
(binary class classification problem). </p></td> -->
     <td>
      <p class="normal">
       Multivariate
      </p>
     </td>
     <td>
      <p class="normal">
       Classification
      </p>
     </td>
     <td>
      <p class="normal">
       Real
      </p>
     </td>
     <td>
      <p class="normal">
       126
      </p>
     </td>
     <td>
      <p class="normal">
       309
      </p>
     </td>
     <td>
      <p class="normal">
       2014
      </p>
     </td>
     <!-- <td><p class="normal">Life </p></td> -->
    </tr>
    <tr>
     <td>
      <table>
       <tr>
        <td>
         <a href="datasets/Dataset+for+ADL+Recognition+with+Wrist-worn+Accelerometer">
          <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
         </a>
        </td>
        <td>
         <p class="normal">
          <b>
           <a href="datasets/Dataset+for+ADL+Recognition+with+Wrist-worn+Accelerometer">
```

```
            <a href="datasets/Dataset for ADL Recognition with Wrist-worn Accelerometer">
             Dataset for ADL Recognition with Wrist-worn Accelerometer
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Recordings of 16 volunteers performing 14 Activities of Daily Liv
ing (ADL) while carrying a single wrist-worn tri-axial accelerometer. </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Clustering
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        3
       </p>
      </td>
      <td>
       <p class="normal">
        2014
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Wilt">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Wilt">
             Wilt
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">High-resolution Remote Sensing data set (Quickbird). Small
number of training samples of diseased trees, large number for other land cover. Testing data set
from stratified random sample of image. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         4889
        </p>
       </td>
       <td>
        <p class="normal">
         6
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/User+Identification+From+Walking+Activity">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/User+Identification+From+Walking+Activity">
              User Identification From Walking Activity
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The dataset collects data from an Android smartphone positioned i
n the chest pocket from 22 participants walking in the wild over a predefined path.
 </p></td> -->
       <td>
        <p class="normal">
         Univariate, Sequential, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
```

```
          <td>
           <a href="datasets/Activity+Recognition+from+Single+Chest-Mounted+Accelerometer">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Activity+Recognition+from+Single+Chest-Mounted+Accelerometer">
              Activity Recognition from Single Chest-Mounted Accelerometer
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The dataset collects data from a wearable accelerometer mounted
on the chest. The dataset is intended for Activity Recognition research purposes. </p></td> -
->
       <td>
        <p class="normal">
         Univariate, Sequential, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Leaf">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Leaf">
              Leaf
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset consists in a collection of shape and texture
features extracted from digital images of leaf specimens originating from a total of 40 different
plant species. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         340
        </p>
       </td>
       <td>
        <p class="normal">
         16
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Dresses_Attribute_Sales">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Dresses_Attribute_Sales">
              Dresses_Attribute_Sales
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset contain Attributes of dresses and their recommendati
ons according to their sales.Sales are monitor on the basis of alternate days.  </p></td> -->
       <td>
        <p class="normal">
         Text
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         501
        </p>
       </td>
       <td>
        <p class="normal">
         13
        </p>
       </td>
       <td>
```

```
        <p class="normal">
         2014
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Tamilnadu+Electricity+Board+Hourly+Readings">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Tamilnadu+Electricity+Board+Hourly+Readings">
              Tamilnadu Electricity Board Hourly Readings
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data can be effectively produced the result to fewer
parameter of the Load profile can be reduced in the Database   </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         45781
        </p>
       </td>
       <td>
        <p class="normal">
         5
        </p>
       </td>
       <td>
        <p class="normal">
         2013
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Airfoil+Self-Noise">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Airfoil+Self-Noise">
              Airfoil Self-Noise
             </a>
            </b>
```

```
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">NASA data set, obtained from a series of aerodynamic and
acoustic tests of two and three-dimensional airfoil blade sections conducted in an anechoic wind t
unnel. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        1503
       </p>
      </td>
      <td>
       <p class="normal">
        6
       </p>
      </td>
      <td>
       <p class="normal">
        2014
       </p>
      </td>
      <!-- <td><p class="normal">Physical </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Wholesale+customers">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Wholesale+customers">
             Wholesale customers
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The data set refers to clients of a wholesale distributor. It
includes the annual spending in monetary units (m.u.) on diverse product categories </p></td>
-->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Clustering
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         440
        </p>
       </td>
       <td>
        <p class="normal">
         8
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
       </td>
       <!-- <td><p class="normal">Business </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Twitter+Data+set+for+Arabic+Sentiment+Analysis">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Twitter+Data+set+for+Arabic+Sentiment+Analysis">
              Twitter Data set for Arabic Sentiment Analysis
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This problem of Sentiment Analysis (SA) has been studied well on
the English language but not Arabic one. Two main approaches have been devised: corpus-based and l
exicon-based.  </p></td> -->
       <td>
        <p class="normal">
         Text
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         2000
        </p>
       </td>
       <td>
        <p class="normal">
         2
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
       </td>
       <!-- <td><p class="normal">Social </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
```

```
          <tr>
           <td>
            <a href="datasets/Combined+Cycle+Power+Plant">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Combined+Cycle+Power+Plant">
               Combined Cycle Power Plant
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The dataset contains 9568 data points collected from a Combined
Cycle Power Plant over 6 years (2006-2011), when the plant was set to work with full load.  <
/p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          9568
         </p>
        </td>
        <td>
         <p class="normal">
          4
         </p>
        </td>
        <td>
         <p class="normal">
          2014
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Urban+Land+Cover">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Urban+Land+Cover">
               Urban Land Cover
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Classification of urban land cover using high resolution aerial i
magery. Intended to assist sustainable urban planning efforts. </p></td> -->
        <td>
```

```
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         168
        </p>
       </td>
       <td>
        <p class="normal">
         148
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
       </td>
       <!-- <td><p class="normal">Physical </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Diabetes+130-US+hospitals+for+years+1999-2008">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Diabetes+130-US+hospitals+for+years+1999-2008">
              Diabetes 130-US hospitals for years 1999-2008
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data has been prepared to analyze factors related to
readmission as well as other

outcomes pertaining to patients with diabetes. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         100000
        </p>
       </td>
       <td>
        <p class="normal">
```

```
          55
         </p>
        </td>
        <td>
         <p class="normal">
          2014
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Bach+Choral+Harmony">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Bach+Choral+Harmony">
               Bach Choral Harmony
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The data set is composed of 60 chorales (5665 events) by J.S.
Bach (1675-1750).
Each event of each chorale is labelled using 1 among 101 chord labels and described
through 14 features. </p></td> -->
        <td>
         <p class="normal">
          Sequential
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          5665
         </p>
        </td>
        <td>
         <p class="normal">
          17
         </p>
        </td>
        <td>
         <p class="normal">
          2014
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/StoneFlakes">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
```

```
            <b>
             <a href="datasets/StoneFlakes">
              StoneFlakes
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Stone flakes are waste products of the stone tool production in
the prehistoric era. The variables are means of geometric and
stylistic features of the flakes contained in different inventories. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering, Causal-Discovery
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         79
        </p>
       </td>
       <td>
        <p class="normal">
         8
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Tennis+Major+Tournament+Match+Statistics">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Tennis+Major+Tournament+Match+Statistics">
              Tennis Major Tournament Match Statistics
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This is a collection of 8 files containing the match statistics f
or both women and men at the four major tennis tournaments of the year 2013. Each file has 42 colu
mns and a minimum of 76 rows. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression, Clustering
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         127
        </p>
       </td>
       <td>
        <p class="normal">
         42
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Parkinson+Speech+Dataset+with++Multiple+Types+of+Sound+Recordings">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Parkinson+Speech+Dataset+with++Multiple+Types+of+Sound+Recordings">
              Parkinson Speech Dataset with  Multiple Types of Sound Recordings
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The training data belongs to 20 Parkinson's Disease (PD)
patients and 20 healthy subjects. From all subjects, multiple types of sound recordings (26) are t
aken. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         1040
        </p>
       </td>
       <td>
        <p class="normal">
         26
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
```

```
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Gesture+Phase+Segmentation">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Gesture+Phase+Segmentation">
             Gesture Phase Segmentation
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The dataset is composed by features extracted from 7 videos with
people gesticulating, aiming at studying Gesture Phase Segmentation. It contains 50 attributes
divided into two files for each video. </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Sequential, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Clustering
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        9900
       </p>
      </td>
      <td>
       <p class="normal">
        50
       </p>
      </td>
      <td>
       <p class="normal">
        2014
       </p>
      </td>
      <!-- <td><p class="normal">Other </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Perfume+Data">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Perfume+Data">
             Perfume Data
            </a>
           </b>
          </p>
         </td>
```

```
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This data consists of odors of 20 different perfumes. Data was
obtained by using a handheld odor meter (OMX-GR sensor) per second for 28 seconds period. </p
></td> -->
        <td>
         <p class="normal">
          Univariate, Domain-Theory
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          560
         </p>
        </td>
        <td>
         <p class="normal">
          2
         </p>
        </td>
        <td>
         <p class="normal">
          2014
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/BlogFeedback">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/BlogFeedback">
               BlogFeedback
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Instances in this dataset contain features extracted from blog
posts. The task associated with the data is to predict how many comments the post will
receive. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
```

```
      <td>
       <p class="normal">
        60021
       </p>
      </td>
      <td>
       <p class="normal">
        281
       </p>
      </td>
      <td>
       <p class="normal">
        2014
       </p>
      </td>
      <!-- <td><p class="normal">Social </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/REALDISP+Activity+Recognition+Dataset">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/REALDISP+Activity+Recognition+Dataset">
             REALDISP Activity Recognition Dataset
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The REALDISP dataset is devised to evaluate techniques dealing wi
th the effects of sensor displacement in wearable activity recognition as well as to benchmark gen
eral activity recognition algorithms  </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        1419
       </p>
      </td>
      <td>
       <p class="normal">
        120
       </p>
      </td>
      <td>
       <p class="normal">
        2014
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
```

```
          <td>
           <a href="datasets/Newspaper+and+magazine+images+segmentation+dataset">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Newspaper+and+magazine+images+segmentation+dataset">
              Newspaper and magazine images segmentation dataset
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Dataset is well suited for segmentation tasks. It contains 101 s
canned pages from different newspapers and magazines in Russian with ground truth pixel-based
masks. </p></td> -->
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         101
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/AAAI+2014+Accepted+Papers">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/AAAI+2014+Accepted+Papers">
              AAAI 2014 Accepted Papers
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data set compromises the metadata for the 2014 AAAI
conference's accepted papers, including paper titles, authors, abstracts, and keywords of varying
granularity. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
```

```
</p>
</td>
<td>
 <p class="normal">
  Clustering
 </p>
</td>
<td>
 <p class="normal">
 </p>
</td>
<td>
 <p class="normal">
  399
 </p>
</td>
<td>
 <p class="normal">
  6
 </p>
</td>
<td>
 <p class="normal">
  2014
 </p>
</td>
<!-- <td><p class="normal">Computer </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/Gas+sensor+array+under+flow+modulation">
      <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/Gas+sensor+array+under+flow+modulation">
        Gas sensor array under flow modulation
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">The data set contains 58 time series acquired from 16 chemical s
ensors under gas flow modulation conditions. The sensors were exposed to different gaseous binary
mixtures of acetone and ethanol. </p></td> -->
 <td>
  <p class="normal">
   Multivariate, Time-Series
  </p>
 </td>
 <td>
  <p class="normal">
   Classification, Regression
  </p>
 </td>
 <td>
  <p class="normal">
   Real
  </p>
 </td>
 <td>
  <p class="normal">
   58
  </p>
 </td>
 <td>
  <p class="normal">
   120432
  </p>
 </td>
 <td>
```

```
      <td>
       <p class="normal">
        2014
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Gas+sensor+array+exposed+to+turbulent+gas+mixtures">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Gas+sensor+array+exposed+to+turbulent+gas+mixtures">
             Gas sensor array exposed to turbulent gas mixtures
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">A chemical detection platform composed of 8 chemoresistive gas s
ensors was exposed to turbulent gas mixtures generated naturally in a wind tunnel. The acquired ti
me series of the sensors are provided. </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        180
       </p>
      </td>
      <td>
       <p class="normal">
        150000
       </p>
      </td>
      <td>
       <p class="normal">
        2014
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/UJIIndoorLoc">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/UJIIndoorLoc">
             UJIIndoorLoc
            </a>
```

```
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The UJIIndoorLoc is a Multi-Building Multi-Floor indoor
localization database to test Indoor Positioning System that rely on WLAN/WiFi fingerprint. <
/p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Integer, Real
       </p>
      </td>
      <td>
       <p class="normal">
        21048
       </p>
      </td>
      <td>
       <p class="normal">
        529
       </p>
      </td>
      <td>
       <p class="normal">
        2014
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Sentence+Classification">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Sentence+Classification">
             Sentence Classification
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Contains sentences from the abstract and introduction of 30
articles annotated with a modified Argumentative Zones annotation scheme. These articles come from
biology, machine learning and psychology. </p></td> -->
      <td>
       <p class="normal">
        Text
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
```

```
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Dow+Jones+Index">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Dow+Jones+Index">
              Dow Jones Index
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset contains weekly data for the Dow Jones Industrial I
ndex.  It has been used in computational investing research. </p></td> -->
       <td>
        <p class="normal">
         Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         750
        </p>
       </td>
       <td>
        <p class="normal">
         16
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
       </td>
       <!-- <td><p class="normal">Business </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
```

```
          <tr>
           <td>
            <a href="datasets/sEMG+for+Basic+Hand+movements">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/sEMG+for+Basic+Hand+movements">
               sEMG for Basic Hand movements
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The “sEMG for Basic Hand movements” includes 2 databases of surfa
ce electromyographic signals of 6 hand movements using Delsys' EMG System. Healthy subjects
conducted six daily life grasps. </p></td> -->
        <td>
         <p class="normal">
          Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          3000
         </p>
        </td>
        <td>
         <p class="normal">
          2500
         </p>
        </td>
        <td>
         <p class="normal">
          2014
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/AAAI+2013+Accepted+Papers">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/AAAI+2013+Accepted+Papers">
               AAAI 2013 Accepted Papers
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This data set compromises the metadata for the 2013 AAAI
conference's accepted papers (main track only), including paper titles, abstracts, and keywords of
varying granularity. </p></td> -->
```

```
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Clustering
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          150
         </p>
        </td>
        <td>
         <p class="normal">
          5
         </p>
        </td>
        <td>
         <p class="normal">
          2014
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Geographical+Original+of+Music">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Geographical+Original+of+Music">
               Geographical Original of Music
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Instances in this dataset contain audio features extracted from
1059 wave files. The task associated with the data is to predict the geographical origin of music.
 </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          1059
         </p>
        </td>
        <td>
         <p class="normal">
```

```
          68
         </p>
        </td>
        <td>
         <p class="normal">
          2014
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Condition+Based+Maintenance+of+Naval+Propulsion+Plants">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Condition+Based+Maintenance+of+Naval+Propulsion+Plants">
               Condition Based Maintenance of Naval Propulsion Plants
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Data have been generated from a sophisticated simulator of a Gas
Turbines (GT), mounted on a Frigate characterized by a COmbined Diesel eLectric And Gas (CODLAG) p
ropulsion plant type. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          11934
         </p>
        </td>
        <td>
         <p class="normal">
          16
         </p>
        </td>
        <td>
         <p class="normal">
          2014
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Grammatical+Facial+Expressions">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
```

```
            <b>
             <a href="datasets/Grammatical+Facial+Expressions">
              Grammatical Facial Expressions
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset supports the development of models that make
possible to interpret Grammatical Facial Expressions from Brazilian Sign Language (Libras). <
/p></td> -->
       <td>
        <p class="normal">
         Multivariate, Sequential
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         27965
        </p>
       </td>
       <td>
        <p class="normal">
         100
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/NoisyOffice">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/NoisyOffice">
              NoisyOffice
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Corpus intended to do cleaning (or binarization) and enhancement
of noisy grayscale printed text images using supervised learning methods. Noisy images and their c
orresponding ground truth provided. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         216
        </p>
       </td>
       <td>
        <p class="normal">
         216
        </p>
       </td>
       <td>
        <p class="normal">
         2015
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/MHEALTH+Dataset">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/MHEALTH+Dataset">
              MHEALTH Dataset
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The MHEALTH (Mobile Health) dataset is devised to benchmark techn
iques dealing with human behavior analysis based on multimodal body sensing. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         120
        </p>
       </td>
       <td>
        <p class="normal">
         23
        </p>
       </td>
       <td>
        <p class="normal">
         2014
        </p>
       </td>
```

```
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Student+Performance">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Student+Performance">
               Student Performance
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Predict student performance in secondary education (high
school).  </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          649
         </p>
        </td>
        <td>
         <p class="normal">
          33
         </p>
        </td>
        <td>
         <p class="normal">
          2014
         </p>
        </td>
        <!-- <td><p class="normal">Social </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/ElectricityLoadDiagrams20112014">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/ElectricityLoadDiagrams20112014">
               ElectricityLoadDiagrams20112014
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
```

```
      </td>
      <!-- <td><p class="normal">This data set contains electricity consumption of 370
points/clients.
 </p></td> -->
      <td>
       <p class="normal">
        Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Regression, Clustering
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        370
       </p>
      </td>
      <td>
       <p class="normal">
        140256
       </p>
      </td>
      <td>
       <p class="normal">
        2015
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Gas+sensor+array+under+dynamic+gas+mixtures">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Gas+sensor+array+under+dynamic+gas+mixtures">
             Gas sensor array under dynamic gas mixtures
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The data set contains the recordings of 16 chemical sensors
exposed to two dynamic gas mixtures at varying concentrations. For each mixture, signals were acqu
ired continuously during 12 hours. </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
```

```
          4178504
         </p>
        </td>
        <td>
         <p class="normal">
          19
         </p>
        </td>
        <td>
         <p class="normal">
          2015
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/microblogPCU">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/microblogPCU">
               microblogPCU
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">MicroblogPCU data is crawled from sina weibo
microblog[http://weibo.com/]. This data can be used to study machine learning methods  as well as
do some social network research.  </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Univariate, Sequential, Text
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Causal-Discovery
         </p>
        </td>
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          221579
         </p>
        </td>
        <td>
         <p class="normal">
          20
         </p>
        </td>
        <td>
         <p class="normal">
          2015
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Firm-Teacher_Clave-Direction_Classification">
```

```
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Firm-Teacher_Clave-Direction_Classification">
              Firm-Teacher_Clave-Direction_Classification
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The data are binary attack-point vectors and their clave-
direction class(es) according to the partido-alto-based paradigm. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         10800
        </p>
       </td>
       <td>
        <p class="normal">
         20
        </p>
       </td>
       <td>
        <p class="normal">
         2015
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Dataset+for+Sensorless+Drive+Diagnosis">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Dataset+for+Sensorless+Drive+Diagnosis">
              Dataset for Sensorless Drive Diagnosis
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Features are extracted from motor current. The motor has intact
and defective components. This results in 11 different classes with different conditions.  </
p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
```

```
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          58509
         </p>
        </td>
        <td>
         <p class="normal">
          49
         </p>
        </td>
        <td>
         <p class="normal">
          2015
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/TV+News+Channel+Commercial+Detection+Dataset">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/TV+News+Channel+Commercial+Detection+Dataset">
               TV News Channel Commercial Detection Dataset
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">TV Commercials data set consists of  standard audio-visual
features of video shots extracted from 150 hours of TV news broadcast of 3 Indian and 2
international news channels ( 30 Hours each).   </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          129685
         </p>
        </td>
        <td>
         <p class="normal">
          12
         </p>
        </td>
        <td>
```

```
        <p class="normal">
         2015
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Phishing+Websites">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Phishing+Websites">
              Phishing Websites
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset collected mainly from: PhishTank archive,
MillerSmiles archive, Google’s searching operators. </p></td> -->
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         2456
        </p>
       </td>
       <td>
        <p class="normal">
         30
        </p>
       </td>
       <td>
        <p class="normal">
         2015
        </p>
       </td>
       <!-- <td><p class="normal">Computer Security </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Greenhouse+Gas+Observing+Network">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Greenhouse+Gas+Observing+Network">
              Greenhouse Gas Observing Network
             </a>
            </b>
           </p>
```

```
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Design an observing network to monitor emissions of a greenhouse
gas (GHG) in California given time series of synthetic observations and tracers from weather model
simulations.
 </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         2921
        </p>
       </td>
       <td>
        <p class="normal">
         5232
        </p>
       </td>
       <td>
        <p class="normal">
         2015
        </p>
       </td>
       <!-- <td><p class="normal">Physical </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Diabetic+Retinopathy+Debrecen+Data+Set">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Diabetic+Retinopathy+Debrecen+Data+Set">
              Diabetic Retinopathy Debrecen Data Set
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset contains features extracted from the Messidor image
set to predict whether an image contains signs of diabetic retinopathy or not.  </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
```

```
</p>
</td>
<td>
 <p class="normal">
  1151
 </p>
</td>
<td>
 <p class="normal">
  20
 </p>
</td>
<td>
 <p class="normal">
  2014
 </p>
</td>
<!-- <td><p class="normal">Life </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
 <td>
  <table>
   <tr>
    <td>
     <a href="datasets/HIV-1+protease+cleavage">
      <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a href="datasets/HIV-1+protease+cleavage">
        HIV-1 protease cleavage
       </a>
      </b>
     </p>
    </td>
   </tr>
  </table>
 </td>
 <!-- <td><p class="normal">The data contains lists of octamers (8 amino acids) and a flag
(-1 or 1) depending on whether HIV-1 protease will cleave in the central position (between amino a
cids 4 and 5). </p></td> -->
 <td>
  <p class="normal">
   Multivariate
  </p>
 </td>
 <td>
  <p class="normal">
   Classification
  </p>
 </td>
 <td>
  <p class="normal">
   Categorical
  </p>
 </td>
 <td>
  <p class="normal">
   6590
  </p>
 </td>
 <td>
  <p class="normal">
   1
  </p>
 </td>
 <td>
  <p class="normal">
   2015
  </p>
 </td>
 <!-- <td><p class="normal">Life </p></td> -->
</tr>
<tr>
 <td>
  <table>
```

```
        <tr>
         <td>
          <a href="datasets/Sentiment+Labelled+Sentences">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Sentiment+Labelled+Sentences">
             Sentiment Labelled Sentences
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The dataset contains sentences labelled with positive or
negative sentiment. </p></td> -->
      <td>
       <p class="normal">
        Text
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        3000
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        2015
       </p>
      </td>
      <!-- <td><p class="normal">Other </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Online+News+Popularity">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Online+News+Popularity">
             Online News Popularity
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This dataset summarizes a heterogeneous set of features about
articles published by Mashable in a period of two years. The goal is to predict the number of shar
es in social networks (popularity). </p></td> -->
      <td>
       <p class="normal">
        Multivariate
```

```
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          39797
         </p>
        </td>
        <td>
         <p class="normal">
          61
         </p>
        </td>
        <td>
         <p class="normal">
          2015
         </p>
        </td>
        <!-- <td><p class="normal">Business </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Forest+type+mapping">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Forest+type+mapping">
               Forest type mapping
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Multi-temporal remote sensing data of a forested area in Japan.
The goal is to map different forest types using spectral data. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          326
         </p>
        </td>
        <td>
         <p class="normal">
          27
         </p>
        </td>
        <td>
```

```
      <td>
       <p class="normal">
        2015
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/wiki4HE">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/wiki4HE">
             wiki4HE
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Survey of faculty members from two Spanish universities on
teaching uses of Wikipedia </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Regression, Clustering, Causal-Discovery
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        913
       </p>
      </td>
      <td>
       <p class="normal">
        53
       </p>
      </td>
      <td>
       <p class="normal">
        2015
       </p>
      </td>
      <!-- <td><p class="normal">Social </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Online+Video+Characteristics+and+Transcoding+Time+Dataset">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Online+Video+Characteristics+and+Transcoding+Time+Dataset">
             Online Video Characteristics and Transcoding Time Dataset
            </a>
           </b>
```

```
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The dataset contains a million randomly sampled video instances
listing 10 fundamental video characteristics along with the YouTube video ID.  </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Integer, Real
       </p>
      </td>
      <td>
       <p class="normal">
        168286
       </p>
      </td>
      <td>
       <p class="normal">
        11
       </p>
      </td>
      <td>
       <p class="normal">
        2015
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Chronic_Kidney_Disease">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Chronic_Kidney_Disease">
             Chronic_Kidney_Disease
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This dataset can be used to predict the chronic kidney disease
and it can be collected from the hospital nearly 2 months of period. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
```

```
        <td>
         <p class="normal">
          400
         </p>
        </td>
        <td>
         <p class="normal">
          25
         </p>
        </td>
        <td>
         <p class="normal">
          2015
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Machine+Learning+based+ZZAlpha+Ltd.+Stock+Recommendations+2012-2014">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Machine+Learning+based+ZZAlpha+Ltd.+Stock+Recommendations+2012-2014"
>
               Machine Learning based ZZAlpha Ltd. Stock Recommendations 2012-2014
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The data here are the ZZAlpha® machine learning recommendations
made for various US traded stock portfolios the morning of each day during the 3 year period Jan 1
, 2012 - Dec 31, 2014.   </p></td> -->
        <td>
         <p class="normal">
          Sequential, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          314080
         </p>
        </td>
        <td>
         <p class="normal">
          0
         </p>
        </td>
        <td>
         <p class="normal">
          2015
         </p>
        </td>
        <!-- <td><p class="normal">Business </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
```

```
          <tr>
           <td>
            <a href="datasets/Folio">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Folio">
               Folio
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">20 photos of leaves for each of 32 different species. </p><
/td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          637
         </p>
        </td>
        <td>
         <p class="normal">
          20
         </p>
        </td>
        <td>
         <p class="normal">
          2015
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Taxi+Service+Trajectory+-+Prediction+Challenge%2C+ECML+PKDD+2015">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Taxi+Service+Trajectory+-+Prediction+Challenge%2C+ECML+PKDD+2015">
               Taxi Service Trajectory - Prediction Challenge, ECML PKDD 2015
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">An accurate dataset describing trajectories performed by all the
442 taxis running in the city of Porto, in Portugal.
 </p></td> -->
        <td>
         <p class="normal">
```

```
          Multivariate, Sequential, Time-Series, Domain-Theory
         </p>
        </td>
        <td>
         <p class="normal">
          Clustering, Causal-Discovery
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          1710671
         </p>
        </td>
        <td>
         <p class="normal">
          9
         </p>
        </td>
        <td>
         <p class="normal">
          2015
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Cuff-Less+Blood+Pressure+Estimation">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Cuff-Less+Blood+Pressure+Estimation">
               Cuff-Less Blood Pressure Estimation
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This Data set provides preprocessed and cleaned vital signals
which can be used in designing algorithms for cuff-less estimation of the blood pressure. </p
></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          12000
         </p>
        </td>
        <td>
         <p class="normal">
          3
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         2015
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Smartphone-
Based+Recognition+of+Human+Activities+and+Postural+Transitions">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Smartphone-
Based+Recognition+of+Human+Activities+and+Postural+Transitions">
              Smartphone-Based Recognition of Human Activities and Postural Transitions
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Activity recognition data set built from the recordings of 30
subjects performing basic activities and postural transitions while carrying a waist-mounted smart
phone with embedded inertial sensors.
 </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         10929
        </p>
       </td>
       <td>
        <p class="normal">
         561
        </p>
       </td>
       <td>
        <p class="normal">
         2015
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Mice+Protein+Expression">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
```

```
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Mice+Protein+Expression">
              Mice Protein Expression
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Expression levels of 77 proteins measured in the cerebral cortex
of 8 classes of control and Down syndrome mice exposed to context fear conditioning, a task used t
o assess associative learning. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         1080
        </p>
       </td>
       <td>
        <p class="normal">
         82
        </p>
       </td>
       <td>
        <p class="normal">
         2015
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/UJIIndoorLoc-Mag">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/UJIIndoorLoc-Mag">
              UJIIndoorLoc-Mag
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The UJIIndoorLoc-Mag is an indoor localization database to test I
ndoor Positioning System that rely on Earth's magnetic field variations. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Sequential, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
```

```
          Classification, Regression, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          40000
         </p>
        </td>
        <td>
         <p class="normal">
          13
         </p>
        </td>
        <td>
         <p class="normal">
          2015
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Heterogeneity+Activity+Recognition">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Heterogeneity+Activity+Recognition">
               Heterogeneity Activity Recognition
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The Heterogeneity Human Activity Recognition (HHAR) dataset from
Smartphones and Smartwatches is a dataset devised to benchmark human activity recognition
algorithms (classification, automatic data segmentation, sensor fusion, feature extraction, etc.)
in real-world contexts; specifically, the dataset is gathered with a variety of different device m
odels and use-scenarios, in order to reflect sensing heterogeneities to be expected in real
deployments. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          43930257
         </p>
        </td>
        <td>
         <p class="normal">
          16
         </p>
        </td>
```

```
        <td>
         <p class="normal">
          2015
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a
href="datasets/Educational+Process+Mining+%28EPM%29%3A+A+Learning+Analytics+Data+Set">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a
href="datasets/Educational+Process+Mining+%28EPM%29%3A+A+Learning+Analytics+Data+Set">
               Educational Process Mining (EPM): A Learning Analytics Data Set
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Educational Process Mining data set is built from the recordings
of 115 subjects' activities through a logging application while learning with an educational simul
ator. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Sequential, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Regression, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          230318
         </p>
        </td>
        <td>
         <p class="normal">
          13
         </p>
        </td>
        <td>
         <p class="normal">
          2015
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/HEPMASS">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
```

```
            <a href="datasets/HEPMASS">
             HEPMASS
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The search for exotic particles requires sorting through a large
number of collisions to find the events of interest. This data set challenges one to detect a new
particle of unknown mass. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        10500000
       </p>
      </td>
      <td>
       <p class="normal">
        28
       </p>
      </td>
      <td>
       <p class="normal">
        2016
       </p>
      </td>
      <!-- <td><p class="normal">Physical </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Indoor+User+Movement+Prediction+from+RSS+data">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Indoor+User+Movement+Prediction+from+RSS+data">
             Indoor User Movement Prediction from RSS data
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This dataset contains temporal data from a Wireless Sensor
Network deployed in real-world office environments. The task is intended as real-life benchmark in
the area of Ambient Assisted Living. </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Sequential, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
```

```
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          13197
         </p>
        </td>
        <td>
         <p class="normal">
          4
         </p>
        </td>
        <td>
         <p class="normal">
          2016
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Open+University+Learning+Analytics+dataset">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Open+University+Learning+Analytics+dataset">
               Open University Learning Analytics dataset
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Open University Learning Analytics Dataset contains data about
courses, students and their interactions with Virtual Learning Environment for seven selected cour
ses and more than 30000 students. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Sequential, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Regression, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          2015
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
```

```
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/default+of+credit+card+clients">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/default+of+credit+card+clients">
              default of credit card clients
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This research aimed at the case of customers' default payments
in Taiwan and compares the predictive accuracy of probability of default among six data mining met
hods. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         30000
        </p>
       </td>
       <td>
        <p class="normal">
         24
        </p>
       </td>
       <td>
        <p class="normal">
         2016
        </p>
       </td>
       <!-- <td><p class="normal">Business </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Mesothelioma%E2%80%99s+disease+data+set+">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Mesothelioma%E2%80%99s+disease+data+set+">
              Mesothelioma's disease data set
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
```

```
      <!-- <td><p class="normal">Mesothelioma’s disease data set were prepared at Dicle
University Faculty of Medicine in Turkey.
Three hundred and twenty-four Mesothelioma patient data. In the dataset, all samples have 34 featu
res. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        324
       </p>
      </td>
      <td>
       <p class="normal">
        34
       </p>
      </td>
      <td>
       <p class="normal">
        2016
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Online+Retail">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Online+Retail">
             Online Retail
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This is a transnational data set which contains all the
transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store o
nline retail. </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Sequential, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Clustering
       </p>
      </td>
      <td>
       <p class="normal">
        Integer, Real
       </p>
      </td>
      <td>
       <p class="normal">
```

```
        541909
       </p>
      </td>
      <td>
       <p class="normal">
        8
       </p>
      </td>
      <td>
       <p class="normal">
        2015
       </p>
      </td>
      <!-- <td><p class="normal">Business </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/SIFT10M">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/SIFT10M">
             SIFT10M
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">In SIFT10M, each data point is a SIFT feature which is extracted
from Caltech-256 by the open source VLFeat library. The corresponding patches of the SIFT features
are provided. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Causal-Discovery
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
       <p class="normal">
        11164866
       </p>
      </td>
      <td>
       <p class="normal">
        128
       </p>
      </td>
      <td>
       <p class="normal">
        2016
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/GPS+Trajectories">
```

```
          <a href="datasets/GPS+Trajectories">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/GPS+Trajectories">
             GPS Trajectories
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The dataset has been feed by Android app called Go!Track. It is
available at Goolge Play Store(https://play.google.com/store/apps/details?id=com.go.router).  
;</p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        163
       </p>
      </td>
      <td>
       <p class="normal">
        15
       </p>
      </td>
      <td>
       <p class="normal">
        2016
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Detect+Malacious+Executable%28AntiVirus%29">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Detect+Malacious+Executable%28AntiVirus%29">
             Detect Malacious Executable(AntiVirus)
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">I extract features from malacious and non-malacious and create a
nd training dataset to teach svm classifier.Dataset made of unknown executable to detect if it is
virus or normal safe executable. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
```

```
      Multivariate
     </p>
    </td>
    <td>
     <p class="normal">
      Classification
     </p>
    </td>
    <td>
     <p class="normal">
      Real
     </p>
    </td>
    <td>
     <p class="normal">
      373
     </p>
    </td>
    <td>
     <p class="normal">
      513
     </p>
    </td>
    <td>
     <p class="normal">
      2016
     </p>
    </td>
    <!-- <td><p class="normal">Computer </p></td> -->
   </tr>
   <tr>
    <td>
     <table>
      <tr>
       <td>
        <a href="datasets/Occupancy+Detection+">
         <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
        </a>
       </td>
       <td>
        <p class="normal">
         <b>
          <a href="datasets/Occupancy+Detection+">
           Occupancy Detection
          </a>
         </b>
        </p>
       </td>
      </tr>
     </table>
    </td>
    <!-- <td><p class="normal">Experimental data used for binary classification (room
occupancy) from Temperature,Humidity,Light and CO2. Ground-truth occupancy was obtained from time
stamped pictures that were taken every minute. </p></td> -->
    <td>
     <p class="normal">
      Multivariate, Time-Series
     </p>
    </td>
    <td>
     <p class="normal">
      Classification
     </p>
    </td>
    <td>
     <p class="normal">
      Real
     </p>
    </td>
    <td>
     <p class="normal">
      20560
     </p>
    </td>
    <td>
     <p class="normal">
      7
     </p>
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         2016
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a
href="datasets/Improved+Spiral+Test+Using+Digitized+Graphics+Tablet+for+Monitoring+Parkinson%E2%80%
Disease">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a
href="datasets/Improved+Spiral+Test+Using+Digitized+Graphics+Tablet+for+Monitoring+Parkinson%E2%80%
Disease">
              Improved Spiral Test Using Digitized Graphics Tablet for Monitoring Parkinson’s Disea
se
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Handwriting database consists of 25 PWP(People with Parkinson)
and 15 healthy individuals.Three types of recordings (Static Spiral Test, Dynamic Spiral Test and
Stability Test) are taken. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         40
        </p>
       </td>
       <td>
        <p class="normal">
         7
        </p>
       </td>
       <td>
        <p class="normal">
         2016
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/News+Aggregator">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
```

```
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/News+Aggregator">
             News Aggregator
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">References to news pages collected from an web aggregator in the period from 10-March-2014 to 10-August-2014. The resources are grouped into clusters that represen t pages discussing the same story. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Clustering
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        422937
       </p>
      </td>
      <td>
       <p class="normal">
        5
       </p>
      </td>
      <td>
       <p class="normal">
        2016
       </p>
      </td>
      <!-- <td><p class="normal">Other </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Air+Quality">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Air+Quality">
             Air Quality
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Contains the responses of a gas multisensor device deployed on t he field in an Italian city. Hourly responses averages are recorded along with gas concentrations references from a certified analyzer.  </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Time-Series
       </p>
      </td>
      <td>
```

```
        <td>
         <p class="normal">
          Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          9358
         </p>
        </td>
        <td>
         <p class="normal">
          15
         </p>
        </td>
        <td>
         <p class="normal">
          2016
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Twin+gas+sensor+arrays">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Twin+gas+sensor+arrays">
               Twin gas sensor arrays
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">5 replicates of an 8-MOX gas sensor array were exposed to differe
nt gas conditions (4 volatiles at 10 concentration levels each). </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Time-Series, Domain-Theory
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          640
         </p>
        </td>
        <td>
         <p class="normal">
          480000
         </p>
        </td>
        <td>
         <p class="normal">
```

```
          2016
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Gas+sensors+for+home+activity+monitoring">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Gas+sensors+for+home+activity+monitoring">
               Gas sensors for home activity monitoring
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">100 recordings of a sensor array under different conditions in a
home setting: background, wine and banana presentations. The array includes 8 MOX gas sensors, and
humidity and temperature sensors.
 </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          919438
         </p>
        </td>
        <td>
         <p class="normal">
          11
         </p>
        </td>
        <td>
         <p class="normal">
          2016
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Facebook+Comment+Volume+Dataset">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Facebook+Comment+Volume+Dataset">
               Facebook Comment Volume Dataset
              </a>
```

```
            </p>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Instances in this dataset contain features extracted from
facebook posts. The task associated with the data is to predict how many comments the post will re
ceive. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         40949
        </p>
       </td>
       <td>
        <p class="normal">
         54
        </p>
       </td>
       <td>
        <p class="normal">
         2016
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a
href="datasets/Smartphone+Dataset+for+Human+Activity+Recognition+%28HAR%29+in+Ambient+Assisted+Livi
28AAL%29">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a
href="datasets/Smartphone+Dataset+for+Human+Activity+Recognition+%28HAR%29+in+Ambient+Assisted+Livi
28AAL%29">
              Smartphone Dataset for Human Activity Recognition (HAR) in Ambient Assisted Living (A
AL)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This data is an addition to an existing dataset on UCI. We
collected more data to improve the accuracy of our human activity recognition algorithms applied i
n the domain of Ambient Assisted Living.   </p></td> -->
       <td>
        <p class="normal">
         Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
```

```
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          5744
         </p>
        </td>
        <td>
         <p class="normal">
          561
         </p>
        </td>
        <td>
         <p class="normal">
          2016
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Polish+companies+bankruptcy+data">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Polish+companies+bankruptcy+data">
               Polish companies bankruptcy data
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The dataset is about bankruptcy prediction of Polish companies.Th
e bankrupt companies were analyzed in the period 2000-2012, while the still operating companies we
re evaluated from 2007 to 2013. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          10503
         </p>
        </td>
        <td>
         <p class="normal">
          64
         </p>
        </td>
        <td>
         <p class="normal">
          2016
```

```
        </p>
       </td>
       <!-- <td><p class="normal">Business </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a
href="datasets/Activity+Recognition+system+based+on+Multisensor+data+fusion+%28AReM%29">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a
href="datasets/Activity+Recognition+system+based+on+Multisensor+data+fusion+%28AReM%29">
              Activity Recognition system based on Multisensor data fusion (AReM)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset contains temporal data from a Wireless Sensor
Network worn by an actor performing the activities: bending, cycling, lying down, sitting,
standing, walking. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Sequential, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         42240
        </p>
       </td>
       <td>
        <p class="normal">
         6
        </p>
       </td>
       <td>
        <p class="normal">
         2016
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Dota2+Games+Results">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Dota2+Games+Results">
              Dota2 Games Results
             </a>
```

```
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Dota 2 is a popular computer game with two teams of 5 players. A
t the start of the game each player chooses a unique hero with different strengths and
weaknesses. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         102944
        </p>
       </td>
       <td>
        <p class="normal">
         116
        </p>
       </td>
       <td>
        <p class="normal">
         2016
        </p>
       </td>
       <!-- <td><p class="normal">Game </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Facebook+metrics">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Facebook+metrics">
              Facebook metrics
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Facebook performance metrics of a renowned cosmetic's brand Faceb
ook page. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
```

```
        </td>
        <td>
         <p class="normal">
          500
         </p>
        </td>
        <td>
         <p class="normal">
          19
         </p>
        </td>
        <td>
         <p class="normal">
          2016
         </p>
        </td>
        <!-- <td><p class="normal">Business </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/UbiqLog+%28smartphone+lifelogging%29">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/UbiqLog+%28smartphone+lifelogging%29">
               UbiqLog (smartphone lifelogging)
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">UbiqLog is the smartphone lifelogging tool that runs on the
smartphone of 35 users for about 2 months.
 </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Causal-Discovery
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          9782222
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          2016
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
```

```
          <a href="datasets/NIPS+Conference+Papers+1987-2015">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/NIPS+Conference+Papers+1987-2015">
             NIPS Conference Papers 1987-2015
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This data set contains the distribution of words in the full
text of the NIPS conference papers published from 1987 to 2015. </p></td> -->
      <td>
       <p class="normal">
        Text
       </p>
      </td>
      <td>
       <p class="normal">
        Clustering
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
       <p class="normal">
        11463
       </p>
      </td>
      <td>
       <p class="normal">
        5812
       </p>
      </td>
      <td>
       <p class="normal">
        2016
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/HTRU2">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/HTRU2">
             HTRU2
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Pulsar candidates collected during the HTRU survey. Pulsars are
a type of star, of considerable scientific interest. Candidates must be classified in to pulsar
and non-pulsar classes to aid discovery. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         17898
        </p>
       </td>
       <td>
        <p class="normal">
         9
        </p>
       </td>
       <td>
        <p class="normal">
         2017
        </p>
       </td>
       <!-- <td><p class="normal">Physical </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Drug+consumption+%28quantified%29">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Drug+consumption+%28quantified%29">
              Drug consumption (quantified)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Classify type of drug consumer by personality data </p>
</td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         1885
        </p>
       </td>
       <td>
        <p class="normal">
         32
        </p>
       </td>
```

```
        <td>
         <p class="normal">
          2016
         </p>
        </td>
        <!-- <td><p class="normal">Social </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Appliances+energy+prediction">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Appliances+energy+prediction">
               Appliances energy prediction
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Experimental data used to create regression models of appliances
energy use in a low energy building. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          19735
         </p>
        </td>
        <td>
         <p class="normal">
          29
         </p>
        </td>
        <td>
         <p class="normal">
          2017
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Miskolc+IIS+Hybrid+IPS">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Miskolc+IIS+Hybrid+IPS">
               Miskolc IIS Hybrid IPS
              </a>
```

```
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The dataset was created for the comparison and evaluation of
hybrid indoor positioning methods. The dataset presented contains data from W-LAN and Bluetooth in
terfaces, and Magnetometer.  </p></td> -->
       <td>
        <p class="normal">
         Text
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering, Causal-Discovery
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         1540
        </p>
       </td>
       <td>
        <p class="normal">
         67
        </p>
       </td>
       <td>
        <p class="normal">
         2016
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/KDC-4007+dataset+Collection">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/KDC-4007+dataset+Collection">
              KDC-4007 dataset Collection
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">KDC-4007 dataset Collection is the Kurdish Documents
Classification text used in categories regarding Kurdish Sorani news and articles. </p></td>
-->
       <td>
        <p class="normal">
         Multivariate, Text
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression
        </p>
       </td>
       <td>
        <p class="normal">
```

```
          Integer
        </p>
       </td>
       <td>
        <p class="normal">
         4007
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         2017
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Geo-
Magnetic+field+and+WLAN+dataset+for+indoor+localisation+from+wristband+and+smartphone">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Geo-
Magnetic+field+and+WLAN+dataset+for+indoor+localisation+from+wristband+and+smartphone">
              Geo-Magnetic field and WLAN dataset for indoor localisation from wristband and smartp
hone
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">A multisource and multivariate dataset for indoor localisation m
ethods based on WLAN and Geo-Magnetic field fingerprinting </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Sequential, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         153540
        </p>
       </td>
       <td>
        <p class="normal">
         25
        </p>
       </td>
       <td>
        <p class="normal">
         2017
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
```

```
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/DrivFace">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/DrivFace">
              DrivFace
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The DrivFace contains images sequences of subjects while driving
in real scenarios. It is composed of 606 samples of 640×480, acquired over different days from 4 d
rivers with several facial features. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         606
        </p>
       </td>
       <td>
        <p class="normal">
         6400
        </p>
       </td>
       <td>
        <p class="normal">
         2016
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Website+Phishing">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Website+Phishing">
              Website Phishing
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
```

```
      <!-- <td><p class="normal">

 </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
       <p class="normal">
        1353
       </p>
      </td>
      <td>
       <p class="normal">
        10
       </p>
      </td>
      <td>
       <p class="normal">
        2016
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/YouTube+Spam+Collection">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/YouTube+Spam+Collection">
             YouTube Spam Collection
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">It is a public set of comments collected for spam research. It h
as five datasets composed by 1,956 real messages extracted from five videos that were among the 10
most viewed on the collection period. </p></td> -->
      <td>
       <p class="normal">
        Text
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        1956
       </p>
```

```
</p>
</td>
<td>
<p class="normal">
5
</p>
</td>
<td>
<p class="normal">
2017
</p>
</td>
<!-- <td><p class="normal">Computer </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
<td>
<table>
<tr>
<td>
<a href="datasets/Beijing+PM2.5+Data">
<img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>
</td>
<td>
<p class="normal">
<b>
<a href="datasets/Beijing+PM2.5+Data">
Beijing PM2.5 Data
</a>
</b>
</p>
</td>
</tr>
</table>
</td>
<!-- <td><p class="normal">This hourly data set contains the PM2.5 data of US Embassy in
Beijing. Meanwhile, meteorological data from Beijing Capital International Airport are also
included.  </p></td> -->
<td>
<p class="normal">
Multivariate, Time-Series
</p>
</td>
<td>
<p class="normal">
Regression
</p>
</td>
<td>
<p class="normal">
Integer, Real
</p>
</td>
<td>
<p class="normal">
43824
</p>
</td>
<td>
<p class="normal">
13
</p>
</td>
<td>
<p class="normal">
2017
</p>
</td>
<!-- <td><p class="normal">Physical </p></td> -->
</tr>
<tr>
<td>
<table>
<tr>
<td>
<a href="datasets/Cargo+2000+Freight+Tracking+and+Tracing">
<img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>
```

```
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Cargo+2000+Freight+Tracking+and+Tracing">
             Cargo 2000 Freight Tracking and Tracing
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Sanitized and anonymized Cargo 2000 (C2K) airfreight tracking
and tracing events, covering five months of business execution (3,942 process instances, 7,932 tra
nsport legs, 56,082 activities).  </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Sequential
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
       <p class="normal">
        3942
       </p>
      </td>
      <td>
       <p class="normal">
        98
       </p>
      </td>
      <td>
       <p class="normal">
        2016
       </p>
      </td>
      <!-- <td><p class="normal">Business </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Cervical+cancer+%28Risk+Factors%29">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Cervical+cancer+%28Risk+Factors%29">
             Cervical cancer (Risk Factors)
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This dataset focuses on the prediction of indicators/diagnosis of
cervical cancer. The features cover demographic information, habits, and historic medical
records. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
```

```
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer, Real
       </p>
      </td>
      <td>
       <p class="normal">
        858
       </p>
      </td>
      <td>
       <p class="normal">
        36
       </p>
      </td>
      <td>
       <p class="normal">
        2017
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Quality+Assessment+of+Digital+Colposcopies">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Quality+Assessment+of+Digital+Colposcopies">
             Quality Assessment of Digital Colposcopies
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This dataset explores the subjective quality assessment of
digital colposcopies. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        287
       </p>
      </td>
      <td>
       <p class="normal">
        69
       </p>
      </td>
      <td>
       <p class="normal">
```

```
        <p class="normal">
         2017
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/KASANDR">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/KASANDR">
              KASANDR
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">KASANDR is a novel, publicly available collection for
recommendation systems that records the behavior of customers of the European leader in e-Commerce
advertising, Kelkoo.  </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Causal-Discovery
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         17764280
        </p>
       </td>
       <td>
        <p class="normal">
         2158859
        </p>
       </td>
       <td>
        <p class="normal">
         2017
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/FMA%3A+A+Dataset+For+Music+Analysis">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/FMA%3A+A+Dataset+For+Music+Analysis">
              FMA: A Dataset For Music Analysis
             </a>
```

```
            </p>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">FMA features 106,574 tracks and includes song title, album,
artist, genres; play counts, favorites, comments; description, biography, tags; together with
audio (343 days, 917 GiB) and features. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         106574
        </p>
       </td>
       <td>
        <p class="normal">
         518
        </p>
       </td>
       <td>
        <p class="normal">
         2017
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Air+quality">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Air+quality">
              Air quality
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal"> Contains the responses of a gas multisensor device deployed on
the field in an Italian city.  </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Real
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         9358
        </p>
       </td>
       <td>
        <p class="normal">
         15
        </p>
       </td>
       <td>
        <p class="normal">
         2016
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Epileptic+Seizure+Recognition">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Epileptic+Seizure+Recognition">
              Epileptic Seizure Recognition
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset is a pre-processed and re-structured/reshaped
version of a very commonly used dataset featuring epileptic seizure detection.  </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         11500
        </p>
       </td>
       <td>
        <p class="normal">
         179
        </p>
       </td>
       <td>
        <p class="normal">
         2017
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
```

```
          <tr>
           <td>
            <a href="datasets/Devanagari+Handwritten+Character+Dataset">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Devanagari+Handwritten+Character+Dataset">
               Devanagari Handwritten Character Dataset
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This is an image database of Handwritten Devanagari characters.
There are 46 classes of characters with 2000 examples each. The dataset is split into training set
(85%) and testing set(15%).  </p></td> -->
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          92000
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          2016
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Stock+portfolio+performance">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Stock+portfolio+performance">
               Stock portfolio performance
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The data set of performances of weighted scoring stock portfolios
are obtained with mixture design from the US stock market historical database. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         315
        </p>
       </td>
       <td>
        <p class="normal">
         12
        </p>
       </td>
       <td>
        <p class="normal">
         2016
        </p>
       </td>
       <!-- <td><p class="normal">Business </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/MoCap+Hand+Postures">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/MoCap+Hand+Postures">
              MoCap Hand Postures
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">5 types of hand postures from 12 users were recorded using
unlabeled markers attached to fingers of a glove in a motion capture environment. Due to
resolution and occlusion, missing values are common. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         78095
        </p>
       </td>
       <td>
        <p class="normal">
         38
        </p>
```

```
      </td>
      <td>
       <p class="normal">
        2016
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a
href="datasets/Early+biomarkers+of+Parkinson%92s+disease+based+on+natural+connected+speech">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a
href="datasets/Early+biomarkers+of+Parkinson%92s+disease+based+on+natural+connected+speech">
             Early biomarkers of Parkinson's disease based on natural connected speech
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Predict a pattern of neurodegeneration in the dataset of speech
features obtained from patients with early untreated Parkinson's disease and patients at high risk
developing Parkinson's disease. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Integer, Real
       </p>
      </td>
      <td>
       <p class="normal">
        130
       </p>
      </td>
      <td>
       <p class="normal">
        65
       </p>
      </td>
      <td>
       <p class="normal">
        2017
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a
href="datasets/Data+for+Software+Engineering+Teamwork+Assessment+in+Education+Setting">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
```

```
          <p class="normal">
           <b>
            <a
href="datasets/Data+for+Software+Engineering+Teamwork+Assessment+in+Education+Setting">
             Data for Software Engineering Teamwork Assessment in Education Setting
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Data include over 100 Team Activity Measures and outcomes (ML cl
asses) obtained from activities of 74 student teams during the creation of final class project in
SW Eng. classes  at SFSU, Fulda, FAU </p></td> -->
      <td>
       <p class="normal">
        Sequential, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer, Real
       </p>
      </td>
      <td>
       <p class="normal">
        74
       </p>
      </td>
      <td>
       <p class="normal">
        102
       </p>
      </td>
      <td>
       <p class="normal">
        2017
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/PM2.5+Data+of+Five+Chinese+Cities">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/PM2.5+Data+of+Five+Chinese+Cities">
             PM2.5 Data of Five Chinese Cities
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This hourly data set contains the PM2.5 data in Beijing,
Shanghai, Guangzhou, Chengdu and Shenyang. Meanwhile, meteorological data for each city are also i
ncluded. </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Time-Series
       </p>
      </td>
      <td>
```

```
        <p class="normal">
         Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         52854
        </p>
       </td>
       <td>
        <p class="normal">
         86
        </p>
       </td>
       <td>
        <p class="normal">
         2017
        </p>
       </td>
       <!-- <td><p class="normal">Physical </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Parkinson+Disease+Spiral+Drawings+Using+Digitized+Graphics+Tablet">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Parkinson+Disease+Spiral+Drawings+Using+Digitized+Graphics+Tablet">
              Parkinson Disease Spiral Drawings Using Digitized Graphics Tablet
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Handwriting database consists of 62 PWP(People with Parkinson)
and 15 healthy individuals. Three types of recordings (Static Spiral Test, Dynamic Spiral Test and
Stability Test) are taken. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         77
        </p>
       </td>
       <td>
        <p class="normal">
         7
        </p>
       </td>
       <td>
        <p class="normal">
```

```
          2017
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Sales_Transactions_Dataset_Weekly">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Sales_Transactions_Dataset_Weekly">
               Sales_Transactions_Dataset_Weekly
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Contains weekly purchased quantities of 800 over products over
52 weeks. Normalised values are provided too. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Clustering
         </p>
        </td>
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          811
         </p>
        </td>
        <td>
         <p class="normal">
          53
         </p>
        </td>
        <td>
         <p class="normal">
          2017
         </p>
        </td>
        <!-- <td><p class="normal"> </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Las+Vegas+Strip">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Las+Vegas+Strip">
               Las Vegas Strip
              </a>
             </b>
            </p>
```

```
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset includes quantitative and categorical features from
online reviews from 21 hotels located in Las Vegas Strip, extracted from TripAdvisor
(http://www.tripadvisor.com). </p></td> -->
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         504
        </p>
       </td>
       <td>
        <p class="normal">
         20
        </p>
       </td>
       <td>
        <p class="normal">
         2017
        </p>
       </td>
       <!-- <td><p class="normal">Business </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Eco-hotel">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Eco-hotel">
              Eco-hotel
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset includes Online Textual Reviews from both online
(e.g., TripAdvisor) and offline (e.g., Guests' book) sources from the Areias do Seixo Eco-
Resort. </p></td> -->
       <td>
        <p class="normal">
         Text
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
```

```
-
401
</p>
</td>
<td>
<p class="normal">
1
</p>
</td>
<td>
<p class="normal">
2017
</p>
</td>
<!-- <td><p class="normal">Business </p></td> -->
</tr>
<tr bgcolor="DDEEFF">
<td>
<table>
<tr>
<td>
<a href="datasets/MEU-Mobile+KSD">
<img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>
</td>
<td>
<p class="normal">
<b>
<a href="datasets/MEU-Mobile+KSD">
MEU-Mobile KSD
</a>
</b>
</p>
</td>
</tr>
</table>
</td>
<!-- <td><p class="normal">This dataset contains keystroke dynamics data collected on a
touch mobile device (Nexus 7). The dataset contains 2856 records, 51 records per subject for 56 su
bjects.  </p></td> -->
<td>
<p class="normal">
Multivariate
</p>
</td>
<td>
<p class="normal">
Classification
</p>
</td>
<td>
<p class="normal">
Integer, Real
</p>
</td>
<td>
<p class="normal">
2856
</p>
</td>
<td>
<p class="normal">
71
</p>
</td>
<td>
<p class="normal">
2016
</p>
</td>
<!-- <td><p class="normal">Computer </p></td> -->
</tr>
<tr>
<td>
<table>
<tr>
<td>
<a href="datasets/Crowdsourced+Mapping">
```

```
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Crowdsourced+Mapping">
              Crowdsourced Mapping
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Crowdsourced data from OpenStreetMap is used to automate the
classification of satellite images into different land cover classes (impervious, farm, forest,
grass, orchard, water).  </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         10546
        </p>
       </td>
       <td>
        <p class="normal">
         29
        </p>
       </td>
       <td>
        <p class="normal">
         2016
        </p>
       </td>
       <!-- <td><p class="normal">Physical </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/gene+expression+cancer+RNA-Seq">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/gene+expression+cancer+RNA-Seq">
              gene expression cancer RNA-Seq
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This collection of data is part of the RNA-Seq (HiSeq) PANCAN
data set, it is a random extraction of gene expressions of patients having different types of tumo
r: BRCA, KIRC, COAD, LUAD and PRAD. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
```

```
      </td>
      <td>
       <p class="normal">
        Classification, Clustering
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        801
       </p>
      </td>
      <td>
       <p class="normal">
        20531
       </p>
      </td>
      <td>
       <p class="normal">
        2016
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a
href="datasets/Hybrid+Indoor+Positioning+Dataset+from+WiFi+RSSI%2C+Bluetooth+and+magnetometer">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a
href="datasets/Hybrid+Indoor+Positioning+Dataset+from+WiFi+RSSI%2C+Bluetooth+and+magnetometer">
             Hybrid Indoor Positioning Dataset from WiFi RSSI, Bluetooth and magnetometer
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The dataset was created for the comparison and evaluation of
hybrid indoor positioning methods. The dataset presented contains data from W-LAN and Bluetooth in
terfaces, and Magnetometer.  </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Sequential, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        1540
       </p>
      </td>
      <td>
       <p class="normal">
        65
```

```
        </p>
      </td>
      <td>
       <p class="normal">
        2016
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/chestnut+%E2%80%93+LARVIC">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/chestnut+%E2%80%93+LARVIC">
             chestnut - LARVIC
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The research project presents this database, shows the images of
chestnuts that will be processed to determine the presence or absence of defects </p></td> --
>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Clustering
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        1451
       </p>
      </td>
      <td>
       <p class="normal">
        3
       </p>
      </td>
      <td>
       <p class="normal">
        2017
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a
href="datasets/Burst+Header+Packet+%28BHP%29+flooding+attack+on+Optical+Burst+Switching+%28OBS%29+N
rk">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
```

```
            <a
href="datasets/Burst+Header+Packet+%28BHP%29+flooding+attack+on+Optical+Burst+Switching+%28OBS%29+N
rk">
              Burst Header Packet (BHP) flooding attack on Optical Burst Switching (OBS) Network
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">One of the primary challenges in identifying the risks of the
Burst Header Packet (BHP) flood attacks in Optical Burst Switching networks (OBS) is the scarcity
of reliable historical data.   </p></td> -->
      <td>
       <p class="normal">
        Text
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
       <p class="normal">
        1075
       </p>
      </td>
      <td>
       <p class="normal">
        22
       </p>
      </td>
      <td>
       <p class="normal">
        2017
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Motion+Capture+Hand+Postures">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Motion+Capture+Hand+Postures">
             Motion Capture Hand Postures
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">5 types of hand postures from 12 users were recorded using
unlabeled markers on fingers of a glove in a motion capture environment. Due to resolution and occ
lusion, missing values are common. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
```

```
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         78095
        </p>
       </td>
       <td>
        <p class="normal">
         38
        </p>
       </td>
       <td>
        <p class="normal">
         2017
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Anuran+Calls+%28MFCCs%29">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Anuran+Calls+%28MFCCs%29">
              Anuran Calls (MFCCs)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Acoustic features extracted from syllables of anuran (frogs)
calls, including the family, the genus, and the species labels (multilabel).  </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         7195
        </p>
       </td>
       <td>
        <p class="normal">
         22
        </p>
       </td>
       <td>
        <p class="normal">
         2017
        </p>
```

```
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/TTC-3600%3A+Benchmark+dataset+for+Turkish+text+categorization">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/TTC-3600%3A+Benchmark+dataset+for+Turkish+text+categorization">
              TTC-3600: Benchmark dataset for Turkish text categorization
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The TTC-3600 data set is a collection of Turkish news and
articles including categorized 3,600 documents from 6 well-known portals in Turkey. It has 4
different forms in ARFF Weka format. </p></td> -->
       <td>
        <p class="normal">
         Text
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         3600
        </p>
       </td>
       <td>
        <p class="normal">
         4814
        </p>
       </td>
       <td>
        <p class="normal">
         2017
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Gastrointestinal+Lesions+in+Regular+Colonoscopy">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Gastrointestinal+Lesions+in+Regular+Colonoscopy">
              Gastrointestinal Lesions in Regular Colonoscopy
             </a>
            </b>
           </p>
          </td>
```

```
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset contains features extracted from colonoscopy videos
used to detect gastrointestinal lesions. It contains 76 lesions: 15 serrated adenomas, 21 hyperpla
stic lesions and 40 adenoma.  </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         76
        </p>
       </td>
       <td>
        <p class="normal">
         698
        </p>
       </td>
       <td>
        <p class="normal">
         2016
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Daily+Demand+Forecasting+Orders">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Daily+Demand+Forecasting+Orders">
              Daily Demand Forecasting Orders
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The dataset was collected during 60 days, this is a real
database of a brazilian logistics company. </p></td> -->
       <td>
        <p class="normal">
         Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
```

```
        <td>
         <p class="normal">
          60
         </p>
        </td>
        <td>
         <p class="normal">
          13
         </p>
        </td>
        <td>
         <p class="normal">
          2017
         </p>
        </td>
        <!-- <td><p class="normal">Business </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Paper+Reviews">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Paper+Reviews">
               Paper Reviews
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This sentiment analysis data set contains scientific paper
reviews from an international conference on computing and informatics. The task is to predict the
orientation or the evaluation of a review. </p></td> -->
        <td>
         <p class="normal">
          Text
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          405
         </p>
        </td>
        <td>
         <p class="normal">
          10
         </p>
        </td>
        <td>
         <p class="normal">
          2017
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
```

```
      <td>
       <a href="datasets/extention+of+Z-Alizadeh+sani+dataset">
        <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
       </a>
      </td>
      <td>
       <p class="normal">
        <b>
         <a href="datasets/extention+of+Z-Alizadeh+sani+dataset">
          extention of Z-Alizadeh sani dataset
         </a>
        </b>
       </p>
      </td>
     </tr>
    </table>
   </td>
   <!-- <td><p class="normal">It was collected for CAD diagnosis. </p></td> -->
   <td>
    <p class="normal">
    </p>
   </td>
   <td>
    <p class="normal">
     Classification
    </p>
   </td>
   <td>
    <p class="normal">
     Integer, Real
    </p>
   </td>
   <td>
    <p class="normal">
     303
    </p>
   </td>
   <td>
    <p class="normal">
     59
    </p>
   </td>
   <td>
    <p class="normal">
     2017
    </p>
   </td>
   <!-- <td><p class="normal">Life </p></td> -->
  </tr>
  <tr>
   <td>
    <table>
     <tr>
      <td>
       <a href="datasets/Z-Alizadeh+Sani">
        <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
       </a>
      </td>
      <td>
       <p class="normal">
        <b>
         <a href="datasets/Z-Alizadeh+Sani">
          Z-Alizadeh Sani
         </a>
        </b>
       </p>
      </td>
     </tr>
    </table>
   </td>
   <!-- <td><p class="normal">It was collected for CAD diagnosis. </p></td> -->
   <td>
    <p class="normal">
    </p>
   </td>
   <td>
    <p class="normal">
```

```
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          303
         </p>
        </td>
        <td>
         <p class="normal">
          56
         </p>
        </td>
        <td>
         <p class="normal">
          2017
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Dynamic+Features+of+VirusShare+Executables">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Dynamic+Features+of+VirusShare+Executables">
               Dynamic Features of VirusShare Executables
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This dataset contains the dynamic features of 107,888
executables, collected by VirusShare from Nov/2010 to Jul/2014. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          107888
         </p>
        </td>
        <td>
         <p class="normal">
          482
         </p>
        </td>
        <td>
         <p class="normal">
          2017
         </p>
```

```
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/IDA2016Challenge">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/IDA2016Challenge">
               IDA2016Challenge
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The dataset consists of data collected from heavy Scania trucks
in everyday usage.  </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          76000
         </p>
        </td>
        <td>
         <p class="normal">
          171
         </p>
        </td>
        <td>
         <p class="normal">
          2017
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/DSRC+Vehicle+Communications">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/DSRC+Vehicle+Communications">
               DSRC Vehicle Communications
              </a>
             </b>
            </p>
           </td>
          </tr>
```

```
        </table>
       </td>
       <!-- <td><p class="normal">This set Provides data regarding wireless communications between
vehicles and road side units. two separate data sets are provided (normal scenario) and in the pre
sence of attacker (jammer). </p></td> -->
       <td>
        <p class="normal">
         Sequential, Text
        </p>
       </td>
       <td>
        <p class="normal">
         Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         10000
        </p>
       </td>
       <td>
        <p class="normal">
         5
        </p>
       </td>
       <td>
        <p class="normal">
         2017
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Mturk+User-Perceived+Clusters+over+Images">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Mturk+User-Perceived+Clusters+over+Images">
              Mturk User-Perceived Clusters over Images
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">This dataset was collected by Shan-Hung Wu and DataLab members
at NTHU, Taiwan. There're 325 user-perceived clusters from 100 users and their corresponding descr
iptions. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Text
        </p>
       </td>
       <td>
        <p class="normal">
         Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
```

```
        <p class="normal">
         180
        </p>
       </td>
       <td>
        <p class="normal">
         500
        </p>
       </td>
       <td>
        <p class="normal">
         2016
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Character+Font+Images">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Character+Font+Images">
              Character Font Images
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Character images from scanned and computer generated
fonts. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         745000
        </p>
       </td>
       <td>
        <p class="normal">
         411
        </p>
       </td>
       <td>
        <p class="normal">
         2016
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/DeliciousMIL%3A+A+Data+Set+for+Multi-Label+Multi-
```

Instance+Learning+with+Instance+Labels">
<img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>
</td>
<td>

<a href="datasets/DeliciousMIL%3A+A+Data+Set+for+Multi-Label+Multi-Instance+Learning+with+Instance+Labels">
DeliciousMIL: A Data Set for Multi-Label Multi-Instance Learning with Instance Labels
</a>

</td>
</tr>
</table>
</td>
<!-- <td> This dataset includes 1) 12234 documents (8251 training, 3983 test) extracted from DeliciousT140 dataset, 2) class labels for all documents, 3) labels for a sub set of sentences of the test documents. </td> -->
<td>

Text

</td>
<td>

Classification

</td>
<td>

Integer

</td>
<td>

12234

</td>
<td>

8519

</td>
<td>

2016

</td>
<!-- <td> Computer </td> -->
</tr>
<tr bgcolor="DDEEFF">
<td>
<table>
<tr>
<td>
<a href="datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Children++">
<img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>
</td>
<td>

<a href="datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Children++">
Autistic Spectrum Disorder Screening Data for Children
</a>

</td>
</tr>
</table>
</td>
<!-- <td> Children screening data for autism suitable for classification and predictive tasks   </td> -->
<td>

```
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          292
         </p>
        </td>
        <td>
         <p class="normal">
          21
         </p>
        </td>
        <td>
         <p class="normal">
          2017
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Adolescent+++">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Adolescent+++">
               Autistic Spectrum Disorder Screening Data for Adolescent
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Autistic Spectrum Disorder Screening Data for Adolescent. This d
ataset is related to classification and predictive tasks. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          104
         </p>
        </td>
        <td>
         <p class="normal">
          21
         </p>
```

```
        </td>
        <td>
         <p class="normal">
          2017
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/APS+Failure+at+Scania+Trucks">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/APS+Failure+at+Scania+Trucks">
               APS Failure at Scania Trucks
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The datasets' positive class consists of component failures for
a specific component of the APS system. The negative class consists of trucks with failures for co
mponents not related to the APS. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          60000
         </p>
        </td>
        <td>
         <p class="normal">
          171
         </p>
        </td>
        <td>
         <p class="normal">
          2017
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Wireless+Indoor+Localization">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Wireless+Indoor+Localization">
```

```
              Wireless Indoor Localization
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Collected in indoor space by observing signal strengths of seven
WiFi signals visible on a smartphone. The decision variable is one of the four rooms.  </p></
td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         2000
        </p>
       </td>
       <td>
        <p class="normal">
         7
        </p>
       </td>
       <td>
        <p class="normal">
         2017
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/HCC+Survival">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/HCC+Survival">
              HCC Survival
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Hepatocellular Carcinoma dataset (HCC dataset) was collected at
a University Hospital in Portugal. It contains real clinical data of 165 patients diagnosed with H
CC. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
```

```
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          165
         </p>
        </td>
        <td>
         <p class="normal">
          49
         </p>
        </td>
        <td>
         <p class="normal">
          2017
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/CSM+%28Conventional+and+Social+Media+Movies%29+Dataset+2014+and+2015">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a
href="datasets/CSM+%28Conventional+and+Social+Media+Movies%29+Dataset+2014+and+2015">
               CSM (Conventional and Social Media Movies) Dataset 2014 and 2015
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">12 features categorized as conventional and social media
features. Both conventional features, collected from movies databases on Web as well as social med
ia features(YouTube,Twitter). </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          217
         </p>
        </td>
        <td>
         <p class="normal">
          12
         </p>
        </td>
        <td>
         <p class="normal">
          2017
         </p>
        </td>
```

```
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/University+of+Tehran+Question+Dataset+2016+%28UTQD.2016%29">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/University+of+Tehran+Question+Dataset+2016+%28UTQD.2016%29">
             University of Tehran Question Dataset 2016 (UTQD.2016)
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Persian questions gathered from a jeopardy game broadcasted on I
ranian national television.  </p></td> -->
      <td>
       <p class="normal">
        Text
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        1175
       </p>
      </td>
      <td>
       <p class="normal">
        3
       </p>
      </td>
      <td>
       <p class="normal">
        2017
       </p>
      </td>
      <!-- <td><p class="normal">Other </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Autism+Screening+Adult">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Autism+Screening+Adult">
             Autism Screening Adult
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
```

```
      <!-- <td><p class="normal">Autistic Spectrum Disorder Screening Data for Adult. This
dataset is related to classification and predictive tasks. </p></td> -->
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
       <p class="normal">
        704
       </p>
      </td>
      <td>
       <p class="normal">
        21
       </p>
      </td>
      <td>
       <p class="normal">
        2017
       </p>
      </td>
      <!-- <td><p class="normal">Social </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a
href="datasets/Activity+recognition+with+healthy+older+people+using+a+batteryless+wearable+sensor">

           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a
href="datasets/Activity+recognition+with+healthy+older+people+using+a+batteryless+wearable+sensor">

             Activity recognition with healthy older people using a batteryless wearable sensor
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Sequential motion data from 14 healthy older people aged 66 to 86
years old using a batteryless, wearable sensor on top of their clothing for the recognition of act
ivities in clinical environments. </p></td> -->
      <td>
       <p class="normal">
        Sequential
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
```

```
      <td>
       <p class="normal">
        75128
       </p>
      </td>
      <td>
       <p class="normal">
        9
       </p>
      </td>
      <td>
       <p class="normal">
        2016
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Immunotherapy+Dataset">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Immunotherapy+Dataset">
             Immunotherapy Dataset
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This dataset contains information about wart treatment results
of 90 patients using immunotherapy. </p></td> -->
      <td>
       <p class="normal">
        Univariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer, Real
       </p>
      </td>
      <td>
       <p class="normal">
        90
       </p>
      </td>
      <td>
       <p class="normal">
        8
       </p>
      </td>
      <td>
       <p class="normal">
        2018
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Cryotherapy+Dataset+">
```

```
          <a href="datasets/Cryotherapy+Dataset+">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Cryotherapy+Dataset+">
             Cryotherapy Dataset
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This dataset contains information about wart treatment results
of 90 patients using cryotherapy. </p></td> -->
      <td>
       <p class="normal">
        Univariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer, Real
       </p>
      </td>
      <td>
       <p class="normal">
        90
       </p>
      </td>
      <td>
       <p class="normal">
        7
       </p>
      </td>
      <td>
       <p class="normal">
        2018
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/OCT+data+%26+Color+Fundus+Images+of+Left+%26+Right+Eyes">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/OCT+data+%26+Color+Fundus+Images+of+Left+%26+Right+Eyes">
             OCT data &amp; Color Fundus Images of Left &amp; Right Eyes
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This dataset contains OCT data (in mat format) and color fundus
data (in jpg format) of left & right eyes of 50 healthy persons. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
```

```
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        50
       </p>
      </td>
      <td>
       <p class="normal">
        2
       </p>
      </td>
      <td>
       <p class="normal">
        2016
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Discrete+Tone+Image+Dataset">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Discrete+Tone+Image+Dataset">
             Discrete Tone Image Dataset
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Discrete Tone Images(DTI)are available which needs to be analyzed
in detail. Here, we created this dataset for those who do research in DTI.
 </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        71
       </p>
      </td>
      <td>
       <p class="normal">
        11
       </p>
      </td>
      <td>
       <p class="normal">
```

```
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/News+Popularity+in+Multiple+Social+Media+Platforms">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/News+Popularity+in+Multiple+Social+Media+Platforms">
              News Popularity in Multiple Social Media Platforms
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Large data set of news items and their respective social
feedback on multiple platforms: Facebook, Google+ and LinkedIn. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series, Text
        </p>
       </td>
       <td>
        <p class="normal">
         Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         93239
        </p>
       </td>
       <td>
        <p class="normal">
         11
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Ultrasonic+flowmeter+diagnostics">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Ultrasonic+flowmeter+diagnostics">
              Ultrasonic flowmeter diagnostics
             </a>
            </b>
```

```
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Fault diagnosis of four liquid ultrasonic flowmeters </p>
</td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          540
         </p>
        </td>
        <td>
         <p class="normal">
          173
         </p>
        </td>
        <td>
         <p class="normal">
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/ICMLA+2014+Accepted+Papers+Data+Set">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/ICMLA+2014+Accepted+Papers+Data+Set">
               ICMLA 2014 Accepted Papers Data Set
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This data set compromises the metadata for the 2014 ICMLA
conference's accepted papers, including ID, paper titles, author's keywords, abstracts and session
s in which they were exposed. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
```

```
        <td>
         <p class="normal">
          105
         </p>
        </td>
        <td>
         <p class="normal">
          5
         </p>
        </td>
        <td>
         <p class="normal">
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/BLE+RSSI+Dataset+for+Indoor+localization+and+Navigation">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/BLE+RSSI+Dataset+for+Indoor+localization+and+Navigation">
               BLE RSSI Dataset for Indoor localization and Navigation
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This dataset contains RSSI readings gathered from an array of
Bluetooth Low Energy (BLE) iBeacons in a real-world and operational indoor environment for localiz
ation and navigation purposes. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Sequential, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          6611
         </p>
        </td>
        <td>
         <p class="normal">
          15
         </p>
        </td>
        <td>
         <p class="normal">
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
```

```
          <td>
           <a href="datasets/Container+Crane+Controller+Data+Set">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Container+Crane+Controller+Data+Set">
              Container Crane Controller Data Set
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">A container crane has the function of transporting containers
from one point to another point. </p></td> -->
       <td>
        <p class="normal">
         Univariate, Domain-Theory
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         15
        </p>
       </td>
       <td>
        <p class="normal">
         3
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Residential+Building+Data+Set">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Residential+Building+Data+Set">
              Residential Building Data Set
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Data set includes construction cost, sale prices, project
variables, and economic variables corresponding to real estate single-family residential
apartments in Tehran, Iran.  </p></td> -->
       <td>
        <p class="normal">
```

```
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          372
         </p>
        </td>
        <td>
         <p class="normal">
          105
         </p>
        </td>
        <td>
         <p class="normal">
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Health+News+in+Twitter">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Health+News+in+Twitter">
               Health News in Twitter
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The data was collected in 2015 using Twitter API. This dataset cc
ntains health news from more than 15 major health news agencies such as BBC, CNN, and NYT.  <
/p></td> -->
        <td>
         <p class="normal">
          Text
         </p>
        </td>
        <td>
         <p class="normal">
          Clustering
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          58000
         </p>
        </td>
        <td>
         <p class="normal">
          25000
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/chipseq">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/chipseq">
              chipseq
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">ChIP-seq experiments characterize protein modifications or bindir
g at
specific genomic locations in specific samples. The machine learning
problem in these data is structured binary classification. </p></td> -->
       <td>
        <p class="normal">
         Sequential
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         4960
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/SGEMM+GPU+kernel+performance">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
```

```
            <a href="datasets/SGEMM+GPU+kernel+performance">
             SGEMM GPU kernel performance
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Running times for multiplying two 2048 x 2048 matrices using a G
PU OpenCL SGEMM kernel with varying parameters (using the library 'CLTune'). </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
       <p class="normal">
        241600
       </p>
      </td>
      <td>
       <p class="normal">
        18
       </p>
      </td>
      <td>
       <p class="normal">
        2018
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Repeat+Consumption+Matrices">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Repeat+Consumption+Matrices">
             Repeat Consumption Matrices
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The dataset contains 7 datasets of User - Item matrices, where
each entry represents how many times a user consumed an item. Item is used as an umbrella term for
various categories. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Clustering
       </p>
      </td>
```

```
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          130000
         </p>
        </td>
        <td>
         <p class="normal">
          21000
         </p>
        </td>
        <td>
         <p class="normal">
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/detection_of_IoT_botnet_attacks_N_BaIoT">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/detection_of_IoT_botnet_attacks_N_BaIoT">
               detection_of_IoT_botnet_attacks_N_BaIoT
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">This dataset addresses the lack of public botnet datasets,
especially for the IoT. It suggests *real* traffic data, gathered from 9 commercial IoT devices au
thentically infected by Mirai and BASHLITE. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Sequential
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          7062606
         </p>
        </td>
        <td>
         <p class="normal">
          115
         </p>
        </td>
        <td>
         <p class="normal">
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
```

```
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Absenteeism+at+work">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Absenteeism+at+work">
              Absenteeism at work
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The database was created with records of absenteeism at work
from July 2007 to July 2010 at a courier company in Brazil. </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         740
        </p>
       </td>
       <td>
        <p class="normal">
         21
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Business </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/SCADI">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/SCADI">
              SCADI
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
```

```
      <!-- <td><p class="normal">First self-care activities dataset based on ICF-CY. </p>
</td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Clustering
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        70
       </p>
      </td>
      <td>
       <p class="normal">
        206
       </p>
      </td>
      <td>
       <p class="normal">
        2018
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Condition+monitoring+of+hydraulic+systems">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Condition+monitoring+of+hydraulic+systems">
             Condition monitoring of hydraulic systems
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The data set addresses the condition assessment of a hydraulic
test rig based on multi sensor data. Four fault types are superimposed with several severity
grades impeding selective quantification. </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        2205
       </p>
      </td>
```

```
      <td>
       <p class="normal">
        43680
       </p>
      </td>
      <td>
       <p class="normal">
        2018
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Carbon+Nanotubes">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Carbon+Nanotubes">
             Carbon Nanotubes
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This dataset contains 10721 initial and calculated atomic
coordinates of carbon nanotubes. </p></td> -->
      <td>
       <p class="normal">
        Univariate
       </p>
      </td>
      <td>
       <p class="normal">
        Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        10721
       </p>
      </td>
      <td>
       <p class="normal">
        8
       </p>
      </td>
      <td>
       <p class="normal">
        2018
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Optical+Interconnection+Network+">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
```

```
          <p class="normal">
           <b>
            <a href="datasets/Optical+Interconnection+Network+">
             Optical Interconnection Network
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">This dataset contains 640 performance measurements from a
simulation of 2-Dimensional Multiprocessor Optical Interconnection Network.  </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Integer, Real
       </p>
      </td>
      <td>
       <p class="normal">
        640
       </p>
      </td>
      <td>
       <p class="normal">
        10
       </p>
      </td>
      <td>
       <p class="normal">
        2018
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Sports+articles+for+objectivity+analysis">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Sports+articles+for+objectivity+analysis">
             Sports articles for objectivity analysis
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">1000 sports articles were labeled using Amazon Mechanical Turk
as objective or subjective. The raw texts, extracted features, and the URLs from which the article
s were retrieved are provided. </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Text
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         1000
        </p>
       </td>
       <td>
        <p class="normal">
         59
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Social </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Breast+Cancer+Coimbra">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Breast+Cancer+Coimbra">
              Breast Cancer Coimbra
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Clinical features were observed or measured for 64 patients with
breast cancer and 52 healthy controls.  </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         116
        </p>
       </td>
       <td>
        <p class="normal">
         10
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
```

```
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data">
             GNFUV Unmanned Surface Vehicles Sensor Data
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The data-set contains four (4) sets of mobile sensor readings
data (humidity, temperature) corresponding to a swarm of four (4) Unmanned Surface Vehicles (USVs)
in a test-bed in Athens (Greece).  </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Regression
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        1672
       </p>
      </td>
      <td>
       <p class="normal">
        5
       </p>
      </td>
      <td>
       <p class="normal">
        2018
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Dishonest+Internet+users+Dataset">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Dishonest+Internet+users+Dataset">
             Dishonest Internet users Dataset
            </a>
           </b>
          </p>
         </td>
        </tr>
```

```
        </table>
       </td>
       <!-- <td><p class="normal">The dataset was used to test an architecture based on a trust mod
el capable to cope with the evaluation of the trustworthiness of users interacting in pervasive en
vironments. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         322
        </p>
       </td>
       <td>
        <p class="normal">
         5
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Victorian+Era+Authorship+Attribution">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Victorian+Era+Authorship+Attribution">
              Victorian Era Authorship Attribution
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">To create the largest authorship attribution dataset, we
extracted works of 50 well-known authors. To have a non-exhaustive learning, in training there are
45 authors whereas, in the testing, it's 50 </p></td> -->
       <td>
        <p class="normal">
         Text
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         93600
```

```
        95000
       </p>
      </td>
      <td>
       <p class="normal">
        1000
       </p>
      </td>
      <td>
       <p class="normal">
        2018
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Simulated+Falls+and+Daily+Living+Activities+Data+Set">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Simulated+Falls+and+Daily+Living+Activities+Data+Set">
             Simulated Falls and Daily Living Activities Data Set
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">20 falls and 16 daily living activities were performed by 17 volu
nteers with 5 repetitions while wearing 6 sensors (3.060 instances) that attached to their head, c
hest, waist, wrist, thigh and ankle. </p></td> -->
      <td>
       <p class="normal">
        Time-Series
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
       <p class="normal">
        3060
       </p>
      </td>
      <td>
       <p class="normal">
        138
       </p>
      </td>
      <td>
       <p class="normal">
        2018
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Multimodal+Damage+Identification+for+Humanitarian+Computing">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
```

```
          <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
         </a>
        </td>
        <td>
         <p class="normal">
          <b>
           <a href="datasets/Multimodal+Damage+Identification+for+Humanitarian+Computing">
            Multimodal Damage Identification for Humanitarian Computing
           </a>
          </b>
         </p>
        </td>
       </tr>
      </table>
     </td>
     <!-- <td><p class="normal">5879 captioned images (image and text) from social media related
to damage during natural disasters/wars, and belong to 6 classes: Fires, Floods, Natural
landscape, Infrastructural, Human, Non-damage. </p></td> -->
     <td>
      <p class="normal">
       Multivariate, Text
      </p>
     </td>
     <td>
      <p class="normal">
       Classification
      </p>
     </td>
     <td>
      <p class="normal">
       Integer
      </p>
     </td>
     <td>
      <p class="normal">
       5879
      </p>
     </td>
     <td>
      <p class="normal">
      </p>
     </td>
     <td>
      <p class="normal">
       2018
      </p>
     </td>
     <!-- <td><p class="normal">Social </p></td> -->
    </tr>
    <tr bgcolor="DDEEFF">
     <td>
      <table>
       <tr>
        <td>
         <a href="datasets/EEG+Steady-State+Visual+Evoked+Potential+Signals">
          <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
         </a>
        </td>
        <td>
         <p class="normal">
          <b>
           <a href="datasets/EEG+Steady-State+Visual+Evoked+Potential+Signals">
            EEG Steady-State Visual Evoked Potential Signals
           </a>
          </b>
         </p>
        </td>
       </tr>
      </table>
     </td>
     <!-- <td><p class="normal">This database consists on 30 subjects performing Brain Computer
Interface for Steady State Visual Evoked Potentials (BCI-SSVEP).  </p></td> -->
     <td>
      <p class="normal">
       Multivariate, Time-Series
      </p>
     </td>
     <td>
```

```
        <td>
         <p class="normal">
          Classification, Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          9200
         </p>
        </td>
        <td>
         <p class="normal">
          16
         </p>
        </td>
        <td>
         <p class="normal">
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Roman+Urdu+Data+Set">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Roman+Urdu+Data+Set">
               Roman Urdu Data Set
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Roman Urdu (the scripting style for Urdu language) is one of the
limited resource languages.A data corpus comprising of more than 20000 records was
collected. </p></td> -->
        <td>
         <p class="normal">
          Text
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          20000
         </p>
        </td>
        <td>
         <p class="normal">
          2
         </p>
        </td>
        <td>
         <p class="normal">
```

```
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Avila">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Avila">
               Avila
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The Avila data set has been extracted from 800 images of the
'Avila Bible', an XII century giant Latin copy of the Bible.  The prediction task consists in
associating each pattern to a copyist. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          20867
         </p>
        </td>
        <td>
         <p class="normal">
          10
         </p>
        </td>
        <td>
         <p class="normal">
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/PANDOR">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/PANDOR">
               PANDOR
              </a>
             </b>
```

```
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">PANDOR is a novel and publicly available dataset for online
recommendation provided by Purch (http://www.purch.com/).  </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Recommendation
       </p>
      </td>
      <td>
       <p class="normal">
        Categorical
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        2018
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Drug+Review+Dataset+%28Druglib.com%29">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Drug+Review+Dataset+%28Druglib.com%29">
             Drug Review Dataset (Druglib.com)
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The dataset provides patient reviews on specific drugs along with
related conditions. Reviews and ratings are grouped into reports on the three aspects benefits, si
de effects and overall comment. </p></td> -->
      <td>
       <p class="normal">
        Multivariate, Text
       </p>
      </td>
      <td>
       <p class="normal">
        Classification, Regression, Clustering
       </p>
      </td>
      <td>
       <p class="normal">
        Integer
       </p>
      </td>
      <td>
```

```
        <p class="normal">
         4143
        </p>
       </td>
       <td>
        <p class="normal">
         8
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal"> </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Drug+Review+Dataset+%28Drugs.com%29">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Drug+Review+Dataset+%28Drugs.com%29">
              Drug Review Dataset (Drugs.com)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The dataset provides patient reviews on specific drugs along with
related conditions and a 10 star patient rating reflecting overall patient satisfaction. </p>
</td> -->
       <td>
        <p class="normal">
         Multivariate, Text
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         215063
        </p>
       </td>
       <td>
        <p class="normal">
         6
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
```

```
            <a href="datasets/Physical+Unclonable+Functions">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Physical+Unclonable+Functions">
               Physical Unclonable Functions
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The dataset is generated from Physical Unclonable Functions (PUFs
) simulation, specifically XOR Arbiter PUFs. PUFs are used for authentication purposes. For more i
nfo, refer to our paper below. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
        </td>
        <td>
         <p class="normal">
          6000000
         </p>
        </td>
        <td>
         <p class="normal">
          129
         </p>
        </td>
        <td>
         <p class="normal">
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Superconductivty+Data">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Superconductivty+Data">
               Superconductivty Data
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Two file s contain data on 21263 superconductors and their releva
nt features. </p></td> -->
        <td>
         <p class="normal">
          Multivariate
```

```
        </p>
       </td>
       <td>
        <p class="normal">
         Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         21263
        </p>
       </td>
       <td>
        <p class="normal">
         81
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Physical </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/WESAD+%28Wearable+Stress+and+Affect+Detection%29">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/WESAD+%28Wearable+Stress+and+Affect+Detection%29">
              WESAD (Wearable Stress and Affect Detection)
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">WESAD (Wearable Stress and Affect Detection) contains data of 15
subjects during a stress-affect lab study, while wearing physiological and motion sensors. </
p></td> -->
       <td>
        <p class="normal">
         Multivariate, Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Regression
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         63000000
        </p>
       </td>
       <td>
        <p class="normal">
         12
        </p>
```

```
        </td>
        <td>
         <p class="normal">
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data+Set+2">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data+Set+2">
               GNFUV Unmanned Surface Vehicles Sensor Data Set 2
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The data-set contains eight (2x4) data-sets of mobile sensor reac
ings data (humidity, temperature) corresponding to a swarm of four Unmanned Surface Vehicles
(USVs) in a test-bed, Athens, Greece. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Sequential, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          10190
         </p>
        </td>
        <td>
         <p class="normal">
          6
         </p>
        </td>
        <td>
         <p class="normal">
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Student+Academics+Performance">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Student+Academics+Performance">
```

```
              Student Academics Performance
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The dataset tried to find the end semester percentage prediction
based on different social, economic and academic attributes.  </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         300
        </p>
       </td>
       <td>
        <p class="normal">
         22
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Online+Shoppers+Purchasing+Intention+Dataset">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Online+Shoppers+Purchasing+Intention+Dataset">
              Online Shoppers Purchasing Intention Dataset
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Of the 12,330 sessions in the dataset,
84.5% (10,422) were negative class samples that did not
end with shopping, and the rest (1908) were positive class
samples ending with shopping. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
```

```
        <p class="normal">
         Integer, Real
        </p>
       </td>
       <td>
        <p class="normal">
         12330
        </p>
       </td>
       <td>
        <p class="normal">
         18
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Business </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/PMU-UD">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/PMU-UD">
              PMU-UD
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">The handwritten dataset was collected from 170 participants with
a total of 5,180 numeral patterns. The dataset is named Prince Mohammad Bin Fahd University -
Urdu/Arabic Database (PMU-UD).  </p></td> -->
       <td>
        <p class="normal">
         Univariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         5180
        </p>
       </td>
       <td>
        <p class="normal">
         9
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr>
```

```
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Parkinson%27s+Disease+Classification">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Parkinson%27s+Disease+Classification">
             Parkinson's Disease Classification
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The data used in this study were gathered from 188 patients with
PD (107 men and 81 women) with ages ranging from 33 to 87 (65.1±10.9). </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Integer, Real
       </p>
      </td>
      <td>
       <p class="normal">
        756
       </p>
      </td>
      <td>
       <p class="normal">
        754
       </p>
      </td>
      <td>
       <p class="normal">
        2018
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/Electrical+Grid+Stability+Simulated+Data+">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/Electrical+Grid+Stability+Simulated+Data+">
             Electrical Grid Stability Simulated Data
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">The local stability analysis of the 4-node star system
(electricity producer is in the center) implementing Decentral Smart Grid Control concept.  <
```

/p></td> -->
<td>

Multivariate

</td>
<td>

Classification, Regression

</td>
<td>

Real

</td>
<td>

10000

</td>
<td>

14

</td>
<td>

2018

</td>
<!-- <td> Physical </td> -->
</tr>
<tr>
<td>
<table>
<tr>
<td>
<a href="datasets/Caesarian+Section+Classification+Dataset">
<img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>
</td>
<td>

<a href="datasets/Caesarian+Section+Classification+Dataset">
Caesarian Section Classification Dataset
</a>

</td>
</tr>
</table>
</td>
<!-- <td> This dataset contains information about caesarian section results of 80 pregnant women with the most important characteristics of delivery problems in the medical field. </td> -->
<td>

Univariate

</td>
<td>

Classification

</td>
<td>

Integer

</td>
<td>

80

</td>

```
        <td>
         <p class="normal">
          5
         </p>
        </td>
        <td>
         <p class="normal">
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Life </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/BAUM-1">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/BAUM-1">
               BAUM-1
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">BAUM-1 dataset contains 1184 multimodal facial video clips collec
ted from 31 subjects. The 1184 video clips contain spontaneous facial expressions and speech of 13
emotional and mental states. </p></td> -->
        <td>
         <p class="normal">
          Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          1184
         </p>
        </td>
        <td>
         <p class="normal">
         </p>
        </td>
        <td>
         <p class="normal">
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/BAUM-2">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
```

```
            <b>
             <a href="datasets/BAUM-2">
              BAUM-2
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">A multilingual audio-visual affective face database consisting
of 1047 video clips of 286 subjects.  </p></td> -->
       <td>
        <p class="normal">
         Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         1047
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Computer </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Audit+Data">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Audit+Data">
              Audit Data
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Exhaustive one year non-confidential data in the year 2015 to
2016 of firms is collected from the Auditor Office of India to build a predictor for classifying s
uspicious firms. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
```

```
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         777
        </p>
       </td>
       <td>
        <p class="normal">
         18
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/BuddyMove+Data+Set">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/BuddyMove+Data+Set">
              BuddyMove Data Set
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">User interest information extracted from user reviews published
in holidayiq.com about various types of point of interests in South India </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Text
        </p>
       </td>
       <td>
        <p class="normal">
         Classification, Clustering
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         249
        </p>
       </td>
       <td>
        <p class="normal">
         7
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Other </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
```

```
<td>
 <table>
  <tr>
   <td>
    <a href="datasets/Real+estate+valuation+data+set">
     <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
    </a>
   </td>
   <td>
    <p class="normal">
     <b>
      <a href="datasets/Real+estate+valuation+data+set">
       Real estate valuation data set
      </a>
     </b>
    </p>
   </td>
  </tr>
 </table>
</td>
<!-- <td><p class="normal">The “real estate valuation” is a regression problem. The market
historical data set of real estate valuation are collected from Sindian Dist., New Taipei City, Ta
iwan.  </p></td> -->
<td>
 <p class="normal">
  Multivariate
 </p>
</td>
<td>
 <p class="normal">
  Regression
 </p>
</td>
<td>
 <p class="normal">
  Integer, Real
 </p>
</td>
<td>
 <p class="normal">
  414
 </p>
</td>
<td>
 <p class="normal">
  7
 </p>
</td>
<td>
 <p class="normal">
  2018
 </p>
</td>
<!-- <td><p class="normal">Business </p></td> -->
</tr>
<tr>
 <td>
  <table>
   <tr>
    <td>
     <a
href="datasets/Early+biomarkers+of+Parkinson%E2%80%99s+disease+based+on+natural+connected+speech+Da
et+">
      <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
     </a>
    </td>
    <td>
     <p class="normal">
      <b>
       <a
href="datasets/Early+biomarkers+of+Parkinson%E2%80%99s+disease+based+on+natural+connected+speech+Da
et+">
        Early biomarkers of Parkinson’s disease based on natural connected speech Data Set
       </a>
      </b>
     </p>
    </td>
```

```
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">. </p></td> -->
       <td>
        <p class="normal">
         Multivariate
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         2018
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr bgcolor="DDEEFF">
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Somerville+Happiness+Survey">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Somerville+Happiness+Survey">
              Somerville Happiness Survey
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">A data extract of a non-federal dataset posted here
https://catalog.data.gov/dataset/somerville-happiness-survey-responses-2011-2013-2015 </p></t
d> -->
       <td>
        <p class="normal">
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Integer
        </p>
       </td>
       <td>
        <p class="normal">
         143
        </p>
       </td>
```

```
      </td>
      <td>
       <p class="normal">
        7
       </p>
      </td>
      <td>
       <p class="normal">
        2018
       </p>
      </td>
      <!-- <td><p class="normal">Life </p></td> -->
     </tr>
     <tr>
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/2.4+GHZ+Indoor+Channel+Measurements">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
         <td>
          <p class="normal">
           <b>
            <a href="datasets/2.4+GHZ+Indoor+Channel+Measurements">
             2.4 GHZ Indoor Channel Measurements
            </a>
           </b>
          </p>
         </td>
        </tr>
       </table>
      </td>
      <!-- <td><p class="normal">Measurement of the S21,consists of 10 sweeps, each sweep
contains 601 frequency points with spacing of 0.167MHz to cover a 100MHz band centered at
2.4GHz. </p></td> -->
      <td>
       <p class="normal">
        Multivariate
       </p>
      </td>
      <td>
       <p class="normal">
        Classification
       </p>
      </td>
      <td>
       <p class="normal">
        Real
       </p>
      </td>
      <td>
       <p class="normal">
        7840
       </p>
      </td>
      <td>
       <p class="normal">
        5
       </p>
      </td>
      <td>
       <p class="normal">
        2018
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
     <tr bgcolor="DDEEFF">
      <td>
       <table>
        <tr>
         <td>
          <a href="datasets/EMG+data+for+gestures">
           <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
          </a>
         </td>
```

```
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/EMG+data+for+gestures">
              EMG data for gestures
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">These are files of raw EMG data recorded by MYO Thalmic
bracelet </p></td> -->
       <td>
        <p class="normal">
         Time-Series
        </p>
       </td>
       <td>
        <p class="normal">
         Classification
        </p>
       </td>
       <td>
        <p class="normal">
         Real
        </p>
       </td>
       <td>
        <p class="normal">
         30000
        </p>
       </td>
       <td>
        <p class="normal">
         6
        </p>
       </td>
       <td>
        <p class="normal">
         2019
        </p>
       </td>
       <!-- <td><p class="normal">Life </p></td> -->
      </tr>
      <tr>
       <td>
        <table>
         <tr>
          <td>
           <a href="datasets/Parking+Birmingham">
            <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
           </a>
          </td>
          <td>
           <p class="normal">
            <b>
             <a href="datasets/Parking+Birmingham">
              Parking Birmingham
             </a>
            </b>
           </p>
          </td>
         </tr>
        </table>
       </td>
       <!-- <td><p class="normal">Data collected from car parks in Birmingham that are operated by
NCP from
Birmingham City Council. UK Open Government Licence (OGL).
https://data.birmingham.gov.uk/dataset/birmingham-parking </p></td> -->
       <td>
        <p class="normal">
         Multivariate, Univariate, Sequential, Time-Series
        </p>
       </td>
```

```
        <td>
         <p class="normal">
          Classification, Regression, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          35717
         </p>
        </td>
        <td>
         <p class="normal">
          4
         </p>
        </td>
        <td>
         <p class="normal">
          2019
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Behavior+of+the+urban+traffic+of+the+city+of+Sao+Paulo+in+Brazil">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Behavior+of+the+urban+traffic+of+the+city+of+Sao+Paulo+in+Brazil">
               Behavior of the urban traffic of the city of Sao Paulo in Brazil
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">The database was created with records of behavior of the urban
traffic of the city of Sao Paulo in Brazil. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Time-Series
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Regression
         </p>
        </td>
        <td>
         <p class="normal">
          Integer, Real
         </p>
        </td>
        <td>
         <p class="normal">
          135
         </p>
        </td>
        <td>
         <p class="normal">
          18
         </p>
        </td>
        <td>
         <p class="normal">
```

```
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Computer </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Travel+Reviews">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Travel+Reviews">
               Travel Reviews
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Reviews on destinations in 10 categories mentioned across East As
ia. Each traveler rating is mapped as Excellent(4), Very Good(3), Average(2), Poor(1), and Terribl
e(0) and average rating is used. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Text
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          980
         </p>
        </td>
        <td>
         <p class="normal">
          11
         </p>
        </td>
        <td>
         <p class="normal">
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr bgcolor="DDEEFF">
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Tarvel+Review+Ratings">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Tarvel+Review+Ratings">
               Tarvel Review Ratings
              </a>
             </b>
```

```
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">Google reviews on attractions from 24 categories across Europe
are considered. Google user rating ranges from 1 to 5 and average user rating per category is calc
ulated. </p></td> -->
        <td>
         <p class="normal">
          Multivariate, Text
         </p>
        </td>
        <td>
         <p class="normal">
          Classification, Clustering
         </p>
        </td>
        <td>
         <p class="normal">
          Real
         </p>
        </td>
        <td>
         <p class="normal">
          5456
         </p>
        </td>
        <td>
         <p class="normal">
          25
         </p>
        </td>
        <td>
         <p class="normal">
          2018
         </p>
        </td>
        <!-- <td><p class="normal">Other </p></td> -->
       </tr>
       <tr>
        <td>
         <table>
          <tr>
           <td>
            <a href="datasets/Rice+Leaf+Diseases">
             <img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
            </a>
           </td>
           <td>
            <p class="normal">
             <b>
              <a href="datasets/Rice+Leaf+Diseases">
               Rice Leaf Diseases
              </a>
             </b>
            </p>
           </td>
          </tr>
         </table>
        </td>
        <!-- <td><p class="normal">There are three classes/diseases: Bacterial leaf blight, Brown sp
ot, and Leaf smut, each having 40 images. The format of all images is jpg.   </p></td> -->
        <td>
         <p class="normal">
          Multivariate
         </p>
        </td>
        <td>
         <p class="normal">
          Classification
         </p>
        </td>
        <td>
         <p class="normal">
          Integer
         </p>
```

```
      </td>
      <td>
       <p class="normal">
        120
       </p>
      </td>
      <td>
       <p class="normal">
       </p>
      </td>
      <td>
       <p class="normal">
        2019
       </p>
      </td>
      <!-- <td><p class="normal">Computer </p></td> -->
     </tr>
    </table>
   </td>
  </tr>
 </table>
 <table align="center" cellpadding="5">
  <tr valign="center">
   <td>
    <p class="normal">
     Supported By:
    </p>
   </td>
   <td>
    <img height="60" src="assets/nsfe.gif"/>
   </td>
   <td>
    <p class="normal">
     In Collaboration With:
    </p>
   </td>
   <td>
    <img src="assets/rexaSmall.jpg"/>
   </td>
  </tr>
 </table>
 <center>
  <span class="normal">
   <a href="about.html">
    About
   </a>
   ||
   <a href="citation_policy.html">
    Citation Policy
   </a>
   ||
   <a href="donation_policy.html">
    Donation Policy
   </a>
   ||
   <a href="contact.html">
    Contact
   </a>
   ||
   <a href="http://cml.ics.uci.edu">
    CML
   </a>
  </span>
 </center>
</body>
</html>
```

Link a la base de datos

In [11]:

```
links = soup.find_all('a')
```

In [12]:

```
links
```

Out[12]:

```
[<a alt="Home" href="index.html"><img border="0" src="assets/logo.gif"/></a>,
 <a href="http://cml.ics.uci.edu"><font color="FFDD33">Center for Machine Learning and Intelligent
Systems</font></a>,
 <a href="about.html">About</a>,
 <a href="citation_policy.html">Citation Policy</a>,
 <a href="donation_policy.html">Donate a Data Set</a>,
 <a href="contact.html">Contact</a>,
 <a href="http://www.google.com/search"><img align="middle" alt="Google" border="0" height="27" sr
c="http://www.google.com/logos/Logo_25blk.gif"/></a>,
 <a href="datasets.php"><font color="#FFDD33" size="3"><b>View ALL Data Sets</b></font></a>,
 <a href="datasets.php?
format=&amp;task=cla&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view=
e">Classification</a>,
 <a href="datasets.php?
format=&amp;task=reg&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view=
e">Regression</a>,
 <a href="datasets.php?
format=&amp;task=clu&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view=
e">Clustering</a>,
 <a href="datasets.php?
format=&amp;task=other&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;vie
ble">Other</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=cat&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view=
e">Categorical</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=num&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view=
e">Numerical</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=mix&amp;area=&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view=
e">Mixed</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=mvar&amp;sort=nameUp&amp;view
le">Multivariate</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=uvar&amp;sort=nameUp&amp;view
le">Univariate</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=seq&amp;sort=nameUp&amp;view=
e">Sequential</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=ts&amp;sort=nameUp&amp;view=t
">Time-Series</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=text&amp;sort=nameUp&amp;view
le">Text</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=dt&amp;sort=nameUp&amp;view=t
">Domain-Theory</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=&amp;numAtt=&amp;numIns=&amp;type=other&amp;sort=nameUp&amp;vie
ble">Other</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=life&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view
le">Life Sciences</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=phys&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view
le">Physical Sciences</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=comp&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view
le">CS / Engineering</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=soc&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view=
e">Social Sciences</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=bus&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view=
e">Business</a>,
 <a href="datasets.php?
format=&amp;task=&amp;att=&amp;area=game&amp;numAtt=&amp;numIns=&amp;type=&amp;sort=nameUp&amp;view
le">Game</a>,
 <a href="datasets.php?
```

format=&task=&att=&area=other&numAtt=&numIns=&type=&sort=nameUp&vie
ble">Other</a>,
<a href="datasets.php?
format=&task=&att=&area=&numAtt=less10&numIns=&type=&sort=nameUp&vi
able">Less than 10</a>,
<a href="datasets.php?
format=&task=&att=&area=&numAtt=10to100&numIns=&type=&sort=nameUp&v
table">10 to 100</a>,
<a href="datasets.php?
format=&task=&att=&area=&numAtt=greater100&numIns=&type=&sort=nameUp&am
ew=table">Greater than 100</a>,
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=less100&type=&sort=nameUp&v
table">Less than 100</a>,
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=100to1000&type=&sort=nameUp&amp
w=table">100 to 1000</a>,
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=greater1000&type=&sort=nameUp&a
iew=table">Greater than 1000</a>,
<a href="datasets.php?
format=mat&task=&att=&area=&numAtt=&numIns=&type=&sort=nameUp&view=
e">Matrix</a>,
<a href="datasets.php?
format=nonmat&task=&att=&area=&numAtt=&numIns=&type=&sort=nameUp&vi
able">Non-Matrix</a>,
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=nameUp&view=lis
ist View</a>,
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=nameDown&view=t
">Name</a>,
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=typeUp&view=tab
Data Types</a>,
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=taskUp&view=tab
Default Task</a>,
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=attTypeUp&view=
e">Attribute Types</a>,
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=instUp&view=tab
# Instances</a>,
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=attUp&view=tabl
# Attributes</a>,
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=dateUp&view=tab
Year</a>,
<a href="datasets/Abalone"><img border="1" src="assets/MLimages/SmallLarge1.jpg"/></a>,
<a href="datasets/Abalone">Abalone</a>,
<a href="datasets/Adult"><img border="1" src="assets/MLimages/SmallLarge2.jpg"/></a>,
<a href="datasets/Adult">Adult</a>,
<a href="datasets/Annealing"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
<a href="datasets/Annealing">Annealing</a>,
<a href="datasets/Anonymous+Microsoft+Web+Data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
<a href="datasets/Anonymous+Microsoft+Web+Data">Anonymous Microsoft Web Data</a>,
<a href="datasets/Arrhythmia"><img border="1" src="assets/MLimages/SmallLarge5.jpg"/></a>,
<a href="datasets/Arrhythmia">Arrhythmia</a>,
<a href="datasets/Artificial+Characters"><img border="1" src="assets/MLimages/SmallLarge6.jpg"/><
/a>,
<a href="datasets/Artificial+Characters">Artificial Characters</a>,
<a href="datasets/Audiology+%28Original%29"><img border="1"
src="assets/MLimages/SmallLarge7.jpg"/></a>,
<a href="datasets/Audiology+%28Original%29">Audiology (Original)</a>,
<a href="datasets/Audiology+%28Standardized%29"><img border="1"
src="assets/MLimages/SmallLarge7.jpg"/></a>,
<a href="datasets/Audiology+%28Standardized%29">Audiology (Standardized)</a>,
<a href="datasets/Auto+MPG"><img border="1" src="assets/MLimages/SmallLarge9.jpg"/></a>,
<a href="datasets/Auto+MPG">Auto MPG</a>,
<a href="datasets/Automobile"><img border="1" src="assets/MLimages/SmallLarge10.jpg"/></a>,
<a href="datasets/Automobile">Automobile</a>,
<a href="datasets/Badges"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
<a href="datasets/Badges">Badges</a>,
<a href="datasets/Balance+Scale"><img border="1" src="assets/MLimages/SmallLarge12.jpg"/></a>,

```
 <a href="datasets/Balance+Scale">Balance Scale</a>,
 <a href="datasets/Balloons"><img border="1" src="assets/MLimages/SmallLarge13.jpg"/></a>,
 <a href="datasets/Balloons">Balloons</a>,
 <a href="datasets/Breast+Cancer"><img border="1" src="assets/MLimages/SmallLarge14.jpg"/></a>,
 <a href="datasets/Breast+Cancer">Breast Cancer</a>,
 <a href="datasets/Breast+Cancer+Wisconsin+%28Original%29"><img border="1"
src="assets/MLimages/SmallLarge14.jpg"/></a>,
 <a href="datasets/Breast+Cancer+Wisconsin+%28Original%29">Breast Cancer Wisconsin (Original)</a>,
 <a href="datasets/Breast+Cancer+Wisconsin+%28Prognostic%29"><img border="1"
src="assets/MLimages/SmallLarge14.jpg"/></a>,
 <a href="datasets/Breast+Cancer+Wisconsin+%28Prognostic%29">Breast Cancer Wisconsin (Prognostic)<
/a>,
 <a href="datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29"><img border="1"
src="assets/MLimages/SmallLarge14.jpg"/></a>,
 <a href="datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29">Breast Cancer Wisconsin (Diagnostic)<
/a>,
 <a href="datasets/Pittsburgh+Bridges"><img border="1" src="assets/MLimages/SmallLarge18.jpg"/></a
>,
 <a href="datasets/Pittsburgh+Bridges">Pittsburgh Bridges</a>,
 <a href="datasets/Car+Evaluation"><img border="1" src="assets/MLimages/SmallLarge19.jpg"/></a>,
 <a href="datasets/Car+Evaluation">Car Evaluation</a>,
 <a href="datasets/Census+Income"><img border="1" src="assets/MLimages/SmallLarge2.jpg"/></a>,
 <a href="datasets/Census+Income">Census Income</a>,
 <a href="datasets/Chess+%28King-Rook+vs.+King-Knight%29"><img border="1"
src="assets/MLimages/SmallLarge24.jpg"/></a>,
 <a href="datasets/Chess+%28King-Rook+vs.+King-Knight%29">Chess (King-Rook vs. King-Knight)</a>,
 <a href="datasets/Chess+%28King-Rook+vs.+King-Pawn%29"><img border="1"
src="assets/MLimages/SmallLarge24.jpg"/></a>,
 <a href="datasets/Chess+%28King-Rook+vs.+King-Pawn%29">Chess (King-Rook vs. King-Pawn)</a>,
 <a href="datasets/Chess+%28King-Rook+vs.+King%29"><img border="1"
src="assets/MLimages/SmallLarge24.jpg"/></a>,
 <a href="datasets/Chess+%28King-Rook+vs.+King%29">Chess (King-Rook vs. King)</a>,
 <a href="datasets/Chess+%28Domain+Theories%29"><img border="1"
src="assets/MLimages/SmallLarge24.jpg"/></a>,
 <a href="datasets/Chess+%28Domain+Theories%29">Chess (Domain Theories)</a>,
 <a href="datasets/Bach+Chorales"><img border="1" src="assets/MLimages/SmallLarge25.jpg"/></a>,
 <a href="datasets/Bach+Chorales">Bach Chorales</a>,
 <a href="datasets/Connect-4"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Connect-4">Connect-4</a>,
 <a href="datasets/Credit+Approval"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/><
/a>,
 <a href="datasets/Credit+Approval">Credit Approval</a>,
 <a href="datasets/Japanese+Credit+Screening"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Japanese+Credit+Screening">Japanese Credit Screening</a>,
 <a href="datasets/Computer+Hardware"><img border="1" src="assets/MLimages/SmallLarge29.jpg"/></a>
,
 <a href="datasets/Computer+Hardware">Computer Hardware</a>,
 <a href="datasets/Contraceptive+Method+Choice"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Contraceptive+Method+Choice">Contraceptive Method Choice</a>,
 <a href="datasets/Covertype"><img border="1" src="assets/MLimages/SmallLarge31.jpg"/></a>,
 <a href="datasets/Covertype">Covertype</a>,
 <a href="datasets/Cylinder+Bands"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></
a>,
 <a href="datasets/Cylinder+Bands">Cylinder Bands</a>,
 <a href="datasets/Dermatology"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Dermatology">Dermatology</a>,
 <a href="datasets/Diabetes"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Diabetes">Diabetes</a>,
 <a href="datasets/DGP2+-+The+Second+Data+Generation+Program"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/DGP2+-+The+Second+Data+Generation+Program">DGP2 - The Second Data Generation Pr
ogram</a>,
 <a href="datasets/Document+Understanding"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Document+Understanding">Document Understanding</a>,
 <a href="datasets/EBL+Domain+Theories"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/EBL+Domain+Theories">EBL Domain Theories</a>,
 <a href="datasets/Echocardiogram"><img border="1" src="assets/MLimages/SmallLarge38.jpg"/></a>,
 <a href="datasets/Echocardiogram">Echocardiogram</a>,
 <a href="datasets/Ecoli"><img border="1" src="assets/MLimages/SmallLarge120.jpg"/></a>,
 <a href="datasets/Ecoli">Ecoli</a>,
 <a href="datasets/Flags"><img border="1" src="assets/MLimages/SmallLarge40.jpg"/></a>,
 <a href="datasets/Flags">Flags</a>,
 <a href="datasets/Function+Finding"><img border="1" src="assets/MLimages/SmallLarge41.jpg"/></a>,
```

```
 <a href="datasets/Function+Finding">Function Finding</a>,
 <a href="datasets/Glass+Identification"><img border="1" src="assets/MLimages/SmallLarge42.jpg"/><
/a>,
 <a href="datasets/Glass+Identification">Glass Identification</a>,
 <a href="datasets/Haberman%27s+Survival"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Haberman%27s+Survival">Haberman's Survival</a>,
 <a href="datasets/Hayes-Roth"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Hayes-Roth">Hayes-Roth</a>,
 <a href="datasets/Heart+Disease"><img border="1" src="assets/MLimages/SmallLarge45.jpg"/></a>,
 <a href="datasets/Heart+Disease">Heart Disease</a>,
 <a href="datasets/Hepatitis"><img border="1" src="assets/MLimages/SmallLarge46.jpg"/></a>,
 <a href="datasets/Hepatitis">Hepatitis</a>,
 <a href="datasets/Horse+Colic"><img border="1" src="assets/MLimages/SmallLarge47.jpg"/></a>,
 <a href="datasets/Horse+Colic">Horse Colic</a>,
 <a href="datasets/ICU"><img border="1" src="assets/MLimages/SmallLarge49.jpg"/></a>,
 <a href="datasets/ICU">ICU</a>,
 <a href="datasets/Image+Segmentation"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Image+Segmentation">Image Segmentation</a>,
 <a href="datasets/Internet+Advertisements"><img border="1"
src="assets/MLimages/SmallLarge51.jpg"/></a>,
 <a href="datasets/Internet+Advertisements">Internet Advertisements</a>,
 <a href="datasets/Ionosphere"><img border="1" src="assets/MLimages/SmallLarge52.jpg"/></a>,
 <a href="datasets/Ionosphere">Ionosphere</a>,
 <a href="datasets/Iris"><img border="1" src="assets/MLimages/SmallLarge53.jpg"/></a>,
 <a href="datasets/Iris">Iris</a>,
 <a href="datasets/ISOLET"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/ISOLET">ISOLET</a>,
 <a href="datasets/Kinship"><img border="1" src="assets/MLimages/SmallLarge55.jpg"/></a>,
 <a href="datasets/Kinship">Kinship</a>,
 <a href="datasets/Labor+Relations"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/><
/a>,
 <a href="datasets/Labor+Relations">Labor Relations</a>,
 <a href="datasets/LED+Display+Domain"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/LED+Display+Domain">LED Display Domain</a>,
 <a href="datasets/Lenses"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Lenses">Lenses</a>,
 <a href="datasets/Letter+Recognition"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Letter+Recognition">Letter Recognition</a>,
 <a href="datasets/Liver+Disorders"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/><
/a>,
 <a href="datasets/Liver+Disorders">Liver Disorders</a>,
 <a href="datasets/Logic+Theorist"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></
a>,
 <a href="datasets/Logic+Theorist">Logic Theorist</a>,
 <a href="datasets/Lung+Cancer"><img border="1" src="assets/MLimages/SmallLarge62.jpg"/></a>,
 <a href="datasets/Lung+Cancer">Lung Cancer</a>,
 <a href="datasets/Lymphography"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>
,
 <a href="datasets/Lymphography">Lymphography</a>,
 <a href="datasets/Mechanical+Analysis"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Mechanical+Analysis">Mechanical Analysis</a>,
 <a href="datasets/Meta-data"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Meta-data">Meta-data</a>,
 <a href="datasets/Mobile+Robots"><img border="1" src="assets/MLimages/SmallLarge66.jpg"/></a>,
 <a href="datasets/Mobile+Robots">Mobile Robots</a>,
 <a href="datasets/Molecular+Biology+%28Promoter+Gene+Sequences%29"><img border="1"
src="assets/MLimages/SmallLarge67.jpg"/></a>,
 <a href="datasets/Molecular+Biology+%28Promoter+Gene+Sequences%29">Molecular Biology (Promoter Ge
ne Sequences)</a>,
 <a href="datasets/Molecular+Biology+%28Protein+Secondary+Structure%29"><img border="1"
src="assets/MLimages/SmallLarge67.jpg"/></a>,
 <a href="datasets/Molecular+Biology+%28Protein+Secondary+Structure%29">Molecular Biology (Protein
Secondary Structure)</a>,
 <a href="datasets/Molecular+Biology+%28Splice-junction+Gene+Sequences%29"><img border="1"
src="assets/MLimages/SmallLarge67.jpg"/></a>,
 <a href="datasets/Molecular+Biology+%28Splice-junction+Gene+Sequences%29">Molecular Biology
(Splice-junction Gene Sequences)</a>,
 <a href="datasets/MONK%27s+Problems"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/MONK%27s+Problems">MONK's Problems</a>,
 <a href="datasets/Moral+Reasoner"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></
a>,
```

```
 <a href="datasets/Moral+Reasoner">Moral Reasoner</a>,
 <a href="datasets/Multiple+Features"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Multiple+Features">Multiple Features</a>,
 <a href="datasets/Mushroom"><img border="1" src="assets/MLimages/SmallLarge73.jpg"/></a>,
 <a href="datasets/Mushroom">Mushroom</a>,
 <a href="datasets/Musk+%28Version+1%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Musk+%28Version+1%29">Musk (Version 1)</a>,
 <a href="datasets/Musk+%28Version+2%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Musk+%28Version+2%29">Musk (Version 2)</a>,
 <a href="datasets/Nursery"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Nursery">Nursery</a>,
 <a href="datasets/Othello+Domain+Theory"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Othello+Domain+Theory">Othello Domain Theory</a>,
 <a href="datasets/Page+Blocks+Classification"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Page+Blocks+Classification">Page Blocks Classification</a>,
 <a href="datasets/Optical+Recognition+of+Handwritten+Digits"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Optical+Recognition+of+Handwritten+Digits">Optical Recognition of Handwritten D
igits</a>,
 <a href="datasets/Pen-Based+Recognition+of+Handwritten+Digits"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Pen-Based+Recognition+of+Handwritten+Digits">Pen-Based Recognition of
Handwritten Digits</a>,
 <a href="datasets/Post-Operative+Patient"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Post-Operative+Patient">Post-Operative Patient</a>,
 <a href="datasets/Primary+Tumor"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a
>,
 <a href="datasets/Primary+Tumor">Primary Tumor</a>,
 <a href="datasets/Prodigy"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Prodigy">Prodigy</a>,
 <a href="datasets/Qualitative+Structure+Activity+Relationships"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Qualitative+Structure+Activity+Relationships">Qualitative Structure Activity Re
lationships</a>,
 <a href="datasets/Quadruped+Mammals"><img border="1" src="assets/MLimages/SmallLarge86.jpg"/></a>
,
 <a href="datasets/Quadruped+Mammals">Quadruped Mammals</a>,
 <a href="datasets/Servo"><img border="1" src="assets/MLimages/SmallLarge87.jpg"/></a>,
 <a href="datasets/Servo">Servo</a>,
 <a href="datasets/Shuttle+Landing+Control"><img border="1"
src="assets/MLimages/SmallLarge92.jpg"/></a>,
 <a href="datasets/Shuttle+Landing+Control">Shuttle Landing Control</a>,
 <a href="datasets/Solar+Flare"><img border="1" src="assets/MLimages/SmallLarge89.jpg"/></a>,
 <a href="datasets/Solar+Flare">Solar Flare</a>,
 <a href="datasets/Soybean+%28Large%29"><img border="1" src="assets/MLimages/SmallLarge90.jpg"/></
a>,
 <a href="datasets/Soybean+%28Large%29">Soybean (Large)</a>,
 <a href="datasets/Soybean+%28Small%29"><img border="1" src="assets/MLimages/SmallLarge90.jpg"/></
a>,
 <a href="datasets/Soybean+%28Small%29">Soybean (Small)</a>,
 <a href="datasets/Challenger+USA+Space+Shuttle+O-Ring"><img border="1"
src="assets/MLimages/SmallLarge92.jpg"/></a>,
 <a href="datasets/Challenger+USA+Space+Shuttle+O-Ring">Challenger USA Space Shuttle O-Ring</a>,
 <a href="datasets/Low+Resolution+Spectrometer"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Low+Resolution+Spectrometer">Low Resolution Spectrometer</a>,
 <a href="datasets/Spambase"><img border="1" src="assets/MLimages/SmallLarge94.jpg"/></a>,
 <a href="datasets/Spambase">Spambase</a>,
 <a href="datasets/SPECT+Heart"><img border="1" src="assets/MLimages/SmallLarge45.jpg"/></a>,
 <a href="datasets/SPECT+Heart">SPECT Heart</a>,
 <a href="datasets/SPECTF+Heart"><img border="1" src="assets/MLimages/SmallLarge45.jpg"/></a>,
 <a href="datasets/SPECTF+Heart">SPECTF Heart</a>,
 <a href="datasets/Sponge"><img border="1" src="assets/MLimages/SmallLarge97.jpg"/></a>,
 <a href="datasets/Sponge">Sponge</a>,
 <a href="datasets/Statlog+Project"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/><
/a>,
 <a href="datasets/Statlog+Project">Statlog Project</a>,
 <a href="datasets/Student+Loan+Relational"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Student+Loan+Relational">Student Loan Relational</a>,
 <a href="datasets/Teaching+Assistant+Evaluation"><img border="1"
```

src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Teaching+Assistant+Evaluation">Teaching Assistant Evaluation</a>,
 <a href="datasets/Tic-Tac-Toe+Endgame"><img border="1" src="assets/MLimages/SmallLarge101.jpg"/><
/a>,
 <a href="datasets/Tic-Tac-Toe+Endgame">Tic-Tac-Toe Endgame</a>,
 <a href="datasets/Thyroid+Disease"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/><
/a>,
 <a href="datasets/Thyroid+Disease">Thyroid Disease</a>,
 <a href="datasets/Trains"><img border="1" src="assets/MLimages/SmallLarge103.jpg"/></a>,
 <a href="datasets/Trains">Trains</a>,
 <a href="datasets/University"><img border="1" src="assets/MLimages/SmallLarge104.jpg"/></a>,
 <a href="datasets/University">University</a>,
 <a href="datasets/Congressional+Voting+Records"><img border="1"
src="assets/MLimages/SmallLarge105.jpg"/></a>,
 <a href="datasets/Congressional+Voting+Records">Congressional Voting Records</a>,
 <a href="datasets/Water+Treatment+Plant"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Water+Treatment+Plant">Water Treatment Plant</a>,
 <a href="datasets/Waveform+Database+Generator+%28Version+1%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Waveform+Database+Generator+%28Version+1%29">Waveform Database Generator
(Version 1)</a>,
 <a href="datasets/Waveform+Database+Generator+%28Version+2%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Waveform+Database+Generator+%28Version+2%29">Waveform Database Generator
(Version 2)</a>,
 <a href="datasets/Wine"><img border="1" src="assets/MLimages/SmallLarge109.jpg"/></a>,
 <a href="datasets/Wine">Wine</a>,
 <a href="datasets/Yeast"><img border="1" src="assets/MLimages/SmallLarge110.jpg"/></a>,
 <a href="datasets/Yeast">Yeast</a>,
 <a href="datasets/Zoo"><img border="1" src="assets/MLimages/SmallLarge111.jpg"/></a>,
 <a href="datasets/Zoo">Zoo</a>,
 <a href="datasets/Undocumented"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>
,
 <a href="datasets/Undocumented">Undocumented</a>,
 <a href="datasets/Twenty+Newsgroups"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Twenty+Newsgroups">Twenty Newsgroups</a>,
 <a href="datasets/Australian+Sign+Language+signs"><img border="1"
src="assets/MLimages/SmallLarge114.jpg"/></a>,
 <a href="datasets/Australian+Sign+Language+signs">Australian Sign Language signs</a>,
 <a href="datasets/Australian+Sign+Language+signs+%28High+Quality%29"><img border="1"
src="assets/MLimages/SmallLarge114.jpg"/></a>,
 <a href="datasets/Australian+Sign+Language+signs+%28High+Quality%29">Australian Sign Language sig
ns (High Quality)</a>,
 <a href="datasets/US+Census+Data+%281990%29"><img border="1"
src="assets/MLimages/SmallLarge2.jpg"/></a>,
 <a href="datasets/US+Census+Data+%281990%29">US Census Data (1990)</a>,
 <a href="datasets/Census-Income+%28KDD%29"><img border="1"
src="assets/MLimages/SmallLarge2.jpg"/></a>,
 <a href="datasets/Census-Income+%28KDD%29">Census-Income (KDD)</a>,
 <a href="datasets/Coil+1999+Competition+Data"><img border="1"
src="assets/MLimages/SmallLarge118.jpg"/></a>,
 <a href="datasets/Coil+1999+Competition+Data">Coil 1999 Competition Data</a>,
 <a href="datasets/Corel+Image+Features"><img border="1" src="assets/MLimages/SmallLarge119.jpg"/>
</a>,
 <a href="datasets/Corel+Image+Features">Corel Image Features</a>,
 <a href="datasets/E.+Coli+Genes"><img border="1" src="assets/MLimages/SmallLarge120.jpg"/></a>,
 <a href="datasets/E.+Coli+Genes">E. Coli Genes</a>,
 <a href="datasets/EEG+Database"><img border="1" src="assets/MLimages/SmallLarge121.jpg"/></a>,
 <a href="datasets/EEG+Database">EEG Database</a>,
 <a href="datasets/El+Nino"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/El+Nino">El Nino</a>,
 <a href="datasets/Entree+Chicago+Recommendation+Data"><img border="1"
src="assets/MLimages/SmallLarge123.jpg"/></a>,
 <a href="datasets/Entree+Chicago+Recommendation+Data">Entree Chicago Recommendation Data</a>,
 <a href="datasets/CMU+Face+Images"><img border="1" src="assets/MLimages/SmallLarge124.jpg"/></a>,
 <a href="datasets/CMU+Face+Images">CMU Face Images</a>,
 <a href="datasets/Insurance+Company+Benchmark+%28COIL+2000%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Insurance+Company+Benchmark+%28COIL+2000%29">Insurance Company Benchmark (COIL
2000)</a>,
 <a href="datasets/Internet+Usage+Data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Internet+Usage+Data">Internet Usage Data</a>,
 <a href="datasets/IPUMS+Census+Database"><img border="1" src="assets/MLimages/SmallLarge2.jpg"/><
/a>,

```
 <a href="datasets/IPUMS+Census+Database">IPUMS Census Database</a>,
 <a href="datasets/Japanese+Vowels"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/><
/a>,
 <a href="datasets/Japanese+Vowels">Japanese Vowels</a>,
 <a href="datasets/KDD+Cup+1998+Data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/KDD+Cup+1998+Data">KDD Cup 1998 Data</a>,
 <a href="datasets/KDD+Cup+1999+Data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/KDD+Cup+1999+Data">KDD Cup 1999 Data</a>,
 <a href="datasets/M.+Tuberculosis+Genes"><img border="1"
src="assets/MLimages/SmallLarge131.jpg"/></a>,
 <a href="datasets/M.+Tuberculosis+Genes">M. Tuberculosis Genes</a>,
 <a href="datasets/Movie"><img border="1" src="assets/MLimages/SmallLarge132.jpg"/></a>,
 <a href="datasets/Movie">Movie</a>,
 <a href="datasets/MSNBC.com+Anonymous+Web+Data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/MSNBC.com+Anonymous+Web+Data">MSNBC.com Anonymous Web Data</a>,
 <a href="datasets/NSF+Research+Award+Abstracts+1990-2003"><img border="1"
src="assets/MLimages/SmallLarge134.jpg"/></a>,
 <a href="datasets/NSF+Research+Award+Abstracts+1990-2003">NSF Research Award Abstracts 1990-2003<
/a>,
 <a href="datasets/Pioneer-1+Mobile+Robot+Data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Pioneer-1+Mobile+Robot+Data">Pioneer-1 Mobile Robot Data</a>,
 <a href="datasets/Pseudo+Periodic+Synthetic+Time+Series"><img border="1"
src="assets/MLimages/SmallLarge136.jpg"/></a>,
 <a href="datasets/Pseudo+Periodic+Synthetic+Time+Series">Pseudo Periodic Synthetic Time
Series</a>,
 <a href="datasets/Reuters-21578+Text+Categorization+Collection"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Reuters-21578+Text+Categorization+Collection">Reuters-21578 Text Categorization
Collection</a>,
 <a href="datasets/Robot+Execution+Failures"><img border="1"
src="assets/MLimages/SmallLarge138.jpg"/></a>,
 <a href="datasets/Robot+Execution+Failures">Robot Execution Failures</a>,
 <a href="datasets/Synthetic+Control+Chart+Time+Series"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Synthetic+Control+Chart+Time+Series">Synthetic Control Chart Time Series</a>,
 <a href="datasets/Syskill+and+Webert+Web+Page+Ratings"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Syskill+and+Webert+Web+Page+Ratings">Syskill and Webert Web Page Ratings</a>,
 <a href="datasets/UNIX+User+Data"><img border="1" src="assets/MLimages/SmallLarge141.jpg"/></a>,
 <a href="datasets/UNIX+User+Data">UNIX User Data</a>,
 <a href="datasets/Volcanoes+on+Venus+-+JARtool+experiment"><img border="1"
src="assets/MLimages/SmallLarge142.jpg"/></a>,
 <a href="datasets/Volcanoes+on+Venus+-+JARtool+experiment">Volcanoes on Venus - JARtool
experiment</a>,
 <a href="datasets/Statlog+%28Australian+Credit+Approval%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Statlog+%28Australian+Credit+Approval%29">Statlog (Australian Credit Approval)<
/a>,
 <a href="datasets/Statlog+%28German+Credit+Data%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Statlog+%28German+Credit+Data%29">Statlog (German Credit Data)</a>,
 <a href="datasets/Statlog+%28Heart%29"><img border="1" src="assets/MLimages/SmallLarge45.jpg"/></
a>,
 <a href="datasets/Statlog+%28Heart%29">Statlog (Heart)</a>,
 <a href="datasets/Statlog+%28Landsat+Satellite%29"><img border="1"
src="assets/MLimages/SmallLarge146.jpg"/></a>,
 <a href="datasets/Statlog+%28Landsat+Satellite%29">Statlog (Landsat Satellite)</a>,
 <a href="datasets/Statlog+%28Image+Segmentation%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Statlog+%28Image+Segmentation%29">Statlog (Image Segmentation)</a>,
 <a href="datasets/Statlog+%28Shuttle%29"><img border="1" src="assets/MLimages/SmallLarge92.jpg"/>
</a>,
 <a href="datasets/Statlog+%28Shuttle%29">Statlog (Shuttle)</a>,
 <a href="datasets/Statlog+%28Vehicle+Silhouettes%29"><img border="1"
src="assets/MLimages/SmallLarge149.jpg"/></a>,
 <a href="datasets/Statlog+%28Vehicle+Silhouettes%29">Statlog (Vehicle Silhouettes)</a>,
 <a href="datasets/Connectionist+Bench+%28Nettalk+Corpus%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Connectionist+Bench+%28Nettalk+Corpus%29">Connectionist Bench (Nettalk Corpus)<
/a>,
 <a href="datasets/Connectionist+Bench+%28Sonar%2C+Mines+vs.+Rocks%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Connectionist+Bench+%28Sonar%2C+Mines+vs.+Rocks%29">Connectionist Bench (Sonar,
```

```
 <a href="datasets/Connectionist+Bench+%28Sonar,+Mines+vs.+Rocks%29">Connectionist Bench (Sonar,
Mines vs. Rocks)</a>,
 <a href="datasets/Connectionist+Bench+%28Vowel+Recognition+-+Deterding+Data%29"><img border="1" s
rc="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Connectionist+Bench+%28Vowel+Recognition+-+Deterding+Data%29">Connectionist Ben
ch (Vowel Recognition - Deterding Data)</a>,
 <a href="datasets/Economic+Sanctions"><img border="1" src="assets/MLimages/SmallLarge153.jpg"/></
a>,
 <a href="datasets/Economic+Sanctions">Economic Sanctions</a>,
 <a href="datasets/Protein+Data"><img border="1" src="assets/MLimages/SmallLarge154.jpg"/></a>,
 <a href="datasets/Protein+Data">Protein Data</a>,
 <a href="datasets/Cloud"><img border="1" src="assets/MLimages/SmallLarge155.jpg"/></a>,
 <a href="datasets/Cloud">Cloud</a>,
 <a href="datasets/CalIt2+Building+People+Counts"><img border="1"
src="assets/MLimages/SmallLarge156.jpg"/></a>,
 <a href="datasets/CalIt2+Building+People+Counts">CalIt2 Building People Counts</a>,
 <a href="datasets/Dodgers+Loop+Sensor"><img border="1" src="assets/MLimages/SmallLarge157.jpg"/><
/a>,
 <a href="datasets/Dodgers+Loop+Sensor">Dodgers Loop Sensor</a>,
 <a href="datasets/Poker+Hand"><img border="1" src="assets/MLimages/SmallLarge158.jpg"/></a>,
 <a href="datasets/Poker+Hand">Poker Hand</a>,
 <a href="datasets/MAGIC+Gamma+Telescope"><img border="1"
src="assets/MLimages/SmallLarge159.jpg"/></a>,
 <a href="datasets/MAGIC+Gamma+Telescope">MAGIC Gamma Telescope</a>,
 <a href="datasets/UJI+Pen+Characters"><img border="1" src="assets/MLimages/SmallLarge160.jpg"/></
a>,
 <a href="datasets/UJI+Pen+Characters">UJI Pen Characters</a>,
 <a href="datasets/Mammographic+Mass"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Mammographic+Mass">Mammographic Mass</a>,
 <a href="datasets/Forest+Fires"><img border="1" src="assets/MLimages/SmallLarge162.jpg"/></a>,
 <a href="datasets/Forest+Fires">Forest Fires</a>,
 <a href="datasets/Reuters+Transcribed+Subset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Reuters+Transcribed+Subset">Reuters Transcribed Subset</a>,
 <a href="datasets/Bag+of+Words"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>
,
 <a href="datasets/Bag+of+Words">Bag of Words</a>,
 <a href="datasets/Concrete+Compressive+Strength"><img border="1"
src="assets/MLimages/SmallLarge165.jpg"/></a>,
 <a href="datasets/Concrete+Compressive+Strength">Concrete Compressive Strength</a>,
 <a href="datasets/Hill-Valley"><img border="1" src="assets/MLimages/SmallLarge166.jpg"/></a>,
 <a href="datasets/Hill-Valley">Hill-Valley</a>,
 <a href="datasets/Arcene"><img border="1" src="assets/MLimages/SmallLarge167.jpg"/></a>,
 <a href="datasets/Arcene">Arcene</a>,
 <a href="datasets/Dexter"><img border="1" src="assets/MLimages/SmallLarge168.jpg"/></a>,
 <a href="datasets/Dexter">Dexter</a>,
 <a href="datasets/Dorothea"><img border="1" src="assets/MLimages/SmallLarge169.jpg"/></a>,
 <a href="datasets/Dorothea">Dorothea</a>,
 <a href="datasets/Gisette"><img border="1" src="assets/MLimages/SmallLarge170.jpg"/></a>,
 <a href="datasets/Gisette">Gisette</a>,
 <a href="datasets/Madelon"><img border="1" src="assets/MLimages/SmallLarge171.jpg"/></a>,
 <a href="datasets/Madelon">Madelon</a>,
 <a href="datasets/Ozone+Level+Detection"><img border="1"
src="assets/MLimages/SmallLarge172.jpg"/></a>,
 <a href="datasets/Ozone+Level+Detection">Ozone Level Detection</a>,
 <a href="datasets/Abscisic+Acid+Signaling+Network"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Abscisic+Acid+Signaling+Network">Abscisic Acid Signaling Network</a>,
 <a href="datasets/Parkinsons"><img border="1" src="assets/MLimages/SmallLarge174.jpg"/></a>,
 <a href="datasets/Parkinsons">Parkinsons</a>,
 <a href="datasets/Character+Trajectories"><img border="1"
src="assets/MLimages/SmallLarge175.jpg"/></a>,
 <a href="datasets/Character+Trajectories">Character Trajectories</a>,
 <a href="datasets/Blood+Transfusion+Service+Center"><img border="1"
src="assets/MLimages/SmallLarge176.jpg"/></a>,
 <a href="datasets/Blood+Transfusion+Service+Center">Blood Transfusion Service Center</a>,
 <a href="datasets/UJI+Pen+Characters+%28Version+2%29"><img border="1"
src="assets/MLimages/SmallLarge160.jpg"/></a>,
 <a href="datasets/UJI+Pen+Characters+%28Version+2%29">UJI Pen Characters (Version 2)</a>,
 <a href="datasets/Semeion+Handwritten+Digit"><img border="1"
src="assets/MLimages/SmallLarge178.jpg"/></a>,
 <a href="datasets/Semeion+Handwritten+Digit">Semeion Handwritten Digit</a>,
 <a href="datasets/SECOM"><img border="1" src="assets/MLimages/SmallLarge179.jpg"/></a>,
 <a href="datasets/SECOM">SECOM</a>,
 <a href="datasets/Plants"><img border="1" src="assets/MLimages/SmallLarge180.jpg"/></a>,
 <a href="datasets/Plants">Plants</a>,
 <a href="datasets/Libras+Movement"><img border="1" src="assets/MLimages/SmallLarge181.jpg"/></a>,
```

```
 <a href="datasets/Libras+Movement"><img border="1" src="assets/MLimages/SmallLarge161.jpg"/></a>,
 <a href="datasets/Libras+Movement">Libras Movement</a>,
 <a href="datasets/Concrete+Slump+Test"><img border="1" src="assets/MLimages/SmallLarge165.jpg"/><
/a>,
 <a href="datasets/Concrete+Slump+Test">Concrete Slump Test</a>,
 <a href="datasets/Communities+and+Crime"><img border="1"
src="assets/MLimages/SmallLarge183.jpg"/></a>,
 <a href="datasets/Communities+and+Crime">Communities and Crime</a>,
 <a href="datasets/Acute+Inflammations"><img border="1" src="assets/MLimages/SmallLarge184.jpg"/><
/a>,
 <a href="datasets/Acute+Inflammations">Acute Inflammations</a>,
 <a href="datasets/Wine+Quality"><img border="1" src="assets/MLimages/SmallLarge186.jpg"/></a>,
 <a href="datasets/Wine+Quality">Wine Quality</a>,
 <a href="datasets/URL+Reputation"><img border="1" src="assets/MLimages/SmallLarge187.jpg"/></a>,
 <a href="datasets/URL+Reputation">URL Reputation</a>,
 <a href="datasets/p53+Mutants"><img border="1" src="assets/MLimages/SmallLarge188.jpg"/></a>,
 <a href="datasets/p53+Mutants">p53 Mutants</a>,
 <a href="datasets/Parkinsons+Telemonitoring"><img border="1"
src="assets/MLimages/SmallLarge174.jpg"/></a>,
 <a href="datasets/Parkinsons+Telemonitoring">Parkinsons Telemonitoring</a>,
 <a href="datasets/Demospongiae"><img border="1" src="assets/MLimages/SmallLarge190.jpg"/></a>,
 <a href="datasets/Demospongiae">Demospongiae</a>,
 <a href="datasets/Opinosis+Opinion+%26frasl%3B+Review"><img border="1"
src="assets/MLimages/SmallLarge191.jpg"/></a>,
 <a href="datasets/Opinosis+Opinion+%26frasl%3B+Review">Opinosis Opinion ⁄ Review</a>,
 <a href="datasets/Breast+Tissue"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a
>,
 <a href="datasets/Breast+Tissue">Breast Tissue</a>,
 <a href="datasets/Cardiotocography"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>,
 <a href="datasets/Cardiotocography">Cardiotocography</a>,
 <a href="datasets/Wall-Following+Robot+Navigation+Data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Wall-Following+Robot+Navigation+Data">Wall-Following Robot Navigation Data</a>,
 <a href="datasets/Spoken+Arabic+Digit"><img border="1" src="assets/MLimages/SmallLarge195.jpg"/><
/a>,
 <a href="datasets/Spoken+Arabic+Digit">Spoken Arabic Digit</a>,
 <a href="datasets/Localization+Data+for+Person+Activity"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Localization+Data+for+Person+Activity">Localization Data for Person
Activity</a>,
 <a href="datasets/AutoUniv"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/AutoUniv">AutoUniv</a>,
 <a href="datasets/Steel+Plates+Faults"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Steel+Plates+Faults">Steel Plates Faults</a>,
 <a href="datasets/MiniBooNE+particle+identification"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/MiniBooNE+particle+identification">MiniBooNE particle identification</a>,
 <a href="datasets/YearPredictionMSD"><img border="1" src="assets/MLimages/SmallLarge203.jpg"/></a
>,
 <a href="datasets/YearPredictionMSD">YearPredictionMSD</a>,
 <a href="datasets/PEMS-SF"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/PEMS-SF">PEMS-SF</a>,
 <a href="datasets/OpinRank+Review+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/OpinRank+Review+Dataset">OpinRank Review Dataset</a>,
 <a href="datasets/Relative+location+of+CT+slices+on+axial+axis"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Relative+location+of+CT+slices+on+axial+axis">Relative location of CT slices on
axial axis</a>,
 <a href="datasets/Online+Handwritten+Assamese+Characters+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Online+Handwritten+Assamese+Characters+Dataset">Online Handwritten Assamese Cha
racters Dataset</a>,
 <a href="datasets/PubChem+Bioassay+Data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/PubChem+Bioassay+Data">PubChem Bioassay Data</a>,
 <a href="datasets/Record+Linkage+Comparison+Patterns"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Record+Linkage+Comparison+Patterns">Record Linkage Comparison Patterns</a>,
 <a href="datasets/Communities+and+Crime+Unnormalized"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Communities+and+Crime+Unnormalized">Communities and Crime Unnormalized</a>,
 <a href="datasets/Vertebral+Column"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>,
 <a href="datasets/Vertebral+Column">Vertebral Column</a>,
 <a href="datasets/EMG+Physical+Action+Data+Set"><img border="1"
```

```
 <a href="datasets/EMG+Physical+Action+Data+Set"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/EMG+Physical+Action+Data+Set">EMG Physical Action Data Set</a>,
 <a href="datasets/Vicon+Physical+Action+Data+Set"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Vicon+Physical+Action+Data+Set">Vicon Physical Action Data Set</a>,
 <a href="datasets/Amazon+Commerce+reviews+set"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Amazon+Commerce+reviews+set">Amazon Commerce reviews set</a>,
 <a href="datasets/Amazon+Access+Samples"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Amazon+Access+Samples">Amazon Access Samples</a>,
 <a href="datasets/Reuter_50_50"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>
,
 <a href="datasets/Reuter_50_50">Reuter_50_50</a>,
 <a href="datasets/Farm+Ads"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Farm+Ads">Farm Ads</a>,
 <a href="datasets/DBWorld+e-mails"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/><
/a>,
 <a href="datasets/DBWorld+e-mails">DBWorld e-mails</a>,
 <a href="datasets/KEGG+Metabolic+Relation+Network+%28Directed%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/KEGG+Metabolic+Relation+Network+%28Directed%29">KEGG Metabolic Relation Network
(Directed)</a>,
 <a href="datasets/KEGG+Metabolic+Reaction+Network+%28Undirected%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/KEGG+Metabolic+Reaction+Network+%28Undirected%29">KEGG Metabolic Reaction
Network (Undirected)</a>,
 <a href="datasets/Bank+Marketing"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></
a>,
 <a href="datasets/Bank+Marketing">Bank Marketing</a>,
 <a href="datasets/YouTube+Comedy+Slam+Preference+Data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/YouTube+Comedy+Slam+Preference+Data">YouTube Comedy Slam Preference Data</a>,
 <a href="datasets/Gas+Sensor+Array+Drift+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Gas+Sensor+Array+Drift+Dataset">Gas Sensor Array Drift Dataset</a>,
 <a href="datasets/ILPD+%28Indian+Liver+Patient+Dataset%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/ILPD+%28Indian+Liver+Patient+Dataset%29">ILPD (Indian Liver Patient Dataset)</a
>,
 <a href="datasets/OPPORTUNITY+Activity+Recognition"><img border="1"
src="assets/MLimages/SmallLarge226.jpg"/></a>,
 <a href="datasets/OPPORTUNITY+Activity+Recognition">OPPORTUNITY Activity Recognition</a>,
 <a href="datasets/Nomao"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Nomao">Nomao</a>,
 <a href="datasets/SMS+Spam+Collection"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/SMS+Spam+Collection">SMS Spam Collection</a>,
 <a href="datasets/Skin+Segmentation"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Skin+Segmentation">Skin Segmentation</a>,
 <a href="datasets/Planning+Relax"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></
a>,
 <a href="datasets/Planning+Relax">Planning Relax</a>,
 <a href="datasets/PAMAP2+Physical+Activity+Monitoring"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/PAMAP2+Physical+Activity+Monitoring">PAMAP2 Physical Activity Monitoring</a>,
 <a href="datasets/Restaurant+%26+consumer+data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Restaurant+%26+consumer+data">Restaurant &amp; consumer data</a>,
 <a href="datasets/CNAE-9"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/CNAE-9">CNAE-9</a>,
 <a href="datasets/Individual+household+electric+power+consumption"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Individual+household+electric+power+consumption">Individual household electric
power consumption</a>,
 <a href="datasets/seeds"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/seeds">seeds</a>,
 <a href="datasets/Northix"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Northix">Northix</a>,
 <a href="datasets/QtyT40I10D100K"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></
a>,
 <a href="datasets/QtyT40I10D100K">QtyT40I10D100K</a>,
 <a href="datasets/Legal+Case+Reports"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Legal+Case+Reports">Legal Case Reports</a>,
 <a href="datasets/Human+Activity+Recognition+Using+Smartphones"><img border="1"
```

```
<a href="datasets/Human+Activity+Recognition+Using+Smartphones"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Human+Activity+Recognition+Using+Smartphones">Human Activity Recognition Using
Smartphones</a>,
 <a href="datasets/One-hundred+plant+species+leaves+data+set"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/One-hundred+plant+species+leaves+data+set">One-hundred plant species leaves dat
a set</a>,
 <a href="datasets/Energy+efficiency"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Energy+efficiency">Energy efficiency</a>,
 <a href="datasets/Yacht+Hydrodynamics"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Yacht+Hydrodynamics">Yacht Hydrodynamics</a>,
 <a href="datasets/Fertility"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Fertility">Fertility</a>,
 <a href="datasets/Daphnet+Freezing+of+Gait"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Daphnet+Freezing+of+Gait">Daphnet Freezing of Gait</a>,
 <a href="datasets/3D+Road+Network+%28North+Jutland%2C+Denmark%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/3D+Road+Network+%28North+Jutland%2C+Denmark%29">3D Road Network (North Jutland,
Denmark)</a>,
 <a href="datasets/ISTANBUL+STOCK+EXCHANGE"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/ISTANBUL+STOCK+EXCHANGE">ISTANBUL STOCK EXCHANGE</a>,
 <a href="datasets/Buzz+in+social+media+"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Buzz+in+social+media+">Buzz in social media </a>,
 <a href="datasets/First-order+theorem+proving"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/First-order+theorem+proving">First-order theorem proving</a>,
 <a href="datasets/Wearable+Computing%3A+Classification+of+Body+Postures+and+Movements+%28PUC-Rio%
29"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Wearable+Computing%3A+Classification+of+Body+Postures+and+Movements+%28PUC-Rio%
29">Wearable Computing: Classification of Body Postures and Movements (PUC-Rio)</a>,
 <a href="datasets/Gas+sensor+arrays+in+open+sampling+settings"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Gas+sensor+arrays+in+open+sampling+settings">Gas sensor arrays in open sampling
settings</a>,
 <a href="datasets/Climate+Model+Simulation+Crashes"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Climate+Model+Simulation+Crashes">Climate Model Simulation Crashes</a>,
 <a href="datasets/MicroMass"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/MicroMass">MicroMass</a>,
 <a href="datasets/QSAR+biodegradation"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/QSAR+biodegradation">QSAR biodegradation</a>,
 <a href="datasets/BLOGGER"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/BLOGGER">BLOGGER</a>,
 <a href="datasets/Daily+and+Sports+Activities"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Daily+and+Sports+Activities">Daily and Sports Activities</a>,
 <a href="datasets/User+Knowledge+Modeling"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/User+Knowledge+Modeling">User Knowledge Modeling</a>,
 <a
href="datasets/Reuters+RCV1+RCV2+Multilingual%2C+Multiview+Text+Categorization+Test+collection"><i
mg border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a
href="datasets/Reuters+RCV1+RCV2+Multilingual%2C+Multiview+Text+Categorization+Test+collection">Reu
RCV1 RCV2 Multilingual, Multiview Text Categorization Test collection</a>,
 <a href="datasets/NYSK"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/NYSK">NYSK</a>,
 <a href="datasets/Turkiye+Student+Evaluation"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Turkiye+Student+Evaluation">Turkiye Student Evaluation</a>,
 <a
href="datasets/ser+Knowledge+Modeling+Data+%28Students%27+Knowledge+Levels+on+DC+Electrical+Machine
"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a
href="datasets/ser+Knowledge+Modeling+Data+%28Students%27+Knowledge+Levels+on+DC+Electrical+Machine
">ser Knowledge Modeling Data (Students' Knowledge Levels on DC Electrical Machines)</a>,
 <a href="datasets/EEG+Eye+State"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a
>,
 <a href="datasets/EEG+Eye+State">EEG Eye State</a>,
 <a href="datasets/Physicochemical+Properties+of+Protein+Tertiary+Structure"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>
```

src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Physicochemical+Properties+of+Protein+Tertiary+Structure">Physicochemical
Properties of Protein Tertiary Structure</a>,
 <a href="datasets/seismic-bumps"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a
>,
 <a href="datasets/seismic-bumps">seismic-bumps</a>,
 <a href="datasets/banknote+authentication"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/banknote+authentication">banknote authentication</a>,
 <a href="datasets/USPTO+Algorithm+Challenge%2C+run+by+NASA-
Harvard+Tournament+Lab+and+TopCoder++++Problem%3A+Pat"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/USPTO+Algorithm+Challenge%2C+run+by+NASA-
Harvard+Tournament+Lab+and+TopCoder++++Problem%3A+Pat">USPTO Algorithm Challenge, run by NASA-
Harvard Tournament Lab and TopCoder   Problem: Pat</a>,
 <a href="datasets/YouTube+Multiview+Video+Games+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/YouTube+Multiview+Video+Games+Dataset">YouTube Multiview Video Games
Dataset</a>,
 <a href="datasets/Gas+Sensor+Array+Drift+Dataset+at+Different+Concentrations"><img border="1" src
="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Gas+Sensor+Array+Drift+Dataset+at+Different+Concentrations">Gas Sensor Array Dr
ift Dataset at Different Concentrations</a>,
 <a href="datasets/Activities+of+Daily+Living+%28ADLs%29+Recognition+Using+Binary+Sensors"><img bo
rder="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a
href="datasets/Activities+of+Daily+Living+%28ADLs%29+Recognition+Using+Binary+Sensors">Activities
of Daily Living (ADLs) Recognition Using Binary Sensors</a>,
 <a href="datasets/SkillCraft1+Master+Table+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/SkillCraft1+Master+Table+Dataset">SkillCraft1 Master Table Dataset</a>,
 <a href="datasets/Weight+Lifting+Exercises+monitored+with+Inertial+Measurement+Units"><img
border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Weight+Lifting+Exercises+monitored+with+Inertial+Measurement+Units">Weight
Lifting Exercises monitored with Inertial Measurement Units</a>,
 <a href="datasets/SML2010"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/SML2010">SML2010</a>,
 <a href="datasets/Bike+Sharing+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Bike+Sharing+Dataset">Bike Sharing Dataset</a>,
 <a href="datasets/Predict+keywords+activities+in+a+online+social+media"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Predict+keywords+activities+in+a+online+social+media">Predict keywords
activities in a online social media</a>,
 <a href="datasets/Thoracic+Surgery+Data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Thoracic+Surgery+Data">Thoracic Surgery Data</a>,
 <a href="datasets/EMG+dataset+in+Lower+Limb"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/EMG+dataset+in+Lower+Limb">EMG dataset in Lower Limb</a>,
 <a href="datasets/SUSY"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/SUSY">SUSY</a>,
 <a href="datasets/HIGGS"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/HIGGS">HIGGS</a>,
 <a href="datasets/Qualitative_Bankruptcy"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Qualitative_Bankruptcy">Qualitative_Bankruptcy</a>,
 <a href="datasets/LSVT+Voice+Rehabilitation"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/LSVT+Voice+Rehabilitation">LSVT Voice Rehabilitation</a>,
 <a href="datasets/Dataset+for+ADL+Recognition+with+Wrist-worn+Accelerometer"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Dataset+for+ADL+Recognition+with+Wrist-worn+Accelerometer">Dataset for ADL Reco
gnition with Wrist-worn Accelerometer</a>,
 <a href="datasets/Wilt"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Wilt">Wilt</a>,
 <a href="datasets/User+Identification+From+Walking+Activity"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/User+Identification+From+Walking+Activity">User Identification From Walking Act
ivity</a>,
 <a href="datasets/Activity+Recognition+from+Single+Chest-Mounted+Accelerometer"><img border="1" s
rc="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Activity+Recognition+from+Single+Chest-Mounted+Accelerometer">Activity
Recognition from Single Chest-Mounted Accelerometer</a>,
 <a href="datasets/Leaf"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Leaf">Leaf</a>,
 <a href="datasets/Dresses_Attribute_Sales"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,

src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Dresses_Attribute_Sales">Dresses_Attribute_Sales</a>,
 <a href="datasets/Tamilnadu+Electricity+Board+Hourly+Readings"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Tamilnadu+Electricity+Board+Hourly+Readings">Tamilnadu Electricity Board Hourly
Readings</a>,
 <a href="datasets/Airfoil+Self-Noise"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Airfoil+Self-Noise">Airfoil Self-Noise</a>,
 <a href="datasets/Wholesale+customers"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Wholesale+customers">Wholesale customers</a>,
 <a href="datasets/Twitter+Data+set+for+Arabic+Sentiment+Analysis"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Twitter+Data+set+for+Arabic+Sentiment+Analysis">Twitter Data set for Arabic Sen
timent Analysis</a>,
 <a href="datasets/Combined+Cycle+Power+Plant"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Combined+Cycle+Power+Plant">Combined Cycle Power Plant</a>,
 <a href="datasets/Urban+Land+Cover"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>,
 <a href="datasets/Urban+Land+Cover">Urban Land Cover</a>,
 <a href="datasets/Diabetes+130-US+hospitals+for+years+1999-2008"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Diabetes+130-US+hospitals+for+years+1999-2008">Diabetes 130-US hospitals for ye
ars 1999-2008</a>,
 <a href="datasets/Bach+Choral+Harmony"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Bach+Choral+Harmony">Bach Choral Harmony</a>,
 <a href="datasets/StoneFlakes"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/StoneFlakes">StoneFlakes</a>,
 <a href="datasets/Tennis+Major+Tournament+Match+Statistics"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Tennis+Major+Tournament+Match+Statistics">Tennis Major Tournament Match
Statistics</a>,
 <a href="datasets/Parkinson+Speech+Dataset+with++Multiple+Types+of+Sound+Recordings"><img
border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Parkinson+Speech+Dataset+with++Multiple+Types+of+Sound+Recordings">Parkinson Sp
eech Dataset with Multiple Types of Sound Recordings</a>,
 <a href="datasets/Gesture+Phase+Segmentation"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Gesture+Phase+Segmentation">Gesture Phase Segmentation</a>,
 <a href="datasets/Perfume+Data"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>
,
 <a href="datasets/Perfume+Data">Perfume Data</a>,
 <a href="datasets/BlogFeedback"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>
,
 <a href="datasets/BlogFeedback">BlogFeedback</a>,
 <a href="datasets/REALDISP+Activity+Recognition+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/REALDISP+Activity+Recognition+Dataset">REALDISP Activity Recognition
Dataset</a>,
 <a href="datasets/Newspaper+and+magazine+images+segmentation+dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Newspaper+and+magazine+images+segmentation+dataset">Newspaper and magazine imag
es segmentation dataset</a>,
 <a href="datasets/AAAI+2014+Accepted+Papers"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/AAAI+2014+Accepted+Papers">AAAI 2014 Accepted Papers</a>,
 <a href="datasets/Gas+sensor+array+under+flow+modulation"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Gas+sensor+array+under+flow+modulation">Gas sensor array under flow
modulation</a>,
 <a href="datasets/Gas+sensor+array+exposed+to+turbulent+gas+mixtures"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Gas+sensor+array+exposed+to+turbulent+gas+mixtures">Gas sensor array exposed to
turbulent gas mixtures</a>,
 <a href="datasets/UJIIndoorLoc"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>
,
 <a href="datasets/UJIIndoorLoc">UJIIndoorLoc</a>,
 <a href="datasets/Sentence+Classification"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Sentence+Classification">Sentence Classification</a>,
 <a href="datasets/Dow+Jones+Index"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/><
/a>,
 <a href="datasets/Dow+Jones+Index">Dow Jones Index</a>,
 <a href="datasets/sEMG+for+Basic+Hand+movements"><img border="1"

src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/sEMG+for+Basic+Hand+movements">sEMG for Basic Hand movements</a>,
 <a href="datasets/AAAI+2013+Accepted+Papers"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/AAAI+2013+Accepted+Papers">AAAI 2013 Accepted Papers</a>,
 <a href="datasets/Geographical+Original+of+Music"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Geographical+Original+of+Music">Geographical Original of Music</a>,
 <a href="datasets/Condition+Based+Maintenance+of+Naval+Propulsion+Plants"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Condition+Based+Maintenance+of+Naval+Propulsion+Plants">Condition Based
Maintenance of Naval Propulsion Plants</a>,
 <a href="datasets/Grammatical+Facial+Expressions"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Grammatical+Facial+Expressions">Grammatical Facial Expressions</a>,
 <a href="datasets/NoisyOffice"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/NoisyOffice">NoisyOffice</a>,
 <a href="datasets/MHEALTH+Dataset"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/><
/a>,
 <a href="datasets/MHEALTH+Dataset">MHEALTH Dataset</a>,
 <a href="datasets/Student+Performance"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Student+Performance">Student Performance</a>,
 <a href="datasets/ElectricityLoadDiagrams20112014"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/ElectricityLoadDiagrams20112014">ElectricityLoadDiagrams20112014</a>,
 <a href="datasets/Gas+sensor+array+under+dynamic+gas+mixtures"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Gas+sensor+array+under+dynamic+gas+mixtures">Gas sensor array under dynamic gas
mixtures</a>,
 <a href="datasets/microblogPCU"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>
,
 <a href="datasets/microblogPCU">microblogPCU</a>,
 <a href="datasets/Firm-Teacher_Clave-Direction_Classification"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Firm-Teacher_Clave-Direction_Classification">Firm-Teacher_Clave-
Direction_Classification</a>,
 <a href="datasets/Dataset+for+Sensorless+Drive+Diagnosis"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Dataset+for+Sensorless+Drive+Diagnosis">Dataset for Sensorless Drive
Diagnosis</a>,
 <a href="datasets/TV+News+Channel+Commercial+Detection+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/TV+News+Channel+Commercial+Detection+Dataset">TV News Channel Commercial
Detection Dataset</a>,
 <a href="datasets/Phishing+Websites"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Phishing+Websites">Phishing Websites</a>,
 <a href="datasets/Greenhouse+Gas+Observing+Network"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Greenhouse+Gas+Observing+Network">Greenhouse Gas Observing Network</a>,
 <a href="datasets/Diabetic+Retinopathy+Debrecen+Data+Set"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Diabetic+Retinopathy+Debrecen+Data+Set">Diabetic Retinopathy Debrecen Data Set<
/a>,
 <a href="datasets/HIV-1+protease+cleavage"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/HIV-1+protease+cleavage">HIV-1 protease cleavage</a>,
 <a href="datasets/Sentiment+Labelled+Sentences"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Sentiment+Labelled+Sentences">Sentiment Labelled Sentences</a>,
 <a href="datasets/Online+News+Popularity"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Online+News+Popularity">Online News Popularity</a>,
 <a href="datasets/Forest+type+mapping"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Forest+type+mapping">Forest type mapping</a>,
 <a href="datasets/wiki4HE"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/wiki4HE">wiki4HE</a>,
 <a href="datasets/Online+Video+Characteristics+and+Transcoding+Time+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Online+Video+Characteristics+and+Transcoding+Time+Dataset">Online Video
Characteristics and Transcoding Time Dataset</a>,
 <a href="datasets/Chronic_Kidney_Disease"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Chronic_Kidney_Disease">Chronic_Kidney_Disease</a>,
 <a href="datasets/Machine+Learning+based+ZZAlpha+Ltd.+Stock+Recommendations+2012-2014"><img

```
border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Machine+Learning+based+ZZAlpha+Ltd.+Stock+Recommendations+2012-2014">Machine Le
arning based ZZAlpha Ltd. Stock Recommendations 2012-2014</a>,
 <a href="datasets/Folio"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Folio">Folio</a>,
 <a href="datasets/Taxi+Service+Trajectory+-+Prediction+Challenge%2C+ECML+PKDD+2015"><img
border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Taxi+Service+Trajectory+-+Prediction+Challenge%2C+ECML+PKDD+2015">Taxi Service
Trajectory - Prediction Challenge, ECML PKDD 2015</a>,
 <a href="datasets/Cuff-Less+Blood+Pressure+Estimation"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Cuff-Less+Blood+Pressure+Estimation">Cuff-Less Blood Pressure Estimation</a>,
 <a href="datasets/Smartphone-Based+Recognition+of+Human+Activities+and+Postural+Transitions"><img
border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Smartphone-
Based+Recognition+of+Human+Activities+and+Postural+Transitions">Smartphone-Based Recognition of Hu
man Activities and Postural Transitions</a>,
 <a href="datasets/Mice+Protein+Expression"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Mice+Protein+Expression">Mice Protein Expression</a>,
 <a href="datasets/UJIIndoorLoc-Mag"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>,
 <a href="datasets/UJIIndoorLoc-Mag">UJIIndoorLoc-Mag</a>,
 <a href="datasets/Heterogeneity+Activity+Recognition"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Heterogeneity+Activity+Recognition">Heterogeneity Activity Recognition</a>,
 <a href="datasets/Educational+Process+Mining+%28EPM%29%3A+A+Learning+Analytics+Data+Set"><img bor
der="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a
href="datasets/Educational+Process+Mining+%28EPM%29%3A+A+Learning+Analytics+Data+Set">Educational
Process Mining (EPM): A Learning Analytics Data Set</a>,
 <a href="datasets/HEPMASS"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/HEPMASS">HEPMASS</a>,
 <a href="datasets/Indoor+User+Movement+Prediction+from+RSS+data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Indoor+User+Movement+Prediction+from+RSS+data">Indoor User Movement Prediction
from RSS data</a>,
 <a href="datasets/Open+University+Learning+Analytics+dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Open+University+Learning+Analytics+dataset">Open University Learning Analytics
dataset</a>,
 <a href="datasets/default+of+credit+card+clients"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/default+of+credit+card+clients">default of credit card clients</a>,
 <a href="datasets/Mesothelioma%E2%80%99s+disease+data+set+"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Mesothelioma%E2%80%99s+disease+data+set+">Mesothelioma’s disease data set </a>,
 <a href="datasets/Online+Retail"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a
>,
 <a href="datasets/Online+Retail">Online Retail</a>,
 <a href="datasets/SIFT10M"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/SIFT10M">SIFT10M</a>,
 <a href="datasets/GPS+Trajectories"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>,
 <a href="datasets/GPS+Trajectories">GPS Trajectories</a>,
 <a href="datasets/Detect+Malacious+Executable%28AntiVirus%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Detect+Malacious+Executable%28AntiVirus%29">Detect Malacious
Executable(AntiVirus)</a>,
 <a href="datasets/Occupancy+Detection+"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Occupancy+Detection+">Occupancy Detection </a>,
 <a
href="datasets/Improved+Spiral+Test+Using+Digitized+Graphics+Tablet+for+Monitoring+Parkinson%E2%80%
Disease"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a
href="datasets/Improved+Spiral+Test+Using+Digitized+Graphics+Tablet+for+Monitoring+Parkinson%E2%80%
Disease">Improved Spiral Test Using Digitized Graphics Tablet for Monitoring Parkinson’s
Disease</a>,
 <a href="datasets/News+Aggregator"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/><
/a>,
 <a href="datasets/News+Aggregator">News Aggregator</a>,
 <a href="datasets/Air+Quality"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Air+Quality">Air Quality</a>,
 <a href="datasets/Twin+gas+sensor+arrays"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Twin+gas+sensor+arrays">Twin gas sensor arrays</a>,
```

```
 <a href="datasets/Gas+sensors+for+home+activity+monitoring"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Gas+sensors+for+home+activity+monitoring">Gas sensors for home activity
monitoring</a>,
 <a href="datasets/Facebook+Comment+Volume+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Facebook+Comment+Volume+Dataset">Facebook Comment Volume Dataset</a>,
 <a
href="datasets/Smartphone+Dataset+for+Human+Activity+Recognition+%28HAR%29+in+Ambient+Assisted+Livi
28AAL%29"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a
href="datasets/Smartphone+Dataset+for+Human+Activity+Recognition+%28HAR%29+in+Ambient+Assisted+Livi
28AAL%29">Smartphone Dataset for Human Activity Recognition (HAR) in Ambient Assisted Living (AAL)
</a>,
 <a href="datasets/Polish+companies+bankruptcy+data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Polish+companies+bankruptcy+data">Polish companies bankruptcy data</a>,
 <a href="datasets/Activity+Recognition+system+based+on+Multisensor+data+fusion+%28AReM%29"><img b
order="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a
href="datasets/Activity+Recognition+system+based+on+Multisensor+data+fusion+%28AReM%29">Activity R
ecognition system based on Multisensor data fusion (AReM)</a>,
 <a href="datasets/Dota2+Games+Results"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Dota2+Games+Results">Dota2 Games Results</a>,
 <a href="datasets/Facebook+metrics"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>,
 <a href="datasets/Facebook+metrics">Facebook metrics</a>,
 <a href="datasets/UbiqLog+%28smartphone+lifelogging%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/UbiqLog+%28smartphone+lifelogging%29">UbiqLog (smartphone lifelogging)</a>,
 <a href="datasets/NIPS+Conference+Papers+1987-2015"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/NIPS+Conference+Papers+1987-2015">NIPS Conference Papers 1987-2015</a>,
 <a href="datasets/HTRU2"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/HTRU2">HTRU2</a>,
 <a href="datasets/Drug+consumption+%28quantified%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Drug+consumption+%28quantified%29">Drug consumption (quantified)</a>,
 <a href="datasets/Appliances+energy+prediction"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Appliances+energy+prediction">Appliances energy prediction</a>,
 <a href="datasets/Miskolc+IIS+Hybrid+IPS"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Miskolc+IIS+Hybrid+IPS">Miskolc IIS Hybrid IPS</a>,
 <a href="datasets/KDC-4007+dataset+Collection"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/KDC-4007+dataset+Collection">KDC-4007 dataset Collection</a>,
 <a href="datasets/Geo-
Magnetic+field+and+WLAN+dataset+for+indoor+localisation+from+wristband+and+smartphone"><img
border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Geo-
Magnetic+field+and+WLAN+dataset+for+indoor+localisation+from+wristband+and+smartphone">Geo-
Magnetic field and WLAN dataset for indoor localisation from wristband and smartphone</a>,
 <a href="datasets/DrivFace"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/DrivFace">DrivFace</a>,
 <a href="datasets/Website+Phishing"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>,
 <a href="datasets/Website+Phishing">Website Phishing</a>,
 <a href="datasets/YouTube+Spam+Collection"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/YouTube+Spam+Collection">YouTube Spam Collection</a>,
 <a href="datasets/Beijing+PM2.5+Data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Beijing+PM2.5+Data">Beijing PM2.5 Data</a>,
 <a href="datasets/Cargo+2000+Freight+Tracking+and+Tracing"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Cargo+2000+Freight+Tracking+and+Tracing">Cargo 2000 Freight Tracking and
Tracing</a>,
 <a href="datasets/Cervical+cancer+%28Risk+Factors%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Cervical+cancer+%28Risk+Factors%29">Cervical cancer (Risk Factors)</a>,
 <a href="datasets/Quality+Assessment+of+Digital+Colposcopies"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Quality+Assessment+of+Digital+Colposcopies">Quality Assessment of Digital
Colposcopies</a>,
 <a href="datasets/KASANDR"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
```

```
 <a href="datasets/KASANDR">KASANDR</a>,
 <a href="datasets/FMA%3A+A+Dataset+For+Music+Analysis"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/FMA%3A+A+Dataset+For+Music+Analysis">FMA: A Dataset For Music Analysis</a>,
 <a href="datasets/Air+quality"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Air+quality">Air quality</a>,
 <a href="datasets/Epileptic+Seizure+Recognition"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Epileptic+Seizure+Recognition">Epileptic Seizure Recognition</a>,
 <a href="datasets/Devanagari+Handwritten+Character+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Devanagari+Handwritten+Character+Dataset">Devanagari Handwritten Character Data
set</a>,
 <a href="datasets/Stock+portfolio+performance"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Stock+portfolio+performance">Stock portfolio performance</a>,
 <a href="datasets/MoCap+Hand+Postures"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/MoCap+Hand+Postures">MoCap Hand Postures</a>,
 <a href="datasets/Early+biomarkers+of+Parkinson%92s+disease+based+on+natural+connected+speech"><i
mg border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a
href="datasets/Early+biomarkers+of+Parkinson%92s+disease+based+on+natural+connected+speech">Early
biomarkers of Parkinson’s disease based on natural connected speech</a>,
 <a href="datasets/Data+for+Software+Engineering+Teamwork+Assessment+in+Education+Setting"><img bo
rder="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Data+for+Software+Engineering+Teamwork+Assessment+in+Education+Setting">Data fo
r Software Engineering Teamwork Assessment in Education Setting</a>,
 <a href="datasets/PM2.5+Data+of+Five+Chinese+Cities"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/PM2.5+Data+of+Five+Chinese+Cities">PM2.5 Data of Five Chinese Cities</a>,
 <a href="datasets/Parkinson+Disease+Spiral+Drawings+Using+Digitized+Graphics+Tablet"><img
border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Parkinson+Disease+Spiral+Drawings+Using+Digitized+Graphics+Tablet">Parkinson Di
sease Spiral Drawings Using Digitized Graphics Tablet</a>,
 <a href="datasets/Sales_Transactions_Dataset_Weekly"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Sales_Transactions_Dataset_Weekly">Sales_Transactions_Dataset_Weekly</a>,
 <a href="datasets/Las+Vegas+Strip"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/><
/a>,
 <a href="datasets/Las+Vegas+Strip">Las Vegas Strip</a>,
 <a href="datasets/Eco-hotel"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Eco-hotel">Eco-hotel</a>,
 <a href="datasets/MEU-Mobile+KSD"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></
a>,
 <a href="datasets/MEU-Mobile+KSD">MEU-Mobile KSD</a>,
 <a href="datasets/Crowdsourced+Mapping"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Crowdsourced+Mapping">Crowdsourced Mapping</a>,
 <a href="datasets/gene+expression+cancer+RNA-Seq"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/gene+expression+cancer+RNA-Seq">gene expression cancer RNA-Seq</a>,
 <a
href="datasets/Hybrid+Indoor+Positioning+Dataset+from+WiFi+RSSI%2C+Bluetooth+and+magnetometer"><im
g border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a
href="datasets/Hybrid+Indoor+Positioning+Dataset+from+WiFi+RSSI%2C+Bluetooth+and+magnetometer">Hybr
Indoor Positioning Dataset from WiFi RSSI, Bluetooth and magnetometer</a>,
 <a href="datasets/chestnut+%E2%80%93+LARVIC"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/chestnut+%E2%80%93+LARVIC">chestnut – LARVIC</a>,
 <a
href="datasets/Burst+Header+Packet+%28BHP%29+flooding+attack+on+Optical+Burst+Switching+%28OBS%29+N
rk"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a
href="datasets/Burst+Header+Packet+%28BHP%29+flooding+attack+on+Optical+Burst+Switching+%28OBS%29+N
rk">Burst Header Packet (BHP) flooding attack on Optical Burst Switching (OBS) Network</a>,
 <a href="datasets/Motion+Capture+Hand+Postures"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Motion+Capture+Hand+Postures">Motion Capture Hand Postures</a>,
 <a href="datasets/Anuran+Calls+%28MFCCs%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Anuran+Calls+%28MFCCs%29">Anuran Calls (MFCCs)</a>,
 <a href="datasets/TTC-3600%3A+Benchmark+dataset+for+Turkish+text+categorization"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/TTC-3600%3A+Benchmark+dataset+for+Turkish+text+categorization">TTC-3600:
Benchmark dataset for Turkish text categorization</a>,
```

```
 <a href="datasets/Gastrointestinal+Lesions+in+Regular+Colonoscopy"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Gastrointestinal+Lesions+in+Regular+Colonoscopy">Gastrointestinal Lesions in Re
gular Colonoscopy</a>,
 <a href="datasets/Daily+Demand+Forecasting+Orders"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Daily+Demand+Forecasting+Orders">Daily Demand Forecasting Orders</a>,
 <a href="datasets/Paper+Reviews"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a
>,
 <a href="datasets/Paper+Reviews">Paper Reviews</a>,
 <a href="datasets/extention+of+Z-Alizadeh+sani+dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/extention+of+Z-Alizadeh+sani+dataset">extention of Z-Alizadeh sani dataset</a>,
 <a href="datasets/Z-Alizadeh+Sani"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/><
/a>,
 <a href="datasets/Z-Alizadeh+Sani">Z-Alizadeh Sani</a>,
 <a href="datasets/Dynamic+Features+of+VirusShare+Executables"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Dynamic+Features+of+VirusShare+Executables">Dynamic Features of VirusShare
Executables</a>,
 <a href="datasets/IDA2016Challenge"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>,
 <a href="datasets/IDA2016Challenge">IDA2016Challenge</a>,
 <a href="datasets/DSRC+Vehicle+Communications"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/DSRC+Vehicle+Communications">DSRC Vehicle Communications</a>,
 <a href="datasets/Mturk+User-Perceived+Clusters+over+Images"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Mturk+User-Perceived+Clusters+over+Images">Mturk User-Perceived Clusters over I
mages</a>,
 <a href="datasets/Character+Font+Images"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Character+Font+Images">Character Font Images</a>,
 <a href="datasets/DeliciousMIL%3A+A+Data+Set+for+Multi-Label+Multi-
Instance+Learning+with+Instance+Labels"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/DeliciousMIL%3A+A+Data+Set+for+Multi-Label+Multi-
Instance+Learning+with+Instance+Labels">DeliciousMIL: A Data Set for Multi-Label Multi-Instance Le
arning with Instance Labels</a>,
 <a href="datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Children++"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Children++">Autistic Spectrum Dis
order Screening Data for Children  </a>,
 <a href="datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Adolescent+++"><img border="1" sr
c="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Adolescent+++">Autistic Spectrum
Disorder Screening Data for Adolescent   </a>,
 <a href="datasets/APS+Failure+at+Scania+Trucks"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/APS+Failure+at+Scania+Trucks">APS Failure at Scania Trucks</a>,
 <a href="datasets/Wireless+Indoor+Localization"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Wireless+Indoor+Localization">Wireless Indoor Localization</a>,
 <a href="datasets/HCC+Survival"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>
,
 <a href="datasets/HCC+Survival">HCC Survival</a>,
 <a href="datasets/CSM+%28Conventional+and+Social+Media+Movies%29+Dataset+2014+and+2015"><img bord
er="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/CSM+%28Conventional+and+Social+Media+Movies%29+Dataset+2014+and+2015">CSM
(Conventional and Social Media Movies) Dataset 2014 and 2015</a>,
 <a href="datasets/University+of+Tehran+Question+Dataset+2016+%28UTQD.2016%29"><img border="1" src
="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/University+of+Tehran+Question+Dataset+2016+%28UTQD.2016%29">University of
Tehran Question Dataset 2016 (UTQD.2016)</a>,
 <a href="datasets/Autism+Screening+Adult"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Autism+Screening+Adult">Autism Screening Adult</a>,
 <a
href="datasets/Activity+recognition+with+healthy+older+people+using+a+batteryless+wearable+sensor">
<img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a
href="datasets/Activity+recognition+with+healthy+older+people+using+a+batteryless+wearable+sensor">
vity recognition with healthy older people using a batteryless wearable sensor</a>,
 <a href="datasets/Immunotherapy+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Immunotherapy+Dataset">Immunotherapy Dataset</a>,
 <a href="datasets/Cryotherapy+Dataset+"><img border="1"
```

src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Cryotherapy+Dataset+">Cryotherapy Dataset </a>,
 <a href="datasets/OCT+data+%26+Color+Fundus+Images+of+Left+%26+Right+Eyes"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/OCT+data+%26+Color+Fundus+Images+of+Left+%26+Right+Eyes">OCT data & Color F
undus Images of Left & Right Eyes</a>,
 <a href="datasets/Discrete+Tone+Image+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Discrete+Tone+Image+Dataset">Discrete Tone Image Dataset</a>,
 <a href="datasets/News+Popularity+in+Multiple+Social+Media+Platforms"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/News+Popularity+in+Multiple+Social+Media+Platforms">News Popularity in Multiple
Social Media Platforms</a>,
 <a href="datasets/Ultrasonic+flowmeter+diagnostics"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Ultrasonic+flowmeter+diagnostics">Ultrasonic flowmeter diagnostics</a>,
 <a href="datasets/ICMLA+2014+Accepted+Papers+Data+Set"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/ICMLA+2014+Accepted+Papers+Data+Set">ICMLA 2014 Accepted Papers Data Set</a>,
 <a href="datasets/BLE+RSSI+Dataset+for+Indoor+localization+and+Navigation"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/BLE+RSSI+Dataset+for+Indoor+localization+and+Navigation">BLE RSSI Dataset for I
ndoor localization and Navigation</a>,
 <a href="datasets/Container+Crane+Controller+Data+Set"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Container+Crane+Controller+Data+Set">Container Crane Controller Data Set</a>,
 <a href="datasets/Residential+Building+Data+Set"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Residential+Building+Data+Set">Residential Building Data Set</a>,
 <a href="datasets/Health+News+in+Twitter"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Health+News+in+Twitter">Health News in Twitter</a>,
 <a href="datasets/chipseq"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/chipseq">chipseq</a>,
 <a href="datasets/SGEMM+GPU+kernel+performance"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/SGEMM+GPU+kernel+performance">SGEMM GPU kernel performance</a>,
 <a href="datasets/Repeat+Consumption+Matrices"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Repeat+Consumption+Matrices">Repeat Consumption Matrices</a>,
 <a href="datasets/detection_of_IoT_botnet_attacks_N_BaIoT"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a
href="datasets/detection_of_IoT_botnet_attacks_N_BaIoT">detection_of_IoT_botnet_attacks_N_BaIoT</a>

 <a href="datasets/Absenteeism+at+work"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Absenteeism+at+work">Absenteeism at work</a>,
 <a href="datasets/SCADI"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/SCADI">SCADI</a>,
 <a href="datasets/Condition+monitoring+of+hydraulic+systems"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Condition+monitoring+of+hydraulic+systems">Condition monitoring of hydraulic sy
stems</a>,
 <a href="datasets/Carbon+Nanotubes"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/>
</a>,
 <a href="datasets/Carbon+Nanotubes">Carbon Nanotubes</a>,
 <a href="datasets/Optical+Interconnection+Network+"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Optical+Interconnection+Network+">Optical Interconnection Network </a>,
 <a href="datasets/Sports+articles+for+objectivity+analysis"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Sports+articles+for+objectivity+analysis">Sports articles for objectivity analy
sis</a>,
 <a href="datasets/Breast+Cancer+Coimbra"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Breast+Cancer+Coimbra">Breast Cancer Coimbra</a>,
 <a href="datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data">GNFUV Unmanned Surface Vehicles Se
nsor Data</a>,
 <a href="datasets/Dishonest+Internet+users+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Dishonest+Internet+users+Dataset">Dishonest Internet users Dataset</a>,
 <a href="datasets/Victorian+Era+Authorship+Attribution"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Victorian+Era+Authorship+Attribution">Victorian Era Authorship Attribution</a>,

```
 <a href="datasets/Simulated+Falls+and+Daily+Living+Activities+Data+Set"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Simulated+Falls+and+Daily+Living+Activities+Data+Set">Simulated Falls and Daily
Living Activities Data Set</a>,
 <a href="datasets/Multimodal+Damage+Identification+for+Humanitarian+Computing"><img border="1" sr
c="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Multimodal+Damage+Identification+for+Humanitarian+Computing">Multimodal Damage
Identification for Humanitarian Computing</a>,
 <a href="datasets/EEG+Steady-State+Visual+Evoked+Potential+Signals"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/EEG+Steady-State+Visual+Evoked+Potential+Signals">EEG Steady-State Visual
Evoked Potential Signals</a>,
 <a href="datasets/Roman+Urdu+Data+Set"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Roman+Urdu+Data+Set">Roman Urdu Data Set</a>,
 <a href="datasets/Avila"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Avila">Avila</a>,
 <a href="datasets/PANDOR"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/PANDOR">PANDOR</a>,
 <a href="datasets/Drug+Review+Dataset+%28Druglib.com%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Drug+Review+Dataset+%28Druglib.com%29">Drug Review Dataset (Druglib.com)</a>,
 <a href="datasets/Drug+Review+Dataset+%28Drugs.com%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Drug+Review+Dataset+%28Drugs.com%29">Drug Review Dataset (Drugs.com)</a>,
 <a href="datasets/Physical+Unclonable+Functions"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Physical+Unclonable+Functions">Physical Unclonable Functions</a>,
 <a href="datasets/Superconductivty+Data"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Superconductivty+Data">Superconductivty Data</a>,
 <a href="datasets/WESAD+%28Wearable+Stress+and+Affect+Detection%29"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/WESAD+%28Wearable+Stress+and+Affect+Detection%29">WESAD (Wearable Stress and Af
fect Detection)</a>,
 <a href="datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data+Set+2"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data+Set+2">GNFUV Unmanned Surface
Vehicles Sensor Data Set 2</a>,
 <a href="datasets/Student+Academics+Performance"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Student+Academics+Performance">Student Academics Performance</a>,
 <a href="datasets/Online+Shoppers+Purchasing+Intention+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Online+Shoppers+Purchasing+Intention+Dataset">Online Shoppers Purchasing
Intention Dataset</a>,
 <a href="datasets/PMU-UD"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/PMU-UD">PMU-UD</a>,
 <a href="datasets/Parkinson%27s+Disease+Classification"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Parkinson%27s+Disease+Classification">Parkinson's Disease Classification</a>,
 <a href="datasets/Electrical+Grid+Stability+Simulated+Data+"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Electrical+Grid+Stability+Simulated+Data+">Electrical Grid Stability Simulated
Data </a>,
 <a href="datasets/Caesarian+Section+Classification+Dataset"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Caesarian+Section+Classification+Dataset">Caesarian Section Classification
Dataset</a>,
 <a href="datasets/BAUM-1"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/BAUM-1">BAUM-1</a>,
 <a href="datasets/BAUM-2"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/BAUM-2">BAUM-2</a>,
 <a href="datasets/Audit+Data"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Audit+Data">Audit Data</a>,
 <a href="datasets/BuddyMove+Data+Set"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/BuddyMove+Data+Set">BuddyMove Data Set</a>,
 <a href="datasets/Real+estate+valuation+data+set"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Real+estate+valuation+data+set">Real estate valuation data set</a>,
 <a
href="datasets/Early+biomarkers+of+Parkinson%E2%80%99s+disease+based+on+natural+connected+speech+Da
et+"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a
href="datasets/Early+biomarkers+of+Parkinson%E2%80%99s+disease+based+on+natural+connected+speech+Da
et+">Early biomarkers of Parkinson’s disease based on natural connected speech Data Set </a>,
```

```
 <a href="datasets/Somerville+Happiness+Survey"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Somerville+Happiness+Survey">Somerville Happiness Survey</a>,
 <a href="datasets/2.4+GHZ+Indoor+Channel+Measurements"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/2.4+GHZ+Indoor+Channel+Measurements">2.4 GHZ Indoor Channel Measurements</a>,
 <a href="datasets/EMG+data+for+gestures"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/EMG+data+for+gestures">EMG data for gestures</a>,
 <a href="datasets/Parking+Birmingham"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Parking+Birmingham">Parking Birmingham</a>,
 <a href="datasets/Behavior+of+the+urban+traffic+of+the+city+of+Sao+Paulo+in+Brazil"><img
border="1" src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Behavior+of+the+urban+traffic+of+the+city+of+Sao+Paulo+in+Brazil">Behavior of t
he urban traffic of the city of Sao Paulo in Brazil</a>,
 <a href="datasets/Travel+Reviews"><img border="1" src="assets/MLimages/SmallLargedefault.jpg"/></
a>,
 <a href="datasets/Travel+Reviews">Travel Reviews</a>,
 <a href="datasets/Tarvel+Review+Ratings"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Tarvel+Review+Ratings">Tarvel Review Ratings</a>,
 <a href="datasets/Rice+Leaf+Diseases"><img border="1"
src="assets/MLimages/SmallLargedefault.jpg"/></a>,
 <a href="datasets/Rice+Leaf+Diseases">Rice Leaf Diseases</a>,
 <a href="about.html">About</a>,
 <a href="citation_policy.html">Citation Policy</a>,
 <a href="donation_policy.html">Donation Policy</a>,
 <a href="contact.html">Contact</a>,
 <a href="http://cml.ics.uci.edu">CML</a>]
```

In [262]:

```
len(links)
```

Out[262]:

```
988
```

In [263]:

```
import re
```

In [264]:

```
re.findall('"(datasets/.+?)"', str(links))
```

Out[264]:

```
['datasets/Abalone',
 'datasets/Abalone',
 'datasets/Adult',
 'datasets/Adult',
 'datasets/Annealing',
 'datasets/Annealing',
 'datasets/Anonymous+Microsoft+Web+Data',
 'datasets/Anonymous+Microsoft+Web+Data',
 'datasets/Arrhythmia',
 'datasets/Arrhythmia',
 'datasets/Artificial+Characters',
 'datasets/Artificial+Characters',
 'datasets/Audiology+%28Original%29',
 'datasets/Audiology+%28Original%29',
 'datasets/Audiology+%28Standardized%29',
 'datasets/Audiology+%28Standardized%29',
 'datasets/Auto+MPG',
 'datasets/Auto+MPG',
 'datasets/Automobile',
 'datasets/Automobile',
 'datasets/Badges',
 'datasets/Badges',
 'datasets/Balance+Scale',
 'datasets/Balance+Scale',
```

```
'datasets/Balloons',
'datasets/Balloons',
'datasets/Breast+Cancer',
'datasets/Breast+Cancer',
'datasets/Breast+Cancer+Wisconsin+%28Original%29',
'datasets/Breast+Cancer+Wisconsin+%28Original%29',
'datasets/Breast+Cancer+Wisconsin+%28Prognostic%29',
'datasets/Breast+Cancer+Wisconsin+%28Prognostic%29',
'datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29',
'datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29',
'datasets/Pittsburgh+Bridges',
'datasets/Pittsburgh+Bridges',
'datasets/Car+Evaluation',
'datasets/Car+Evaluation',
'datasets/Census+Income',
'datasets/Census+Income',
'datasets/Chess+%28King-Rook+vs.+King-Knight%29',
'datasets/Chess+%28King-Rook+vs.+King-Knight%29',
'datasets/Chess+%28King-Rook+vs.+King-Pawn%29',
'datasets/Chess+%28King-Rook+vs.+King-Pawn%29',
'datasets/Chess+%28King-Rook+vs.+King%29',
'datasets/Chess+%28King-Rook+vs.+King%29',
'datasets/Chess+%28Domain+Theories%29',
'datasets/Chess+%28Domain+Theories%29',
'datasets/Bach+Chorales',
'datasets/Bach+Chorales',
'datasets/Connect-4',
'datasets/Connect-4',
'datasets/Credit+Approval',
'datasets/Credit+Approval',
'datasets/Japanese+Credit+Screening',
'datasets/Japanese+Credit+Screening',
'datasets/Computer+Hardware',
'datasets/Computer+Hardware',
'datasets/Contraceptive+Method+Choice',
'datasets/Contraceptive+Method+Choice',
'datasets/Covertype',
'datasets/Covertype',
'datasets/Cylinder+Bands',
'datasets/Cylinder+Bands',
'datasets/Dermatology',
'datasets/Dermatology',
'datasets/Diabetes',
'datasets/Diabetes',
'datasets/DGP2+-+The+Second+Data+Generation+Program',
'datasets/DGP2+-+The+Second+Data+Generation+Program',
'datasets/Document+Understanding',
'datasets/Document+Understanding',
'datasets/EBL+Domain+Theories',
'datasets/EBL+Domain+Theories',
'datasets/Echocardiogram',
'datasets/Echocardiogram',
'datasets/Ecoli',
'datasets/Ecoli',
'datasets/Flags',
'datasets/Flags',
'datasets/Function+Finding',
'datasets/Function+Finding',
'datasets/Glass+Identification',
'datasets/Glass+Identification',
'datasets/Haberman%27s+Survival',
'datasets/Haberman%27s+Survival',
'datasets/Hayes-Roth',
'datasets/Hayes-Roth',
'datasets/Heart+Disease',
'datasets/Heart+Disease',
'datasets/Hepatitis',
'datasets/Hepatitis',
'datasets/Horse+Colic',
'datasets/Horse+Colic',
'datasets/ICU',
'datasets/ICU',
'datasets/Image+Segmentation',
'datasets/Image+Segmentation',
'datasets/Internet+Advertisements',
'datasets/Internet+Advertisements',
'datasets/Ionosphere',
```

```
'datasets/Ionosphere',
'datasets/Iris',
'datasets/Iris',
'datasets/ISOLET',
'datasets/ISOLET',
'datasets/Kinship',
'datasets/Kinship',
'datasets/Labor+Relations',
'datasets/Labor+Relations',
'datasets/LED+Display+Domain',
'datasets/LED+Display+Domain',
'datasets/Lenses',
'datasets/Lenses',
'datasets/Letter+Recognition',
'datasets/Letter+Recognition',
'datasets/Liver+Disorders',
'datasets/Liver+Disorders',
'datasets/Logic+Theorist',
'datasets/Logic+Theorist',
'datasets/Lung+Cancer',
'datasets/Lung+Cancer',
'datasets/Lymphography',
'datasets/Lymphography',
'datasets/Mechanical+Analysis',
'datasets/Mechanical+Analysis',
'datasets/Meta-data',
'datasets/Meta-data',
'datasets/Mobile+Robots',
'datasets/Mobile+Robots',
'datasets/Molecular+Biology+%28Promoter+Gene+Sequences%29',
'datasets/Molecular+Biology+%28Promoter+Gene+Sequences%29',
'datasets/Molecular+Biology+%28Protein+Secondary+Structure%29',
'datasets/Molecular+Biology+%28Protein+Secondary+Structure%29',
'datasets/Molecular+Biology+%28Splice-junction+Gene+Sequences%29',
'datasets/Molecular+Biology+%28Splice-junction+Gene+Sequences%29',
'datasets/MONK%27s+Problems',
'datasets/MONK%27s+Problems',
'datasets/Moral+Reasoner',
'datasets/Moral+Reasoner',
'datasets/Multiple+Features',
'datasets/Multiple+Features',
'datasets/Mushroom',
'datasets/Mushroom',
'datasets/Musk+%28Version+1%29',
'datasets/Musk+%28Version+1%29',
'datasets/Musk+%28Version+2%29',
'datasets/Musk+%28Version+2%29',
'datasets/Nursery',
'datasets/Nursery',
'datasets/Othello+Domain+Theory',
'datasets/Othello+Domain+Theory',
'datasets/Page+Blocks+Classification',
'datasets/Page+Blocks+Classification',
'datasets/Optical+Recognition+of+Handwritten+Digits',
'datasets/Optical+Recognition+of+Handwritten+Digits',
'datasets/Pen-Based+Recognition+of+Handwritten+Digits',
'datasets/Pen-Based+Recognition+of+Handwritten+Digits',
'datasets/Post-Operative+Patient',
'datasets/Post-Operative+Patient',
'datasets/Primary+Tumor',
'datasets/Primary+Tumor',
'datasets/Prodigy',
'datasets/Prodigy',
'datasets/Qualitative+Structure+Activity+Relationships',
'datasets/Qualitative+Structure+Activity+Relationships',
'datasets/Quadruped+Mammals',
'datasets/Quadruped+Mammals',
'datasets/Servo',
'datasets/Servo',
'datasets/Shuttle+Landing+Control',
'datasets/Shuttle+Landing+Control',
'datasets/Solar+Flare',
'datasets/Solar+Flare',
'datasets/Soybean+%28Large%29',
'datasets/Soybean+%28Large%29',
'datasets/Soybean+%28Small%29',
'datasets/Soybean+%28Small%29',
```

```
'datasets/Challenger+USA+Space+Shuttle+O-Ring',
'datasets/Challenger+USA+Space+Shuttle+O-Ring',
'datasets/Low+Resolution+Spectrometer',
'datasets/Low+Resolution+Spectrometer',
'datasets/Spambase',
'datasets/Spambase',
'datasets/SPECT+Heart',
'datasets/SPECT+Heart',
'datasets/SPECTF+Heart',
'datasets/SPECTF+Heart',
'datasets/Sponge',
'datasets/Sponge',
'datasets/Statlog+Project',
'datasets/Statlog+Project',
'datasets/Student+Loan+Relational',
'datasets/Student+Loan+Relational',
'datasets/Teaching+Assistant+Evaluation',
'datasets/Teaching+Assistant+Evaluation',
'datasets/Tic-Tac-Toe+Endgame',
'datasets/Tic-Tac-Toe+Endgame',
'datasets/Thyroid+Disease',
'datasets/Thyroid+Disease',
'datasets/Trains',
'datasets/Trains',
'datasets/University',
'datasets/University',
'datasets/Congressional+Voting+Records',
'datasets/Congressional+Voting+Records',
'datasets/Water+Treatment+Plant',
'datasets/Water+Treatment+Plant',
'datasets/Waveform+Database+Generator+%28Version+1%29',
'datasets/Waveform+Database+Generator+%28Version+1%29',
'datasets/Waveform+Database+Generator+%28Version+2%29',
'datasets/Waveform+Database+Generator+%28Version+2%29',
'datasets/Wine',
'datasets/Wine',
'datasets/Yeast',
'datasets/Yeast',
'datasets/Zoo',
'datasets/Zoo',
'datasets/Undocumented',
'datasets/Undocumented',
'datasets/Twenty+Newsgroups',
'datasets/Twenty+Newsgroups',
'datasets/Australian+Sign+Language+signs',
'datasets/Australian+Sign+Language+signs',
'datasets/Australian+Sign+Language+signs+%28High+Quality%29',
'datasets/Australian+Sign+Language+signs+%28High+Quality%29',
'datasets/US+Census+Data+%281990%29',
'datasets/US+Census+Data+%281990%29',
'datasets/Census-Income+%28KDD%29',
'datasets/Census-Income+%28KDD%29',
'datasets/Coil+1999+Competition+Data',
'datasets/Coil+1999+Competition+Data',
'datasets/Corel+Image+Features',
'datasets/Corel+Image+Features',
'datasets/E.+Coli+Genes',
'datasets/E.+Coli+Genes',
'datasets/EEG+Database',
'datasets/EEG+Database',
'datasets/El+Nino',
'datasets/El+Nino',
'datasets/Entree+Chicago+Recommendation+Data',
'datasets/Entree+Chicago+Recommendation+Data',
'datasets/CMU+Face+Images',
'datasets/CMU+Face+Images',
'datasets/Insurance+Company+Benchmark+%28COIL+2000%29',
'datasets/Insurance+Company+Benchmark+%28COIL+2000%29',
'datasets/Internet+Usage+Data',
'datasets/Internet+Usage+Data',
'datasets/IPUMS+Census+Database',
'datasets/IPUMS+Census+Database',
'datasets/Japanese+Vowels',
'datasets/Japanese+Vowels',
'datasets/KDD+Cup+1998+Data',
'datasets/KDD+Cup+1998+Data',
'datasets/KDD+Cup+1999+Data',
```

```
'datasets/KDD+Cup+1999+Data',
'datasets/M.+Tuberculosis+Genes',
'datasets/M.+Tuberculosis+Genes',
'datasets/Movie',
'datasets/Movie',
'datasets/MSNBC.com+Anonymous+Web+Data',
'datasets/MSNBC.com+Anonymous+Web+Data',
'datasets/NSF+Research+Award+Abstracts+1990-2003',
'datasets/NSF+Research+Award+Abstracts+1990-2003',
'datasets/Pioneer-1+Mobile+Robot+Data',
'datasets/Pioneer-1+Mobile+Robot+Data',
'datasets/Pseudo+Periodic+Synthetic+Time+Series',
'datasets/Pseudo+Periodic+Synthetic+Time+Series',
'datasets/Reuters-21578+Text+Categorization+Collection',
'datasets/Reuters-21578+Text+Categorization+Collection',
'datasets/Robot+Execution+Failures',
'datasets/Robot+Execution+Failures',
'datasets/Synthetic+Control+Chart+Time+Series',
'datasets/Synthetic+Control+Chart+Time+Series',
'datasets/Syskill+and+Webert+Web+Page+Ratings',
'datasets/Syskill+and+Webert+Web+Page+Ratings',
'datasets/UNIX+User+Data',
'datasets/UNIX+User+Data',
'datasets/Volcanoes+on+Venus+-+JARtool+experiment',
'datasets/Volcanoes+on+Venus+-+JARtool+experiment',
'datasets/Statlog+%28Australian+Credit+Approval%29',
'datasets/Statlog+%28Australian+Credit+Approval%29',
'datasets/Statlog+%28German+Credit+Data%29',
'datasets/Statlog+%28German+Credit+Data%29',
'datasets/Statlog+%28Heart%29',
'datasets/Statlog+%28Heart%29',
'datasets/Statlog+%28Landsat+Satellite%29',
'datasets/Statlog+%28Landsat+Satellite%29',
'datasets/Statlog+%28Image+Segmentation%29',
'datasets/Statlog+%28Image+Segmentation%29',
'datasets/Statlog+%28Shuttle%29',
'datasets/Statlog+%28Shuttle%29',
'datasets/Statlog+%28Vehicle+Silhouettes%29',
'datasets/Statlog+%28Vehicle+Silhouettes%29',
'datasets/Connectionist+Bench+%28Nettalk+Corpus%29',
'datasets/Connectionist+Bench+%28Nettalk+Corpus%29',
'datasets/Connectionist+Bench+%28Sonar%2C+Mines+vs.+Rocks%29',
'datasets/Connectionist+Bench+%28Sonar%2C+Mines+vs.+Rocks%29',
'datasets/Connectionist+Bench+%28Vowel+Recognition+-+Deterding+Data%29',
'datasets/Connectionist+Bench+%28Vowel+Recognition+-+Deterding+Data%29',
'datasets/Economic+Sanctions',
'datasets/Economic+Sanctions',
'datasets/Protein+Data',
'datasets/Protein+Data',
'datasets/Cloud',
'datasets/Cloud',
'datasets/CalIt2+Building+People+Counts',
'datasets/CalIt2+Building+People+Counts',
'datasets/Dodgers+Loop+Sensor',
'datasets/Dodgers+Loop+Sensor',
'datasets/Poker+Hand',
'datasets/Poker+Hand',
'datasets/MAGIC+Gamma+Telescope',
'datasets/MAGIC+Gamma+Telescope',
'datasets/UJI+Pen+Characters',
'datasets/UJI+Pen+Characters',
'datasets/Mammographic+Mass',
'datasets/Mammographic+Mass',
'datasets/Forest+Fires',
'datasets/Forest+Fires',
'datasets/Reuters+Transcribed+Subset',
'datasets/Reuters+Transcribed+Subset',
'datasets/Bag+of+Words',
'datasets/Bag+of+Words',
'datasets/Concrete+Compressive+Strength',
'datasets/Concrete+Compressive+Strength',
'datasets/Hill-Valley',
'datasets/Hill-Valley',
'datasets/Arcene',
'datasets/Arcene',
'datasets/Dexter',
'datasets/Dexter',
```

```
'datasets/Dorothea',
'datasets/Dorothea',
'datasets/Gisette',
'datasets/Gisette',
'datasets/Madelon',
'datasets/Madelon',
'datasets/Ozone+Level+Detection',
'datasets/Ozone+Level+Detection',
'datasets/Abscisic+Acid+Signaling+Network',
'datasets/Abscisic+Acid+Signaling+Network',
'datasets/Parkinsons',
'datasets/Parkinsons',
'datasets/Character+Trajectories',
'datasets/Character+Trajectories',
'datasets/Blood+Transfusion+Service+Center',
'datasets/Blood+Transfusion+Service+Center',
'datasets/UJI+Pen+Characters+%28Version+2%29',
'datasets/UJI+Pen+Characters+%28Version+2%29',
'datasets/Semeion+Handwritten+Digit',
'datasets/Semeion+Handwritten+Digit',
'datasets/SECOM',
'datasets/SECOM',
'datasets/Plants',
'datasets/Plants',
'datasets/Libras+Movement',
'datasets/Libras+Movement',
'datasets/Concrete+Slump+Test',
'datasets/Concrete+Slump+Test',
'datasets/Communities+and+Crime',
'datasets/Communities+and+Crime',
'datasets/Acute+Inflammations',
'datasets/Acute+Inflammations',
'datasets/Wine+Quality',
'datasets/Wine+Quality',
'datasets/URL+Reputation',
'datasets/URL+Reputation',
'datasets/p53+Mutants',
'datasets/p53+Mutants',
'datasets/Parkinsons+Telemonitoring',
'datasets/Parkinsons+Telemonitoring',
'datasets/Demospongiae',
'datasets/Demospongiae',
'datasets/Opinosis+Opinion+%26frasl%3B+Review',
'datasets/Opinosis+Opinion+%26frasl%3B+Review',
'datasets/Breast+Tissue',
'datasets/Breast+Tissue',
'datasets/Cardiotocography',
'datasets/Cardiotocography',
'datasets/Wall-Following+Robot+Navigation+Data',
'datasets/Wall-Following+Robot+Navigation+Data',
'datasets/Spoken+Arabic+Digit',
'datasets/Spoken+Arabic+Digit',
'datasets/Localization+Data+for+Person+Activity',
'datasets/Localization+Data+for+Person+Activity',
'datasets/AutoUniv',
'datasets/AutoUniv',
'datasets/Steel+Plates+Faults',
'datasets/Steel+Plates+Faults',
'datasets/MiniBooNE+particle+identification',
'datasets/MiniBooNE+particle+identification',
'datasets/YearPredictionMSD',
'datasets/YearPredictionMSD',
'datasets/PEMS-SF',
'datasets/PEMS-SF',
'datasets/OpinRank+Review+Dataset',
'datasets/OpinRank+Review+Dataset',
'datasets/Relative+location+of+CT+slices+on+axial+axis',
'datasets/Relative+location+of+CT+slices+on+axial+axis',
'datasets/Online+Handwritten+Assamese+Characters+Dataset',
'datasets/Online+Handwritten+Assamese+Characters+Dataset',
'datasets/PubChem+Bioassay+Data',
'datasets/PubChem+Bioassay+Data',
'datasets/Record+Linkage+Comparison+Patterns',
'datasets/Record+Linkage+Comparison+Patterns',
'datasets/Communities+and+Crime+Unnormalized',
'datasets/Communities+and+Crime+Unnormalized',
'datasets/Vertebral+Column',
```

```
'datasets/Vertebral+Column',
'datasets/EMG+Physical+Action+Data+Set',
'datasets/EMG+Physical+Action+Data+Set',
'datasets/Vicon+Physical+Action+Data+Set',
'datasets/Vicon+Physical+Action+Data+Set',
'datasets/Amazon+Commerce+reviews+set',
'datasets/Amazon+Commerce+reviews+set',
'datasets/Amazon+Access+Samples',
'datasets/Amazon+Access+Samples',
'datasets/Reuter_50_50',
'datasets/Reuter_50_50',
'datasets/Farm+Ads',
'datasets/Farm+Ads',
'datasets/DBWorld+e-mails',
'datasets/DBWorld+e-mails',
'datasets/KEGG+Metabolic+Relation+Network+%28Directed%29',
'datasets/KEGG+Metabolic+Relation+Network+%28Directed%29',
'datasets/KEGG+Metabolic+Reaction+Network+%28Undirected%29',
'datasets/KEGG+Metabolic+Reaction+Network+%28Undirected%29',
'datasets/Bank+Marketing',
'datasets/Bank+Marketing',
'datasets/YouTube+Comedy+Slam+Preference+Data',
'datasets/YouTube+Comedy+Slam+Preference+Data',
'datasets/Gas+Sensor+Array+Drift+Dataset',
'datasets/Gas+Sensor+Array+Drift+Dataset',
'datasets/ILPD+%28Indian+Liver+Patient+Dataset%29',
'datasets/ILPD+%28Indian+Liver+Patient+Dataset%29',
'datasets/OPPORTUNITY+Activity+Recognition',
'datasets/OPPORTUNITY+Activity+Recognition',
'datasets/Nomao',
'datasets/Nomao',
'datasets/SMS+Spam+Collection',
'datasets/SMS+Spam+Collection',
'datasets/Skin+Segmentation',
'datasets/Skin+Segmentation',
'datasets/Planning+Relax',
'datasets/Planning+Relax',
'datasets/PAMAP2+Physical+Activity+Monitoring',
'datasets/PAMAP2+Physical+Activity+Monitoring',
'datasets/Restaurant+%26+consumer+data',
'datasets/Restaurant+%26+consumer+data',
'datasets/CNAE-9',
'datasets/CNAE-9',
'datasets/Individual+household+electric+power+consumption',
'datasets/Individual+household+electric+power+consumption',
'datasets/seeds',
'datasets/seeds',
'datasets/Northix',
'datasets/Northix',
'datasets/QtyT40I10D100K',
'datasets/QtyT40I10D100K',
'datasets/Legal+Case+Reports',
'datasets/Legal+Case+Reports',
'datasets/Human+Activity+Recognition+Using+Smartphones',
'datasets/Human+Activity+Recognition+Using+Smartphones',
'datasets/One-hundred+plant+species+leaves+data+set',
'datasets/One-hundred+plant+species+leaves+data+set',
'datasets/Energy+efficiency',
'datasets/Energy+efficiency',
'datasets/Yacht+Hydrodynamics',
'datasets/Yacht+Hydrodynamics',
'datasets/Fertility',
'datasets/Fertility',
'datasets/Daphnet+Freezing+of+Gait',
'datasets/Daphnet+Freezing+of+Gait',
'datasets/3D+Road+Network+%28North+Jutland%2C+Denmark%29',
'datasets/3D+Road+Network+%28North+Jutland%2C+Denmark%29',
'datasets/ISTANBUL+STOCK+EXCHANGE',
'datasets/ISTANBUL+STOCK+EXCHANGE',
'datasets/Buzz+in+social+media+',
'datasets/Buzz+in+social+media+',
'datasets/First-order+theorem+proving',
'datasets/First-order+theorem+proving',
'datasets/Wearable+Computing%3A+Classification+of+Body+Postures+and+Movements+%28PUC-Rio%29',
'datasets/Wearable+Computing%3A+Classification+of+Body+Postures+and+Movements+%28PUC-Rio%29',
'datasets/Gas+sensor+arrays+in+open+sampling+settings',
'datasets/Gas+sensor+arrays+in+open+sampling+settings',
```

```
 'datasets/Climate+Model+Simulation+Crashes',
 'datasets/Climate+Model+Simulation+Crashes',
 'datasets/MicroMass',
 'datasets/MicroMass',
 'datasets/QSAR+biodegradation',
 'datasets/QSAR+biodegradation',
 'datasets/BLOGGER',
 'datasets/BLOGGER',
 'datasets/Daily+and+Sports+Activities',
 'datasets/Daily+and+Sports+Activities',
 'datasets/User+Knowledge+Modeling',
 'datasets/User+Knowledge+Modeling',
 'datasets/Reuters+RCV1+RCV2+Multilingual%2C+Multiview+Text+Categorization+Test+collection',
 'datasets/Reuters+RCV1+RCV2+Multilingual%2C+Multiview+Text+Categorization+Test+collection',
 'datasets/NYSK',
 'datasets/NYSK',
 'datasets/Turkiye+Student+Evaluation',
 'datasets/Turkiye+Student+Evaluation',

'datasets/ser+Knowledge+Modeling+Data+%28Students%27+Knowledge+Levels+on+DC+Electrical+Machines%29'


'datasets/ser+Knowledge+Modeling+Data+%28Students%27+Knowledge+Levels+on+DC+Electrical+Machines%29'

 'datasets/EEG+Eye+State',
 'datasets/EEG+Eye+State',
 'datasets/Physicochemical+Properties+of+Protein+Tertiary+Structure',
 'datasets/Physicochemical+Properties+of+Protein+Tertiary+Structure',
 'datasets/seismic-bumps',
 'datasets/seismic-bumps',
 'datasets/banknote+authentication',
 'datasets/banknote+authentication',
 'datasets/USPTO+Algorithm+Challenge%2C+run+by+NASA-
Harvard+Tournament+Lab+and+TopCoder++++Problem%3A+Pat',
 'datasets/USPTO+Algorithm+Challenge%2C+run+by+NASA-
Harvard+Tournament+Lab+and+TopCoder++++Problem%3A+Pat',
 'datasets/YouTube+Multiview+Video+Games+Dataset',
 'datasets/YouTube+Multiview+Video+Games+Dataset',
 'datasets/Gas+Sensor+Array+Drift+Dataset+at+Different+Concentrations',
 'datasets/Gas+Sensor+Array+Drift+Dataset+at+Different+Concentrations',
 'datasets/Activities+of+Daily+Living+%28ADLs%29+Recognition+Using+Binary+Sensors',
 'datasets/Activities+of+Daily+Living+%28ADLs%29+Recognition+Using+Binary+Sensors',
 'datasets/SkillCraft1+Master+Table+Dataset',
 'datasets/SkillCraft1+Master+Table+Dataset',
 'datasets/Weight+Lifting+Exercises+monitored+with+Inertial+Measurement+Units',
 'datasets/Weight+Lifting+Exercises+monitored+with+Inertial+Measurement+Units',
 'datasets/SML2010',
 'datasets/SML2010',
 'datasets/Bike+Sharing+Dataset',
 'datasets/Bike+Sharing+Dataset',
 'datasets/Predict+keywords+activities+in+a+online+social+media',
 'datasets/Predict+keywords+activities+in+a+online+social+media',
 'datasets/Thoracic+Surgery+Data',
 'datasets/Thoracic+Surgery+Data',
 'datasets/EMG+dataset+in+Lower+Limb',
 'datasets/EMG+dataset+in+Lower+Limb',
 'datasets/SUSY',
 'datasets/SUSY',
 'datasets/HIGGS',
 'datasets/HIGGS',
 'datasets/Qualitative_Bankruptcy',
 'datasets/Qualitative_Bankruptcy',
 'datasets/LSVT+Voice+Rehabilitation',
 'datasets/LSVT+Voice+Rehabilitation',
 'datasets/Dataset+for+ADL+Recognition+with+Wrist-worn+Accelerometer',
 'datasets/Dataset+for+ADL+Recognition+with+Wrist-worn+Accelerometer',
 'datasets/Wilt',
 'datasets/Wilt',
 'datasets/User+Identification+From+Walking+Activity',
 'datasets/User+Identification+From+Walking+Activity',
 'datasets/Activity+Recognition+from+Single+Chest-Mounted+Accelerometer',
 'datasets/Activity+Recognition+from+Single+Chest-Mounted+Accelerometer',
 'datasets/Leaf',
 'datasets/Leaf',
 'datasets/Dresses_Attribute_Sales',
 'datasets/Dresses_Attribute_Sales',
 'datasets/Tamilnadu+Electricity+Board+Hourly+Readings',
```

```
'datasets/Tamilnadu+Electricity+Board+Hourly+Readings',
'datasets/Airfoil+Self-Noise',
'datasets/Airfoil+Self-Noise',
'datasets/Wholesale+customers',
'datasets/Wholesale+customers',
'datasets/Twitter+Data+set+for+Arabic+Sentiment+Analysis',
'datasets/Twitter+Data+set+for+Arabic+Sentiment+Analysis',
'datasets/Combined+Cycle+Power+Plant',
'datasets/Combined+Cycle+Power+Plant',
'datasets/Urban+Land+Cover',
'datasets/Urban+Land+Cover',
'datasets/Diabetes+130-US+hospitals+for+years+1999-2008',
'datasets/Diabetes+130-US+hospitals+for+years+1999-2008',
'datasets/Bach+Choral+Harmony',
'datasets/Bach+Choral+Harmony',
'datasets/StoneFlakes',
'datasets/StoneFlakes',
'datasets/Tennis+Major+Tournament+Match+Statistics',
'datasets/Tennis+Major+Tournament+Match+Statistics',
'datasets/Parkinson+Speech+Dataset+with++Multiple+Types+of+Sound+Recordings',
'datasets/Parkinson+Speech+Dataset+with++Multiple+Types+of+Sound+Recordings',
'datasets/Gesture+Phase+Segmentation',
'datasets/Gesture+Phase+Segmentation',
'datasets/Perfume+Data',
'datasets/Perfume+Data',
'datasets/BlogFeedback',
'datasets/BlogFeedback',
'datasets/REALDISP+Activity+Recognition+Dataset',
'datasets/REALDISP+Activity+Recognition+Dataset',
'datasets/Newspaper+and+magazine+images+segmentation+dataset',
'datasets/Newspaper+and+magazine+images+segmentation+dataset',
'datasets/AAAI+2014+Accepted+Papers',
'datasets/AAAI+2014+Accepted+Papers',
'datasets/Gas+sensor+array+under+flow+modulation',
'datasets/Gas+sensor+array+under+flow+modulation',
'datasets/Gas+sensor+array+exposed+to+turbulent+gas+mixtures',
'datasets/Gas+sensor+array+exposed+to+turbulent+gas+mixtures',
'datasets/UJIIndoorLoc',
'datasets/UJIIndoorLoc',
'datasets/Sentence+Classification',
'datasets/Sentence+Classification',
'datasets/Dow+Jones+Index',
'datasets/Dow+Jones+Index',
'datasets/sEMG+for+Basic+Hand+movements',
'datasets/sEMG+for+Basic+Hand+movements',
'datasets/AAAI+2013+Accepted+Papers',
'datasets/AAAI+2013+Accepted+Papers',
'datasets/Geographical+Original+of+Music',
'datasets/Geographical+Original+of+Music',
'datasets/Condition+Based+Maintenance+of+Naval+Propulsion+Plants',
'datasets/Condition+Based+Maintenance+of+Naval+Propulsion+Plants',
'datasets/Grammatical+Facial+Expressions',
'datasets/Grammatical+Facial+Expressions',
'datasets/NoisyOffice',
'datasets/NoisyOffice',
'datasets/MHEALTH+Dataset',
'datasets/MHEALTH+Dataset',
'datasets/Student+Performance',
'datasets/Student+Performance',
'datasets/ElectricityLoadDiagrams20112014',
'datasets/ElectricityLoadDiagrams20112014',
'datasets/Gas+sensor+array+under+dynamic+gas+mixtures',
'datasets/Gas+sensor+array+under+dynamic+gas+mixtures',
'datasets/microblogPCU',
'datasets/microblogPCU',
'datasets/Firm-Teacher_Clave-Direction_Classification',
'datasets/Firm-Teacher_Clave-Direction_Classification',
'datasets/Dataset+for+Sensorless+Drive+Diagnosis',
'datasets/Dataset+for+Sensorless+Drive+Diagnosis',
'datasets/TV+News+Channel+Commercial+Detection+Dataset',
'datasets/TV+News+Channel+Commercial+Detection+Dataset',
'datasets/Phishing+Websites',
'datasets/Phishing+Websites',
'datasets/Greenhouse+Gas+Observing+Network',
'datasets/Greenhouse+Gas+Observing+Network',
'datasets/Diabetic+Retinopathy+Debrecen+Data+Set',
'datasets/Diabetic+Retinopathy+Debrecen+Data+Set',
```

```
 'datasets/HIV-1+protease+cleavage',
 'datasets/HIV-1+protease+cleavage',
 'datasets/Sentiment+Labelled+Sentences',
 'datasets/Sentiment+Labelled+Sentences',
 'datasets/Online+News+Popularity',
 'datasets/Online+News+Popularity',
 'datasets/Forest+type+mapping',
 'datasets/Forest+type+mapping',
 'datasets/wiki4HE',
 'datasets/wiki4HE',
 'datasets/Online+Video+Characteristics+and+Transcoding+Time+Dataset',
 'datasets/Online+Video+Characteristics+and+Transcoding+Time+Dataset',
 'datasets/Chronic_Kidney_Disease',
 'datasets/Chronic_Kidney_Disease',
 'datasets/Machine+Learning+based+ZZAlpha+Ltd.+Stock+Recommendations+2012-2014',
 'datasets/Machine+Learning+based+ZZAlpha+Ltd.+Stock+Recommendations+2012-2014',
 'datasets/Folio',
 'datasets/Folio',
 'datasets/Taxi+Service+Trajectory+-+Prediction+Challenge%2C+ECML+PKDD+2015',
 'datasets/Taxi+Service+Trajectory+-+Prediction+Challenge%2C+ECML+PKDD+2015',
 'datasets/Cuff-Less+Blood+Pressure+Estimation',
 'datasets/Cuff-Less+Blood+Pressure+Estimation',
 'datasets/Smartphone-Based+Recognition+of+Human+Activities+and+Postural+Transitions',
 'datasets/Smartphone-Based+Recognition+of+Human+Activities+and+Postural+Transitions',
 'datasets/Mice+Protein+Expression',
 'datasets/Mice+Protein+Expression',
 'datasets/UJIIndoorLoc-Mag',
 'datasets/UJIIndoorLoc-Mag',
 'datasets/Heterogeneity+Activity+Recognition',
 'datasets/Heterogeneity+Activity+Recognition',
 'datasets/Educational+Process+Mining+%28EPM%29%3A+A+Learning+Analytics+Data+Set',
 'datasets/Educational+Process+Mining+%28EPM%29%3A+A+Learning+Analytics+Data+Set',
 'datasets/HEPMASS',
 'datasets/HEPMASS',
 'datasets/Indoor+User+Movement+Prediction+from+RSS+data',
 'datasets/Indoor+User+Movement+Prediction+from+RSS+data',
 'datasets/Open+University+Learning+Analytics+dataset',
 'datasets/Open+University+Learning+Analytics+dataset',
 'datasets/default+of+credit+card+clients',
 'datasets/default+of+credit+card+clients',
 'datasets/Mesothelioma%E2%80%99s+disease+data+set+',
 'datasets/Mesothelioma%E2%80%99s+disease+data+set+',
 'datasets/Online+Retail',
 'datasets/Online+Retail',
 'datasets/SIFT10M',
 'datasets/SIFT10M',
 'datasets/GPS+Trajectories',
 'datasets/GPS+Trajectories',
 'datasets/Detect+Malacious+Executable%28AntiVirus%29',
 'datasets/Detect+Malacious+Executable%28AntiVirus%29',
 'datasets/Occupancy+Detection+',
 'datasets/Occupancy+Detection+',

'datasets/Improved+Spiral+Test+Using+Digitized+Graphics+Tablet+for+Monitoring+Parkinson%E2%80%99s+D
se',

'datasets/Improved+Spiral+Test+Using+Digitized+Graphics+Tablet+for+Monitoring+Parkinson%E2%80%99s+D
se',
 'datasets/News+Aggregator',
 'datasets/News+Aggregator',
 'datasets/Air+Quality',
 'datasets/Air+Quality',
 'datasets/Twin+gas+sensor+arrays',
 'datasets/Twin+gas+sensor+arrays',
 'datasets/Gas+sensors+for+home+activity+monitoring',
 'datasets/Gas+sensors+for+home+activity+monitoring',
 'datasets/Facebook+Comment+Volume+Dataset',
 'datasets/Facebook+Comment+Volume+Dataset',

'datasets/Smartphone+Dataset+for+Human+Activity+Recognition+%28HAR%29+in+Ambient+Assisted+Living+%2
%29',

'datasets/Smartphone+Dataset+for+Human+Activity+Recognition+%28HAR%29+in+Ambient+Assisted+Living+%2
%29',
 'datasets/Polish+companies+bankruptcy+data',
 'datasets/Polish+companies+bankruptcy+data',
 'datasets/Activity+Recognition+system+based+on+Multisensor+data+fusion+%28AReM%29',
```

```
 'datasets/Activity+Recognition+system+based+on+Multisensor+data+fusion+%28AReM%29',
 'datasets/Dota2+Games+Results',
 'datasets/Dota2+Games+Results',
 'datasets/Facebook+metrics',
 'datasets/Facebook+metrics',
 'datasets/UbiqLog+%28smartphone+lifelogging%29',
 'datasets/UbiqLog+%28smartphone+lifelogging%29',
 'datasets/NIPS+Conference+Papers+1987-2015',
 'datasets/NIPS+Conference+Papers+1987-2015',
 'datasets/HTRU2',
 'datasets/HTRU2',
 'datasets/Drug+consumption+%28quantified%29',
 'datasets/Drug+consumption+%28quantified%29',
 'datasets/Appliances+energy+prediction',
 'datasets/Appliances+energy+prediction',
 'datasets/Miskolc+IIS+Hybrid+IPS',
 'datasets/Miskolc+IIS+Hybrid+IPS',
 'datasets/KDC-4007+dataset+Collection',
 'datasets/KDC-4007+dataset+Collection',
 'datasets/Geo-
Magnetic+field+and+WLAN+dataset+for+indoor+localisation+from+wristband+and+smartphone',
 'datasets/Geo-
Magnetic+field+and+WLAN+dataset+for+indoor+localisation+from+wristband+and+smartphone',
 'datasets/DrivFace',
 'datasets/DrivFace',
 'datasets/Website+Phishing',
 'datasets/Website+Phishing',
 'datasets/YouTube+Spam+Collection',
 'datasets/YouTube+Spam+Collection',
 'datasets/Beijing+PM2.5+Data',
 'datasets/Beijing+PM2.5+Data',
 'datasets/Cargo+2000+Freight+Tracking+and+Tracing',
 'datasets/Cargo+2000+Freight+Tracking+and+Tracing',
 'datasets/Cervical+cancer+%28Risk+Factors%29',
 'datasets/Cervical+cancer+%28Risk+Factors%29',
 'datasets/Quality+Assessment+of+Digital+Colposcopies',
 'datasets/Quality+Assessment+of+Digital+Colposcopies',
 'datasets/KASANDR',
 'datasets/KASANDR',
 'datasets/FMA%3A+A+Dataset+For+Music+Analysis',
 'datasets/FMA%3A+A+Dataset+For+Music+Analysis',
 'datasets/Air+quality',
 'datasets/Air+quality',
 'datasets/Epileptic+Seizure+Recognition',
 'datasets/Epileptic+Seizure+Recognition',
 'datasets/Devanagari+Handwritten+Character+Dataset',
 'datasets/Devanagari+Handwritten+Character+Dataset',
 'datasets/Stock+portfolio+performance',
 'datasets/Stock+portfolio+performance',
 'datasets/MoCap+Hand+Postures',
 'datasets/MoCap+Hand+Postures',
 'datasets/Early+biomarkers+of+Parkinson%92s+disease+based+on+natural+connected+speech',
 'datasets/Early+biomarkers+of+Parkinson%92s+disease+based+on+natural+connected+speech',
 'datasets/Data+for+Software+Engineering+Teamwork+Assessment+in+Education+Setting',
 'datasets/Data+for+Software+Engineering+Teamwork+Assessment+in+Education+Setting',
 'datasets/PM2.5+Data+of+Five+Chinese+Cities',
 'datasets/PM2.5+Data+of+Five+Chinese+Cities',
 'datasets/Parkinson+Disease+Spiral+Drawings+Using+Digitized+Graphics+Tablet',
 'datasets/Parkinson+Disease+Spiral+Drawings+Using+Digitized+Graphics+Tablet',
 'datasets/Sales_Transactions_Dataset_Weekly',
 'datasets/Sales_Transactions_Dataset_Weekly',
 'datasets/Las+Vegas+Strip',
 'datasets/Las+Vegas+Strip',
 'datasets/Eco-hotel',
 'datasets/Eco-hotel',
 'datasets/MEU-Mobile+KSD',
 'datasets/MEU-Mobile+KSD',
 'datasets/Crowdsourced+Mapping',
 'datasets/Crowdsourced+Mapping',
 'datasets/gene+expression+cancer+RNA-Seq',
 'datasets/gene+expression+cancer+RNA-Seq',
 'datasets/Hybrid+Indoor+Positioning+Dataset+from+WiFi+RSSI%2C+Bluetooth+and+magnetometer',
 'datasets/Hybrid+Indoor+Positioning+Dataset+from+WiFi+RSSI%2C+Bluetooth+and+magnetometer',
 'datasets/chestnut+%E2%80%93+LARVIC',
 'datasets/chestnut+%E2%80%93+LARVIC',

'datasets/Burst+Header+Packet+%28BHP%29+flooding+attack+on+Optical+Burst+Switching+%28OBS%29+Networ
```

```
'datasets/Burst+Header+Packet+%28BHP%29+flooding+attack+on+Optical+Burst+Switching+%28OBS%29+Networ

 'datasets/Motion+Capture+Hand+Postures',
 'datasets/Motion+Capture+Hand+Postures',
 'datasets/Anuran+Calls+%28MFCCs%29',
 'datasets/Anuran+Calls+%28MFCCs%29',
 'datasets/TTC-3600%3A+Benchmark+dataset+for+Turkish+text+categorization',
 'datasets/TTC-3600%3A+Benchmark+dataset+for+Turkish+text+categorization',
 'datasets/Gastrointestinal+Lesions+in+Regular+Colonoscopy',
 'datasets/Gastrointestinal+Lesions+in+Regular+Colonoscopy',
 'datasets/Daily+Demand+Forecasting+Orders',
 'datasets/Daily+Demand+Forecasting+Orders',
 'datasets/Paper+Reviews',
 'datasets/Paper+Reviews',
 'datasets/extention+of+Z-Alizadeh+sani+dataset',
 'datasets/extention+of+Z-Alizadeh+sani+dataset',
 'datasets/Z-Alizadeh+Sani',
 'datasets/Z-Alizadeh+Sani',
 'datasets/Dynamic+Features+of+VirusShare+Executables',
 'datasets/Dynamic+Features+of+VirusShare+Executables',
 'datasets/IDA2016Challenge',
 'datasets/IDA2016Challenge',
 'datasets/DSRC+Vehicle+Communications',
 'datasets/DSRC+Vehicle+Communications',
 'datasets/Mturk+User-Perceived+Clusters+over+Images',
 'datasets/Mturk+User-Perceived+Clusters+over+Images',
 'datasets/Character+Font+Images',
 'datasets/Character+Font+Images',
 'datasets/DeliciousMIL%3A+A+Data+Set+for+Multi-Label+Multi-
Instance+Learning+with+Instance+Labels',
 'datasets/DeliciousMIL%3A+A+Data+Set+for+Multi-Label+Multi-
Instance+Learning+with+Instance+Labels',
 'datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Children++',
 'datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Children++',
 'datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Adolescent+++',
 'datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Adolescent+++',
 'datasets/APS+Failure+at+Scania+Trucks',
 'datasets/APS+Failure+at+Scania+Trucks',
 'datasets/Wireless+Indoor+Localization',
 'datasets/Wireless+Indoor+Localization',
 'datasets/HCC+Survival',
 'datasets/HCC+Survival',
 'datasets/CSM+%28Conventional+and+Social+Media+Movies%29+Dataset+2014+and+2015',
 'datasets/CSM+%28Conventional+and+Social+Media+Movies%29+Dataset+2014+and+2015',
 'datasets/University+of+Tehran+Question+Dataset+2016+%28UTQD.2016%29',
 'datasets/University+of+Tehran+Question+Dataset+2016+%28UTQD.2016%29',
 'datasets/Autism+Screening+Adult',
 'datasets/Autism+Screening+Adult',
 'datasets/Activity+recognition+with+healthy+older+people+using+a+batteryless+wearable+sensor',
 'datasets/Activity+recognition+with+healthy+older+people+using+a+batteryless+wearable+sensor',
 'datasets/Immunotherapy+Dataset',
 'datasets/Immunotherapy+Dataset',
 'datasets/Cryotherapy+Dataset+',
 'datasets/Cryotherapy+Dataset+',
 'datasets/OCT+data+%26+Color+Fundus+Images+of+Left+%26+Right+Eyes',
 'datasets/OCT+data+%26+Color+Fundus+Images+of+Left+%26+Right+Eyes',
 'datasets/Discrete+Tone+Image+Dataset',
 'datasets/Discrete+Tone+Image+Dataset',
 'datasets/News+Popularity+in+Multiple+Social+Media+Platforms',
 'datasets/News+Popularity+in+Multiple+Social+Media+Platforms',
 'datasets/Ultrasonic+flowmeter+diagnostics',
 'datasets/Ultrasonic+flowmeter+diagnostics',
 'datasets/ICMLA+2014+Accepted+Papers+Data+Set',
 'datasets/ICMLA+2014+Accepted+Papers+Data+Set',
 'datasets/BLE+RSSI+Dataset+for+Indoor+localization+and+Navigation',
 'datasets/BLE+RSSI+Dataset+for+Indoor+localization+and+Navigation',
 'datasets/Container+Crane+Controller+Data+Set',
 'datasets/Container+Crane+Controller+Data+Set',
 'datasets/Residential+Building+Data+Set',
 'datasets/Residential+Building+Data+Set',
 'datasets/Health+News+in+Twitter',
 'datasets/Health+News+in+Twitter',
 'datasets/chipseq',
 'datasets/chipseq',
 'datasets/SGEMM+GPU+kernel+performance',
```

```
 'datasets/SGEMM+GPU+kernel+performance',
 'datasets/Repeat+Consumption+Matrices',
 'datasets/Repeat+Consumption+Matrices',
 'datasets/detection_of_IoT_botnet_attacks_N_BaIoT',
 'datasets/detection_of_IoT_botnet_attacks_N_BaIoT',
 'datasets/Absenteeism+at+work',
 'datasets/Absenteeism+at+work',
 'datasets/SCADI',
 'datasets/SCADI',
 'datasets/Condition+monitoring+of+hydraulic+systems',
 'datasets/Condition+monitoring+of+hydraulic+systems',
 'datasets/Carbon+Nanotubes',
 'datasets/Carbon+Nanotubes',
 'datasets/Optical+Interconnection+Network+',
 'datasets/Optical+Interconnection+Network+',
 'datasets/Sports+articles+for+objectivity+analysis',
 'datasets/Sports+articles+for+objectivity+analysis',
 'datasets/Breast+Cancer+Coimbra',
 'datasets/Breast+Cancer+Coimbra',
 'datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data',
 'datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data',
 'datasets/Dishonest+Internet+users+Dataset',
 'datasets/Dishonest+Internet+users+Dataset',
 'datasets/Victorian+Era+Authorship+Attribution',
 'datasets/Victorian+Era+Authorship+Attribution',
 'datasets/Simulated+Falls+and+Daily+Living+Activities+Data+Set',
 'datasets/Simulated+Falls+and+Daily+Living+Activities+Data+Set',
 'datasets/Multimodal+Damage+Identification+for+Humanitarian+Computing',
 'datasets/Multimodal+Damage+Identification+for+Humanitarian+Computing',
 'datasets/EEG+Steady-State+Visual+Evoked+Potential+Signals',
 'datasets/EEG+Steady-State+Visual+Evoked+Potential+Signals',
 'datasets/Roman+Urdu+Data+Set',
 'datasets/Roman+Urdu+Data+Set',
 'datasets/Avila',
 'datasets/Avila',
 'datasets/PANDOR',
 'datasets/PANDOR',
 'datasets/Drug+Review+Dataset+%28Druglib.com%29',
 'datasets/Drug+Review+Dataset+%28Druglib.com%29',
 'datasets/Drug+Review+Dataset+%28Drugs.com%29',
 'datasets/Drug+Review+Dataset+%28Drugs.com%29',
 'datasets/Physical+Unclonable+Functions',
 'datasets/Physical+Unclonable+Functions',
 'datasets/Superconductivty+Data',
 'datasets/Superconductivty+Data',
 'datasets/WESAD+%28Wearable+Stress+and+Affect+Detection%29',
 'datasets/WESAD+%28Wearable+Stress+and+Affect+Detection%29',
 'datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data+Set+2',
 'datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data+Set+2',
 'datasets/Student+Academics+Performance',
 'datasets/Student+Academics+Performance',
 'datasets/Online+Shoppers+Purchasing+Intention+Dataset',
 'datasets/Online+Shoppers+Purchasing+Intention+Dataset',
 'datasets/PMU-UD',
 'datasets/PMU-UD',
 'datasets/Parkinson%27s+Disease+Classification',
 'datasets/Parkinson%27s+Disease+Classification',
 'datasets/Electrical+Grid+Stability+Simulated+Data+',
 'datasets/Electrical+Grid+Stability+Simulated+Data+',
 'datasets/Caesarian+Section+Classification+Dataset',
 'datasets/Caesarian+Section+Classification+Dataset',
 'datasets/BAUM-1',
 'datasets/BAUM-1',
 'datasets/BAUM-2',
 'datasets/BAUM-2',
 'datasets/Audit+Data',
 'datasets/Audit+Data',
 'datasets/BuddyMove+Data+Set',
 'datasets/BuddyMove+Data+Set',
 'datasets/Real+estate+valuation+data+set',
 'datasets/Real+estate+valuation+data+set',

'datasets/Early+biomarkers+of+Parkinson%E2%80%99s+disease+based+on+natural+connected+speech+Data+Se


'datasets/Early+biomarkers+of+Parkinson%E2%80%99s+disease+based+on+natural+connected+speech+Data+Se
```

```
 'datasets/Somerville+Happiness+Survey',
 'datasets/Somerville+Happiness+Survey',
 'datasets/2.4+GHZ+Indoor+Channel+Measurements',
 'datasets/2.4+GHZ+Indoor+Channel+Measurements',
 'datasets/EMG+data+for+gestures',
 'datasets/EMG+data+for+gestures',
 'datasets/Parking+Birmingham',
 'datasets/Parking+Birmingham',
 'datasets/Behavior+of+the+urban+traffic+of+the+city+of+Sao+Paulo+in+Brazil',
 'datasets/Behavior+of+the+urban+traffic+of+the+city+of+Sao+Paulo+in+Brazil',
 'datasets/Travel+Reviews',
 'datasets/Travel+Reviews',
 'datasets/Tarvel+Review+Ratings',
 'datasets/Tarvel+Review+Ratings',
 'datasets/Rice+Leaf+Diseases',
 'datasets/Rice+Leaf+Diseases']
```

In [265]:

```
lista_links= (re.findall('"(datasets/.+?)"', str(links)))
```

In [266]:

```
lista_links1 = []
for i in lista_links:
    if i not in lista_links1:
        lista_links1.append(i)
```

In [267]:

```
lista_links1
```

Out[267]:

```
['datasets/Abalone',
 'datasets/Adult',
 'datasets/Annealing',
 'datasets/Anonymous+Microsoft+Web+Data',
 'datasets/Arrhythmia',
 'datasets/Artificial+Characters',
 'datasets/Audiology+%28Original%29',
 'datasets/Audiology+%28Standardized%29',
 'datasets/Auto+MPG',
 'datasets/Automobile',
 'datasets/Badges',
 'datasets/Balance+Scale',
 'datasets/Balloons',
 'datasets/Breast+Cancer',
 'datasets/Breast+Cancer+Wisconsin+%28Original%29',
 'datasets/Breast+Cancer+Wisconsin+%28Prognostic%29',
 'datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29',
 'datasets/Pittsburgh+Bridges',
 'datasets/Car+Evaluation',
 'datasets/Census+Income',
 'datasets/Chess+%28King-Rook+vs.+King-Knight%29',
 'datasets/Chess+%28King-Rook+vs.+King-Pawn%29',
 'datasets/Chess+%28King-Rook+vs.+King%29',
 'datasets/Chess+%28Domain+Theories%29',
 'datasets/Bach+Chorales',
 'datasets/Connect-4',
 'datasets/Credit+Approval',
 'datasets/Japanese+Credit+Screening',
 'datasets/Computer+Hardware',
 'datasets/Contraceptive+Method+Choice',
 'datasets/Covertype',
 'datasets/Cylinder+Bands',
 'datasets/Dermatology',
 'datasets/Diabetes',
 'datasets/DGP2+-+The+Second+Data+Generation+Program',
 'datasets/Document+Understanding',
 'datasets/EBL+Domain+Theories',
 'datasets/Echocardiogram',
 'datasets/Ecoli',
 'datasets/Flags',
```

```
'datasets/Flags',
'datasets/Function+Finding',
'datasets/Glass+Identification',
'datasets/Haberman%27s+Survival',
'datasets/Hayes-Roth',
'datasets/Heart+Disease',
'datasets/Hepatitis',
'datasets/Horse+Colic',
'datasets/ICU',
'datasets/Image+Segmentation',
'datasets/Internet+Advertisements',
'datasets/Ionosphere',
'datasets/Iris',
'datasets/ISOLET',
'datasets/Kinship',
'datasets/Labor+Relations',
'datasets/LED+Display+Domain',
'datasets/Lenses',
'datasets/Letter+Recognition',
'datasets/Liver+Disorders',
'datasets/Logic+Theorist',
'datasets/Lung+Cancer',
'datasets/Lymphography',
'datasets/Mechanical+Analysis',
'datasets/Meta-data',
'datasets/Mobile+Robots',
'datasets/Molecular+Biology+%28Promoter+Gene+Sequences%29',
'datasets/Molecular+Biology+%28Protein+Secondary+Structure%29',
'datasets/Molecular+Biology+%28Splice-junction+Gene+Sequences%29',
'datasets/MONK%27s+Problems',
'datasets/Moral+Reasoner',
'datasets/Multiple+Features',
'datasets/Mushroom',
'datasets/Musk+%28Version+1%29',
'datasets/Musk+%28Version+2%29',
'datasets/Nursery',
'datasets/Othello+Domain+Theory',
'datasets/Page+Blocks+Classification',
'datasets/Optical+Recognition+of+Handwritten+Digits',
'datasets/Pen-Based+Recognition+of+Handwritten+Digits',
'datasets/Post-Operative+Patient',
'datasets/Primary+Tumor',
'datasets/Prodigy',
'datasets/Qualitative+Structure+Activity+Relationships',
'datasets/Quadruped+Mammals',
'datasets/Servo',
'datasets/Shuttle+Landing+Control',
'datasets/Solar+Flare',
'datasets/Soybean+%28Large%29',
'datasets/Soybean+%28Small%29',
'datasets/Challenger+USA+Space+Shuttle+O-Ring',
'datasets/Low+Resolution+Spectrometer',
'datasets/Spambase',
'datasets/SPECT+Heart',
'datasets/SPECTF+Heart',
'datasets/Sponge',
'datasets/Statlog+Project',
'datasets/Student+Loan+Relational',
'datasets/Teaching+Assistant+Evaluation',
'datasets/Tic-Tac-Toe+Endgame',
'datasets/Thyroid+Disease',
'datasets/Trains',
'datasets/University',
'datasets/Congressional+Voting+Records',
'datasets/Water+Treatment+Plant',
'datasets/Waveform+Database+Generator+%28Version+1%29',
'datasets/Waveform+Database+Generator+%28Version+2%29',
'datasets/Wine',
'datasets/Yeast',
'datasets/Zoo',
'datasets/Undocumented',
'datasets/Twenty+Newsgroups',
'datasets/Australian+Sign+Language+signs',
'datasets/Australian+Sign+Language+signs+%28High+Quality%29',
'datasets/US+Census+Data+%281990%29',
'datasets/Census-Income+%28KDD%29',
'datasets/Coil+1999+Competition+Data',
```

```
'datasets/Corel+Image+Features',
'datasets/E.+Coli+Genes',
'datasets/EEG+Database',
'datasets/El+Nino',
'datasets/Entree+Chicago+Recommendation+Data',
'datasets/CMU+Face+Images',
'datasets/Insurance+Company+Benchmark+%28COIL+2000%29',
'datasets/Internet+Usage+Data',
'datasets/IPUMS+Census+Database',
'datasets/Japanese+Vowels',
'datasets/KDD+Cup+1998+Data',
'datasets/KDD+Cup+1999+Data',
'datasets/M.+Tuberculosis+Genes',
'datasets/Movie',
'datasets/MSNBC.com+Anonymous+Web+Data',
'datasets/NSF+Research+Award+Abstracts+1990-2003',
'datasets/Pioneer-1+Mobile+Robot+Data',
'datasets/Pseudo+Periodic+Synthetic+Time+Series',
'datasets/Reuters-21578+Text+Categorization+Collection',
'datasets/Robot+Execution+Failures',
'datasets/Synthetic+Control+Chart+Time+Series',
'datasets/Syskill+and+Webert+Web+Page+Ratings',
'datasets/UNIX+User+Data',
'datasets/Volcanoes+on+Venus+-+JARtool+experiment',
'datasets/Statlog+%28Australian+Credit+Approval%29',
'datasets/Statlog+%28German+Credit+Data%29',
'datasets/Statlog+%28Heart%29',
'datasets/Statlog+%28Landsat+Satellite%29',
'datasets/Statlog+%28Image+Segmentation%29',
'datasets/Statlog+%28Shuttle%29',
'datasets/Statlog+%28Vehicle+Silhouettes%29',
'datasets/Connectionist+Bench+%28Nettalk+Corpus%29',
'datasets/Connectionist+Bench+%28Sonar%2C+Mines+vs.+Rocks%29',
'datasets/Connectionist+Bench+%28Vowel+Recognition+-+Deterding+Data%29',
'datasets/Economic+Sanctions',
'datasets/Protein+Data',
'datasets/Cloud',
'datasets/CalIt2+Building+People+Counts',
'datasets/Dodgers+Loop+Sensor',
'datasets/Poker+Hand',
'datasets/MAGIC+Gamma+Telescope',
'datasets/UJI+Pen+Characters',
'datasets/Mammographic+Mass',
'datasets/Forest+Fires',
'datasets/Reuters+Transcribed+Subset',
'datasets/Bag+of+Words',
'datasets/Concrete+Compressive+Strength',
'datasets/Hill-Valley',
'datasets/Arcene',
'datasets/Dexter',
'datasets/Dorothea',
'datasets/Gisette',
'datasets/Madelon',
'datasets/Ozone+Level+Detection',
'datasets/Abscisic+Acid+Signaling+Network',
'datasets/Parkinsons',
'datasets/Character+Trajectories',
'datasets/Blood+Transfusion+Service+Center',
'datasets/UJI+Pen+Characters+%28Version+2%29',
'datasets/Semeion+Handwritten+Digit',
'datasets/SECOM',
'datasets/Plants',
'datasets/Libras+Movement',
'datasets/Concrete+Slump+Test',
'datasets/Communities+and+Crime',
'datasets/Acute+Inflammations',
'datasets/Wine+Quality',
'datasets/URL+Reputation',
'datasets/p53+Mutants',
'datasets/Parkinsons+Telemonitoring',
'datasets/Demospongiae',
'datasets/Opinosis+Opinion+%26frasl%3B+Review',
'datasets/Breast+Tissue',
'datasets/Cardiotocography',
'datasets/Wall-Following+Robot+Navigation+Data',
'datasets/Spoken+Arabic+Digit',
'datasets/Localization+Data+for+Person+Activity',
```

```
 'datasets/AutoUniv',
 'datasets/Steel+Plates+Faults',
 'datasets/MiniBooNE+particle+identification',
 'datasets/YearPredictionMSD',
 'datasets/PEMS-SF',
 'datasets/OpinRank+Review+Dataset',
 'datasets/Relative+location+of+CT+slices+on+axial+axis',
 'datasets/Online+Handwritten+Assamese+Characters+Dataset',
 'datasets/PubChem+Bioassay+Data',
 'datasets/Record+Linkage+Comparison+Patterns',
 'datasets/Communities+and+Crime+Unnormalized',
 'datasets/Vertebral+Column',
 'datasets/EMG+Physical+Action+Data+Set',
 'datasets/Vicon+Physical+Action+Data+Set',
 'datasets/Amazon+Commerce+reviews+set',
 'datasets/Amazon+Access+Samples',
 'datasets/Reuter_50_50',
 'datasets/Farm+Ads',
 'datasets/DBWorld+e-mails',
 'datasets/KEGG+Metabolic+Relation+Network+%28Directed%29',
 'datasets/KEGG+Metabolic+Reaction+Network+%28Undirected%29',
 'datasets/Bank+Marketing',
 'datasets/YouTube+Comedy+Slam+Preference+Data',
 'datasets/Gas+Sensor+Array+Drift+Dataset',
 'datasets/ILPD+%28Indian+Liver+Patient+Dataset%29',
 'datasets/OPPORTUNITY+Activity+Recognition',
 'datasets/Nomao',
 'datasets/SMS+Spam+Collection',
 'datasets/Skin+Segmentation',
 'datasets/Planning+Relax',
 'datasets/PAMAP2+Physical+Activity+Monitoring',
 'datasets/Restaurant+%26+consumer+data',
 'datasets/CNAE-9',
 'datasets/Individual+household+electric+power+consumption',
 'datasets/seeds',
 'datasets/Northix',
 'datasets/QtyT40I10D100K',
 'datasets/Legal+Case+Reports',
 'datasets/Human+Activity+Recognition+Using+Smartphones',
 'datasets/One-hundred+plant+species+leaves+data+set',
 'datasets/Energy+efficiency',
 'datasets/Yacht+Hydrodynamics',
 'datasets/Fertility',
 'datasets/Daphnet+Freezing+of+Gait',
 'datasets/3D+Road+Network+%28North+Jutland%2C+Denmark%29',
 'datasets/ISTANBUL+STOCK+EXCHANGE',
 'datasets/Buzz+in+social+media+',
 'datasets/First-order+theorem+proving',
 'datasets/Wearable+Computing%3A+Classification+of+Body+Postures+and+Movements+%28PUC-Rio%29',
 'datasets/Gas+sensor+arrays+in+open+sampling+settings',
 'datasets/Climate+Model+Simulation+Crashes',
 'datasets/MicroMass',
 'datasets/QSAR+biodegradation',
 'datasets/BLOGGER',
 'datasets/Daily+and+Sports+Activities',
 'datasets/User+Knowledge+Modeling',
 'datasets/Reuters+RCV1+RCV2+Multilingual%2C+Multiview+Text+Categorization+Test+collection',
 'datasets/NYSK',
 'datasets/Turkiye+Student+Evaluation',

'datasets/ser+Knowledge+Modeling+Data+%28Students%27+Knowledge+Levels+on+DC+Electrical+Machines%29'

 'datasets/EEG+Eye+State',
 'datasets/Physicochemical+Properties+of+Protein+Tertiary+Structure',
 'datasets/seismic-bumps',
 'datasets/banknote+authentication',
 'datasets/USPTO+Algorithm+Challenge%2C+run+by+NASA-
Harvard+Tournament+Lab+and+TopCoder++++Problem%3A+Pat',
 'datasets/YouTube+Multiview+Video+Games+Dataset',
 'datasets/Gas+Sensor+Array+Drift+Dataset+at+Different+Concentrations',
 'datasets/Activities+of+Daily+Living+%28ADLs%29+Recognition+Using+Binary+Sensors',
 'datasets/SkillCraft1+Master+Table+Dataset',
 'datasets/Weight+Lifting+Exercises+monitored+with+Inertial+Measurement+Units',
 'datasets/SML2010',
 'datasets/Bike+Sharing+Dataset',
 'datasets/Predict+keywords+activities+in+a+online+social+media',
 'datasets/Thoracic+Surgery+Data',
```

```
'datasets/EMG+dataset+in+Lower+Limb',
'datasets/SUSY',
'datasets/HIGGS',
'datasets/Qualitative_Bankruptcy',
'datasets/LSVT+Voice+Rehabilitation',
'datasets/Dataset+for+ADL+Recognition+with+Wrist-worn+Accelerometer',
'datasets/Wilt',
'datasets/User+Identification+From+Walking+Activity',
'datasets/Activity+Recognition+from+Single+Chest-Mounted+Accelerometer',
'datasets/Leaf',
'datasets/Dresses_Attribute_Sales',
'datasets/Tamilnadu+Electricity+Board+Hourly+Readings',
'datasets/Airfoil+Self-Noise',
'datasets/Wholesale+customers',
'datasets/Twitter+Data+set+for+Arabic+Sentiment+Analysis',
'datasets/Combined+Cycle+Power+Plant',
'datasets/Urban+Land+Cover',
'datasets/Diabetes+130-US+hospitals+for+years+1999-2008',
'datasets/Bach+Choral+Harmony',
'datasets/StoneFlakes',
'datasets/Tennis+Major+Tournament+Match+Statistics',
'datasets/Parkinson+Speech+Dataset+with++Multiple+Types+of+Sound+Recordings',
'datasets/Gesture+Phase+Segmentation',
'datasets/Perfume+Data',
'datasets/BlogFeedback',
'datasets/REALDISP+Activity+Recognition+Dataset',
'datasets/Newspaper+and+magazine+images+segmentation+dataset',
'datasets/AAAI+2014+Accepted+Papers',
'datasets/Gas+sensor+array+under+flow+modulation',
'datasets/Gas+sensor+array+exposed+to+turbulent+gas+mixtures',
'datasets/UJIIndoorLoc',
'datasets/Sentence+Classification',
'datasets/Dow+Jones+Index',
'datasets/sEMG+for+Basic+Hand+movements',
'datasets/AAAI+2013+Accepted+Papers',
'datasets/Geographical+Original+of+Music',
'datasets/Condition+Based+Maintenance+of+Naval+Propulsion+Plants',
'datasets/Grammatical+Facial+Expressions',
'datasets/NoisyOffice',
'datasets/MHEALTH+Dataset',
'datasets/Student+Performance',
'datasets/ElectricityLoadDiagrams20112014',
'datasets/Gas+sensor+array+under+dynamic+gas+mixtures',
'datasets/microblogPCU',
'datasets/Firm-Teacher_Clave-Direction_Classification',
'datasets/Dataset+for+Sensorless+Drive+Diagnosis',
'datasets/TV+News+Channel+Commercial+Detection+Dataset',
'datasets/Phishing+Websites',
'datasets/Greenhouse+Gas+Observing+Network',
'datasets/Diabetic+Retinopathy+Debrecen+Data+Set',
'datasets/HIV-1+protease+cleavage',
'datasets/Sentiment+Labelled+Sentences',
'datasets/Online+News+Popularity',
'datasets/Forest+type+mapping',
'datasets/wiki4HE',
'datasets/Online+Video+Characteristics+and+Transcoding+Time+Dataset',
'datasets/Chronic_Kidney_Disease',
'datasets/Machine+Learning+based+ZZAlpha+Ltd.+Stock+Recommendations+2012-2014',
'datasets/Folio',
'datasets/Taxi+Service+Trajectory+-+Prediction+Challenge%2C+ECML+PKDD+2015',
'datasets/Cuff-Less+Blood+Pressure+Estimation',
'datasets/Smartphone-Based+Recognition+of+Human+Activities+and+Postural+Transitions',
'datasets/Mice+Protein+Expression',
'datasets/UJIIndoorLoc-Mag',
'datasets/Heterogeneity+Activity+Recognition',
'datasets/Educational+Process+Mining+%28EPM%29%3A+A+Learning+Analytics+Data+Set',
'datasets/HEPMASS',
'datasets/Indoor+User+Movement+Prediction+from+RSS+data',
'datasets/Open+University+Learning+Analytics+dataset',
'datasets/default+of+credit+card+clients',
'datasets/Mesothelioma%E2%80%99s+disease+data+set+',
'datasets/Online+Retail',
'datasets/SIFT10M',
'datasets/GPS+Trajectories',
'datasets/Detect+Malacious+Executable%28AntiVirus%29',
'datasets/Occupancy+Detection+',
```

```
'datasets/Improved+Spiral+Test+Using+Digitized+Graphics+Tablet+for+Monitoring+Parkinson%E2%80%99s+D
se',
 'datasets/News+Aggregator',
 'datasets/Air+Quality',
 'datasets/Twin+gas+sensor+arrays',
 'datasets/Gas+sensors+for+home+activity+monitoring',
 'datasets/Facebook+Comment+Volume+Dataset',

'datasets/Smartphone+Dataset+for+Human+Activity+Recognition+%28HAR%29+in+Ambient+Assisted+Living+%2
%29',
 'datasets/Polish+companies+bankruptcy+data',
 'datasets/Activity+Recognition+system+based+on+Multisensor+data+fusion+%28AReM%29',
 'datasets/Dota2+Games+Results',
 'datasets/Facebook+metrics',
 'datasets/UbiqLog+%28smartphone+lifelogging%29',
 'datasets/NIPS+Conference+Papers+1987-2015',
 'datasets/HTRU2',
 'datasets/Drug+consumption+%28quantified%29',
 'datasets/Appliances+energy+prediction',
 'datasets/Miskolc+IIS+Hybrid+IPS',
 'datasets/KDC-4007+dataset+Collection',
 'datasets/Geo-
Magnetic+field+and+WLAN+dataset+for+indoor+localisation+from+wristband+and+smartphone',
 'datasets/DrivFace',
 'datasets/Website+Phishing',
 'datasets/YouTube+Spam+Collection',
 'datasets/Beijing+PM2.5+Data',
 'datasets/Cargo+2000+Freight+Tracking+and+Tracing',
 'datasets/Cervical+cancer+%28Risk+Factors%29',
 'datasets/Quality+Assessment+of+Digital+Colposcopies',
 'datasets/KASANDR',
 'datasets/FMA%3A+A+Dataset+For+Music+Analysis',
 'datasets/Air+quality',
 'datasets/Epileptic+Seizure+Recognition',
 'datasets/Devanagari+Handwritten+Character+Dataset',
 'datasets/Stock+portfolio+performance',
 'datasets/MoCap+Hand+Postures',
 'datasets/Early+biomarkers+of+Parkinson%92s+disease+based+on+natural+connected+speech',
 'datasets/Data+for+Software+Engineering+Teamwork+Assessment+in+Education+Setting',
 'datasets/PM2.5+Data+of+Five+Chinese+Cities',
 'datasets/Parkinson+Disease+Spiral+Drawings+Using+Digitized+Graphics+Tablet',
 'datasets/Sales_Transactions_Dataset_Weekly',
 'datasets/Las+Vegas+Strip',
 'datasets/Eco-hotel',
 'datasets/MEU-Mobile+KSD',
 'datasets/Crowdsourced+Mapping',
 'datasets/gene+expression+cancer+RNA-Seq',
 'datasets/Hybrid+Indoor+Positioning+Dataset+from+WiFi+RSSI%2C+Bluetooth+and+magnetometer',
 'datasets/chestnut+%E2%80%93+LARVIC',

'datasets/Burst+Header+Packet+%28BHP%29+flooding+attack+on+Optical+Burst+Switching+%28OBS%29+Networ

 'datasets/Motion+Capture+Hand+Postures',
 'datasets/Anuran+Calls+%28MFCCs%29',
 'datasets/TTC-3600%3A+Benchmark+dataset+for+Turkish+text+categorization',
 'datasets/Gastrointestinal+Lesions+in+Regular+Colonoscopy',
 'datasets/Daily+Demand+Forecasting+Orders',
 'datasets/Paper+Reviews',
 'datasets/extention+of+Z-Alizadeh+sani+dataset',
 'datasets/Z-Alizadeh+Sani',
 'datasets/Dynamic+Features+of+VirusShare+Executables',
 'datasets/IDA2016Challenge',
 'datasets/DSRC+Vehicle+Communications',
 'datasets/Mturk+User-Perceived+Clusters+over+Images',
 'datasets/Character+Font+Images',
 'datasets/DeliciousMIL%3A+A+Data+Set+for+Multi-Label+Multi-
Instance+Learning+with+Instance+Labels',
 'datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Children++',
 'datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Adolescent+++',
 'datasets/APS+Failure+at+Scania+Trucks',
 'datasets/Wireless+Indoor+Localization',
 'datasets/HCC+Survival',
 'datasets/CSM+%28Conventional+and+Social+Media+Movies%29+Dataset+2014+and+2015',
 'datasets/University+of+Tehran+Question+Dataset+2016+%28UTQD.2016%29',
 'datasets/Autism+Screening+Adult',
 'datasets/Activity+recognition+with+healthy+older+people+using+a+batteryless+wearable+sensor',
 'datasets/Immunotherapy+Dataset',
```

```
 'datasets/Cryotherapy+Dataset+',
 'datasets/OCT+data+%26+Color+Fundus+Images+of+Left+%26+Right+Eyes',
 'datasets/Discrete+Tone+Image+Dataset',
 'datasets/News+Popularity+in+Multiple+Social+Media+Platforms',
 'datasets/Ultrasonic+flowmeter+diagnostics',
 'datasets/ICMLA+2014+Accepted+Papers+Data+Set',
 'datasets/BLE+RSSI+Dataset+for+Indoor+localization+and+Navigation',
 'datasets/Container+Crane+Controller+Data+Set',
 'datasets/Residential+Building+Data+Set',
 'datasets/Health+News+in+Twitter',
 'datasets/chipseq',
 'datasets/SGEMM+GPU+kernel+performance',
 'datasets/Repeat+Consumption+Matrices',
 'datasets/detection_of_IoT_botnet_attacks_N_BaIoT',
 'datasets/Absenteeism+at+work',
 'datasets/SCADI',
 'datasets/Condition+monitoring+of+hydraulic+systems',
 'datasets/Carbon+Nanotubes',
 'datasets/Optical+Interconnection+Network+',
 'datasets/Sports+articles+for+objectivity+analysis',
 'datasets/Breast+Cancer+Coimbra',
 'datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data',
 'datasets/Dishonest+Internet+users+Dataset',
 'datasets/Victorian+Era+Authorship+Attribution',
 'datasets/Simulated+Falls+and+Daily+Living+Activities+Data+Set',
 'datasets/Multimodal+Damage+Identification+for+Humanitarian+Computing',
 'datasets/EEG+Steady-State+Visual+Evoked+Potential+Signals',
 'datasets/Roman+Urdu+Data+Set',
 'datasets/Avila',
 'datasets/PANDOR',
 'datasets/Drug+Review+Dataset+%28Druglib.com%29',
 'datasets/Drug+Review+Dataset+%28Drugs.com%29',
 'datasets/Physical+Unclonable+Functions',
 'datasets/Superconductivty+Data',
 'datasets/WESAD+%28Wearable+Stress+and+Affect+Detection%29',
 'datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data+Set+2',
 'datasets/Student+Academics+Performance',
 'datasets/Online+Shoppers+Purchasing+Intention+Dataset',
 'datasets/PMU-UD',
 'datasets/Parkinson%27s+Disease+Classification',
 'datasets/Electrical+Grid+Stability+Simulated+Data+',
 'datasets/Caesarian+Section+Classification+Dataset',
 'datasets/BAUM-1',
 'datasets/BAUM-2',
 'datasets/Audit+Data',
 'datasets/BuddyMove+Data+Set',
 'datasets/Real+estate+valuation+data+set',

'datasets/Early+biomarkers+of+Parkinson%E2%80%99s+disease+based+on+natural+connected+speech+Data+Se

 'datasets/Somerville+Happiness+Survey',
 'datasets/2.4+GHZ+Indoor+Channel+Measurements',
 'datasets/EMG+data+for+gestures',
 'datasets/Parking+Birmingham',
 'datasets/Behavior+of+the+urban+traffic+of+the+city+of+Sao+Paulo+in+Brazil',
 'datasets/Travel+Reviews',
 'datasets/Tarvel+Review+Ratings',
 'datasets/Rice+Leaf+Diseases']
```

In [268]:

```
links1=[]
for x in lista_links1:
    links1.append ("http://archive.ics.uci.edu/ml/" + str(x))
```

In [269]:

```
links1
```

Out[269]:

```
['http://archive.ics.uci.edu/ml/datasets/Abalone',
 'http://archive.ics.uci.edu/ml/datasets/Adult',
 'http://archive.ics.uci.edu/ml/datasets/Annealing',
 'http://archive.ics.uci.edu/ml/datasets/Anonymous+Microsoft+Web+Data',
```

```
'http://archive.ics.uci.edu/ml/datasets/Arrhythmia',
'http://archive.ics.uci.edu/ml/datasets/Artificial+Characters',
'http://archive.ics.uci.edu/ml/datasets/Audiology+%28Original%29',
'http://archive.ics.uci.edu/ml/datasets/Audiology+%28Standardized%29',
'http://archive.ics.uci.edu/ml/datasets/Auto+MPG',
'http://archive.ics.uci.edu/ml/datasets/Automobile',
'http://archive.ics.uci.edu/ml/datasets/Badges',
'http://archive.ics.uci.edu/ml/datasets/Balance+Scale',
'http://archive.ics.uci.edu/ml/datasets/Balloons',
'http://archive.ics.uci.edu/ml/datasets/Breast+Cancer',
'http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Original%29',
'http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Prognostic%29',
'http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29',
'http://archive.ics.uci.edu/ml/datasets/Pittsburgh+Bridges',
'http://archive.ics.uci.edu/ml/datasets/Car+Evaluation',
'http://archive.ics.uci.edu/ml/datasets/Census+Income',
'http://archive.ics.uci.edu/ml/datasets/Chess+%28King-Rook+vs.+King-Knight%29',
'http://archive.ics.uci.edu/ml/datasets/Chess+%28King-Rook+vs.+King-Pawn%29',
'http://archive.ics.uci.edu/ml/datasets/Chess+%28King-Rook+vs.+King%29',
'http://archive.ics.uci.edu/ml/datasets/Chess+%28Domain+Theories%29',
'http://archive.ics.uci.edu/ml/datasets/Bach+Chorales',
'http://archive.ics.uci.edu/ml/datasets/Connect-4',
'http://archive.ics.uci.edu/ml/datasets/Credit+Approval',
'http://archive.ics.uci.edu/ml/datasets/Japanese+Credit+Screening',
'http://archive.ics.uci.edu/ml/datasets/Computer+Hardware',
'http://archive.ics.uci.edu/ml/datasets/Contraceptive+Method+Choice',
'http://archive.ics.uci.edu/ml/datasets/Covertype',
'http://archive.ics.uci.edu/ml/datasets/Cylinder+Bands',
'http://archive.ics.uci.edu/ml/datasets/Dermatology',
'http://archive.ics.uci.edu/ml/datasets/Diabetes',
'http://archive.ics.uci.edu/ml/datasets/DGP2+-+The+Second+Data+Generation+Program',
'http://archive.ics.uci.edu/ml/datasets/Document+Understanding',
'http://archive.ics.uci.edu/ml/datasets/EBL+Domain+Theories',
'http://archive.ics.uci.edu/ml/datasets/Echocardiogram',
'http://archive.ics.uci.edu/ml/datasets/Ecoli',
'http://archive.ics.uci.edu/ml/datasets/Flags',
'http://archive.ics.uci.edu/ml/datasets/Function+Finding',
'http://archive.ics.uci.edu/ml/datasets/Glass+Identification',
'http://archive.ics.uci.edu/ml/datasets/Haberman%27s+Survival',
'http://archive.ics.uci.edu/ml/datasets/Hayes-Roth',
'http://archive.ics.uci.edu/ml/datasets/Heart+Disease',
'http://archive.ics.uci.edu/ml/datasets/Hepatitis',
'http://archive.ics.uci.edu/ml/datasets/Horse+Colic',
'http://archive.ics.uci.edu/ml/datasets/ICU',
'http://archive.ics.uci.edu/ml/datasets/Image+Segmentation',
'http://archive.ics.uci.edu/ml/datasets/Internet+Advertisements',
'http://archive.ics.uci.edu/ml/datasets/Ionosphere',
'http://archive.ics.uci.edu/ml/datasets/Iris',
'http://archive.ics.uci.edu/ml/datasets/ISOLET',
'http://archive.ics.uci.edu/ml/datasets/Kinship',
'http://archive.ics.uci.edu/ml/datasets/Labor+Relations',
'http://archive.ics.uci.edu/ml/datasets/LED+Display+Domain',
'http://archive.ics.uci.edu/ml/datasets/Lenses',
'http://archive.ics.uci.edu/ml/datasets/Letter+Recognition',
'http://archive.ics.uci.edu/ml/datasets/Liver+Disorders',
'http://archive.ics.uci.edu/ml/datasets/Logic+Theorist',
'http://archive.ics.uci.edu/ml/datasets/Lung+Cancer',
'http://archive.ics.uci.edu/ml/datasets/Lymphography',
'http://archive.ics.uci.edu/ml/datasets/Mechanical+Analysis',
'http://archive.ics.uci.edu/ml/datasets/Meta-data',
'http://archive.ics.uci.edu/ml/datasets/Mobile+Robots',
'http://archive.ics.uci.edu/ml/datasets/Molecular+Biology+%28Promoter+Gene+Sequences%29',
'http://archive.ics.uci.edu/ml/datasets/Molecular+Biology+%28Protein+Secondary+Structure%29',
'http://archive.ics.uci.edu/ml/datasets/Molecular+Biology+%28Splice-junction+Gene+Sequences%29',
'http://archive.ics.uci.edu/ml/datasets/MONK%27s+Problems',
'http://archive.ics.uci.edu/ml/datasets/Moral+Reasoner',
'http://archive.ics.uci.edu/ml/datasets/Multiple+Features',
'http://archive.ics.uci.edu/ml/datasets/Mushroom',
'http://archive.ics.uci.edu/ml/datasets/Musk+%28Version+1%29',
'http://archive.ics.uci.edu/ml/datasets/Musk+%28Version+2%29',
'http://archive.ics.uci.edu/ml/datasets/Nursery',
'http://archive.ics.uci.edu/ml/datasets/Othello+Domain+Theory',
'http://archive.ics.uci.edu/ml/datasets/Page+Blocks+Classification',
'http://archive.ics.uci.edu/ml/datasets/Optical+Recognition+of+Handwritten+Digits',
'http://archive.ics.uci.edu/ml/datasets/Pen-Based+Recognition+of+Handwritten+Digits',
'http://archive.ics.uci.edu/ml/datasets/Post-Operative+Patient',
'http://archive.ics.uci.edu/ml/datasets/Primary+Tumor',
```

```
 'http://archive.ics.uci.edu/ml/datasets/Prodigy',
 'http://archive.ics.uci.edu/ml/datasets/Qualitative+Structure+Activity+Relationships',
 'http://archive.ics.uci.edu/ml/datasets/Quadruped+Mammals',
 'http://archive.ics.uci.edu/ml/datasets/Servo',
 'http://archive.ics.uci.edu/ml/datasets/Shuttle+Landing+Control',
 'http://archive.ics.uci.edu/ml/datasets/Solar+Flare',
 'http://archive.ics.uci.edu/ml/datasets/Soybean+%28Large%29',
 'http://archive.ics.uci.edu/ml/datasets/Soybean+%28Small%29',
 'http://archive.ics.uci.edu/ml/datasets/Challenger+USA+Space+Shuttle+O-Ring',
 'http://archive.ics.uci.edu/ml/datasets/Low+Resolution+Spectrometer',
 'http://archive.ics.uci.edu/ml/datasets/Spambase',
 'http://archive.ics.uci.edu/ml/datasets/SPECT+Heart',
 'http://archive.ics.uci.edu/ml/datasets/SPECTF+Heart',
 'http://archive.ics.uci.edu/ml/datasets/Sponge',
 'http://archive.ics.uci.edu/ml/datasets/Statlog+Project',
 'http://archive.ics.uci.edu/ml/datasets/Student+Loan+Relational',
 'http://archive.ics.uci.edu/ml/datasets/Teaching+Assistant+Evaluation',
 'http://archive.ics.uci.edu/ml/datasets/Tic-Tac-Toe+Endgame',
 'http://archive.ics.uci.edu/ml/datasets/Thyroid+Disease',
 'http://archive.ics.uci.edu/ml/datasets/Trains',
 'http://archive.ics.uci.edu/ml/datasets/University',
 'http://archive.ics.uci.edu/ml/datasets/Congressional+Voting+Records',
 'http://archive.ics.uci.edu/ml/datasets/Water+Treatment+Plant',
 'http://archive.ics.uci.edu/ml/datasets/Waveform+Database+Generator+%28Version+1%29',
 'http://archive.ics.uci.edu/ml/datasets/Waveform+Database+Generator+%28Version+2%29',
 'http://archive.ics.uci.edu/ml/datasets/Wine',
 'http://archive.ics.uci.edu/ml/datasets/Yeast',
 'http://archive.ics.uci.edu/ml/datasets/Zoo',
 'http://archive.ics.uci.edu/ml/datasets/Undocumented',
 'http://archive.ics.uci.edu/ml/datasets/Twenty+Newsgroups',
 'http://archive.ics.uci.edu/ml/datasets/Australian+Sign+Language+signs',
 'http://archive.ics.uci.edu/ml/datasets/Australian+Sign+Language+signs+%28High+Quality%29',
 'http://archive.ics.uci.edu/ml/datasets/US+Census+Data+%281990%29',
 'http://archive.ics.uci.edu/ml/datasets/Census-Income+%28KDD%29',
 'http://archive.ics.uci.edu/ml/datasets/Coil+1999+Competition+Data',
 'http://archive.ics.uci.edu/ml/datasets/Corel+Image+Features',
 'http://archive.ics.uci.edu/ml/datasets/E.+Coli+Genes',
 'http://archive.ics.uci.edu/ml/datasets/EEG+Database',
 'http://archive.ics.uci.edu/ml/datasets/El+Nino',
 'http://archive.ics.uci.edu/ml/datasets/Entree+Chicago+Recommendation+Data',
 'http://archive.ics.uci.edu/ml/datasets/CMU+Face+Images',
 'http://archive.ics.uci.edu/ml/datasets/Insurance+Company+Benchmark+%28COIL+2000%29',
 'http://archive.ics.uci.edu/ml/datasets/Internet+Usage+Data',
 'http://archive.ics.uci.edu/ml/datasets/IPUMS+Census+Database',
 'http://archive.ics.uci.edu/ml/datasets/Japanese+Vowels',
 'http://archive.ics.uci.edu/ml/datasets/KDD+Cup+1998+Data',
 'http://archive.ics.uci.edu/ml/datasets/KDD+Cup+1999+Data',
 'http://archive.ics.uci.edu/ml/datasets/M.+Tuberculosis+Genes',
 'http://archive.ics.uci.edu/ml/datasets/Movie',
 'http://archive.ics.uci.edu/ml/datasets/MSNBC.com+Anonymous+Web+Data',
 'http://archive.ics.uci.edu/ml/datasets/NSF+Research+Award+Abstracts+1990-2003',
 'http://archive.ics.uci.edu/ml/datasets/Pioneer-1+Mobile+Robot+Data',
 'http://archive.ics.uci.edu/ml/datasets/Pseudo+Periodic+Synthetic+Time+Series',
 'http://archive.ics.uci.edu/ml/datasets/Reuters-21578+Text+Categorization+Collection',
 'http://archive.ics.uci.edu/ml/datasets/Robot+Execution+Failures',
 'http://archive.ics.uci.edu/ml/datasets/Synthetic+Control+Chart+Time+Series',
 'http://archive.ics.uci.edu/ml/datasets/Syskill+and+Webert+Web+Page+Ratings',
 'http://archive.ics.uci.edu/ml/datasets/UNIX+User+Data',
 'http://archive.ics.uci.edu/ml/datasets/Volcanoes+on+Venus+-+JARtool+experiment',
 'http://archive.ics.uci.edu/ml/datasets/Statlog+%28Australian+Credit+Approval%29',
 'http://archive.ics.uci.edu/ml/datasets/Statlog+%28German+Credit+Data%29',
 'http://archive.ics.uci.edu/ml/datasets/Statlog+%28Heart%29',
 'http://archive.ics.uci.edu/ml/datasets/Statlog+%28Landsat+Satellite%29',
 'http://archive.ics.uci.edu/ml/datasets/Statlog+%28Image+Segmentation%29',
 'http://archive.ics.uci.edu/ml/datasets/Statlog+%28Shuttle%29',
 'http://archive.ics.uci.edu/ml/datasets/Statlog+%28Vehicle+Silhouettes%29',
 'http://archive.ics.uci.edu/ml/datasets/Connectionist+Bench+%28Nettalk+Corpus%29',
 'http://archive.ics.uci.edu/ml/datasets/Connectionist+Bench+%28Sonar%2C+Mines+vs.+Rocks%29',
 'http://archive.ics.uci.edu/ml/datasets/Connectionist+Bench+%28Vowel+Recognition+-
+Deterding+Data%29',
 'http://archive.ics.uci.edu/ml/datasets/Economic+Sanctions',
 'http://archive.ics.uci.edu/ml/datasets/Protein+Data',
 'http://archive.ics.uci.edu/ml/datasets/Cloud',
 'http://archive.ics.uci.edu/ml/datasets/CalIt2+Building+People+Counts',
 'http://archive.ics.uci.edu/ml/datasets/Dodgers+Loop+Sensor',
 'http://archive.ics.uci.edu/ml/datasets/Poker+Hand',
 'http://archive.ics.uci.edu/ml/datasets/MAGIC+Gamma+Telescope',
```

'http://archive.ics.uci.edu/ml/datasets/MAGIC+Gamma+Telescope',
'http://archive.ics.uci.edu/ml/datasets/UJI+Pen+Characters',
'http://archive.ics.uci.edu/ml/datasets/Mammographic+Mass',
'http://archive.ics.uci.edu/ml/datasets/Forest+Fires',
'http://archive.ics.uci.edu/ml/datasets/Reuters+Transcribed+Subset',
'http://archive.ics.uci.edu/ml/datasets/Bag+of+Words',
'http://archive.ics.uci.edu/ml/datasets/Concrete+Compressive+Strength',
'http://archive.ics.uci.edu/ml/datasets/Hill-Valley',
'http://archive.ics.uci.edu/ml/datasets/Arcene',
'http://archive.ics.uci.edu/ml/datasets/Dexter',
'http://archive.ics.uci.edu/ml/datasets/Dorothea',
'http://archive.ics.uci.edu/ml/datasets/Gisette',
'http://archive.ics.uci.edu/ml/datasets/Madelon',
'http://archive.ics.uci.edu/ml/datasets/Ozone+Level+Detection',
'http://archive.ics.uci.edu/ml/datasets/Abscisic+Acid+Signaling+Network',
'http://archive.ics.uci.edu/ml/datasets/Parkinsons',
'http://archive.ics.uci.edu/ml/datasets/Character+Trajectories',
'http://archive.ics.uci.edu/ml/datasets/Blood+Transfusion+Service+Center',
'http://archive.ics.uci.edu/ml/datasets/UJI+Pen+Characters+%28Version+2%29',
'http://archive.ics.uci.edu/ml/datasets/Semeion+Handwritten+Digit',
'http://archive.ics.uci.edu/ml/datasets/SECOM',
'http://archive.ics.uci.edu/ml/datasets/Plants',
'http://archive.ics.uci.edu/ml/datasets/Libras+Movement',
'http://archive.ics.uci.edu/ml/datasets/Concrete+Slump+Test',
'http://archive.ics.uci.edu/ml/datasets/Communities+and+Crime',
'http://archive.ics.uci.edu/ml/datasets/Acute+Inflammations',
'http://archive.ics.uci.edu/ml/datasets/Wine+Quality',
'http://archive.ics.uci.edu/ml/datasets/URL+Reputation',
'http://archive.ics.uci.edu/ml/datasets/p53+Mutants',
'http://archive.ics.uci.edu/ml/datasets/Parkinsons+Telemonitoring',
'http://archive.ics.uci.edu/ml/datasets/Demospongiae',
'http://archive.ics.uci.edu/ml/datasets/Opinosis+Opinion+%26frasl%3B+Review',
'http://archive.ics.uci.edu/ml/datasets/Breast+Tissue',
'http://archive.ics.uci.edu/ml/datasets/Cardiotocography',
'http://archive.ics.uci.edu/ml/datasets/Wall-Following+Robot+Navigation+Data',
'http://archive.ics.uci.edu/ml/datasets/Spoken+Arabic+Digit',
'http://archive.ics.uci.edu/ml/datasets/Localization+Data+for+Person+Activity',
'http://archive.ics.uci.edu/ml/datasets/AutoUniv',
'http://archive.ics.uci.edu/ml/datasets/Steel+Plates+Faults',
'http://archive.ics.uci.edu/ml/datasets/MiniBooNE+particle+identification',
'http://archive.ics.uci.edu/ml/datasets/YearPredictionMSD',
'http://archive.ics.uci.edu/ml/datasets/PEMS-SF',
'http://archive.ics.uci.edu/ml/datasets/OpinRank+Review+Dataset',
'http://archive.ics.uci.edu/ml/datasets/Relative+location+of+CT+slices+on+axial+axis',
'http://archive.ics.uci.edu/ml/datasets/Online+Handwritten+Assamese+Characters+Dataset',
'http://archive.ics.uci.edu/ml/datasets/PubChem+Bioassay+Data',
'http://archive.ics.uci.edu/ml/datasets/Record+Linkage+Comparison+Patterns',
'http://archive.ics.uci.edu/ml/datasets/Communities+and+Crime+Unnormalized',
'http://archive.ics.uci.edu/ml/datasets/Vertebral+Column',
'http://archive.ics.uci.edu/ml/datasets/EMG+Physical+Action+Data+Set',
'http://archive.ics.uci.edu/ml/datasets/Vicon+Physical+Action+Data+Set',
'http://archive.ics.uci.edu/ml/datasets/Amazon+Commerce+reviews+set',
'http://archive.ics.uci.edu/ml/datasets/Amazon+Access+Samples',
'http://archive.ics.uci.edu/ml/datasets/Reuter_50_50',
'http://archive.ics.uci.edu/ml/datasets/Farm+Ads',
'http://archive.ics.uci.edu/ml/datasets/DBWorld+e-mails',
'http://archive.ics.uci.edu/ml/datasets/KEGG+Metabolic+Relation+Network+%28Directed%29',
'http://archive.ics.uci.edu/ml/datasets/KEGG+Metabolic+Reaction+Network+%28Undirected%29',
'http://archive.ics.uci.edu/ml/datasets/Bank+Marketing',
'http://archive.ics.uci.edu/ml/datasets/YouTube+Comedy+Slam+Preference+Data',
'http://archive.ics.uci.edu/ml/datasets/Gas+Sensor+Array+Drift+Dataset',
'http://archive.ics.uci.edu/ml/datasets/ILPD+%28Indian+Liver+Patient+Dataset%29',
'http://archive.ics.uci.edu/ml/datasets/OPPORTUNITY+Activity+Recognition',
'http://archive.ics.uci.edu/ml/datasets/Nomao',
'http://archive.ics.uci.edu/ml/datasets/SMS+Spam+Collection',
'http://archive.ics.uci.edu/ml/datasets/Skin+Segmentation',
'http://archive.ics.uci.edu/ml/datasets/Planning+Relax',
'http://archive.ics.uci.edu/ml/datasets/PAMAP2+Physical+Activity+Monitoring',
'http://archive.ics.uci.edu/ml/datasets/Restaurant+%26+consumer+data',
'http://archive.ics.uci.edu/ml/datasets/CNAE-9',
'http://archive.ics.uci.edu/ml/datasets/Individual+household+electric+power+consumption',
'http://archive.ics.uci.edu/ml/datasets/seeds',
'http://archive.ics.uci.edu/ml/datasets/Northix',
'http://archive.ics.uci.edu/ml/datasets/QtyT40I10D100K',
'http://archive.ics.uci.edu/ml/datasets/Legal+Case+Reports',
'http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones',
'http://archive.ics.uci.edu/ml/datasets/One-hundred+plant+species+leaves+data+set',
'http://archive.ics.uci.edu/ml/datasets/Energy+efficiency',

```
 'http://archive.ics.uci.edu/ml/datasets/Energy+efficiency',
 'http://archive.ics.uci.edu/ml/datasets/Yacht+Hydrodynamics',
 'http://archive.ics.uci.edu/ml/datasets/Fertility',
 'http://archive.ics.uci.edu/ml/datasets/Daphnet+Freezing+of+Gait',
 'http://archive.ics.uci.edu/ml/datasets/3D+Road+Network+%28North+Jutland%2C+Denmark%29',
 'http://archive.ics.uci.edu/ml/datasets/ISTANBUL+STOCK+EXCHANGE',
 'http://archive.ics.uci.edu/ml/datasets/Buzz+in+social+media+',
 'http://archive.ics.uci.edu/ml/datasets/First-order+theorem+proving',

'http://archive.ics.uci.edu/ml/datasets/Wearable+Computing%3A+Classification+of+Body+Postures+and+M
ents+%28PUC-Rio%29',
 'http://archive.ics.uci.edu/ml/datasets/Gas+sensor+arrays+in+open+sampling+settings',
 'http://archive.ics.uci.edu/ml/datasets/Climate+Model+Simulation+Crashes',
 'http://archive.ics.uci.edu/ml/datasets/MicroMass',
 'http://archive.ics.uci.edu/ml/datasets/QSAR+biodegradation',
 'http://archive.ics.uci.edu/ml/datasets/BLOGGER',
 'http://archive.ics.uci.edu/ml/datasets/Daily+and+Sports+Activities',
 'http://archive.ics.uci.edu/ml/datasets/User+Knowledge+Modeling',

'http://archive.ics.uci.edu/ml/datasets/Reuters+RCV1+RCV2+Multilingual%2C+Multiview+Text+Categoriza
+Test+collection',
 'http://archive.ics.uci.edu/ml/datasets/NYSK',
 'http://archive.ics.uci.edu/ml/datasets/Turkiye+Student+Evaluation',

'http://archive.ics.uci.edu/ml/datasets/ser+Knowledge+Modeling+Data+%28Students%27+Knowledge+Levels
DC+Electrical+Machines%29',
 'http://archive.ics.uci.edu/ml/datasets/EEG+Eye+State',

'http://archive.ics.uci.edu/ml/datasets/Physicochemical+Properties+of+Protein+Tertiary+Structure',
 'http://archive.ics.uci.edu/ml/datasets/seismic-bumps',
 'http://archive.ics.uci.edu/ml/datasets/banknote+authentication',
 'http://archive.ics.uci.edu/ml/datasets/USPTO+Algorithm+Challenge%2C+run+by+NASA-
Harvard+Tournament+Lab+and+TopCoder++++Problem%3A+Pat',
 'http://archive.ics.uci.edu/ml/datasets/YouTube+Multiview+Video+Games+Dataset',

'http://archive.ics.uci.edu/ml/datasets/Gas+Sensor+Array+Drift+Dataset+at+Different+Concentrations'


'http://archive.ics.uci.edu/ml/datasets/Activities+of+Daily+Living+%28ADLs%29+Recognition+Using+Bir
Sensors',
 'http://archive.ics.uci.edu/ml/datasets/SkillCraft1+Master+Table+Dataset',

'http://archive.ics.uci.edu/ml/datasets/Weight+Lifting+Exercises+monitored+with+Inertial+Measuremen
its',
 'http://archive.ics.uci.edu/ml/datasets/SML2010',
 'http://archive.ics.uci.edu/ml/datasets/Bike+Sharing+Dataset',
 'http://archive.ics.uci.edu/ml/datasets/Predict+keywords+activities+in+a+online+social+media',
 'http://archive.ics.uci.edu/ml/datasets/Thoracic+Surgery+Data',
 'http://archive.ics.uci.edu/ml/datasets/EMG+dataset+in+Lower+Limb',
 'http://archive.ics.uci.edu/ml/datasets/SUSY',
 'http://archive.ics.uci.edu/ml/datasets/HIGGS',
 'http://archive.ics.uci.edu/ml/datasets/Qualitative_Bankruptcy',
 'http://archive.ics.uci.edu/ml/datasets/LSVT+Voice+Rehabilitation',
 'http://archive.ics.uci.edu/ml/datasets/Dataset+for+ADL+Recognition+with+Wrist-
worn+Accelerometer',
 'http://archive.ics.uci.edu/ml/datasets/Wilt',
 'http://archive.ics.uci.edu/ml/datasets/User+Identification+From+Walking+Activity',
 'http://archive.ics.uci.edu/ml/datasets/Activity+Recognition+from+Single+Chest-
Mounted+Accelerometer',
 'http://archive.ics.uci.edu/ml/datasets/Leaf',
 'http://archive.ics.uci.edu/ml/datasets/Dresses_Attribute_Sales',
 'http://archive.ics.uci.edu/ml/datasets/Tamilnadu+Electricity+Board+Hourly+Readings',
 'http://archive.ics.uci.edu/ml/datasets/Airfoil+Self-Noise',
 'http://archive.ics.uci.edu/ml/datasets/Wholesale+customers',
 'http://archive.ics.uci.edu/ml/datasets/Twitter+Data+set+for+Arabic+Sentiment+Analysis',
 'http://archive.ics.uci.edu/ml/datasets/Combined+Cycle+Power+Plant',
 'http://archive.ics.uci.edu/ml/datasets/Urban+Land+Cover',
 'http://archive.ics.uci.edu/ml/datasets/Diabetes+130-US+hospitals+for+years+1999-2008',
 'http://archive.ics.uci.edu/ml/datasets/Bach+Choral+Harmony',
 'http://archive.ics.uci.edu/ml/datasets/StoneFlakes',
 'http://archive.ics.uci.edu/ml/datasets/Tennis+Major+Tournament+Match+Statistics',

'http://archive.ics.uci.edu/ml/datasets/Parkinson+Speech+Dataset+with++Multiple+Types+of+Sound+Recc
gs',
 'http://archive.ics.uci.edu/ml/datasets/Gesture+Phase+Segmentation',
 'http://archive.ics.uci.edu/ml/datasets/Perfume+Data',
 'http://archive.ics.uci.edu/ml/datasets/BlogFeedback',
 'http://archive.ics.uci.edu/ml/datasets/REALDISP+Activity+Recognition+Dataset',
```

'http://archive.ics.uci.edu/ml/datasets/KEADIDI+Activity+Recognition+Dataset',
'http://archive.ics.uci.edu/ml/datasets/Newspaper+and+magazine+images+segmentation+dataset',
'http://archive.ics.uci.edu/ml/datasets/AAAI+2014+Accepted+Papers',
'http://archive.ics.uci.edu/ml/datasets/Gas+sensor+array+under+flow+modulation',
'http://archive.ics.uci.edu/ml/datasets/Gas+sensor+array+exposed+to+turbulent+gas+mixtures',
'http://archive.ics.uci.edu/ml/datasets/UJIIndoorLoc',
'http://archive.ics.uci.edu/ml/datasets/Sentence+Classification',
'http://archive.ics.uci.edu/ml/datasets/Dow+Jones+Index',
'http://archive.ics.uci.edu/ml/datasets/sEMG+for+Basic+Hand+movements',
'http://archive.ics.uci.edu/ml/datasets/AAAI+2013+Accepted+Papers',
'http://archive.ics.uci.edu/ml/datasets/Geographical+Original+of+Music',
'http://archive.ics.uci.edu/ml/datasets/Condition+Based+Maintenance+of+Naval+Propulsion+Plants',
'http://archive.ics.uci.edu/ml/datasets/Grammatical+Facial+Expressions',
'http://archive.ics.uci.edu/ml/datasets/NoisyOffice',
'http://archive.ics.uci.edu/ml/datasets/MHEALTH+Dataset',
'http://archive.ics.uci.edu/ml/datasets/Student+Performance',
'http://archive.ics.uci.edu/ml/datasets/ElectricityLoadDiagrams20112014',
'http://archive.ics.uci.edu/ml/datasets/Gas+sensor+array+under+dynamic+gas+mixtures',
'http://archive.ics.uci.edu/ml/datasets/microblogPCU',
'http://archive.ics.uci.edu/ml/datasets/Firm-Teacher_Clave-Direction_Classification',
'http://archive.ics.uci.edu/ml/datasets/Dataset+for+Sensorless+Drive+Diagnosis',
'http://archive.ics.uci.edu/ml/datasets/TV+News+Channel+Commercial+Detection+Dataset',
'http://archive.ics.uci.edu/ml/datasets/Phishing+Websites',
'http://archive.ics.uci.edu/ml/datasets/Greenhouse+Gas+Observing+Network',
'http://archive.ics.uci.edu/ml/datasets/Diabetic+Retinopathy+Debrecen+Data+Set',
'http://archive.ics.uci.edu/ml/datasets/HIV-1+protease+cleavage',
'http://archive.ics.uci.edu/ml/datasets/Sentiment+Labelled+Sentences',
'http://archive.ics.uci.edu/ml/datasets/Online+News+Popularity',
'http://archive.ics.uci.edu/ml/datasets/Forest+type+mapping',
'http://archive.ics.uci.edu/ml/datasets/wiki4HE',

'http://archive.ics.uci.edu/ml/datasets/Online+Video+Characteristics+and+Transcoding+Time+Dataset',

'http://archive.ics.uci.edu/ml/datasets/Chronic_Kidney_Disease',

'http://archive.ics.uci.edu/ml/datasets/Machine+Learning+based+ZZAlpha+Ltd.+Stock+Recommendations+2
2014',
'http://archive.ics.uci.edu/ml/datasets/Folio',
'http://archive.ics.uci.edu/ml/datasets/Taxi+Service+Trajectory+-
+Prediction+Challenge%2C+ECML+PKDD+2015',
'http://archive.ics.uci.edu/ml/datasets/Cuff-Less+Blood+Pressure+Estimation',
'http://archive.ics.uci.edu/ml/datasets/Smartphone-
Based+Recognition+of+Human+Activities+and+Postural+Transitions',
'http://archive.ics.uci.edu/ml/datasets/Mice+Protein+Expression',
'http://archive.ics.uci.edu/ml/datasets/UJIIndoorLoc-Mag',
'http://archive.ics.uci.edu/ml/datasets/Heterogeneity+Activity+Recognition',

'http://archive.ics.uci.edu/ml/datasets/Educational+Process+Mining+%28EPM%29%3A+A+Learning+Analytic
ta+Set',
'http://archive.ics.uci.edu/ml/datasets/HEPMASS',
'http://archive.ics.uci.edu/ml/datasets/Indoor+User+Movement+Prediction+from+RSS+data',
'http://archive.ics.uci.edu/ml/datasets/Open+University+Learning+Analytics+dataset',
'http://archive.ics.uci.edu/ml/datasets/default+of+credit+card+clients',
'http://archive.ics.uci.edu/ml/datasets/Mesothelioma%E2%80%99s+disease+data+set+',
'http://archive.ics.uci.edu/ml/datasets/Online+Retail',
'http://archive.ics.uci.edu/ml/datasets/SIFT10M',
'http://archive.ics.uci.edu/ml/datasets/GPS+Trajectories',
'http://archive.ics.uci.edu/ml/datasets/Detect+Malacious+Executable%28AntiVirus%29',
'http://archive.ics.uci.edu/ml/datasets/Occupancy+Detection+',

'http://archive.ics.uci.edu/ml/datasets/Improved+Spiral+Test+Using+Digitized+Graphics+Tablet+for+Mo
ring+Parkinson%E2%80%99s+Disease',
'http://archive.ics.uci.edu/ml/datasets/News+Aggregator',
'http://archive.ics.uci.edu/ml/datasets/Air+Quality',
'http://archive.ics.uci.edu/ml/datasets/Twin+gas+sensor+arrays',
'http://archive.ics.uci.edu/ml/datasets/Gas+sensors+for+home+activity+monitoring',
'http://archive.ics.uci.edu/ml/datasets/Facebook+Comment+Volume+Dataset',

'http://archive.ics.uci.edu/ml/datasets/Smartphone+Dataset+for+Human+Activity+Recognition+%28HAR%29
Ambient+Assisted+Living+%28AAL%29',
'http://archive.ics.uci.edu/ml/datasets/Polish+companies+bankruptcy+data',

'http://archive.ics.uci.edu/ml/datasets/Activity+Recognition+system+based+on+Multisensor+data+fusio
8AReM%29',
'http://archive.ics.uci.edu/ml/datasets/Dota2+Games+Results',
'http://archive.ics.uci.edu/ml/datasets/Facebook+metrics',
'http://archive.ics.uci.edu/ml/datasets/UbiqLog+%28smartphone+lifelogging%29',
'http://archive.ics.uci.edu/ml/datasets/NIPS+Conference+Papers+1987-2015',

'http://archive.ics.uci.edu/ml/datasets/NIPS+Conference+Papers+1987-2015',
'http://archive.ics.uci.edu/ml/datasets/HTRU2',
'http://archive.ics.uci.edu/ml/datasets/Drug+consumption+%28quantified%29',
'http://archive.ics.uci.edu/ml/datasets/Appliances+energy+prediction',
'http://archive.ics.uci.edu/ml/datasets/Miskolc+IIS+Hybrid+IPS',
'http://archive.ics.uci.edu/ml/datasets/KDC-4007+dataset+Collection',
'http://archive.ics.uci.edu/ml/datasets/Geo-Magnetic+field+and+WLAN+dataset+for+indoor+localisation+from+wristband+and+smartphone',
'http://archive.ics.uci.edu/ml/datasets/DrivFace',
'http://archive.ics.uci.edu/ml/datasets/Website+Phishing',
'http://archive.ics.uci.edu/ml/datasets/YouTube+Spam+Collection',
'http://archive.ics.uci.edu/ml/datasets/Beijing+PM2.5+Data',
'http://archive.ics.uci.edu/ml/datasets/Cargo+2000+Freight+Tracking+and+Tracing',
'http://archive.ics.uci.edu/ml/datasets/Cervical+cancer+%28Risk+Factors%29',
'http://archive.ics.uci.edu/ml/datasets/Quality+Assessment+of+Digital+Colposcopies',
'http://archive.ics.uci.edu/ml/datasets/KASANDR',
'http://archive.ics.uci.edu/ml/datasets/FMA%3A+A+Dataset+For+Music+Analysis',
'http://archive.ics.uci.edu/ml/datasets/Air+quality',
'http://archive.ics.uci.edu/ml/datasets/Epileptic+Seizure+Recognition',
'http://archive.ics.uci.edu/ml/datasets/Devanagari+Handwritten+Character+Dataset',
'http://archive.ics.uci.edu/ml/datasets/Stock+portfolio+performance',
'http://archive.ics.uci.edu/ml/datasets/MoCap+Hand+Postures',

'http://archive.ics.uci.edu/ml/datasets/Early+biomarkers+of+Parkinson%92s+disease+based+on+natural+
ected+speech',

'http://archive.ics.uci.edu/ml/datasets/Data+for+Software+Engineering+Teamwork+Assessment+in+Educat
Setting',
'http://archive.ics.uci.edu/ml/datasets/PM2.5+Data+of+Five+Chinese+Cities',

'http://archive.ics.uci.edu/ml/datasets/Parkinson+Disease+Spiral+Drawings+Using+Digitized+Graphics+
et',
'http://archive.ics.uci.edu/ml/datasets/Sales_Transactions_Dataset_Weekly',
'http://archive.ics.uci.edu/ml/datasets/Las+Vegas+Strip',
'http://archive.ics.uci.edu/ml/datasets/Eco-hotel',
'http://archive.ics.uci.edu/ml/datasets/MEU-Mobile+KSD',
'http://archive.ics.uci.edu/ml/datasets/Crowdsourced+Mapping',
'http://archive.ics.uci.edu/ml/datasets/gene+expression+cancer+RNA-Seq',

'http://archive.ics.uci.edu/ml/datasets/Hybrid+Indoor+Positioning+Dataset+from+WiFi+RSSI%2C+Bluetoo
nd+magnetometer',
'http://archive.ics.uci.edu/ml/datasets/chestnut+%E2%80%93+LARVIC',

'http://archive.ics.uci.edu/ml/datasets/Burst+Header+Packet+%28BHP%29+flooding+attack+on+Optical+Bu
Switching+%28OBS%29+Network',
'http://archive.ics.uci.edu/ml/datasets/Motion+Capture+Hand+Postures',
'http://archive.ics.uci.edu/ml/datasets/Anuran+Calls+%28MFCCs%29',
'http://archive.ics.uci.edu/ml/datasets/TTC-3600%3A+Benchmark+dataset+for+Turkish+text+categorization',
'http://archive.ics.uci.edu/ml/datasets/Gastrointestinal+Lesions+in+Regular+Colonoscopy',
'http://archive.ics.uci.edu/ml/datasets/Daily+Demand+Forecasting+Orders',
'http://archive.ics.uci.edu/ml/datasets/Paper+Reviews',
'http://archive.ics.uci.edu/ml/datasets/extention+of+Z-Alizadeh+sani+dataset',
'http://archive.ics.uci.edu/ml/datasets/Z-Alizadeh+Sani',
'http://archive.ics.uci.edu/ml/datasets/Dynamic+Features+of+VirusShare+Executables',
'http://archive.ics.uci.edu/ml/datasets/IDA2016Challenge',
'http://archive.ics.uci.edu/ml/datasets/DSRC+Vehicle+Communications',
'http://archive.ics.uci.edu/ml/datasets/Mturk+User-Perceived+Clusters+over+Images',
'http://archive.ics.uci.edu/ml/datasets/Character+Font+Images',
'http://archive.ics.uci.edu/ml/datasets/DeliciousMIL%3A+A+Data+Set+for+Multi-Label+Multi-Instance+Learning+with+Instance+Labels',

'http://archive.ics.uci.edu/ml/datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Children++',

'http://archive.ics.uci.edu/ml/datasets/Autistic+Spectrum+Disorder+Screening+Data+for+Adolescent+++

'http://archive.ics.uci.edu/ml/datasets/APS+Failure+at+Scania+Trucks',
'http://archive.ics.uci.edu/ml/datasets/Wireless+Indoor+Localization',
'http://archive.ics.uci.edu/ml/datasets/HCC+Survival',

'http://archive.ics.uci.edu/ml/datasets/CSM+%28Conventional+and+Social+Media+Movies%29+Dataset+2014
+2015',

'http://archive.ics.uci.edu/ml/datasets/University+of+Tehran+Question+Dataset+2016+%28UTQD.2016%29'

'http://archive.ics.uci.edu/ml/datasets/Autism+Screening+Adult',

'http://archive.ics.uci.edu/ml/datasets/Activity+recognition+with+healthy+older+people+using+a+batt

```
'http://archive.ics.uci.edu/ml/datasets/Activity+recognition+with+healthy+older+people+using+a+batt
ess+wearable+sensor',
 'http://archive.ics.uci.edu/ml/datasets/Immunotherapy+Dataset',
 'http://archive.ics.uci.edu/ml/datasets/Cryotherapy+Dataset+',
 'http://archive.ics.uci.edu/ml/datasets/OCT+data+%26+Color+Fundus+Images+of+Left+%26+Right+Eyes',
 'http://archive.ics.uci.edu/ml/datasets/Discrete+Tone+Image+Dataset',
 'http://archive.ics.uci.edu/ml/datasets/News+Popularity+in+Multiple+Social+Media+Platforms',
 'http://archive.ics.uci.edu/ml/datasets/Ultrasonic+flowmeter+diagnostics',
 'http://archive.ics.uci.edu/ml/datasets/ICMLA+2014+Accepted+Papers+Data+Set',
 'http://archive.ics.uci.edu/ml/datasets/BLE+RSSI+Dataset+for+Indoor+localization+and+Navigation',
 'http://archive.ics.uci.edu/ml/datasets/Container+Crane+Controller+Data+Set',
 'http://archive.ics.uci.edu/ml/datasets/Residential+Building+Data+Set',
 'http://archive.ics.uci.edu/ml/datasets/Health+News+in+Twitter',
 'http://archive.ics.uci.edu/ml/datasets/chipseq',
 'http://archive.ics.uci.edu/ml/datasets/SGEMM+GPU+kernel+performance',
 'http://archive.ics.uci.edu/ml/datasets/Repeat+Consumption+Matrices',
 'http://archive.ics.uci.edu/ml/datasets/detection_of_IoT_botnet_attacks_N_BaIoT',
 'http://archive.ics.uci.edu/ml/datasets/Absenteeism+at+work',
 'http://archive.ics.uci.edu/ml/datasets/SCADI',
 'http://archive.ics.uci.edu/ml/datasets/Condition+monitoring+of+hydraulic+systems',
 'http://archive.ics.uci.edu/ml/datasets/Carbon+Nanotubes',
 'http://archive.ics.uci.edu/ml/datasets/Optical+Interconnection+Network+',
 'http://archive.ics.uci.edu/ml/datasets/Sports+articles+for+objectivity+analysis',
 'http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Coimbra',
 'http://archive.ics.uci.edu/ml/datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data',
 'http://archive.ics.uci.edu/ml/datasets/Dishonest+Internet+users+Dataset',
 'http://archive.ics.uci.edu/ml/datasets/Victorian+Era+Authorship+Attribution',
 'http://archive.ics.uci.edu/ml/datasets/Simulated+Falls+and+Daily+Living+Activities+Data+Set',

'http://archive.ics.uci.edu/ml/datasets/Multimodal+Damage+Identification+for+Humanitarian+Computing

 'http://archive.ics.uci.edu/ml/datasets/EEG+Steady-State+Visual+Evoked+Potential+Signals',
 'http://archive.ics.uci.edu/ml/datasets/Roman+Urdu+Data+Set',
 'http://archive.ics.uci.edu/ml/datasets/Avila',
 'http://archive.ics.uci.edu/ml/datasets/PANDOR',
 'http://archive.ics.uci.edu/ml/datasets/Drug+Review+Dataset+%28Druglib.com%29',
 'http://archive.ics.uci.edu/ml/datasets/Drug+Review+Dataset+%28Drugs.com%29',
 'http://archive.ics.uci.edu/ml/datasets/Physical+Unclonable+Functions',
 'http://archive.ics.uci.edu/ml/datasets/Superconductivty+Data',
 'http://archive.ics.uci.edu/ml/datasets/WESAD+%28Wearable+Stress+and+Affect+Detection%29',
 'http://archive.ics.uci.edu/ml/datasets/GNFUV+Unmanned+Surface+Vehicles+Sensor+Data+Set+2',
 'http://archive.ics.uci.edu/ml/datasets/Student+Academics+Performance',
 'http://archive.ics.uci.edu/ml/datasets/Online+Shoppers+Purchasing+Intention+Dataset',
 'http://archive.ics.uci.edu/ml/datasets/PMU-UD',
 'http://archive.ics.uci.edu/ml/datasets/Parkinson%27s+Disease+Classification',
 'http://archive.ics.uci.edu/ml/datasets/Electrical+Grid+Stability+Simulated+Data+',
 'http://archive.ics.uci.edu/ml/datasets/Caesarian+Section+Classification+Dataset',
 'http://archive.ics.uci.edu/ml/datasets/BAUM-1',
 'http://archive.ics.uci.edu/ml/datasets/BAUM-2',
 'http://archive.ics.uci.edu/ml/datasets/Audit+Data',
 'http://archive.ics.uci.edu/ml/datasets/BuddyMove+Data+Set',
 'http://archive.ics.uci.edu/ml/datasets/Real+estate+valuation+data+set',

'http://archive.ics.uci.edu/ml/datasets/Early+biomarkers+of+Parkinson%E2%80%99s+disease+based+on+na
l+connected+speech+Data+Set+',
 'http://archive.ics.uci.edu/ml/datasets/Somerville+Happiness+Survey',
 'http://archive.ics.uci.edu/ml/datasets/2.4+GHZ+Indoor+Channel+Measurements',
 'http://archive.ics.uci.edu/ml/datasets/EMG+data+for+gestures',
 'http://archive.ics.uci.edu/ml/datasets/Parking+Birmingham',

'http://archive.ics.uci.edu/ml/datasets/Behavior+of+the+urban+traffic+of+the+city+of+Sao+Paulo+in+B
l',
 'http://archive.ics.uci.edu/ml/datasets/Travel+Reviews',
 'http://archive.ics.uci.edu/ml/datasets/Tarvel+Review+Ratings',
 'http://archive.ics.uci.edu/ml/datasets/Rice+Leaf+Diseases']
```

In [270]:

```
datos = soup.find_all('p')
```

In [271]:

```
datos
```

Out[271]:

["-//W3C//DTD HTML 4.01 Transitional//EN\">

 ,
 Browse Through: ,
 Default Task ,
 <a href="datasets.php?
format=&task=cla&att=&area=&numAtt=&numIns=&type=&sort=nameUp&view=
e">Classification</a> <font color="red">(350)</font>
<a href="datasets.php?
format=&task=reg&att=&area=&numAtt=&numIns=&type=&sort=nameUp&view=
e">Regression</a> <font color="red">(96)</font>
<a href="datasets.php?
format=&task=clu&att=&area=&numAtt=&numIns=&type=&sort=nameUp&view=
e">Clustering</a> <font color="red">(84)</font>
<a href="datasets.php?
format=&task=other&att=&area=&numAtt=&numIns=&type=&sort=nameUp&vie
ble">Other</a> <font color="red">(55)</font> ,
 Attribute Type ,
 <a href="datasets.php?
format=&task=&att=cat&area=&numAtt=&numIns=&type=&sort=nameUp&view=
e">Categorical</a> <font color="red">(38)</font>
<a href="datasets.php?
format=&task=&att=num&area=&numAtt=&numIns=&type=&sort=nameUp&view=
e">Numerical</a> <font color="red">(307)</font>
<a href="datasets.php?
format=&task=&att=mix&area=&numAtt=&numIns=&type=&sort=nameUp&view=
e">Mixed</a> <font color="red">(55)</font> ,
 Data Type ,
 <a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=mvar&sort=nameUp&view
le">Multivariate</a> <font color="red">(357)</font>
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=uvar&sort=nameUp&view
le">Univariate</a> <font color="red">(23)</font>
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=seq&sort=nameUp&view=
e">Sequential</a> <font color="red">(47)</font>
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=ts&sort=nameUp&view=t
">Time-Series</a> <font color="red">(91)</font>
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=text&sort=nameUp&view
le">Text</a> <font color="red">(53)</font>
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=dt&sort=nameUp&view=t
">Domain-Theory</a> <font color="red">(23)</font>
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=other&sort=nameUp&vie
ble">Other</a> <font color="red">(21)</font>
 ,
 Area ,
 <a href="datasets.php?
format=&task=&att=&area=life&numAtt=&numIns=&type=&sort=nameUp&view
le">Life Sciences</a> <font color="red">(107)</font>
<a href="datasets.php?
format=&task=&att=&area=phys&numAtt=&numIns=&type=&sort=nameUp&view
le">Physical Sciences</a> <font color="red">(49)</font>
<a href="datasets.php?
format=&task=&att=&area=comp&numAtt=&numIns=&type=&sort=nameUp&view
le">CS / Engineering</a> <font color="red">(170)</font>
<a href="datasets.php?
format=&task=&att=&area=soc&numAtt=&numIns=&type=&sort=nameUp&view=
e">Social Sciences</a> <font color="red">(26)</font>
<a href="datasets.php?
format=&task=&att=&area=bus&numAtt=&numIns=&type=&sort=nameUp&view=
e">Business</a> <font color="red">(29)</font>
<a href="datasets.php?
format=&task=&att=&area=game&numAtt=&numIns=&type=&sort=nameUp&view
le">Game</a> <font color="red">(10)</font>
<a href="datasets.php?
format=&task=&att=&area=other&numAtt=&numIns=&type=&sort=nameUp&vie
ble">Other</a> <font color="red">(73)</font> ,
 # Attributes ,
 <a href="datasets.php?
format=&task=&att=&area=&numAtt=less10&numIns=&type=&sort=nameUp&vi
able">Less than 10</a> <font color="red">(113)</font>
<a href="datasets.php?
format=&task=&att=&area=&numAtt=10to100&numIns=&type=&sort=nameUp&v
table">10 to 100</a> <font color="red">(210)</font>
<a href="datasets.php?
format=&task=&att=&area=&numAtt=greater100&numIns=&type=&sort=nameUp&am
ew=table">Greater than 100</a> <font color="red">(84)</font> ,
 # Instances ,

 <a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=less100&type=&sort=nameUp&v
table">Less than 100</a> <font color="red">(27)</font>
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=100to1000&type=&sort=nameUp&amp
w=table">100 to 1000</a> <font color="red">(162)</font>
<a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=greater1000&type=&sort=nameUp&a
iew=table">Greater than 1000</a> <font color="red">(246)</font> ,
 Format Type ,
 <a href="datasets.php?
format=mat&task=&att=&area=&numAtt=&numIns=&type=&sort=nameUp&view=
e">Matrix</a> <font color="red">(324)</font>
<a href="datasets.php?
format=nonmat&task=&att=&area=&numAtt=&numIns=&type=&sort=nameUp&vi

able">Non-Matrix</a> <font color="red">(145)</font> ,
 469 Data Sets ,
 <font color="gray">Table View</font> <a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=nameUp&view=lis
ist View</a> ,
 <a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=nameDown&view=t
">Name</a> ,
 <a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=typeUp&view=tab
Data Types</a> ,
 <a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=taskUp&view=tab
Default Task</a> ,
 <a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=attTypeUp&view=
e">Attribute Types</a> ,
 <a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=instUp&view=tab
# Instances</a> ,
 <a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=attUp&view=tabl
# Attributes</a> ,
 <a href="datasets.php?
format=&task=&att=&area=&numAtt=&numIns=&type=&sort=dateUp&view=tab
Year</a> ,
 <a href="datasets/Abalone">Abalone</a> ,
 Multivariate ,
 Classification ,
 Categorical, Integer, Real ,
 4177 ,
 8 ,
 1995 ,
 <a href="datasets/Adult">Adult</a> ,
 Multivariate ,
 Classification ,
 Categorical, Integer ,
 48842 ,
 14 ,
 1996 ,
 <a href="datasets/Annealing">Annealing</a> ,
 Multivariate ,
 Classification ,
 Categorical, Integer, Real ,
 798 ,
 38 ,
  ,
 <a href="datasets/Anonymous+Microsoft+Web+Data">Anonymous Microsoft Web
Data</a> ,
  ,
 Recommender-Systems ,
 Categorical ,
 37711 ,
 294 ,
 1998 ,
 <a href="datasets/Arrhythmia">Arrhythmia</a> ,
 Multivariate ,
 Classification ,
 Categorical, Integer, Real ,
 452 ,
 279 ,
 1998 ,
 <a href="datasets/Artificial+Characters">Artificial Characters</a> ,
 Multivariate ,
 Classification ,
 Categorical, Integer, Real ,
 6000 ,
 7 ,
 1992 ,
 <a href="datasets/Audiology+%28Original%29">Audiology (Original)</a>
,
 Multivariate ,
 Classification ,
 Categorical ,
 226 ,
   ,
 1987 ,

```
 <p class="normal"><b><a href="datasets/Audiology+%28Standardized%29">Audiology (Standardized)</a>
</b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">226 </p>,
 <p class="normal">69 </p>,
 <p class="normal">1992 </p>,
 <p class="normal"><b><a href="datasets/Auto+MPG">Auto MPG</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Categorical, Real </p>,
 <p class="normal">398 </p>,
 <p class="normal">8 </p>,
 <p class="normal">1993 </p>,
 <p class="normal"><b><a href="datasets/Automobile">Automobile</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">205 </p>,
 <p class="normal">26 </p>,
 <p class="normal">1987 </p>,
 <p class="normal">><b><a href="datasets/Badges">Badges</a></b></p>,
 <p class="normal">Univariate, Text </p>,
 <p class="normal">Classification </p>,
 <p class="normal"> </p>,
 <p class="normal">294 </p>,
 <p class="normal">1 </p>,
 <p class="normal">1994 </p>,
 <p class="normal">><b><a href="datasets/Balance+Scale">Balance Scale</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">625 </p>,
 <p class="normal">4 </p>,
 <p class="normal">1994 </p>,
 <p class="normal">><b><a href="datasets/Balloons">Balloons</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">16 </p>,
 <p class="normal">4 </p>,
 <p class="normal"> </p>,
 <p class="normal">><b><a href="datasets/Breast+Cancer">Breast Cancer</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">286 </p>,
 <p class="normal">9 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">><b><a href="datasets/Breast+Cancer+Wisconsin+%28Original%29">Breast Cancer
Wisconsin (Original)</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Integer </p>,
 <p class="normal">699 </p>,
 <p class="normal">10 </p>,
 <p class="normal">1992 </p>,
 <p class="normal">><b><a href="datasets/Breast+Cancer+Wisconsin+%28Prognostic%29">Breast Cancer Wi
sconsin (Prognostic)</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification, Regression </p>,
 <p class="normal">Real </p>,
 <p class="normal">198 </p>,
 <p class="normal">34 </p>,
 <p class="normal">1995 </p>,
 <p class="normal">><b><a href="datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29">Breast Cancer Wi
sconsin (Diagnostic)</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Real </p>,
 <p class="normal">569 </p>,
 <p class="normal">32 </p>,
 <p class="normal">1995 </p>,
 <p class="normal">><b><a href="datasets/Pittsburgh+Bridges">Pittsburgh Bridges</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
```

```
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">108 </p>,
 <p class="normal">13 </p>,
 <p class="normal">1990 </p>,
 <p class="normal"><b><a href="datasets/Car+Evaluation">Car Evaluation</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">1728 </p>,
 <p class="normal">6 </p>,
 <p class="normal">1997 </p>,
 <p class="normal">><b><a href="datasets/Census+Income">Census Income</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">48842 </p>,
 <p class="normal">14 </p>,
 <p class="normal">1996 </p>,
 <p class="normal">><b><a href="datasets/Chess+%28King-Rook+vs.+King-Knight%29">Chess (King-Rook vs
. King-Knight)</a></b></p>,
 <p class="normal">Multivariate, Data-Generator </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal"> </p>,
 <p class="normal">22 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">><b><a href="datasets/Chess+%28King-Rook+vs.+King-Pawn%29">Chess (King-Rook vs.
King-Pawn)</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">3196 </p>,
 <p class="normal">36 </p>,
 <p class="normal">1989 </p>,
 <p class="normal">><b><a href="datasets/Chess+%28King-Rook+vs.+King%29">Chess (King-Rook vs. King)
</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">28056 </p>,
 <p class="normal">6 </p>,
 <p class="normal">1994 </p>,
 <p class="normal">><b><a href="datasets/Chess+%28Domain+Theories%29">Chess (Domain Theories)</a></
b></p>,
 <p class="normal">Domain-Theory </p>,
 <p class="normal">> </p>,
 <p class="normal">> </p>,
 <p class="normal">> </p>,
 <p class="normal">> </p>,
 <p class="normal">> </p>,
 <p class="normal">><b><a href="datasets/Bach+Chorales">Bach Chorales</a></b></p>,
 <p class="normal">Univariate, Time-Series </p>,
 <p class="normal">> </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">100 </p>,
 <p class="normal">6 </p>,
 <p class="normal">> </p>,
 <p class="normal">><b><a href="datasets/Connect-4">Connect-4</a></b></p>,
 <p class="normal">Multivariate, Spatial </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">67557 </p>,
 <p class="normal">42 </p>,
 <p class="normal">1995 </p>,
 <p class="normal">><b><a href="datasets/Credit+Approval">Credit Approval</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">690 </p>,
 <p class="normal">15 </p>,
 <p class="normal">> </p>,
 <p class="normal">><b><a href="datasets/Japanese+Credit+Screening">Japanese Credit Screening</a></
b></p>,
 <p class="normal">Multivariate, Domain-Theory </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Real, Integer </p>,
 <p class="normal">125 </p>,
```

```
 <p class="normal"> </p>,
 <p class="normal">1992 </p>,
 <p class="normal"><b><a href="datasets/Computer+Hardware">Computer Hardware</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Integer </p>,
 <p class="normal">209 </p>,
 <p class="normal">9 </p>,
 <p class="normal">1987 </p>,
 <p class="normal"><b><a href="datasets/Contraceptive+Method+Choice">Contraceptive Method
Choice</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">1473 </p>,
 <p class="normal">9 </p>,
 <p class="normal">1997 </p>,
 <p class="normal">&lt;b&gt;&lt;a href="datasets/Covertype"&gt;Covertype&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">581012 </p>,
 <p class="normal">54 </p>,
 <p class="normal">1998 </p>,
 <p class="normal">&lt;b&gt;&lt;a href="datasets/Cylinder+Bands"&gt;Cylinder Bands&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">512 </p>,
 <p class="normal">39 </p>,
 <p class="normal">1995 </p>,
 <p class="normal">&lt;b&gt;&lt;a href="datasets/Dermatology"&gt;Dermatology&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">366 </p>,
 <p class="normal">33 </p>,
 <p class="normal">1998 </p>,
 <p class="normal">&lt;b&gt;&lt;a href="datasets/Diabetes"&gt;Diabetes&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
 <p class="normal">Multivariate, Time-Series </p>,
 <p class="normal"> </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal"> </p>,
 <p class="normal">20 </p>,
 <p class="normal"> </p>,
 <p class="normal">&lt;b&gt;&lt;a href="datasets/DGP2+-+The+Second+Data+Generation+Program"&gt;DGP2 - The Seco
nd Data Generation Program&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
 <p class="normal">Data-Generator </p>,
 <p class="normal"> </p>,
 <p class="normal">Real </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal">&lt;b&gt;&lt;a href="datasets/Document+Understanding"&gt;Document Understanding&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;
,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">132 </p>,
 <p class="normal">12 </p>,
 <p class="normal">1989 </p>,
 <p class="normal">&lt;b&gt;&lt;a href="datasets/Ecoli"&gt;Ecoli&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
 <p class="normal">Multivariate </p>,
```

```
<p class="normal">Classification </p>,
<p class="normal">Real </p>,
<p class="normal">336 </p>,
<p class="normal">8 </p>,
<p class="normal">1996 </p>,
<p class="normal"><b><a href="datasets/Flags">Flags</a></b></p>,
<p class="normal">Multivariate </p>,
<p class="normal">Classification </p>,
<p class="normal">Categorical, Integer </p>,
<p class="normal">194 </p>,
<p class="normal">30 </p>,
<p class="normal">1990 </p>,
<p class="normal"><b><a href="datasets/Function+Finding">Function Finding</a></b></p>,
<p class="normal"> </p>,
<p class="normal">Function-Learning </p>,
<p class="normal">Real </p>,
<p class="normal">352 </p>,
<p class="normal"> </p>,
<p class="normal">1990 </p>,
<p class="normal"><b><a href="datasets/Glass+Identification">Glass Identification</a></b></p>,
<p class="normal">Multivariate </p>,
<p class="normal">Classification </p>,
<p class="normal">Real </p>,
<p class="normal">214 </p>,
<p class="normal">10 </p>,
<p class="normal">1987 </p>,
<p class="normal"><b><a href="datasets/Haberman%27s+Survival">Haberman's Survival</a></b></p>,
<p class="normal">Multivariate </p>,
<p class="normal">Classification </p>,
<p class="normal">Integer </p>,
<p class="normal">306 </p>,
<p class="normal">3 </p>,
<p class="normal">1999 </p>,
<p class="normal"><b><a href="datasets/Hayes-Roth">Hayes-Roth</a></b></p>,
<p class="normal">Multivariate </p>,
<p class="normal">Classification </p>,
<p class="normal">Categorical </p>,
<p class="normal">160 </p>,
<p class="normal">5 </p>,
<p class="normal">1989 </p>,
<p class="normal"><b><a href="datasets/Heart+Disease">Heart Disease</a></b></p>,
<p class="normal">Multivariate </p>,
<p class="normal">Classification </p>,
<p class="normal">Categorical, Integer, Real </p>,
<p class="normal">303 </p>,
<p class="normal">75 </p>,
<p class="normal">1988 </p>,
<p class="normal"><b><a href="datasets/Hepatitis">Hepatitis</a></b></p>,
<p class="normal">Multivariate </p>,
<p class="normal">Classification </p>,
<p class="normal">Categorical, Integer, Real </p>,
<p class="normal">155 </p>,
<p class="normal">19 </p>,
<p class="normal">1988 </p>,
<p class="normal">><b><a href="datasets/Horse+Colic">Horse Colic</a></b></p>,
<p class="normal">Multivariate </p>,
<p class="normal">Classification </p>,
<p class="normal">Categorical, Integer, Real </p>,
<p class="normal">368 </p>,
<p class="normal">27 </p>,
<p class="normal">1989 </p>,
<p class="normal">><b><a href="datasets/ICU">ICU</a></b></p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">> </p>,
<p class="normal">Real </p>,
<p class="normal">> </p>,
<p class="normal">> </p>,
<p class="normal">> </p>,
<p class="normal">><b><a href="datasets/Image+Segmentation">Image Segmentation</a></b></p>,
<p class="normal">Multivariate </p>,
<p class="normal">Classification </p>,
<p class="normal">Real </p>,
<p class="normal">2310 </p>,
<p class="normal">19 </p>,
<p class="normal">1990 </p>,
<p class="normal">><b><a href="datasets/Internet+Advertisements">Internet Advertisements</a></b></
p>,
```

```
<p class="normal">Multivariate </p>,
<p class="normal">Classification </p>,
<p class="normal">Categorical, Integer, Real </p>,
<p class="normal">3279 </p>,
<p class="normal">1558 </p>,
<p class="normal">1998 </p>,
<p class="normal"><b><a href="datasets/Ionosphere">Ionosphere</a></b></p>,
<p class="normal">Multivariate </p>,
<p class="normal">Classification </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">351 </p>,
<p class="normal">34 </p>,
<p class="normal">1989 </p>,
<p class="normal"><b><a href="datasets/Iris">Iris</a></b></p>,
<p class="normal">Multivariate </p>,
<p class="normal">Classification </p>,
<p class="normal">Real </p>,
<p class="normal">150 </p>,
<p class="normal">4 </p>,
<p class="normal">1988 </p>,
<p class="normal">&lt;b&gt;&lt;a href="datasets/ISOLET"&gt;ISOLET&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
<p class="normal">Multivariate </p>,
<p class="normal">Classification </p>,
<p class="normal">Real </p>,
<p class="normal">7797 </p>,
<p class="normal">617 </p>,
<p class="normal">1994 </p>,
<p class="normal">&lt;b&gt;&lt;a href="datasets/Kinship"&gt;Kinship&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
<p class="normal">Relational </p>,
<p class="normal">Relational-Learning </p>,
<p class="normal">Categorical </p>,
<p class="normal">104 </p>,
<p class="normal">12 </p>,
<p class="normal">1990 </p>,
<p class="normal">&lt;b&gt;&lt;a href="datasets/Labor+Relations"&gt;Labor Relations&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
<p class="normal">Multivariate </p>,
<p class="normal"> </p>,
<p class="normal">Categorical, Integer, Real </p>,
<p class="normal">57 </p>,
<p class="normal">16 </p>,
<p class="normal">1988 </p>,
<p class="normal">&lt;b&gt;&lt;a href="datasets/LED+Display+Domain"&gt;LED Display Domain&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
<p class="normal">Multivariate, Data-Generator </p>,
<p class="normal">Classification </p>,
<p class="normal">Categorical </p>,
<p class="normal"> </p>,
<p class="normal">7 </p>,
<p class="normal">1988 </p>,
<p class="normal">&lt;b&gt;&lt;a href="datasets/Lenses"&gt;Lenses&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
<p class="normal">Multivariate </p>,
<p class="normal">Classification </p>,
<p class="normal">Categorical </p>,
<p class="normal">24 </p>,
<p class="normal">4 </p>,
<p class="normal">1990 </p>,
<p class="normal">&lt;b&gt;&lt;a href="datasets/Letter+Recognition"&gt;Letter Recognition&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
<p class="normal">Multivariate </p>,
<p class="normal">Classification </p>,
<p class="normal">Integer </p>,
<p class="normal">20000 </p>,
<p class="normal">16 </p>,
<p class="normal">1991 </p>,
<p class="normal">&lt;b&gt;&lt;a href="datasets/Liver+Disorders"&gt;Liver Disorders&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
<p class="normal">Multivariate </p>,
<p class="normal"> </p>,
<p class="normal">Categorical, Integer, Real </p>,
<p class="normal">345 </p>,
<p class="normal">7 </p>,
<p class="normal">1990 </p>,
<p class="normal">&lt;b&gt;&lt;a href="datasets/Logic+Theorist"&gt;Logic Theorist&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
<p class="normal">Domain-Theory </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal">&lt;b&gt;&lt;a href="datasets/Lung+Cancer"&gt;Lung Cancer&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
```

```
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Integer </p>,
 <p class="normal">32 </p>,
 <p class="normal">56 </p>,
 <p class="normal">1992 </p>,
 <p class="normal"><b><a href="datasets/Lymphography">Lymphography</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">148 </p>,
 <p class="normal">18 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">><b><a href="datasets/Mechanical+Analysis">Mechanical Analysis</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">209 </p>,
 <p class="normal">8 </p>,
 <p class="normal">1990 </p>,
 <p class="normal">><b><a href="datasets/Meta-data">Meta-data</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">528 </p>,
 <p class="normal">22 </p>,
 <p class="normal">1996 </p>,
 <p class="normal">><b><a href="datasets/Mobile+Robots">Mobile Robots</a></b></p>,
 <p class="normal">Domain-Theory </p>,
 <p class="normal">> </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">> </p>,
 <p class="normal">> </p>,
 <p class="normal">1995 </p>,
 <p class="normal">><b><a href="datasets/Molecular+Biology+%28Promoter+Gene+Sequences%29">Molecular
Biology (Promoter Gene Sequences)</a></b></p>,
 <p class="normal">Sequential, Domain-Theory </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">106 </p>,
 <p class="normal">58 </p>,
 <p class="normal">1990 </p>,
 <p class="normal">><b><a
href="datasets/Molecular+Biology+%28Protein+Secondary+Structure%29">Molecular Biology (Protein Sec
ondary Structure)</a></b></p>,
 <p class="normal">Sequential </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">128 </p>,
 <p class="normal">> </p>,
 <p class="normal">> </p>,
 <p class="normal">><b><a href="datasets/Molecular+Biology+%28Splice-
junction+Gene+Sequences%29">Molecular Biology (Splice-junction Gene Sequences)</a></b></p>,
 <p class="normal">Sequential, Domain-Theory </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">3190 </p>,
 <p class="normal">61 </p>,
 <p class="normal">1992 </p>,
 <p class="normal">><b><a href="datasets/MONK%27s+Problems">MONK's Problems</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">432 </p>,
 <p class="normal">7 </p>,
 <p class="normal">1992 </p>,
 <p class="normal">><b><a href="datasets/Moral+Reasoner">Moral Reasoner</a></b></p>,
 <p class="normal">Domain-Theory </p>,
 <p class="normal">> </p>,
 <p class="normal">> </p>,
 <p class="normal">202 </p>,
 <p class="normal">> </p>,
 <p class="normal">1994 </p>,
 <p class="normal">><b><a href="datasets/Multiple+Features">Multiple Features</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Integer, Real </p>,
```

```
 <p class="normal">Integer, Real </p>,
 <p class="normal">2000 </p>,
 <p class="normal">649 </p>,
 <p class="normal"> </p>,
 <p class="normal"><b><a href="datasets/Mushroom">Mushroom</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">8124 </p>,
 <p class="normal">22 </p>,
 <p class="normal">1987 </p>,
 <p class="normal"><b><a href="datasets/Musk+%28Version+1%29">Musk (Version 1)</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Integer </p>,
 <p class="normal">476 </p>,
 <p class="normal">168 </p>,
 <p class="normal">1994 </p>,
 <p class="normal"><b><a href="datasets/Musk+%28Version+2%29">Musk (Version 2)</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Integer </p>,
 <p class="normal">6598 </p>,
 <p class="normal">168 </p>,
 <p class="normal">1994 </p>,
 <p class="normal"><b><a href="datasets/Nursery">Nursery</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">12960 </p>,
 <p class="normal">8 </p>,
 <p class="normal">1997 </p>,
 <p class="normal"><b><a href="datasets/Othello+Domain+Theory">Othello Domain Theory</a></b></p>,
 <p class="normal">Domain-Theory </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal">1991 </p>,
 <p class="normal">
```


Page Blocks Classification


```
</b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Integer, Real </p>,
 <p class="normal">5473 </p>,
 <p class="normal">10 </p>,
 <p class="normal">1995 </p>,
 <p class="normal">
```


Optical Recognition of Handwritten Digits


```
</b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Integer </p>,
 <p class="normal">5620 </p>,
 <p class="normal">64 </p>,
 <p class="normal">1998 </p>,
 <p class="normal">
```


Pen-Based Recognition of Handwritten Digits


```
</b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Integer </p>,
 <p class="normal">10992 </p>,
 <p class="normal">16 </p>,
 <p class="normal">1998 </p>,
 <p class="normal">
```


Post-Operative Patient


```
</b></p>
,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">90 </p>,
 <p class="normal">8 </p>,
 <p class="normal">1993 </p>,
 <p class="normal">
```


Primary Tumor


```
</b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">339 </p>,
 <p class="normal">17 </p>,
 <p class="normal">1988 </p>,
```

```
 <p class="normal">1990 </p>,
 <p class="normal"><b><a href="datasets/Prodigy">Prodigy</a></b></p>,
 <p class="normal">Domain-Theory </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"><b><a href="datasets/Qualitative+Structure+Activity+Relationships">Qualitative
Structure Activity Relationships</a></b></p>,
 <p class="normal">Domain-Theory </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"><b><a href="datasets/Quadruped+Mammals">Quadruped Mammals</a></b></p>,
 <p class="normal">Multivariate, Data-Generator </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Real </p>,
 <p class="normal">72 </p>,
 <p class="normal">1992 </p>,
 <p class="normal">72 </p>,
 <p class="normal">1992 </p>,
 <p class="normal"><b><a href="datasets/Servo">Servo</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">167 </p>,
 <p class="normal">4 </p>,
 <p class="normal">1993 </p>,
 <p class="normal"><b><a href="datasets/Shuttle+Landing+Control">Shuttle Landing Control</a></b></
p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">15 </p>,
 <p class="normal">6 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">15 </p>,
 <p class="normal">6 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">15 </p>,
 <p class="normal">6 </p>,
 <p class="normal">1988 </p>,
 <p class="normal"><b><a href="datasets/Solar+Flare">Solar Flare</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">1389 </p>,
 <p class="normal">10 </p>,
 <p class="normal">1989 </p>,
 <p class="normal">1389 </p>,
 <p class="normal">10 </p>,
 <p class="normal">1989 </p>,
 <p class="normal">1389 </p>,
 <p class="normal">10 </p>,
 <p class="normal">1989 </p>,
 <p class="normal">1389 </p>,
 <p class="normal">10 </p>,
 <p class="normal">1989 </p>,
 <p class="normal"><b><a href="datasets/Soybean+%28Large%29">Soybean (Large)</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">307 </p>,
 <p class="normal">35 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">307 </p>,
 <p class="normal">35 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">307 </p>,
 <p class="normal">35 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">307 </p>,
 <p class="normal">35 </p>,
 <p class="normal">1988 </p>,
 <p class="normal"><b><a href="datasets/Soybean+%28Small%29">Soybean (Small)</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">47 </p>,
 <p class="normal">35 </p>,
 <p class="normal">1987 </p>,
 <p class="normal">47 </p>,
 <p class="normal">35 </p>,
 <p class="normal">1987 </p>,
 <p class="normal">47 </p>,
 <p class="normal">35 </p>,
 <p class="normal">1987 </p>,
 <p class="normal">47 </p>,
 <p class="normal">35 </p>,
 <p class="normal">1987 </p>,
 <p class="normal"><b><a href="datasets/Challenger+USA+Space+Shuttle+O-Ring">Challenger USA Space
Shuttle O-Ring</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Integer </p>,
 <p class="normal">23 </p>,
 <p class="normal">4 </p>,
 <p class="normal">1993 </p>,
 <p class="normal">23 </p>,
 <p class="normal">4 </p>,
 <p class="normal">1993 </p>,
 <p class="normal">23 </p>,
 <p class="normal">4 </p>,
 <p class="normal">1993 </p>,
 <p class="normal">23 </p>,
 <p class="normal">4 </p>,
 <p class="normal">1993 </p>,
 <p class="normal"><b><a href="datasets/Low+Resolution+Spectrometer">Low Resolution
Spectrometer</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Integer, Real </p>,
 <p class="normal">531 </p>,
 <p class="normal">102 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">531 </p>,
 <p class="normal">102 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">531 </p>,
 <p class="normal">102 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">531 </p>,
 <p class="normal">102 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">531 </p>,
 <p class="normal">102 </p>,
 <p class="normal">1988 </p>,
 <p class="normal"><b><a href="datasets/Spambase">Spambase</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
```

```
 <p class="normal">Classification </p>,
 <p class="normal">Integer, Real </p>,
 <p class="normal">4601 </p>,
 <p class="normal">57 </p>,
 <p class="normal">1999 </p>,
 <p class="normal"><b><a href="datasets/SPECT+Heart">SPECT Heart</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">267 </p>,
 <p class="normal">22 </p>,
 <p class="normal">2001 </p>,
 <p class="normal"><b><a href="datasets/SPECTF+Heart">SPECTF Heart</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Integer </p>,
 <p class="normal">267 </p>,
 <p class="normal">44 </p>,
 <p class="normal">2001 </p>,
 <p class="normal"><b><a href="datasets/Sponge">Sponge</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Clustering </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">76 </p>,
 <p class="normal">45 </p>,
 <p class="normal"> </p>,
 <p class="normal"><b><a href="datasets/Statlog+Project">Statlog Project</a></b></p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal">1992 </p>,
 <p class="normal">
```



```
<b><a href="datasets/Student+Loan+Relational">Student Loan Relational</a></b></
p>,
 <p class="normal">Domain-Theory </p>,
 <p class="normal">
```



```
 <p class="normal"> </p>,
 <p class="normal">1000 </p>,
 <p class="normal"> </p>,
 <p class="normal">1993 </p>,
 <p class="normal">
```



```
<b><a href="datasets/Teaching+Assistant+Evaluation">Teaching Assistant
Evaluation</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">151 </p>,
 <p class="normal">5 </p>,
 <p class="normal">1997 </p>,
 <p class="normal">
```



```
<b><a href="datasets/Tic-Tac-Toe+Endgame">Tic-Tac-Toe Endgame</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">958 </p>,
 <p class="normal">9 </p>,
 <p class="normal">1991 </p>,
 <p class="normal">
```



```
<b><a href="datasets/Thyroid+Disease">Thyroid Disease</a></b></p>,
 <p class="normal">Multivariate, Domain-Theory </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Real </p>,
 <p class="normal">7200 </p>,
 <p class="normal">21 </p>,
 <p class="normal">1987 </p>,
 <p class="normal">
```



```
<b><a href="datasets/Trains">Trains</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">10 </p>,
 <p class="normal">32 </p>,
 <p class="normal">1994 </p>,
 <p class="normal">
```



```
<b><a href="datasets/University">University</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">285 </p>,
 <p class="normal">17 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">
```



```
<b><a href="datasets/Congressional+Voting+Records">Congressional Voting
```

```
 <p class="normal"><b><a href="datasets/Congressional+Voting+Records">Congressional Voting
Records</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">435 </p>,
 <p class="normal">16 </p>,
 <p class="normal">1987 </p>,
 <p class="normal"><b><a href="datasets/Water+Treatment+Plant">Water Treatment Plant</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Clustering </p>,
 <p class="normal">Integer, Real </p>,
 <p class="normal">527 </p>,
 <p class="normal">38 </p>,
 <p class="normal">1993 </p>,
 <p class="normal"><b><a href="datasets/Waveform+Database+Generator+%28Version+1%29">Waveform
Database Generator (Version 1)</a></b></p>,
 <p class="normal">Multivariate, Data-Generator </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Real </p>,
 <p class="normal">5000 </p>,
 <p class="normal">21 </p>,
 <p class="normal">1988 </p>,
 <p class="normal"><b><a href="datasets/Waveform+Database+Generator+%28Version+2%29">Waveform
Database Generator (Version 2)</a></b></p>,
 <p class="normal">Multivariate, Data-Generator </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Real </p>,
 <p class="normal">5000 </p>,
 <p class="normal">40 </p>,
 <p class="normal">1988 </p>,
 <p class="normal"><b><a href="datasets/Wine">Wine</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Integer, Real </p>,
 <p class="normal">178 </p>,
 <p class="normal">13 </p>,
 <p class="normal">1991 </p>,
 <p class="normal"><b><a href="datasets/Yeast">Yeast</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Real </p>,
 <p class="normal">1484 </p>,
 <p class="normal">8 </p>,
 <p class="normal">1996 </p>,
 <p class="normal"><b><a href="datasets/Zoo">Zoo</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">101 </p>,
 <p class="normal">17 </p>,
 <p class="normal">1990 </p>,
 <p class="normal"><b><a href="datasets/Undocumented">Undocumented</a></b></p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal">><b><a href="datasets/Twenty+Newsgroups">Twenty Newsgroups</a></b></p>,
 <p class="normal">Text </p>,
 <p class="normal">> </p>,
 <p class="normal">> </p>,
 <p class="normal">20000 </p>,
 <p class="normal">> </p>,
 <p class="normal">1999 </p>,
 <p class="normal">><b><a href="datasets/Australian+Sign+Language+signs">Australian Sign Language s
igns</a></b></p>,
 <p class="normal">Multivariate, Time-Series </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Real </p>,
 <p class="normal">6650 </p>,
 <p class="normal">15 </p>,
 <p class="normal">1999 </p>,
 <p class="normal">><b><a
href="datasets/Australian+Sign+Language+signs+%28High+Quality%29">Australian Sign Language signs (
High Quality)</a></b></p>,
 <p class="normal">Multivariate, Time-Series </p>,
```

```
 <p class="normal">Multivariate, Time-Series </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Real </p>,
 <p class="normal">2565 </p>,
 <p class="normal">22 </p>,
 <p class="normal">2002 </p>,
 <p class="normal"><b><a href="datasets/US+Census+Data+%281990%29">US Census Data (1990)</a></b></
p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Clustering </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">2458285 </p>,
 <p class="normal">68 </p>,
 <p class="normal"> </p>,
 <p class="normal"><b><a href="datasets/Census-Income+%28KDD%29">Census-Income (KDD)</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">299285 </p>,
 <p class="normal">40 </p>,
 <p class="normal">2000 </p>,
 <p class="normal">&lt;b&gt;&lt;a href="datasets/Coil+1999+Competition+Data"&gt;Coil 1999 Competition Data&lt;/a&gt;
&lt;/b&gt;&lt;/p&gt;,
 <p class="normal">Multivariate </p>,
 <p class="normal">&lt; /p&gt;,
 <p class="normal">Categorical, Real </p>,
 <p class="normal">340 </p>,
 <p class="normal">17 </p>,
 <p class="normal">1999 </p>,
 <p class="normal">&lt;b&gt;&lt;a href="datasets/Corel+Image+Features"&gt;Corel Image Features&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
 <p class="normal">Multivariate </p>,
 <p class="normal">&lt; /p&gt;,
 <p class="normal">Real </p>,
 <p class="normal">68040 </p>,
 <p class="normal">89 </p>,
 <p class="normal">1999 </p>,
 <p class="normal">&lt;b&gt;&lt;a href="datasets/E.+Coli+Genes"&gt;E. Coli Genes&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
 <p class="normal">Relational </p>,
 <p class="normal">&lt; /p&gt;,
 <p class="normal">&lt; /p&gt;,
 <p class="normal">&lt; /p&gt;,
 <p class="normal">&lt; /p&gt;,
 <p class="normal">2001 </p>,
 <p class="normal">&lt;b&gt;&lt;a href="datasets/EEG+Database"&gt;EEG Database&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
 <p class="normal">Multivariate, Time-Series </p>,
 <p class="normal">&lt; /p&gt;,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">122 </p>,
 <p class="normal">4 </p>,
 <p class="normal">1999 </p>,
 <p class="normal">&lt;b&gt;&lt;a href="datasets/El+Nino"&gt;El Nino&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
 <p class="normal">Spatio-temporal </p>,
 <p class="normal">&lt; /p&gt;,
 <p class="normal">Integer, Real </p>,
 <p class="normal">178080 </p>,
 <p class="normal">12 </p>,
 <p class="normal">1999 </p>,
 <p class="normal">&lt;b&gt;&lt;a href="datasets/Entree+Chicago+Recommendation+Data"&gt;Entree Chicago
Recommendation Data&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
 <p class="normal">Transactional, Sequential </p>,
 <p class="normal">Recommender-Systems </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">50672 </p>,
 <p class="normal">&lt; /p&gt;,
 <p class="normal">2000 </p>,
 <p class="normal">&lt;b&gt;&lt;a href="datasets/CMU+Face+Images"&gt;CMU Face Images&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
 <p class="normal">Image </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Integer </p>,
 <p class="normal">640 </p>,
 <p class="normal">&lt; /p&gt;,
 <p class="normal">1999 </p>,
 <p class="normal">&lt;b&gt;&lt;a href="datasets/Insurance+Company+Benchmark+%28COIL+2000%29"&gt;Insurance
Company Benchmark (COIL 2000)&lt;/a&gt;&lt;/b&gt;&lt;/p&gt;,
 <p class="normal">Multivariate </p>,
 <p class="normal">Regression, Description </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">2000 </p>,
```

```
 <p class="normal">9000 </p>,
 <p class="normal">86 </p>,
 <p class="normal">2000 </p>,
 <p class="normal"><b><a href="datasets/Internet+Usage+Data">Internet Usage Data</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">  </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">10104 </p>,
 <p class="normal">72 </p>,
 <p class="normal">1999 </p>,
 <p class="normal"><b><a href="datasets/IPUMS+Census+Database">IPUMS Census Database</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">  </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">256932 </p>,
 <p class="normal">61 </p>,
 <p class="normal">1999 </p>,
 <p class="normal"><b><a href="datasets/Japanese+Vowels">Japanese Vowels</a></b></p>,
 <p class="normal">Multivariate, Time-Series </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Real </p>,
 <p class="normal">640 </p>,
 <p class="normal">12 </p>,
 <p class="normal">  </p>,
 <p class="normal"><b><a href="datasets/KDD+Cup+1998+Data">KDD Cup 1998 Data</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">191779 </p>,
 <p class="normal">481 </p>,
 <p class="normal">1998 </p>,
 <p class="normal"><b><a href="datasets/KDD+Cup+1999+Data">KDD Cup 1999 Data</a></b></p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">4000000 </p>,
 <p class="normal">42 </p>,
 <p class="normal">1999 </p>,
 <p class="normal"><b><a href="datasets/M.+Tuberculosis+Genes">M. Tuberculosis Genes</a></b></p>,
 <p class="normal">Relational </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">2001 </p>,
 <p class="normal">  <b><a href="datasets/Movie">Movie</a></b></p>,
 <p class="normal">Multivariate, Relational </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">10000 </p>,
 <p class="normal">  </p>,
 <p class="normal">1999 </p>,
 <p class="normal">  <b><a href="datasets/MSNBC.com+Anonymous+Web+Data">MSNBC.com Anonymous Web
Data</a></b></p>,
 <p class="normal">Sequential </p>,
 <p class="normal">  </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">989818 </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">  <b><a href="datasets/NSF+Research+Award+Abstracts+1990-2003">NSF Research Award
Abstracts 1990-2003</a></b></p>,
 <p class="normal">Text </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">129000 </p>,
 <p class="normal">  </p>,
 <p class="normal">2003 </p>,
 <p class="normal">  <b><a href="datasets/Pioneer-1+Mobile+Robot+Data">Pioneer-1 Mobile Robot
Data</a></b></p>,
 <p class="normal">Multivariate, Time-Series </p>,
 <p class="normal">  </p>,
 <p class="normal">Categorical, Real </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">1999 </p>,
 <p class="normal">  <b><a href="datasets/Pseudo+Periodic+Synthetic+Time+Series">Pseudo Periodic Syn
```

```
thetic Time Series</a></b></p>,
 <p class="normal">Univariate, Time-Series </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal">100000 </p>,
 <p class="normal"> </p>,
 <p class="normal">1999 </p>,
 <p class="normal"><b><a href="datasets/Reuters-21578+Text+Categorization+Collection">Reuters-
21578 Text Categorization Collection</a></b></p>,
 <p class="normal">Text </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">21578 </p>,
 <p class="normal">5 </p>,
 <p class="normal">1997 </p>,
 <p class="normal"><b><a href="datasets/Robot+Execution+Failures">Robot Execution Failures</a></b>
</p>,
 <p class="normal">Multivariate, Time-Series </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Integer </p>,
 <p class="normal">463 </p>,
 <p class="normal">90 </p>,
 <p class="normal">1999 </p>,
 <p class="normal"><b><a href="datasets/Synthetic+Control+Chart+Time+Series">Synthetic Control
Chart Time Series</a></b></p>,
 <p class="normal">Time-Series </p>,
 <p class="normal">Classification, Clustering </p>,
 <p class="normal">Real </p>,
 <p class="normal">600 </p>,
 <p class="normal">  </p>,
 <p class="normal">1999 </p>,
 <p class="normal">><b><a href="datasets/Syskill+and+Webert+Web+Page+Ratings">Syskill and Webert We
b Page Ratings</a></b></p>,
 <p class="normal">Multivariate, Text </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">332 </p>,
 <p class="normal">5 </p>,
 <p class="normal">1998 </p>,
 <p class="normal">><b><a href="datasets/UNIX+User+Data">UNIX User Data</a></b></p>,
 <p class="normal">Text, Sequential </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">><b><a href="datasets/Volcanoes+on+Venus+-+JARtool+experiment">Volcanoes on
Venus - JARtool experiment</a></b></p>,
 <p class="normal">Image </p>,
 ...]
```

In [430]:

```
datos[25]
```

Out[430]:

```
<p class="normal"><b><a href="datasets/Abalone">Abalone</a></b></p>
```

In [431]:

```
datos[3309]
```

Out[431]:

```
<p class="normal"> In Collaboration With:</p>
```

In [432]:

```
datos1=datos[25:3500]
```

In [433]:

```
datos1=datos[25:3500]
```

In [434]:

```
len(datos1)
```

Out[434]:

```
3285
```

In [435]:

```
tipo_datos=datos1[1:3500:7]
```

In [436]:

```
tipo_datos
```

Out[436]:

```
[<p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal"> </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Univariate, Text </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate, Data-Generator </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Domain-Theory </p>,
 <p class="normal">Univariate, Time-Series </p>,
 <p class="normal">Multivariate, Spatial </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate, Domain-Theory </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate, Time-Series </p>,
 <p class="normal">Data-Generator </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">  </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate, Time-Series </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Relational </p>
```

```
<p class="normal">Relational </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Data-Generator </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Domain-Theory </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Domain-Theory </p>,
<p class="normal">Sequential, Domain-Theory </p>,
<p class="normal">Sequential </p>,
<p class="normal">Sequential, Domain-Theory </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Domain-Theory </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Domain-Theory </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Domain-Theory </p>,
<p class="normal">Domain-Theory </p>,
<p class="normal">Multivariate, Data-Generator </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal"> </p>,
<p class="normal">Domain-Theory </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Domain-Theory </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Data-Generator </p>,
<p class="normal">Multivariate, Data-Generator </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal"> </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Relational </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Spatio-temporal </p>,
<p class="normal">Transactional, Sequential </p>,
<p class="normal">Image </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Relational </p>,
<p class="normal">Multivariate, Relational </p>,
<p class="normal">Sequential </p>,
```

```
<p class="normal">Sequential </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Univariate, Time-Series </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Time-Series </p>,
<p class="normal">Multivariate, Text </p>,
<p class="normal">Text, Sequential </p>,
<p class="normal">Image </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal"> </p>,
<p class="normal">Domain-Theory </p>,
<p class="normal"> </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Text </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Sequential </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Univariate, Sequential, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Text </p>,
<p class="normal">Domain-Theory </p>,
<p class="normal">Multivariate, Sequential </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Time-Series </p>,
<p class="normal">Time-Series </p>,
<p class="normal">Multivariate, Text, Domain-Theory </p>,
```

```
<p class="normal">Multivariate, Text, Domain-Theory </p>,
<p class="normal">Time-Series, Domain-Theory </p>,
<p class="normal">Multivariate, Text, Domain-Theory </p>,
<p class="normal">Text </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate, Univariate, Text </p>,
<p class="normal">Multivariate, Univariate, Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Univariate </p>,
<p class="normal">Multivariate, Text, Domain-Theory </p>,
<p class="normal">Univariate </p>,
<p class="normal">Univariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Text </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Univariate, Text </p>,
<p class="normal">Sequential </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal"> </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Sequential, Text </p>,
<p class="normal">Multivariate, Univariate, Time-Series </p>,
<p class="normal">Time-Series, Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Sequential </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential, Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Domain-Theory </p>,
<p class="normal">Multivariate, Text </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate, Sequential, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential, Time-Series, Text </p>,
<p class="normal">Univariate </p>,
<p class="normal">Multivariate, Sequential, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">> </p>,
<p class="normal">> </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Univariate, Sequential, Time-Series </p>,
<p class="normal">Univariate, Sequential, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
```

```
<p class="normal">Multivariate </p>,
<p class="normal">Sequential </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential, Time-Series </p>,
<p class="normal">Univariate, Domain-Theory </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal"> </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Text </p>,
<p class="normal">Time-Series </p>,
<p class="normal">Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Time-Series </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate, Univariate, Sequential, Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">> </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Sequential, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential, Time-Series, Domain-Theory </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential, Time-Series </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate, Sequential, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential, Time-Series </p>,
<p class="normal">Multivariate, Sequential, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate, Time-Series, Domain-Theory </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate, Text </p>,
```

```
<p class="normal">Multivariate, Sequential, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate, Sequential </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal"> </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Sequential, Time-Series </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">> </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential, Time-Series </p>,
<p class="normal">> </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Time-Series </p>,
<p class="normal">Text </p>,
<p class="normal">> </p>,
<p class="normal">> </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Sequential, Text </p>,
<p class="normal">Multivariate, Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Text </p>,
<p class="normal">> </p>,
<p class="normal">Sequential </p>,
<p class="normal">Univariate </p>,
<p class="normal">Univariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series, Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential, Time-Series </p>,
<p class="normal">Univariate, Domain-Theory </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Text </p>,
<p class="normal">Sequential </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Sequential </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Univariate </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Text </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Multivariate, Time-Series </p>,
<p class="normal">Multivariate </p>,
<p class="normal">Text </p>,
<p class="normal">Time-Series </p>,
```

```
 <p class="normal">Multivariate, Text </p>,
 <p class="normal">Multivariate, Time-Series </p>,
 <p class="normal">Text </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate, Text </p>,
 <p class="normal">Multivariate, Text </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate, Time-Series </p>,
 <p class="normal">Multivariate, Sequential, Time-Series </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Univariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Univariate </p>,
 <p class="normal">Time-Series </p>,
 <p class="normal">Time-Series </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate, Text </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal"> </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal">Time-Series </p>,
 <p class="normal">Multivariate, Univariate, Sequential, Time-Series </p>,
 <p class="normal">Multivariate, Time-Series </p>,
 <p class="normal">Multivariate, Text </p>,
 <p class="normal">Multivariate, Text </p>,
 <p class="normal">Multivariate </p>,
 <p class="normal"> In Collaboration With:</p>]
```

In [437]:

```python
for x in tipo_datos:
    print(type(x))
```

```
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
```

```
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
```

```
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
```

```
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
```

```
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
```

```
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
```

```
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
<class 'bs4.element.Tag'>
```

```
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Domain-Theory </p>
<p class="normal">Univariate, Time-Series </p>
<p class="normal">Multivariate, Spatial </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Domain-Theory </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Data-Generator </p>
<p class="normal"> </p>
<p class="normal"> </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal"> </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Relational </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Data-Generator </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Domain-Theory </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Domain-Theory </p>
<p class="normal">Sequential, Domain-Theory </p>
<p class="normal">Sequential </p>
<p class="normal">Sequential, Domain-Theory </p>
<p class="normal">Multivariate </p>
<p class="normal">Domain-Theory </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Domain-Theory </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Domain-Theory </p>
<p class="normal">Domain-Theory </p>
<p class="normal">Multivariate, Data-Generator </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate
```

```
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Domain-Theory </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Data-Generator </p>
<p class="normal">Multivariate, Data-Generator </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal"> </p>
<p class="normal">Text </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Relational </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Spatio-temporal </p>
<p class="normal">Transactional, Sequential </p>
<p class="normal">Image </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Relational </p>
<p class="normal">Multivariate, Relational </p>
<p class="normal">Sequential </p>
<p class="normal">Text </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Univariate, Time-Series </p>
<p class="normal">Text </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Time-Series </p>
<p class="normal">Multivariate, Text </p>
<p class="normal">Text, Sequential </p>
<p class="normal">Image </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Text </p>
<p class="normal">Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Sequential </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential </p>
<p class="normal">Multivariate </p>
```

```
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Univariate, Sequential, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Text </p>
<p class="normal">Domain-Theory </p>
<p class="normal">Multivariate, Sequential </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Time-Series </p>
<p class="normal">Time-Series </p>
<p class="normal">Multivariate, Text, Domain-Theory </p>
<p class="normal">Time-Series, Domain-Theory </p>
<p class="normal">Multivariate, Text, Domain-Theory </p>
<p class="normal">Text </p>
<p class="normal">Text </p>
<p class="normal">Multivariate, Univariate, Text </p>
<p class="normal">Multivariate, Univariate, Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Univariate </p>
<p class="normal">Multivariate, Text, Domain-Theory </p>
<p class="normal">Univariate </p>
<p class="normal">Univariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Text </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Univariate, Text </p>
<p class="normal">Sequential </p>
<p class="normal">Text </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal"> </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Sequential, Text </p>
<p class="normal">Multivariate, Univariate, Time-Series </p>
<p class="normal">Time-Series, Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Sequential </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential, Text </p>
<p class="normal">Multivariate </p>
```

```
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Domain-Theory </p>
<p class="normal">Multivariate, Text </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate, Sequential, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential, Time-Series, Text </p>
<p class="normal">Univariate </p>
<p class="normal">Multivariate, Sequential, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal"> </p>
<p class="normal"> </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Univariate, Sequential, Time-Series </p>
<p class="normal">Univariate, Sequential, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Sequential </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential, Time-Series </p>
<p class="normal">Univariate, Domain-Theory </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">  </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Text </p>
<p class="normal">Time-Series </p>
<p class="normal">Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Time-Series </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate, Univariate, Sequential, Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">  </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Sequential, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential, Time-Series, Domain-Theory </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
```

```
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential, Time-Series </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate, Sequential, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential, Time-Series </p>
<p class="normal">Multivariate, Sequential, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate, Time-Series, Domain-Theory </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Text </p>
<p class="normal">Multivariate, Text </p>
<p class="normal">Multivariate, Sequential, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Text </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate, Sequential </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal"> </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Sequential, Time-Series </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal"> </p>
<p class="normal">Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential, Time-Series </p>
<p class="normal"> </p>
<p class="normal">Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Time-Series </p>
<p class="normal">Text </p>
<p class="normal"> </p>
<p class="normal"> </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Sequential, Text </p>
<p class="normal">Multivariate, Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
```

```
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Text </p>
<p class="normal"> </p>
<p class="normal">Sequential </p>
<p class="normal">Univariate </p>
<p class="normal">Univariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series, Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential, Time-Series </p>
<p class="normal">Univariate, Domain-Theory </p>
<p class="normal">Multivariate </p>
<p class="normal">Text </p>
<p class="normal">Sequential </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Sequential </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Univariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Text </p>
<p class="normal">Time-Series </p>
<p class="normal">Multivariate, Text </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Text </p>
<p class="normal">Multivariate, Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate, Sequential, Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Univariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal">Univariate </p>
<p class="normal">Time-Series </p>
<p class="normal">Time-Series </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate, Text </p>
<p class="normal">Multivariate </p>
<p class="normal">Multivariate </p>
<p class="normal"> </p>
<p class="normal">Multivariate </p>
<p class="normal">Time-Series </p>
<p class="normal">Multivariate, Univariate, Sequential, Time-Series </p>
<p class="normal">Multivariate, Time-Series </p>
<p class="normal">Multivariate, Text </p>
<p class="normal">Multivariate, Text </p>
<p class="normal">Multivariate </p>
<p class="normal"> In Collaboration With:</p>
```

In [439]:

```
prub = "normal [\w.-]+"
for x in tipo_datos:
    a = re.findall(prub,x)
    print (a)
```

```
---------------------------------------------------------------------------
TypeError                                 Traceback (most recent call last)
<ipython-input-439-8fbe072d5dda> in <module>
      1 prub = "normal [\w.-]+"
```

```
      1 prub = "normal [\w.-]+
      2 for x in tipo_datos:
----> 3     a = re.findall(prub,x)
      4     print (a)

~\Anaconda3\lib\re.py in findall(pattern, string, flags)
    221
    222     Empty matches are included in the result."""
--> 223     return _compile(pattern, flags).findall(string)
    224
    225 def finditer(pattern, string, flags=0):

TypeError: expected string or bytes-like object
```

In [440]:

```
tipo_datos_string=str(tipo_datos)
```

In [441]:

```
tipo_datos_string
```

Out[441]:

```
'[<p class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Univariate, Text\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate, Data-Generator\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Domain-Theory\xa0</p>, <p
class="normal">Univariate, Time-Series\xa0</p>, <p class="normal">Multivariate, Spatial\xa0</p>, <
p class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate, Domain-Theory\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate, Time-Series\xa0</p>, <p
class="normal">Data-Generator\xa0</p>, <p class="normal">\xa0</p>, <p class="normal">\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate, Time-Series\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Relational\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate, Data-Generator\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Domain-Theory\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Domain-Theory\xa0</p>, <p
class="normal">Sequential, Domain-Theory\xa0</p>, <p class="normal">Sequential\xa0</p>, <p
class="normal">Sequential, Domain-Theory\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Domain-Theory\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Domain-Theory\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Domain-Theory\xa0</p>, <p class="normal">Domain-Theory\xa0</p>, <p
class="normal">Multivariate, Data-Generator\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">\xa0</p>, <p class="normal">Domain-Theory\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate, Domain-Theory\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
```

class="normal">Multivariate\xa0 , Multivariate, Data-Generator\xa0 , <p
class="normal">Multivariate, Data-Generator\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">>\xa0 , Text\xa0 , Multivariate, Time-
Series\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Relational\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">Spatio-temporal\xa0 , Transactional, Sequential\xa0 , <p cl
ass="normal">Image\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Relational\xa0 , <p
class="normal">Multivariate, Relational\xa0 , Sequential\xa0 , <p
class="normal">Text\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">Univariate, Time-Series\xa0 , Text\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Time-Series\xa0 , <p
class="normal">Multivariate, Text\xa0 , Text, Sequential\xa0 , <p
class="normal">Image\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">>\xa0 , Domain-Theory\xa0 , >\xa0 , <p c
lass="normal">Multivariate\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate, Sequential\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Text\xa0 , Text\xa0 , <p
class="normal">Multivariate\xa0 , Sequential\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate, Sequential, Time-
Series\xa0 , Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Time-Series\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Sequential\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Sequential\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Text\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Sequential\xa0 , Multivariate, Time-
Series\xa0 , Univariate, Sequential, Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Text\xa0 , <p
class="normal">Domain-Theory\xa0 , Multivariate, Sequential\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Time-Series\xa0 , Time-Series\xa0 , <p
class="normal">Multivariate, Text, Domain-Theory\xa0 , Time-Series, Domain-Th
eory\xa0 , Multivariate, Text, Domain-Theory\xa0 , <p
class="normal">Text\xa0 , Text\xa0 , Multivariate,
Univariate, Text\xa0 , Multivariate, Univariate, Text\xa0 , <p
class="normal">Multivariate\xa0 , Text\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Univariate\xa0 , <p
class="normal">Multivariate, Text, Domain-Theory\xa0 , Univariate\xa0 , <p
class="normal">Univariate\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate, Text\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Univariate, Text\xa0 , Sequential\xa0 , <p cl
ass="normal">Text\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">>\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Sequential, Text\xa0 , <p c
lass="normal">Multivariate, Univariate, Time-Series\xa0 , Time-Series,
Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Sequential\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate, Sequential, Text\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Sequential, Time-Series\xa0 , <p

class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Domain-Theory\xa0 , <p
class="normal">Multivariate, Text\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">Multivariate, Sequential, Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Sequential, Time-Series, Text\xa0 , <p
class="normal">Univariate\xa0 , Multivariate, Sequential, Time-
Series\xa0 , Multivariate\xa0 , Multivariate, Time-
Series\xa0 , >\xa0 , >\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Multivariate\xa0 , <p
class="normal">Univariate, Sequential, Time-Series\xa0 , Univariate,
Sequential, Time-Series\xa0 , Multivariate\xa0 , <p
class="normal">Text\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Text\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Sequential\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Sequential, Time-Series\xa0 , Univariate,
Domain-Theory\xa0 , Multivariate\xa0 , Multivariate,
Time-Series\xa0 , >\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Multivariate, Time-
Series\xa0 , Multivariate\xa0 , Text\xa0 , <p
class="normal">Time-Series\xa0 , Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate, Sequential\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Time-Series\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Multivariate, Univariate,
Sequential, Text\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">>\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Text\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Sequential, Time-Series\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Sequential, Time-Series, Domain-Theory\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate, Sequential, Time-
Series\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">Multivariate, Sequential, Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate, Sequential, Time-
Series\xa0 , Multivariate, Sequential, Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Sequential, Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Multivariate, Time-Series,
Domain-Theory\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Time-Series\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate, Sequential, Time-
Series\xa0 , Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Text\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Text\xa0 , <p
class="normal">Multivariate, Text\xa0 , Multivariate, Sequential, Time-
Series\xa0 , Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Text\xa0 , Multivariate, Time-Series\xa0 , <p
class="normal">Multivariate, Sequential\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Multivariate, Time-
Series\xa0 , Multivariate, Time-Series\xa0 , >\xa0 , <
p class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Sequential, Time-Series\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , >\xa0 , <p
class="normal">Text\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate, Sequential, Time-Series\xa0 , >\xa0 , <p
class="normal">Text\xa0 , Multivariate\xa0 , <p
class="normal">Multivariate\xa0 , Text\xa0 , <p
class="normal">Multivariate\xa0 , Time-Series\xa0 , <p
class="normal">Text\xa0 , >\xa0 , >\xa0 , <p
class="normal">Multivariate, Time-Series\xa0 , Multivariate\xa0 , <p

```
class="normal">Sequential, Text\xa0</p>, <p class="normal">Multivariate, Text\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Text\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Text\xa0</p>, <p class="normal">\xa0</p>, <p class="normal">Sequential\xa0</p>, <p
class="normal">Univariate\xa0</p>, <p class="normal">Univariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate, Time-Series, Text\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate, Sequential, Time-
Series\xa0</p>, <p class="normal">Univariate, Domain-Theory\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Text\xa0</p>, <p
class="normal">Sequential\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate, Sequential\xa0</p>, <p
class="normal">Multivariate, Time-Series\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate, Time-Series\xa0</p>, <p class="normal">Univariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate, Text\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate, Time-Series\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Text\xa0</p>, <p class="normal">Time-
Series\xa0</p>, <p class="normal">Multivariate, Text\xa0</p>, <p class="normal">Multivariate,
Time-Series\xa0</p>, <p class="normal">Text\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate, Text\xa0</p>, <p
class="normal">Multivariate, Text\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate, Time-Series\xa0</p>, <p
class="normal">Multivariate, Sequential, Time-Series\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Univariate\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Univariate\xa0</p>, <p class="normal">Time-
Series\xa0</p>, <p class="normal">Time-Series\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate, Text\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">\xa0</p>, <p
class="normal">Multivariate\xa0</p>, <p class="normal">Time-Series\xa0</p>, <p
class="normal">Multivariate, Univariate, Sequential, Time-Series\xa0</p>, <p
class="normal">Multivariate, Time-Series\xa0</p>, <p class="normal">Multivariate, Text\xa0</p>, <p
class="normal">Multivariate, Text\xa0</p>, <p class="normal">Multivariate\xa0</p>, <p
class="normal">\xa0In Collaboration With:</p>]'
```

In [442]:

```
pat="([A-Z]+[a-z-A-Z]+)"
re.findall(pat,tipo_datos_string)
```

Out[442]:

```
['Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Univariate',
 'Text',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Data-Generator',
 'Multivariate',
 'Multivariate',
 'Domain-Theory',
 'Univariate',
 'Time-Series',
 'Multivariate',
 'Spatial',
 'Multivariate',
 'Multivariate',
 'Domain-Theory',
```

```
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Time-Series',
    'Data-Generator',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Time-Series',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Relational',
    'Multivariate',
    'Multivariate',
    'Data-Generator',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Domain-Theory',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Domain-Theory',
    'Sequential',
    'Domain-Theory',
    'Sequential',
    'Sequential',
    'Domain-Theory',
    'Multivariate',
    'Domain-Theory',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Domain-Theory',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Domain-Theory',
    'Domain-Theory',
    'Multivariate',
    'Data-Generator',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Domain-Theory',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Domain-Theory',
    'Multivariate',
    'Multivariate',
```

```
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Data-Generator',
    'Multivariate',
    'Data-Generator',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Text',
    'Multivariate',
    'Time-Series',
    'Multivariate',
    'Time-Series',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Relational',
    'Multivariate',
    'Time-Series',
    'Spatio-temporal',
    'Transactional',
    'Sequential',
    'Image',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Time-Series',
    'Multivariate',
    'Multivariate',
    'Relational',
    'Multivariate',
    'Relational',
    'Sequential',
    'Text',
    'Multivariate',
    'Time-Series',
    'Univariate',
    'Time-Series',
    'Text',
    'Multivariate',
    'Time-Series',
    'Time-Series',
    'Multivariate',
    'Text',
    'Text',
    'Sequential',
    'Image',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Domain-Theory',
    'Multivariate',
    'Multivariate',
    'Time-Series',
    'Multivariate',
    'Time-Series',
    'Multivariate',
    'Multivariate',
    'Multivariate',
    'Sequential',
    'Multivariate',
    'Multivariate',
    'Text',
    'Text',
    'Multivariate',
    'Sequential',
    'Multivariate',
    'Multivariate',
```

```
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Text',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Multivariate',
 'Time-Series',
 'Univariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Text',
 'Domain-Theory',
 'Multivariate',
 'Sequential',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Time-Series',
 'Multivariate',
 'Text',
 'Domain-Theory',
 'Time-Series',
 'Domain-Theory',
 'Multivariate',
 'Text',
 'Domain-Theory',
 'Text',
 'Text',
 'Multivariate',
 'Univariate',
 'Text',
 'Multivariate',
 'Univariate',
 'Text',
 'Multivariate',
 'Text',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Univariate',
 'Multivariate',
 'Text',
```

```
 'Text',
 'Domain-Theory',
 'Univariate',
 'Univariate',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Text',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Univariate',
 'Text',
 'Sequential',
 'Text',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Sequential',
 'Text',
 'Multivariate',
 'Univariate',
 'Time-Series',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Text',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Domain-Theory',
 'Multivariate',
 'Text',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Text',
 'Univariate',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Multivariate'
```

```
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Univariate',
 'Sequential',
 'Time-Series',
 'Univariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Text',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Text',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Univariate',
 'Domain-Theory',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Text',
 'Time-Series',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Univariate',
 'Sequential',
 'Text',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Text',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Domain-Theory',
 'Multivariate',
 'Multivariate',
 'Time-Series',
```

```
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Time-Series',
 'Domain-Theory',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Text',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Text',
 'Multivariate',
 'Text',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Text',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Sequential',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
```

```
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Text',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Text',
 'Multivariate',
 'Multivariate',
 'Text',
 'Multivariate',
 'Time-Series',
 'Text',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Sequential',
 'Text',
 'Multivariate',
 'Text',
 'Multivariate',
 'Text',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Text',
 'Sequential',
 'Univariate',
 'Univariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Text',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Univariate',
 'Domain-Theory',
 'Multivariate',
 'Text',
 'Sequential',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Sequential',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Univariate',
 'Multivariate',
 'Multivariate',
 'Text',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Text',
 'Time-Series',
 'Multivariate',
 'Text',
 'Multivariate',
 'Time-Series',
 'Text',
 'Multivariate',
```

```
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Text',
 'Multivariate',
 'Text',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Univariate',
 'Multivariate',
 'Multivariate',
 'Univariate',
 'Time-Series',
 'Time-Series',
 'Multivariate',
 'Multivariate',
 'Text',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Univariate',
 'Sequential',
 'Time-Series',
 'Multivariate',
 'Time-Series',
 'Multivariate',
 'Text',
 'Multivariate',
 'Text',
 'Multivariate',
 'In',
 'Collaboration',
 'With']
```

In [443]:

```
prub= "\g<normal> [\w.-]+"
re.findall(prub,tipo_datos_string)
```

```
---------------------------------------------------------------------------
error                                     Traceback (most recent call last)
<ipython-input-443-acd9f63f9983> in <module>
      1 prub= "\g<normal> [\w.-]+"
----> 2 re.findall(prub,tipo_datos_string)

~\Anaconda3\lib\re.py in findall(pattern, string, flags)
    221
    222     Empty matches are included in the result."""
--> 223     return _compile(pattern, flags).findall(string)
    224
    225 def finditer(pattern, string, flags=0):

~\Anaconda3\lib\re.py in _compile(pattern, flags)
    284     if not sre_compile.isstring(pattern):
    285         raise TypeError("first argument must be string or compiled pattern")
--> 286     p = sre_compile.compile(pattern, flags)
    287     if not (flags & DEBUG):
    288         if len(_cache) >= _MAXCACHE:

~\Anaconda3\lib\sre_compile.py in compile(p, flags)
    762     if isstring(p):
    763         pattern = p
--> 764         p = sre_parse.parse(p, flags)
    765     else:
    766         pattern = None

~\Anaconda3\lib\sre_parse.py in parse(str, flags, pattern)
```

```
~\Anaconda3\lib\sre_parse.py in parse(str, flags, pattern)
    928
    929     try:
--> 930         p = _parse_sub(source, pattern, flags & SRE_FLAG_VERBOSE, 0)
    931     except Verbose:
    932         # the VERBOSE flag was switched on inside the pattern.  to be

~\Anaconda3\lib\sre_parse.py in _parse_sub(source, state, verbose, nested)
    424     while True:
    425         itemsappend(_parse(source, state, verbose, nested + 1,
--> 426                            not nested and not items))
    427         if not sourcematch("|"):
    428             break

~\Anaconda3\lib\sre_parse.py in _parse(source, state, verbose, nested, first)
    505
    506         if this[0] == "\\":
--> 507             code = _escape(source, this, state)
    508             subpatternappend(code)
    509

~\Anaconda3\lib\sre_parse.py in _escape(source, escape, state)
    400         if len(escape) == 2:
    401             if c in ASCIILETTERS:
--> 402                 raise source.error("bad escape %s" % escape, len(escape))
    403             return LITERAL, ord(escape[1])
    404     except ValueError:

error: bad escape \g at position 0
```

In [444]:

```
tipo_tarea=datos1[2:3310:7]
```

In [445]:

```
tipo_tarea
```

Out[445]:

```
[<p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Recommender-Systems </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification, Regression </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification
```

```
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Function-Learning </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal"> </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Relational-Learning </p>,
<p class="normal"> </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal"> </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification
```

```
<p class="normal">Classification </p>,
<p class="normal">Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal">Recommender-Systems </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression, Description </p>,
<p class="normal"> </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Causal-Discovery </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification, Causal-Discovery </p>,
<p class="normal">Clustering </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Regression, Description </p>,
<p class="normal">Classification, Regression </p>,
```

```
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification </p>,
<p class="normal"> </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression, Clustering, Causal-Discovery </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Regression, Clustering </p>,
<p class="normal">Classification, Regression, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression, Clustering </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Regression, Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Clustering </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification, Regression, Clustering, Causa </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression </p>,
<p class="normal">Regression </p>,
<p class="normal"> </p>,
```

```
<p class="normal">Classification </p>,
<p class="normal"> </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification, Regression, Clustering </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering, Causal-Discovery </p>,
<p class="normal">Classification, Regression, Clustering </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Clustering </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Clustering </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Regression, Clustering </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification, Causal-Discovery </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression, Clustering, Causal-Discovery </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Clustering, Causal-Discovery </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification, Regression, Clustering </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification, Regression, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Regression, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Causal-Discovery </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
```

```
<p class="normal">Classification, Regression, Clustering </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression </p>,
<p class="normal">Causal-Discovery </p>,
<p class="normal">Clustering </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification, Clustering, Causal-Discovery </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification, Regression, Clustering </p>,
<p class="normal">Classification, Regression, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Causal-Discovery </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification, Regression, Clustering </p>,
<p class="normal">Clustering </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal"> </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Clustering </p>,
<p class="normal">Clustering </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification </p>,
<p class="normal">Regression </p>,
<p class="normal">Classification </p>,
<p class="normal">Classification, Clustering </p>,
<p class="normal">Classification, Clustering </p>,
```

```
 <p class="normal">Classification, Regression </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Clustering </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Clustering </p>,
 <p class="normal">Classification, Clustering </p>,
 <p class="normal">Classification, Clustering </p>,
 <p class="normal">Classification, Clustering </p>,
 <p class="normal">Classification, Regression </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Classification, Regression </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Classification, Clustering </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification, Regression </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Recommendation </p>,
 <p class="normal">Classification, Regression, Clustering </p>,
 <p class="normal">Classification, Regression, Clustering </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Classification, Regression </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification, Clustering </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification, Regression </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification, Clustering </p>,
 <p class="normal">Regression </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification </p>,
 <p class="normal">Classification, Regression, Clustering </p>,
 <p class="normal">Classification, Regression </p>,
 <p class="normal">Classification, Clustering </p>,
 <p class="normal">Classification, Clustering </p>,
 <p class="normal">Classification </p>]
```

In [446]:

```
tipo_tarea_s=str(tipo_tarea)
```

In [447]:

```
tipo_tarea_s
```

Out[447]:

```
'[<p class="normal">Classification\xa0</p>, <p class="normal">Classification\xa0</p>, <p
class="normal">Classification\xa0</p>, <p class="normal">Recommender-Systems\xa0</p>, <p
class="normal">Classification\xa0</p>, <p class="normal">Classification\xa0</p>, <p
class="normal">Classification\xa0</p>, <p class="normal">Classification\xa0</p>, <p
class="normal">Regression\xa0</p>, <p class="normal">Regression\xa0</p>, <p
class="normal">Classification\xa0</p>, <p class="normal">Classification\xa0</p>, <p
class="normal">Classification\xa0</p>, <p class="normal">Classification\xa0</p>, <p
class="normal">Classification\xa0</p>, <p class="normal">Classification, Regression\xa0</p>, <p cl
ass="normal">Classification\xa0</p>, <p class="normal">Classification\xa0</p>, <p
class="normal">Classification\xa0</p>, <p class="normal">Classification\xa0</p>, <p
class="normal">Classification\xa0</p>, <p class="normal">Classification\xa0</p>, <p
class="normal">Classification\xa0</p>, <p class="normal">\xa0</p>, <p class="normal">\xa0</p>, <p
class="normal">Classification\xa0</p>, <p class="normal">Classification\xa0</p>, <p
class="normal">Classification\xa0</p>, <p class="normal">Regression\xa0</p>, <p
class="normal">Classification\xa0</p>, <p class="normal">Classification\xa0</p>, <p
```

class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">\xa0 , \xa0 , \xa0 , <p
class="normal">\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Function-Learning\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , \xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Relational-Learning\xa0 , <p
class="normal">\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">\xa0 , \xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , \xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">\xa0 , \xa0 , Classification\xa0 , <p
class="normal">Regression\xa0 , Classification\xa0 , <p
class="normal">Regression\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Regression\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Clustering\xa0 , \xa0 , \xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Clustering\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">\xa0 , \xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Clustering\xa0 , <p
class="normal">Classification\xa0 , \xa0 , \xa0 , <p
class="normal">\xa0 , \xa0 , \xa0 , <p
class="normal">Recommender-Systems\xa0 , Classification\xa0 , <p
class="normal">Regression, Description\xa0 , \xa0 , <p
class="normal">\xa0 , Classification\xa0 , <p
class="normal">Regression\xa0 , Classification\xa0 , <p
class="normal">\xa0 , \xa0 , \xa0 , <p
class="normal">\xa0 , \xa0 , \xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification, Clustering\xa0 , Classification\xa0 , <p cl
ass="normal">\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , \xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">\xa0 , \xa0 , \xa0 , <p
class="normal">\xa0 , \xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Regression\xa0 , <p
class="normal">Classification\xa0 , Clustering\xa0 , <p
class="normal">Regression\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Causal-Discovery\xa0 , Classification\xa0 , <p
class="normal">Classification, Clustering\xa0 , Classification\xa0 , <p cl
ass="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification, Causal-Discovery\xa0 , Clustering\xa0 , <p
class="normal">Classification, Clustering\xa0 , Regression\xa0 , <p
class="normal">Regression\xa0 , Classification\xa0 , <p
class="normal">Classification, Regression\xa0 , Classification\xa0 , <p cl
ass="normal">Classification\xa0 , Regression\xa0 , <p
class="normal">Classification\xa0 , \xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p

class="normal">Regression\xa0 , Classification\xa0 , <p
class="normal">\xa0 , Regression\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Regression\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Regression, Clustering, Causal-Discovery\xa0 , Classification,
Clustering\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification, Regression,
Clustering\xa0 , Classification, Regression, Clustering\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification, Clustering\xa0 , Classification\xa0 , <p cl
ass="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">\xa0 , Classification\xa0 , Regression, C
lustering\xa0 , Classification, Clustering\xa0 , <p
class="normal">Classification\xa0 , \xa0 , <p
class="normal">Classification\xa0 , Classification, Clustering\xa0 , <p cl
ass="normal">Classification\xa0 , Classification, Regression\xa0 , <p
class="normal">Regression\xa0 , Classification, Regression\xa0 , <p
class="normal">Classification\xa0 , Regression, Clustering\xa0 , <p
class="normal">Classification, Regression\xa0 , Regression,
Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification, Clustering\xa0 , Classification,
Clustering\xa0 , Classification\xa0 , <p
class="normal">Clustering\xa0 , Classification, Clustering\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Regression\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification, Clustering\xa0 , Classification, Regression, C
lustering, Causa\xa0 , Classification, Clustering\xa0 , <p
class="normal">Regression\xa0 , Classification\xa0 , <p
class="normal">Regression\xa0 , Regression\xa0 , <p
class="normal">\xa0 , Classification\xa0 , \xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification, Clustering\xa0 , Classification\xa0 , <p cl
ass="normal">Classification, Clustering\xa0 , Classification,
Clustering\xa0 , Classification\xa0 , Classification,
Clustering\xa0 , Classification, Regression, Clustering\xa0 , <p
class="normal">Regression\xa0 , Classification, Clustering\xa0 , <p
class="normal">Classification\xa0 , Regression\xa0 , <p
class="normal">Classification\xa0 , Classification, Clustering\xa0 , <p cl
ass="normal">Classification\xa0 , Classification, Clustering, Causal-
Discovery\xa0 , Classification, Regression, Clustering\xa0 , <p
class="normal">Classification, Regression\xa0 , Classification,
Clustering\xa0 , Classification, Clustering\xa0 , <p
class="normal">Regression\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Clustering\xa0 , <p
class="normal">Classification, Regression\xa0 , Classification,
Regression\xa0 , Classification, Regression\xa0 , <p
class="normal">Classification\xa0 , Classification, Clustering\xa0 , <p cl
ass="normal">Classification\xa0 , Clustering\xa0 , <p
class="normal">Classification, Regression\xa0 , Regression\xa0 , <p
class="normal">Classification, Clustering\xa0 , Classification,
Regression\xa0 , Classification\xa0 , Classification,
Regression\xa0 , Regression, Clustering\xa0 , <p
class="normal">Classification, Regression\xa0 , Classification, Causal-
Discovery\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification, Clustering\xa0 , <p cl
ass="normal">Classification\xa0 , Regression\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification, Regression\xa0 , <p cl
ass="normal">Classification\xa0 , Regression, Clustering, Causal-
Discovery\xa0 , Regression\xa0 , Classification\xa0 ,
Classification\xa0 , Classification, Clustering\xa0 , <p
class="normal">Clustering, Causal-Discovery\xa0 , Classification,
Regression\xa0 , Classification\xa0 , Classification,
Clustering\xa0 , Classification, Regression, Clustering\xa0 , <p
class="normal">Classification, Clustering\xa0 , Classification, Regression, C
lustering\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification, Regression,
Clustering\xa0 , Classification\xa0 , <p

Clustering\xa0 , Classification\xa0 , Classification, Clustering\xa0 , <p cl
ass="normal">Causal-Discovery\xa0 , Classification, Regression\xa0 , <p cl
ass="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification, Regression, Clustering\xa0 , Classification, C
lustering\xa0 , Regression\xa0 , Classification,
Regression\xa0 , Classification\xa0 , <p
class="normal">Regression\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Regression\xa0 , <p
class="normal">Causal-Discovery\xa0 , Clustering\xa0 , <p
class="normal">Classification, Clustering\xa0 , Classification\xa0 , <p cl
ass="normal">Regression\xa0 , Classification, Clustering, Causal-
Discovery\xa0 , Classification, Regression\xa0 , <p
class="normal">Classification, Regression, Clustering\xa0 , Classification, R
egression, Clustering\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Regression\xa0 , <p
class="normal">Classification, Regression\xa0 , Classification\xa0 , <p cl
ass="normal">Classification\xa0 , Causal-Discovery\xa0 , <p
class="normal">Classification, Clustering\xa0 , Regression\xa0 , <p
class="normal">Classification, Clustering\xa0 , Classification\xa0 , <p cl
ass="normal">Regression\xa0 , Classification, Clustering\xa0 , <p
class="normal">Classification, Regression\xa0 , Classification\xa0 , <p cl
ass="normal">Regression\xa0 , Classification, Regression, Clustering\xa0 ,
Clustering\xa0 , Classification, Regression\xa0 , <p cla
ss="normal">\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification, Clustering\xa0 , <p cl
ass="normal">Classification\xa0 , Classification, Clustering\xa0 , <p
class="normal">Classification\xa0 , Classification, Clustering\xa0 , <p cl
ass="normal">Classification, Clustering\xa0 , Classification,
Clustering\xa0 , Classification\xa0 , <p
class="normal">Regression\xa0 , Classification, Regression\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification, Regression\xa0 , Classification\xa0 , <p cl
ass="normal">Clustering\xa0 , Clustering\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification, Regression\xa0 , <p cl
ass="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Regression\xa0 , <p
class="normal">Classification\xa0 , Classification, Clustering\xa0 , <p cl
ass="normal">Classification, Clustering\xa0 , Classification,
Regression\xa0 , Regression\xa0 , Clustering\xa0 , <p
class="normal">Classification\xa0 , Regression\xa0 , <p
class="normal">Clustering\xa0 , Classification, Clustering\xa0 , <p
class="normal">Classification, Clustering\xa0 , Classification,
Clustering\xa0 , Classification, Regression\xa0 , <p
class="normal">Regression\xa0 , Classification, Regression\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Regression\xa0 , Classification, Clustering\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification, Regression\xa0 , <p cl
ass="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Recommendation\xa0 , Classification, Regression,
Clustering\xa0 , Classification, Regression, Clustering\xa0 , <p
class="normal">Classification\xa0 , Regression\xa0 , <p
class="normal">Classification, Regression\xa0 , Regression\xa0 , <p
class="normal">Classification\xa0 , Classification, Clustering\xa0 , <p cl
ass="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification, Regression\xa0 , Classification\xa0 , <p cl
ass="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification, Clustering\xa0 , <p cl
ass="normal">Regression\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification\xa0 , <p
class="normal">Classification\xa0 , Classification, Regression,
Clustering\xa0 , Classification, Regression\xa0 , <p
class="normal">Classification, Clustering\xa0 , Classification,
Clustering\xa0 , Classification\xa0]'

In [448]:

```
re.findall(pat,tipo_tarea_s)
```

Out[448]:

```
['Classification',
 'Classification',
 'Classification',
 'Recommender-Systems',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Regression',
 'Regression',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Regression',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Regression',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Function-Learning',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Relational-Learning',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Regression',
 'Classification',
 'Regression',
 'Classification',
 'Classification',
```

```
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Recommender-Systems',
    'Classification',
    'Regression',
    'Description',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Regression',
    'Classification',
    'Clustering',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Causal-Discovery',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Causal-Discovery',
    'Clustering',
    'Classification',
    'Clustering',
    'Regression',
    'Regression',
    'Classification',
    'Classification',
    'Regression',
    'Classification',
```

```
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Regression',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Regression',
    'Clustering',
    'Causal-Discovery',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Classification',
    'Regression',
    'Clustering',
    'Classification',
    'Regression',
    'Clustering',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Regression',
    'Clustering',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Regression',
    'Regression',
    'Classification',
    'Regression',
    'Classification',
    'Regression',
    'Clustering',
    'Classification',
    'Regression',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Clustering',
```

```
    'Classification',
    'Clustering',
    'Classification',
    'Clustering',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Regression',
    'Clustering',
    'Causa',
    'Classification',
    'Clustering',
    'Regression',
    'Classification',
    'Regression',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Regression',
    'Clustering',
    'Regression',
    'Classification',
    'Clustering',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Clustering',
    'Causal-Discovery',
    'Classification',
    'Regression',
    'Clustering',
    'Classification',
    'Regression',
    'Classification',
    'Clustering',
    'Classification',
    'Clustering',
    'Regression',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Regression',
    'Classification',
    'Regression',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
```

```
    'Clustering',
    'Classification',
    'Regression',
    'Regression',
    'Classification',
    'Clustering',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Regression',
    'Regression',
    'Clustering',
    'Classification',
    'Regression',
    'Classification',
    'Causal-Discovery',
    'Classification',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Regression',
    'Classification',
    'Regression',
    'Clustering',
    'Causal-Discovery',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Clustering',
    'Clustering',
    'Causal-Discovery',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Regression',
    'Clustering',
    'Classification',
    'Clustering',
    'Classification',
    'Regression',
    'Clustering',
    'Classification',
    'Classification',
    'Classification',
    'Regression',
    'Clustering',
    'Classification',
    'Classification',
    'Classification',
    'Clustering',
    'Causal-Discovery',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Regression',
    'Clustering',
    'Classification',
    'Clustering',
    'Regression',
    'Classification',
    'Regression',
    'Classification',
    'Regression',
    'Classification',
```

```
    'Classification',
    'Classification',
    'Classification',
    'Regression',
    'Causal-Discovery',
    'Clustering',
    'Classification',
    'Clustering',
    'Classification',
    'Regression',
    'Classification',
    'Clustering',
    'Causal-Discovery',
    'Classification',
    'Regression',
    'Classification',
    'Regression',
    'Clustering',
    'Classification',
    'Regression',
    'Clustering',
    'Classification',
    'Classification',
    'Regression',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Causal-Discovery',
    'Classification',
    'Clustering',
    'Regression',
    'Classification',
    'Clustering',
    'Classification',
    'Regression',
    'Classification',
    'Clustering',
    'Classification',
    'Regression',
    'Classification',
    'Regression',
    'Classification',
    'Regression',
    'Clustering',
    'Clustering',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Clustering',
    'Classification',
    'Clustering',
    'Classification',
    'Regression',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Regression',
    'Classification',
    'Clustering',
    'Clustering',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
```

```
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Clustering',
    'Classification',
    'Regression',
    'Regression',
    'Clustering',
    'Classification',
    'Regression',
    'Clustering',
    'Classification',
    'Clustering',
    'Classification',
    'Clustering',
    'Classification',
    'Clustering',
    'Classification',
    'Regression',
    'Regression',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Regression',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Recommendation',
    'Classification',
    'Regression',
    'Clustering',
    'Classification',
    'Regression',
    'Clustering',
    'Classification',
    'Regression',
    'Classification',
    'Regression',
    'Regression',
    'Classification',
    'Classification',
    'Clustering',
    'Classification',
    'Classification',
    'Classification',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Classification',
    'Clustering',
    'Regression',
    'Classification',
    'Classification',
    'Classification',
```

```
 'Classification',
 'Classification',
 'Regression',
 'Clustering',
 'Classification',
 'Regression',
 'Classification',
 'Clustering',
 'Classification',
 'Clustering',
 'Classification']
```

In [449]:

```
tipo_var=datos1[3:3310:7]
```

In [450]:

```
tipo_var
```

Out[450]:

```
[<p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">Categorical, Real </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal"> </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">Integer </p>,
 <p class="normal">Real </p>,
 <p class="normal">Real </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">Categorical, Real, Integer </p>,
 <p class="normal">Integer </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">Real </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">Real </p>,
 <p class="normal">Categorical, Integer </p>,
 <p class="normal">Real </p>,
 <p class="normal">Real </p>,
 <p class="normal">Integer </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">Real </p>,
 <p class="normal">Real </p>,
 <p class="normal">Categorical, Integer, Real </p>,
 <p class="normal">Integer, Real </p>,
 <p class="normal">Real </p>,
 <p class="normal">Real </p>,
```

```
<p class="normal">Real </p>,
<p class="normal">Categorical </p>,
<p class="normal">Categorical, Integer, Real </p>,
<p class="normal">Categorical </p>,
<p class="normal">Categorical </p>,
<p class="normal">Integer </p>,
<p class="normal">Categorical, Integer, Real </p>,
<p class="normal"> </p>,
<p class="normal">Integer </p>,
<p class="normal">Categorical </p>,
<p class="normal">Categorical, Integer, Real </p>,
<p class="normal">Categorical, Integer, Real </p>,
<p class="normal">Categorical, Integer, Real </p>,
<p class="normal">Categorical </p>,
<p class="normal">Categorical </p>,
<p class="normal">Categorical </p>,
<p class="normal">Categorical </p>,
<p class="normal"> </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Categorical </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer </p>,
<p class="normal">Categorical </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer </p>,
<p class="normal">Categorical, Integer </p>,
<p class="normal">Categorical </p>,
<p class="normal">Real </p>,
<p class="normal">Categorical, Integer </p>,
<p class="normal">Categorical </p>,
<p class="normal">Categorical </p>,
<p class="normal">Categorical </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Categorical </p>,
<p class="normal">Integer </p>,
<p class="normal">Categorical, Integer </p>,
<p class="normal">Categorical, Integer </p>,
<p class="normal">Categorical, Real </p>,
<p class="normal">Categorical, Integer </p>,
<p class="normal">Categorical, Integer </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Categorical, Integer, Real </p>,
<p class="normal">Categorical, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Categorical, Integer, Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">
```

```
<p class="normal"> </p>,
<p class="normal">Categorical </p>,
<p class="normal"> </p>,
<p class="normal">Categorical, Real </p>,
<p class="normal"> </p>,
<p class="normal">Categorical </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Categorical </p>,
<p class="normal"> </p>,
<p class="normal">Categorical, Integer, Real </p>,
<p class="normal">Categorical, Integer </p>,
<p class="normal">Categorical, Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer </p>,
<p class="normal">Categorical </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Categorical, Integer </p>,
<p class="normal">Categorical, Integer </p>,
<p class="normal">Categorical, Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Categorical </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Categorical, Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Categorical, Integer, Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>
```

```
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal"> </p>,
<p class="normal">Real </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal"> </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">
```

```
<p class="normal"> </p>,
<p class="normal">Integer </p>,
<p class="normal"> </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Categorical </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
```

```
<p class="normal">Integer </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal"> </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
<p class="normal">Integer </p>,
<p class="normal">Integer </p>,
<p class="normal">Real </p>,
<p class="normal">Real </p>,
<p class="normal">Integer, Real </p>,
```

```
 <p class="normal">Integer </p>,
 <p class="normal">Integer </p>,
 <p class="normal">Integer </p>,
 <p class="normal"> </p>,
 <p class="normal">Real </p>,
 <p class="normal">Categorical </p>,
 <p class="normal">Integer </p>,
 <p class="normal">Integer </p>,
 <p class="normal">Integer </p>,
 <p class="normal">Real </p>,
 <p class="normal">Real </p>,
 <p class="normal">Real </p>,
 <p class="normal">Integer, Real </p>,
 <p class="normal">Integer, Real </p>,
 <p class="normal">Real </p>,
 <p class="normal">Integer </p>,
 <p class="normal">Real </p>,
 <p class="normal">Real </p>,
 <p class="normal">Integer, Real </p>,
 <p class="normal">Real </p>,
 <p class="normal">Integer </p>,
 <p class="normal">Real </p>,
 <p class="normal">Real </p>,
 <p class="normal">Real </p>,
 <p class="normal">Integer, Real </p>,
 <p class="normal">Real </p>,
 <p class="normal">Real </p>,
 <p class="normal">Integer </p>]
```

In [451]:

```
tipo_var_s=str(tipo_var)
tipo_var_s
```

Out[451]:

```
'[<p class="normal">Categorical, Integer, Real\xa0</p>, <p class="normal">Categorical,
Integer\xa0</p>, <p class="normal">Categorical, Integer, Real\xa0</p>, <p
class="normal">Categorical\xa0</p>, <p class="normal">Categorical, Integer, Real\xa0</p>, <p
class="normal">Categorical, Integer, Real\xa0</p>, <p class="normal">Categorical\xa0</p>, <p
class="normal">Categorical\xa0</p>, <p class="normal">Categorical, Real\xa0</p>, <p
class="normal">Categorical, Integer, Real\xa0</p>, <p class="normal">\xa0</p>, <p
class="normal">Categorical\xa0</p>, <p class="normal">Categorical\xa0</p>, <p
class="normal">Categorical\xa0</p>, <p class="normal">Integer\xa0</p>, <p
class="normal">Real\xa0</p>, <p class="normal">Real\xa0</p>, <p class="normal">Categorical,
Integer\xa0</p>, <p class="normal">Categorical\xa0</p>, <p class="normal">Categorical,
Integer\xa0</p>, <p class="normal">Categorical, Integer\xa0</p>, <p
class="normal">Categorical\xa0</p>, <p class="normal">Categorical, Integer\xa0</p>, <p
class="normal">\xa0</p>, <p class="normal">Categorical, Integer\xa0</p>, <p
class="normal">Categorical\xa0</p>, <p class="normal">Categorical, Integer, Real\xa0</p>, <p
class="normal">Categorical, Real, Integer\xa0</p>, <p class="normal">Integer\xa0</p>, <p
class="normal">Categorical, Integer\xa0</p>, <p class="normal">Categorical, Integer\xa0</p>, <p cl
ass="normal">Categorical, Integer, Real\xa0</p>, <p class="normal">Categorical, Integer\xa0</p>, <
p class="normal">Categorical, Integer\xa0</p>, <p class="normal">Real\xa0</p>, <p
class="normal">\xa0</p>, <p class="normal">\xa0</p>, <p class="normal">Categorical, Integer,
Real\xa0</p>, <p class="normal">Real\xa0</p>, <p class="normal">Categorical, Integer\xa0</p>, <p c
lass="normal">Real\xa0</p>, <p class="normal">Real\xa0</p>, <p class="normal">Integer\xa0</p>, <p
class="normal">Categorical\xa0</p>, <p class="normal">Categorical, Integer, Real\xa0</p>, <p
class="normal">Categorical, Integer, Real\xa0</p>, <p class="normal">Categorical, Integer,
Real\xa0</p>, <p class="normal">Real\xa0</p>, <p class="normal">Real\xa0</p>, <p
class="normal">Categorical, Integer, Real\xa0</p>, <p class="normal">Integer, Real\xa0</p>, <p
class="normal">Real\xa0</p>, <p class="normal">Real\xa0</p>, <p
class="normal">Categorical\xa0</p>, <p class="normal">Categorical, Integer, Real\xa0</p>, <p
class="normal">Categorical\xa0</p>, <p class="normal">Categorical\xa0</p>, <p
class="normal">Integer\xa0</p>, <p class="normal">Categorical, Integer, Real\xa0</p>, <p
class="normal">\xa0</p>, <p class="normal">Integer\xa0</p>, <p class="normal">Categorical\xa0</p>,
<p class="normal">Categorical, Integer, Real\xa0</p>, <p class="normal">Categorical, Integer, Real
\xa0</p>, <p class="normal">Categorical, Integer, Real\xa0</p>, <p
class="normal">Categorical\xa0</p>, <p class="normal">Categorical\xa0</p>, <p
class="normal">Categorical\xa0</p>, <p class="normal">Categorical\xa0</p>, <p
class="normal">\xa0</p>, <p class="normal">Integer, Real\xa0</p>, <p
class="normal">Categorical\xa0</p>, <p class="normal">Integer\xa0</p>, <p
```

class="normal">Integer\xa0 , Categorical\xa0 , \xa0 ,
Integer, Real\xa0 , Integer\xa0 , <p
class="normal">Integer\xa0 , Categorical, Integer\xa0 , <p
class="normal">Categorical\xa0 , \xa0 , \xa0 , <p cla
ss="normal">Real\xa0 , Categorical, Integer\xa0 , <p
class="normal">Categorical\xa0 , Categorical\xa0 , <p
class="normal">Categorical\xa0 , Categorical\xa0 , <p
class="normal">Integer\xa0 , Integer, Real\xa0 , <p
class="normal">Integer, Real\xa0 , Categorical\xa0 , <p
class="normal">Integer\xa0 , Categorical, Integer\xa0 , <p
class="normal">\xa0 , \xa0 , Categorical,
Integer\xa0 , Categorical\xa0 , Categorical,
Real\xa0 , Categorical\xa0 , Categorical,
Integer\xa0 , Categorical\xa0 , Integer, Real\xa0 ,
Real\xa0 , Real\xa0 , Integer,
Real\xa0 , Real\xa0 , Categorical, Integer\xa0 , <p c
lass="normal">\xa0 , \xa0 , Categorical, Real\xa0 , <
p class="normal">Real\xa0 , Categorical\xa0 , <p
class="normal">Categorical, Integer\xa0 , Categorical, Real\xa0 , <p
class="normal">Real\xa0 , \xa0 , Categorical, Integer, R
eal\xa0 , Integer, Real\xa0 , Categorical\xa0 , <p
class="normal">Integer\xa0 , Categorical, Integer\xa0 , <p
class="normal">Categorical, Integer\xa0 , Categorical, Integer\xa0 , <p cl
ass="normal">Real\xa0 , Categorical, Integer\xa0 , <p
class="normal">Categorical, Integer\xa0 , \xa0 , <p
class="normal">\xa0 , Categorical\xa0 , \xa0 , <p
class="normal">Categorical, Real\xa0 , \xa0 , <p
class="normal">Categorical\xa0 , Integer\xa0 , <p
class="normal">Real\xa0 , Categorical\xa0 , \xa0 , <p
class="normal">\xa0 , Categorical, Integer, Real\xa0 , <p
class="normal">Categorical, Integer\xa0 , Categorical, Real\xa0 , <p
class="normal">Integer\xa0 , Real\xa0 , Integer\xa0 ,
Integer\xa0 , Categorical\xa0 , <p
class="normal">Real\xa0 , Real\xa0 , \xa0 , <p
class="normal">\xa0 , Real\xa0 , Categorical,
Integer\xa0 , Categorical, Integer\xa0 , Categorical, In
teger\xa0 , Real\xa0 , Integer\xa0 , <p
class="normal">Integer\xa0 , Real\xa0 , \xa0 , <p
class="normal">Integer\xa0 , Real\xa0 , Real\xa0 , <p
class="normal">Real\xa0 , Integer\xa0 , Integer\xa0 ,
Integer\xa0 , Real\xa0 , Real\xa0 ,
Integer\xa0 , Real\xa0 , Real\xa0 ,
Real\xa0 , Integer\xa0 , <p
class="normal">Integer\xa0 , Real\xa0 , <p
class="normal">Categorical\xa0 , Real\xa0 , <p
class="normal">Real\xa0 , Real\xa0 , Categorical,
Integer\xa0 , Real\xa0 , Integer, Real\xa0 , <p
class="normal">Real\xa0 , Integer, Real\xa0 , <p
class="normal">Integer\xa0 , \xa0 , Real\xa0 , <p
class="normal">Real\xa0 , Real\xa0 , Real\xa0 , <p
class="normal">Real\xa0 , Categorical, Integer, Real\xa0 , <p
class="normal">Integer, Real\xa0 , Real\xa0 , <p
class="normal">Real\xa0 , Real\xa0 , \xa0 , <p
class="normal">Real\xa0 , Integer\xa0 , Integer,
Real\xa0 , Real\xa0 , Real\xa0 , <p
class="normal">Real\xa0 , Real\xa0 , Real\xa0 , <p
class="normal">Real\xa0 , \xa0 , Real\xa0 , <p
class="normal">\xa0 , \xa0 , Integer, Real\xa0 , <p
class="normal">Integer, Real\xa0 , Real\xa0 , \xa0 ,
Real\xa0 , Integer, Real\xa0 , <p
class="normal">Real\xa0 , Real\xa0 , Real\xa0 , <p
class="normal">Real\xa0 , Real\xa0 , Real\xa0 , <p
class="normal">\xa0 , Integer\xa0 , Real\xa0 , <p
class="normal">Real\xa0 , Integer, Real\xa0 , <p
class="normal">Integer\xa0 , \xa0 , \xa0 , <p
class="normal">Real\xa0 , Integer, Real\xa0 , <p
class="normal">Real\xa0 , Real\xa0 , Real\xa0 , <p
class="normal">Real\xa0 , Real\xa0 , Integer,
Real\xa0 , Real\xa0 , Integer, Real\xa0 , <p
class="normal">Real\xa0 , Real\xa0 , Real\xa0 , <p
class="normal">Integer, Real\xa0 , \xa0 , Real\xa0 ,
Integer\xa0 , Real\xa0 , \xa0 , <p
class="normal">\xa0 , Real\xa0 , Integer, Real\xa0 ,
Real\xa0 , Real\xa0 , Real\xa0 , <p
class="normal">Integer\xa0 , Integer, Real\xa0 , <p
class="normal">Real\xa0 , \xa0 , Integer, Real\xa0 ,
Real\xa0 , Real\xa0 , Integer,

Real\xa0 , Integer, Real\xa0 , Integer, Real\xa0 , Real\xa0 , Real\xa0 , Real\xa0 , \xa0 , Real\xa0 , \xa0 , \xa0 , Real\xa0 , Real\xa0 , Real\xa0 , \xa0 , Real\xa0 , Real\xa0 , Integer\xa0 , \xa0 , Real\xa0 , \xa0 , Integer\xa0 , \xa0 , Real\xa0 , Integer, Real\xa0 , Integer, Real\xa0 , Real\xa0 , Integer\xa0 , Integer, Real\xa0 , Real\xa0 , \xa0 , \xa0 , Real\xa0 , Real\xa0 , Integer, Real\xa0 , Integer\xa0 , Integer, Real\xa0 , Real\xa0 , \xa0 , Real\xa0 , Real\xa0 , Real\xa0 , Real\xa0 , Real\xa0 , Integer\xa0 , Real\xa0 , Real\xa0 , Integer, Real\xa0 , \xa0 , Real\xa0 , Real\xa0 , Real\xa0 , Integer\xa0 , Real\xa0 , Real\xa0 , Integer, Real\xa0 , Categorical\xa0 , \xa0 , Integer, Real\xa0 , \xa0 , \xa0 , Integer, Real\xa0 , Real\xa0 , Real\xa0 , \xa0 , Real\xa0 , Real\xa0 , Real\xa0 , Real\xa0 , Integer, Real\xa0 , Real\xa0 , Integer\xa0 , Real\xa0 , Real\xa0 , Integer\xa0 , Integer, Real\xa0 , Real\xa0 , Integer, Real\xa0 , Integer\xa0 , Real\xa0 , Real\xa0 , Real\xa0 , Real\xa0 , \xa0 , Real\xa0 , Real\xa0 , Real\xa0 , Integer, Real\xa0 , Real\xa0 , Real\xa0 , Real\xa0 , \xa0 , Integer\xa0 , \xa0 , Integer\xa0 , Real\xa0 , Real\xa0 , Integer\xa0 , Integer\xa0 , Integer, Real\xa0 , Real\xa0 , Integer, Real\xa0 , Integer\xa0 , \xa0 , Integer, Real\xa0 , Integer\xa0 , Integer, Real\xa0 , \xa0 , Integer, Real\xa0 , Integer\xa0 , Real\xa0 , Real\xa0 , Integer, Real\xa0 , Integer\xa0 , Real\xa0 , Real\xa0 , Integer, Real\xa0 , Integer, Real\xa0 , Integer, Real\xa0 , Integer, Real\xa0 , Integer\xa0 , Integer, Real\xa0 , Integer\xa0 , \xa0 , Integer, Real\xa0 , \xa0 , Real\xa0 , Real\xa0 , \xa0 , Integer\xa0 , Real\xa0 , Real\xa0 , Integer\xa0 , Real\xa0 , Integer\xa0 , Integer\xa0 , Integer, Real\xa0 , Integer, Real\xa0 , Integer\xa0 , Integer\xa0 , Real\xa0 , Integer\xa0 , Integer, Real\xa0 , Integer\xa0 , Integer\xa0 , Integer\xa0 , Integer, Real\xa0 , Real\xa0 , Integer, Real\xa0 , Integer\xa0 , \xa0 , Integer\xa0 , Real\xa0 , Integer, Real\xa0 , Integer, Real\xa0 , Real\xa0 , \xa0 , Integer, Real\xa0 , Real\xa0 , \xa0 , Integer\xa0 , Real\xa0 , Real\xa0 , Real\xa0 , Integer\xa0 , Integer\xa0 , Real\xa0 , Real\xa0 , Integer, Real\xa0 , \xa0 , Real\xa0 , Real\xa0 , Integer, Real\xa0 , Integer\xa0 , Real\xa0 , Integer\xa0 , Integer\xa0 , Real\xa0 , Integer\xa0 , Integer, Real\xa0 , Integer\xa0 , Integer\xa0 , Integer\xa0 , Integer, Real\xa0 , Real\xa0 , Integer, Real\xa0 , Integer\xa0 , \xa0 , Integer\xa0 , Real\xa0 , Integer, Real\xa0 , Integer, Real\xa0 , Real\xa0 , \xa0 , Integer, Real\xa0 , Real\xa0 , \xa0 , Integer\xa0 , Real\xa0 , Real\xa0 , Real\xa0 , Integer\xa0 , Integer\xa0 , Real\xa0 , Real\xa0 , Integer, Real\xa0 , \xa0 , \xa0 , Integer\xa0 , Integer\xa0 , Integer\xa0 , \xa0 , Real\xa0 , Categorical\xa0 , Integer\xa0 , Integer\xa0 , Integer\xa0 , Real\xa0 , Real\xa0 , Real\xa0 , \xa0 , Integer, Real\xa0 , \xa0 , Integer, Real\xa0 , Real\xa0 , Integer\xa0 , \xa0 , \xa0 , Real\xa0 , Real\xa0 , Integer, Real\xa0 , Real\xa0 , Integer\xa0 , Real\xa0 , Real\xa0 , Real\xa0 , Integer, Real\xa0 , Real\xa0 , Real\xa0 , Integer\xa0]'

In [452]:

```
pa3="([A-Z]+[a-z]+[A-Z]+)"
re.findall(pa3,tipo_var_s)
```

Out[452]:

```
[]
```

In [453]:

```
numero_obs=datos1[4:3409:7]
```

In [454]:

```
numero_obs
```

Out[454]:

```
[<p class="normal">4177 </p>,
 <p class="normal">48842 </p>,
 <p class="normal">798 </p>,
 <p class="normal">37711 </p>,
 <p class="normal">452 </p>,
 <p class="normal">6000 </p>,
 <p class="normal">226 </p>,
 <p class="normal">226 </p>,
 <p class="normal">398 </p>,
 <p class="normal">205 </p>,
 <p class="normal">294 </p>,
 <p class="normal">625 </p>,
 <p class="normal">16 </p>,
 <p class="normal">286 </p>,
 <p class="normal">699 </p>,
 <p class="normal">198 </p>,
 <p class="normal">569 </p>,
 <p class="normal">108 </p>,
 <p class="normal">1728 </p>,
 <p class="normal">48842 </p>,
 <p class="normal"> </p>,
 <p class="normal">3196 </p>,
 <p class="normal">28056 </p>,
 <p class="normal"> </p>,
 <p class="normal">100 </p>,
 <p class="normal">67557 </p>,
 <p class="normal">690 </p>,
 <p class="normal">125 </p>,
 <p class="normal">209 </p>,
 <p class="normal">1473 </p>,
 <p class="normal">581012 </p>,
 <p class="normal">512 </p>,
 <p class="normal">366 </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">132 </p>,
 <p class="normal">336 </p>,
 <p class="normal">194 </p>,
 <p class="normal">352 </p>,
 <p class="normal">214 </p>,
 <p class="normal">306 </p>,
 <p class="normal">160 </p>,
 <p class="normal">303 </p>,
 <p class="normal">155 </p>,
 <p class="normal">368 </p>,
 <p class="normal">  </p>,
 <p class="normal">2310 </p>,
 <p class="normal">3279 </p>,
 <p class="normal">351 </p>,
 <p class="normal">150 </p>,
 <p class="normal">7797 </p>,
 <p class="normal">104 </p>,
```

```
<p class="normal">104 </p>,
<p class="normal">57 </p>,
<p class="normal"> </p>,
<p class="normal">24 </p>,
<p class="normal">20000 </p>,
<p class="normal">345 </p>,
<p class="normal"> </p>,
<p class="normal">32 </p>,
<p class="normal">148 </p>,
<p class="normal">209 </p>,
<p class="normal">528 </p>,
<p class="normal"> </p>,
<p class="normal">106 </p>,
<p class="normal">128 </p>,
<p class="normal">3190 </p>,
<p class="normal">432 </p>,
<p class="normal">202 </p>,
<p class="normal">2000 </p>,
<p class="normal">8124 </p>,
<p class="normal">476 </p>,
<p class="normal">6598 </p>,
<p class="normal">12960 </p>,
<p class="normal">5473 </p>,
<p class="normal">5620 </p>,
<p class="normal">10992 </p>,
<p class="normal">90 </p>,
<p class="normal">339 </p>,
<p class="normal">167 </p>,
<p class="normal">15 </p>,
<p class="normal">1389 </p>,
<p class="normal">307 </p>,
<p class="normal">47 </p>,
<p class="normal">23 </p>,
<p class="normal">531 </p>,
<p class="normal">4601 </p>,
<p class="normal">267 </p>,
<p class="normal">267 </p>,
<p class="normal">76 </p>,
<p class="normal">1000 </p>,
<p class="normal">151 </p>,
<p class="normal">958 </p>,
<p class="normal">7200 </p>,
<p class="normal">10 </p>,
<p class="normal">285 </p>,
<p class="normal">435 </p>,
<p class="normal">527 </p>,
<p class="normal">5000 </p>,
<p class="normal">5000 </p>,
<p class="normal">178 </p>,
<p class="normal">1484 </p>,
<p class="normal">101 </p>,
<p class="normal">20000 </p>,
<p class="normal">6650 </p>,
<p class="normal">2565 </p>,
<p class="normal">2458285 </p>,
<p class="normal">299285 </p>,
<p class="normal">340 </p>,
<p class="normal">68040 </p>,
<p class="normal">122 </p>,
<p class="normal">178080 </p>,
<p class="normal">50672 </p>,
<p class="normal">640 </p>,
<p class="normal">9000 </p>,
<p class="normal">10104 </p>,
<p class="normal">256932 </p>,
<p class="normal">640 </p>,
<p class="normal">191779 </p>,
<p class="normal">4000000 </p>,
<p class="normal">10000 </p>,
```

```
<p class="normal">989818 </p>,
<p class="normal">129000 </p>,
<p class="normal"> </p>,
<p class="normal">100000 </p>,
<p class="normal">21578 </p>,
<p class="normal">463 </p>,
<p class="normal">600 </p>,
<p class="normal">332 </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal">690 </p>,
<p class="normal">1000 </p>,
<p class="normal">270 </p>,
<p class="normal">6435 </p>,
<p class="normal">2310 </p>,
<p class="normal">58000 </p>,
<p class="normal">946 </p>,
<p class="normal">20008 </p>,
<p class="normal">208 </p>,
<p class="normal">528 </p>,
<p class="normal">> </p>,
<p class="normal">> </p>,
<p class="normal">1024 </p>,
<p class="normal">10080 </p>,
<p class="normal">50400 </p>,
<p class="normal">1025010 </p>,
<p class="normal">19020 </p>,
<p class="normal">1364 </p>,
<p class="normal">961 </p>,
<p class="normal">517 </p>,
<p class="normal">200 </p>,
<p class="normal">8000000 </p>,
<p class="normal">1030 </p>,
<p class="normal">606 </p>,
<p class="normal">900 </p>,
<p class="normal">2600 </p>,
<p class="normal">1950 </p>,
<p class="normal">13500 </p>,
<p class="normal">4400 </p>,
<p class="normal">2536 </p>,
<p class="normal">300 </p>,
<p class="normal">197 </p>,
<p class="normal">2858 </p>,
<p class="normal">748 </p>,
<p class="normal">11640 </p>,
<p class="normal">1593 </p>,
<p class="normal">1567 </p>,
<p class="normal">22632 </p>,
<p class="normal">360 </p>,
<p class="normal">103 </p>,
<p class="normal">1994 </p>,
<p class="normal">120 </p>,
<p class="normal">4898 </p>,
<p class="normal">2396130 </p>,
<p class="normal">16772 </p>,
<p class="normal">5875 </p>,
<p class="normal">503 </p>,
<p class="normal">51 </p>,
<p class="normal">106 </p>,
<p class="normal">2126 </p>,
<p class="normal">5456 </p>,
<p class="normal">8800 </p>,
<p class="normal">164860 </p>,
<p class="normal">> </p>,
<p class="normal">1941 </p>,
<p class="normal">130065 </p>,
<p class="normal">515345 </p>,
<p class="normal">440 </p>,
<p class="normal">> </p>,
<p class="normal">53500 </p>,
<p class="normal">8235 </p>,
<p class="normal">> </p>,
<p class="normal">5749132 </p>,
<p class="normal">2215 </p>,
<p class="normal">310 </p>,
<p class="normal">10000 </p>,
<p class="normal">3000 </p>,
```

```
<p class="normal">1500 </p>,
<p class="normal">30000 </p>,
<p class="normal">2500 </p>,
<p class="normal">4143 </p>,
<p class="normal">64 </p>,
<p class="normal">53414 </p>,
<p class="normal">65554 </p>,
<p class="normal">45211 </p>,
<p class="normal">1138562 </p>,
<p class="normal">13910 </p>,
<p class="normal">583 </p>,
<p class="normal">2551 </p>,
<p class="normal">34465 </p>,
<p class="normal">5574 </p>,
<p class="normal">245057 </p>,
<p class="normal">182 </p>,
<p class="normal">3850505 </p>,
<p class="normal">138 </p>,
<p class="normal">1080 </p>,
<p class="normal">2075259 </p>,
<p class="normal">210 </p>,
<p class="normal">115 </p>,
<p class="normal">3960456 </p>,
<p class="normal"> </p>,
<p class="normal">10299 </p>,
<p class="normal">1600 </p>,
<p class="normal">768 </p>,
<p class="normal">308 </p>,
<p class="normal">100 </p>,
<p class="normal">237 </p>,
<p class="normal">434874 </p>,
<p class="normal">536 </p>,
<p class="normal">140000 </p>,
<p class="normal">6118 </p>,
<p class="normal">165632 </p>,
<p class="normal">18000 </p>,
<p class="normal">540 </p>,
<p class="normal">931 </p>,
<p class="normal">1055 </p>,
<p class="normal">100 </p>,
<p class="normal">9120 </p>,
<p class="normal">403 </p>,
<p class="normal">111740 </p>,
<p class="normal">10421 </p>,
<p class="normal">5820 </p>,
<p class="normal">403 </p>,
<p class="normal">14980 </p>,
<p class="normal">45730 </p>,
<p class="normal">2584 </p>,
<p class="normal">1372 </p>,
<p class="normal">306 </p>,
<p class="normal">120000 </p>,
<p class="normal">13910 </p>,
<p class="normal">2747 </p>,
<p class="normal">3395 </p>,
<p class="normal">39242 </p>,
<p class="normal">4137 </p>,
<p class="normal">17389 </p>,
<p class="normal">51 </p>,
<p class="normal">470 </p>,
<p class="normal">132 </p>,
<p class="normal">5000000 </p>,
<p class="normal">11000000 </p>,
<p class="normal">250 </p>,
<p class="normal">126 </p>,
<p class="normal"> </p>,
<p class="normal">4889 </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal">340 </p>,
<p class="normal">501 </p>,
<p class="normal">45781 </p>,
<p class="normal">1503 </p>,
<p class="normal">440 </p>,
<p class="normal">2000 </p>,
<p class="normal">9568 </p>,
<p class="normal">168 </p>,
```

```
<p class="normal">100000 </p>,
<p class="normal">5665 </p>,
<p class="normal">79 </p>,
<p class="normal">127 </p>,
<p class="normal">1040 </p>,
<p class="normal">9900 </p>,
<p class="normal">560 </p>,
<p class="normal">60021 </p>,
<p class="normal">1419 </p>,
<p class="normal">101 </p>,
<p class="normal">399 </p>,
<p class="normal">58 </p>,
<p class="normal">180 </p>,
<p class="normal">21048 </p>,
<p class="normal"> </p>,
<p class="normal">750 </p>,
<p class="normal">3000 </p>,
<p class="normal">150 </p>,
<p class="normal">1059 </p>,
<p class="normal">11934 </p>,
<p class="normal">27965 </p>,
<p class="normal">216 </p>,
<p class="normal">120 </p>,
<p class="normal">649 </p>,
<p class="normal">370 </p>,
<p class="normal">4178504 </p>,
<p class="normal">221579 </p>,
<p class="normal">10800 </p>,
<p class="normal">58509 </p>,
<p class="normal">129685 </p>,
<p class="normal">2456 </p>,
<p class="normal">2921 </p>,
<p class="normal">1151 </p>,
<p class="normal">6590 </p>,
<p class="normal">3000 </p>,
<p class="normal">39797 </p>,
<p class="normal">326 </p>,
<p class="normal">913 </p>,
<p class="normal">168286 </p>,
<p class="normal">400 </p>,
<p class="normal">314080 </p>,
<p class="normal">637 </p>,
<p class="normal">1710671 </p>,
<p class="normal">12000 </p>,
<p class="normal">10929 </p>,
<p class="normal">1080 </p>,
<p class="normal">40000 </p>,
<p class="normal">43930257 </p>,
<p class="normal">230318 </p>,
<p class="normal">10500000 </p>,
<p class="normal">13197 </p>,
<p class="normal"> </p>,
<p class="normal">30000 </p>,
<p class="normal">324 </p>,
<p class="normal">541909 </p>,
<p class="normal">11164866 </p>,
<p class="normal">163 </p>,
<p class="normal">373 </p>,
<p class="normal">20560 </p>,
<p class="normal">40 </p>,
<p class="normal">422937 </p>,
<p class="normal">9358 </p>,
<p class="normal">640 </p>,
<p class="normal">919438 </p>,
<p class="normal">40949 </p>,
<p class="normal">5744 </p>,
<p class="normal">10503 </p>,
<p class="normal">42240 </p>,
<p class="normal">102944 </p>,
<p class="normal">500 </p>,
<p class="normal">9782222 </p>,
<p class="normal">11463 </p>,
<p class="normal">17898 </p>,
<p class="normal">1885 </p>,
<p class="normal">19735 </p>,
<p class="normal">1540 </p>,
<p class="normal">4007 </p>,
```

```
<p class="normal">153540 </p>,
<p class="normal">606 </p>,
<p class="normal">1353 </p>,
<p class="normal">1956 </p>,
<p class="normal">43824 </p>,
<p class="normal">3942 </p>,
<p class="normal">858 </p>,
<p class="normal">287 </p>,
<p class="normal">17764280 </p>,
<p class="normal">106574 </p>,
<p class="normal">9358 </p>,
<p class="normal">11500 </p>,
<p class="normal">92000 </p>,
<p class="normal">315 </p>,
<p class="normal">78095 </p>,
<p class="normal">130 </p>,
<p class="normal">74 </p>,
<p class="normal">52854 </p>,
<p class="normal">77 </p>,
<p class="normal">811 </p>,
<p class="normal">504 </p>,
<p class="normal">401 </p>,
<p class="normal">2856 </p>,
<p class="normal">10546 </p>,
<p class="normal">801 </p>,
<p class="normal">1540 </p>,
<p class="normal">1451 </p>,
<p class="normal">1075 </p>,
<p class="normal">78095 </p>,
<p class="normal">7195 </p>,
<p class="normal">3600 </p>,
<p class="normal">76 </p>,
<p class="normal">60 </p>,
<p class="normal">405 </p>,
<p class="normal">303 </p>,
<p class="normal">303 </p>,
<p class="normal">107888 </p>,
<p class="normal">76000 </p>,
<p class="normal">10000 </p>,
<p class="normal">180 </p>,
<p class="normal">745000 </p>,
<p class="normal">12234 </p>,
<p class="normal">292 </p>,
<p class="normal">104 </p>,
<p class="normal">60000 </p>,
<p class="normal">2000 </p>,
<p class="normal">165 </p>,
<p class="normal">217 </p>,
<p class="normal">1175 </p>,
<p class="normal">704 </p>,
<p class="normal">75128 </p>,
<p class="normal">90 </p>,
<p class="normal">90 </p>,
<p class="normal">50 </p>,
<p class="normal">71 </p>,
<p class="normal">93239 </p>,
<p class="normal">540 </p>,
<p class="normal">105 </p>,
<p class="normal">6611 </p>,
<p class="normal">15 </p>,
<p class="normal">372 </p>,
<p class="normal">58000 </p>,
<p class="normal">4960 </p>,
<p class="normal">241600 </p>,
<p class="normal">130000 </p>,
<p class="normal">7062606 </p>,
<p class="normal">740 </p>,
<p class="normal">70 </p>,
<p class="normal">2205 </p>,
<p class="normal">10721 </p>,
<p class="normal">640 </p>,
<p class="normal">1000 </p>,
<p class="normal">116 </p>,
<p class="normal">1672 </p>,
<p class="normal">322 </p>,
<p class="normal">93600 </p>,
<p class="normal">3060 </p>,
```

```
 <p class="normal">5879 </p>,
 <p class="normal">9200 </p>,
 <p class="normal">20000 </p>,
 <p class="normal">20867 </p>,
 <p class="normal"> </p>,
 <p class="normal">4143 </p>,
 <p class="normal">215063 </p>,
 <p class="normal">6000000 </p>,
 <p class="normal">21263 </p>,
 <p class="normal">63000000 </p>,
 <p class="normal">10190 </p>,
 <p class="normal">300 </p>,
 <p class="normal">12330 </p>,
 <p class="normal">5180 </p>,
 <p class="normal">756 </p>,
 <p class="normal">10000 </p>,
 <p class="normal">80 </p>,
 <p class="normal">1184 </p>,
 <p class="normal">1047 </p>,
 <p class="normal">777 </p>,
 <p class="normal">249 </p>,
 <p class="normal">414 </p>,
 <p class="normal">> </p>,
 <p class="normal">143 </p>,
 <p class="normal">7840 </p>,
 <p class="normal">30000 </p>,
 <p class="normal">35717 </p>,
 <p class="normal">135 </p>,
 <p class="normal">980 </p>,
 <p class="normal">5456 </p>,
 <p class="normal">120 </p>]
```

In [455]:

```
numero_var=datos1[5:3310:7]
```

In [456]:

```
numero_var
```

Out[456]:

```
[<p class="normal">8 </p>,
 <p class="normal">14 </p>,
 <p class="normal">38 </p>,
 <p class="normal">294 </p>,
 <p class="normal">279 </p>,
 <p class="normal">7 </p>,
 <p class="normal">> </p>,
 <p class="normal">69 </p>,
 <p class="normal">8 </p>,
 <p class="normal">26 </p>,
 <p class="normal">1 </p>,
 <p class="normal">4 </p>,
 <p class="normal">4 </p>,
 <p class="normal">9 </p>,
 <p class="normal">10 </p>,
 <p class="normal">34 </p>,
 <p class="normal">32 </p>,
 <p class="normal">13 </p>,
 <p class="normal">6 </p>,
 <p class="normal">14 </p>,
 <p class="normal">22 </p>,
 <p class="normal">36 </p>,
 <p class="normal">6 </p>,
 <p class="normal">> </p>,
 <p class="normal">6 </p>,
 <p class="normal">42 </p>,
 <p class="normal">15 </p>,
 <p class="normal">> </p>,
 <p class="normal">9 </p>,
 <p class="normal">9 </p>,
 <p class="normal">54 </p>,
 <p class="normal">39 </p>,
 <p class="normal">33 </p>,
```

```
<p class="normal">20 </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal">12 </p>,
<p class="normal">8 </p>,
<p class="normal">30 </p>,
<p class="normal"> </p>,
<p class="normal">10 </p>,
<p class="normal">3 </p>,
<p class="normal">5 </p>,
<p class="normal">75 </p>,
<p class="normal">19 </p>,
<p class="normal">27 </p>,
<p class="normal">  </p>,
<p class="normal">19 </p>,
<p class="normal">1558 </p>,
<p class="normal">34 </p>,
<p class="normal">4 </p>,
<p class="normal">617 </p>,
<p class="normal">12 </p>,
<p class="normal">16 </p>,
<p class="normal">7 </p>,
<p class="normal">4 </p>,
<p class="normal">16 </p>,
<p class="normal">7 </p>,
<p class="normal">  </p>,
<p class="normal">56 </p>,
<p class="normal">18 </p>,
<p class="normal">8 </p>,
<p class="normal">22 </p>,
<p class="normal">  </p>,
<p class="normal">58 </p>,
<p class="normal">  </p>,
<p class="normal">61 </p>,
<p class="normal">7 </p>,
<p class="normal">  </p>,
<p class="normal">649 </p>,
<p class="normal">22 </p>,
<p class="normal">168 </p>,
<p class="normal">168 </p>,
<p class="normal">8 </p>,
<p class="normal">  </p>,
<p class="normal">10 </p>,
<p class="normal">64 </p>,
<p class="normal">16 </p>,
<p class="normal">8 </p>,
<p class="normal">17 </p>,
<p class="normal">  </p>,
<p class="normal">  </p>,
<p class="normal">72 </p>,
<p class="normal">4 </p>,
<p class="normal">6 </p>,
<p class="normal">10 </p>,
<p class="normal">35 </p>,
<p class="normal">35 </p>,
<p class="normal">4 </p>,
<p class="normal">102 </p>,
<p class="normal">57 </p>,
<p class="normal">22 </p>,
<p class="normal">44 </p>,
<p class="normal">45 </p>,
<p class="normal">  </p>,
<p class="normal">  </p>,
<p class="normal">5 </p>,
<p class="normal">9 </p>,
<p class="normal">21 </p>,
<p class="normal">32 </p>,
<p class="normal">17 </p>,
<p class="normal">16 </p>,
<p class="normal">38 </p>,
<p class="normal">21 </p>,
<p class="normal">40 </p>,
<p class="normal">13 </p>,
<p class="normal">8 </p>,
<p class="normal">17 </p>,
<p class="normal">  </p>,
```

```
<p class="normal"> </p>,
<p class="normal">15 </p>,
<p class="normal">22 </p>,
<p class="normal">68 </p>,
<p class="normal">40 </p>,
<p class="normal">17 </p>,
<p class="normal">89 </p>,
<p class="normal"> </p>,
<p class="normal">4 </p>,
<p class="normal">12 </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal">86 </p>,
<p class="normal">72 </p>,
<p class="normal">61 </p>,
<p class="normal">12 </p>,
<p class="normal">481 </p>,
<p class="normal">42 </p>,
<p class="normal">  </p>,
<p class="normal">  </p>,
<p class="normal">  </p>,
<p class="normal">  </p>,
<p class="normal">  </p>,
<p class="normal">  </p>,
<p class="normal">5 </p>,
<p class="normal">90 </p>,
<p class="normal">  </p>,
<p class="normal">5 </p>,
<p class="normal">  </p>,
<p class="normal">  </p>,
<p class="normal">14 </p>,
<p class="normal">20 </p>,
<p class="normal">13 </p>,
<p class="normal">36 </p>,
<p class="normal">19 </p>,
<p class="normal">9 </p>,
<p class="normal">18 </p>,
<p class="normal">4 </p>,
<p class="normal">60 </p>,
<p class="normal">10 </p>,
<p class="normal">  </p>,
<p class="normal">  </p>,
<p class="normal">10 </p>,
<p class="normal">4 </p>,
<p class="normal">3 </p>,
<p class="normal">11 </p>,
<p class="normal">11 </p>,
<p class="normal">  </p>,
<p class="normal">6 </p>,
<p class="normal">13 </p>,
<p class="normal">  </p>,
<p class="normal">100000 </p>,
<p class="normal">9 </p>,
<p class="normal">101 </p>,
<p class="normal">10000 </p>,
<p class="normal">20000 </p>,
<p class="normal">100000 </p>,
<p class="normal">5000 </p>,
<p class="normal">500 </p>,
<p class="normal">73 </p>,
<p class="normal">43 </p>,
<p class="normal">23 </p>,
<p class="normal">3 </p>,
<p class="normal">5 </p>,
<p class="normal">  </p>,
<p class="normal">256 </p>,
<p class="normal">591 </p>,
<p class="normal">70 </p>,
<p class="normal">91 </p>,
<p class="normal">10 </p>,
<p class="normal">128 </p>,
<p class="normal">6 </p>,
<p class="normal">12 </p>,
<p class="normal">3231961 </p>,
<p class="normal">5409 </p>,
<p class="normal">26 </p>,
<p class="normal">  </p>,
```

```
<p class="normal"> </p>,
<p class="normal">10 </p>,
<p class="normal">23 </p>,
<p class="normal">24 </p>,
<p class="normal">13 </p>,
<p class="normal">8 </p>,
<p class="normal"> </p>,
<p class="normal">27 </p>,
<p class="normal">50 </p>,
<p class="normal">90 </p>,
<p class="normal">138672 </p>,
<p class="normal"> </p>,
<p class="normal">386 </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal">12 </p>,
<p class="normal">147 </p>,
<p class="normal">6 </p>,
<p class="normal">8 </p>,
<p class="normal">27 </p>,
<p class="normal">10000 </p>,
<p class="normal">20000 </p>,
<p class="normal">10000 </p>,
<p class="normal">54877 </p>,
<p class="normal">4702 </p>,
<p class="normal">24 </p>,
<p class="normal">29 </p>,
<p class="normal">17 </p>,
<p class="normal">3 </p>,
<p class="normal">128 </p>,
<p class="normal">10 </p>,
<p class="normal">242 </p>,
<p class="normal">120 </p>,
<p class="normal"> </p>,
<p class="normal">4 </p>,
<p class="normal">13 </p>,
<p class="normal">52 </p>,
<p class="normal">47 </p>,
<p class="normal">857 </p>,
<p class="normal">9 </p>,
<p class="normal">7 </p>,
<p class="normal">200 </p>,
<p class="normal">4 </p>,
<p class="normal">561 </p>,
<p class="normal">64 </p>,
<p class="normal">8 </p>,
<p class="normal">7 </p>,
<p class="normal">10 </p>,
<p class="normal">9 </p>,
<p class="normal">4 </p>,
<p class="normal">8 </p>,
<p class="normal">77 </p>,
<p class="normal">51 </p>,
<p class="normal">18 </p>,
<p class="normal">1950000 </p>,
<p class="normal">18 </p>,
<p class="normal">1300 </p>,
<p class="normal">41 </p>,
<p class="normal">6 </p>,
<p class="normal">5625 </p>,
<p class="normal">5 </p>,
<p class="normal">7 </p>,
<p class="normal">33 </p>,
<p class="normal">5 </p>,
<p class="normal">15 </p>,
<p class="normal">9 </p>,
<p class="normal">19 </p>,
<p class="normal">5 </p>,
<p class="normal">5 </p>,
<p class="normal">1000000 </p>,
<p class="normal">129 </p>,
<p class="normal"> </p>,
<p class="normal">20 </p>,
<p class="normal">152 </p>,
<p class="normal">24 </p>,
```

```
<p class="normal">21 </p>,
<p class="normal">16 </p>,
<p class="normal">35 </p>,
<p class="normal">17 </p>,
<p class="normal">5 </p>,
<p class="normal">18 </p>,
<p class="normal">28 </p>,
<p class="normal">7 </p>,
<p class="normal">309 </p>,
<p class="normal">3 </p>,
<p class="normal">6 </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal">16 </p>,
<p class="normal">13 </p>,
<p class="normal">5 </p>,
<p class="normal">6 </p>,
<p class="normal">8 </p>,
<p class="normal">2 </p>,
<p class="normal">4 </p>,
<p class="normal">148 </p>,
<p class="normal">55 </p>,
<p class="normal">17 </p>,
<p class="normal">8 </p>,
<p class="normal">42 </p>,
<p class="normal">26 </p>,
<p class="normal">50 </p>,
<p class="normal">2 </p>,
<p class="normal">281 </p>,
<p class="normal">120 </p>,
<p class="normal">> </p>,
<p class="normal">6 </p>,
<p class="normal">120432 </p>,
<p class="normal">150000 </p>,
<p class="normal">529 </p>,
<p class="normal">> </p>,
<p class="normal">16 </p>,
<p class="normal">2500 </p>,
<p class="normal">5 </p>,
<p class="normal">68 </p>,
<p class="normal">16 </p>,
<p class="normal">100 </p>,
<p class="normal">216 </p>,
<p class="normal">23 </p>,
<p class="normal">33 </p>,
<p class="normal">140256 </p>,
<p class="normal">19 </p>,
<p class="normal">20 </p>,
<p class="normal">20 </p>,
<p class="normal">49 </p>,
<p class="normal">12 </p>,
<p class="normal">30 </p>,
<p class="normal">5232 </p>,
<p class="normal">20 </p>,
<p class="normal">1 </p>,
<p class="normal">> </p>,
<p class="normal">61 </p>,
<p class="normal">27 </p>,
<p class="normal">53 </p>,
<p class="normal">11 </p>,
<p class="normal">25 </p>,
<p class="normal">0 </p>,
<p class="normal">20 </p>,
<p class="normal">9 </p>,
<p class="normal">3 </p>,
<p class="normal">561 </p>,
<p class="normal">82 </p>,
<p class="normal">13 </p>,
<p class="normal">16 </p>,
<p class="normal">13 </p>,
<p class="normal">28 </p>,
<p class="normal">4 </p>,
<p class="normal">> </p>,
<p class="normal">24 </p>,
<p class="normal">34 </p>,
<p class="normal">8 </p>,
<p class="normal">128 </p>,
<p class="normal">15 </p>,
```

```
<p class="normal">513 </p>,
<p class="normal">7 </p>,
<p class="normal">7 </p>,
<p class="normal">5 </p>,
<p class="normal">15 </p>,
<p class="normal">480000 </p>,
<p class="normal">11 </p>,
<p class="normal">54 </p>,
<p class="normal">561 </p>,
<p class="normal">64 </p>,
<p class="normal">6 </p>,
<p class="normal">116 </p>,
<p class="normal">19 </p>,
<p class="normal"> </p>,
<p class="normal">5812 </p>,
<p class="normal">9 </p>,
<p class="normal">32 </p>,
<p class="normal">29 </p>,
<p class="normal">67 </p>,
<p class="normal"> </p>,
<p class="normal">25 </p>,
<p class="normal">6400 </p>,
<p class="normal">10 </p>,
<p class="normal">5 </p>,
<p class="normal">13 </p>,
<p class="normal">98 </p>,
<p class="normal">36 </p>,
<p class="normal">69 </p>,
<p class="normal">2158859 </p>,
<p class="normal">518 </p>,
<p class="normal">15 </p>,
<p class="normal">179 </p>,
<p class="normal">> </p>,
<p class="normal">12 </p>,
<p class="normal">38 </p>,
<p class="normal">65 </p>,
<p class="normal">102 </p>,
<p class="normal">86 </p>,
<p class="normal">7 </p>,
<p class="normal">53 </p>,
<p class="normal">20 </p>,
<p class="normal">1 </p>,
<p class="normal">71 </p>,
<p class="normal">29 </p>,
<p class="normal">20531 </p>,
<p class="normal">65 </p>,
<p class="normal">3 </p>,
<p class="normal">22 </p>,
<p class="normal">38 </p>,
<p class="normal">22 </p>,
<p class="normal">4814 </p>,
<p class="normal">698 </p>,
<p class="normal">13 </p>,
<p class="normal">10 </p>,
<p class="normal">59 </p>,
<p class="normal">56 </p>,
<p class="normal">482 </p>,
<p class="normal">171 </p>,
<p class="normal">5 </p>,
<p class="normal">500 </p>,
<p class="normal">411 </p>,
<p class="normal">8519 </p>,
<p class="normal">21 </p>,
<p class="normal">21 </p>,
<p class="normal">171 </p>,
<p class="normal">7 </p>,
<p class="normal">49 </p>,
<p class="normal">12 </p>,
<p class="normal">3 </p>,
<p class="normal">21 </p>,
<p class="normal">9 </p>,
<p class="normal">8 </p>,
<p class="normal">7 </p>,
<p class="normal">2 </p>,
<p class="normal">11 </p>,
<p class="normal">11 </p>,
<p class="normal">173 </p>
```

```
 <p class="normal">175 </p>,
 <p class="normal">5 </p>,
 <p class="normal">15 </p>,
 <p class="normal">3 </p>,
 <p class="normal">105 </p>,
 <p class="normal">25000 </p>,
 <p class="normal"> </p>,
 <p class="normal">18 </p>,
 <p class="normal">21000 </p>,
 <p class="normal">115 </p>,
 <p class="normal">21 </p>,
 <p class="normal">206 </p>,
 <p class="normal">43680 </p>,
 <p class="normal">8 </p>,
 <p class="normal">10 </p>,
 <p class="normal">59 </p>,
 <p class="normal">10 </p>,
 <p class="normal">5 </p>,
 <p class="normal">5 </p>,
 <p class="normal">1000 </p>,
 <p class="normal">138 </p>,
 <p class="normal">  </p>,
 <p class="normal">16 </p>,
 <p class="normal">2 </p>,
 <p class="normal">10 </p>,
 <p class="normal">  </p>,
 <p class="normal">8 </p>,
 <p class="normal">6 </p>,
 <p class="normal">129 </p>,
 <p class="normal">81 </p>,
 <p class="normal">12 </p>,
 <p class="normal">6 </p>,
 <p class="normal">22 </p>,
 <p class="normal">18 </p>,
 <p class="normal">9 </p>,
 <p class="normal">754 </p>,
 <p class="normal">14 </p>,
 <p class="normal">5 </p>,
 <p class="normal">  </p>,
 <p class="normal">  </p>,
 <p class="normal">18 </p>,
 <p class="normal">7 </p>,
 <p class="normal">7 </p>,
 <p class="normal">  </p>,
 <p class="normal">7 </p>,
 <p class="normal">5 </p>,
 <p class="normal">6 </p>,
 <p class="normal">4 </p>,
 <p class="normal">18 </p>,
 <p class="normal">11 </p>,
 <p class="normal">25 </p>,
 <p class="normal">  </p>]
```

In [457]:

```
numero_var_s=str(numero_var)
pat3= "[0-9]+"
re.findall(pat3,numero_var_s)
```

Out[457]:

```
['8',
 '14',
 '38',
 '294',
 '279',
 '7',
 '69',
 '8',
 '26',
 '1',
 '4',
 '4',
 '9',
 '10',
 '34',
```

```
    '32',
    '13',
    '6',
    '14',
    '22',
    '36',
    '6',
    '6',
    '42',
    '15',
    '9',
    '9',
    '54',
    '39',
    '33',
    '20',
    '12',
    '8',
    '30',
    '10',
    '3',
    '5',
    '75',
    '19',
    '27',
    '19',
    '1558',
    '34',
    '4',
    '617',
    '12',
    '16',
    '7',
    '4',
    '16',
    '7',
    '56',
    '18',
    '8',
    '22',
    '58',
    '61',
    '7',
    '649',
    '22',
    '168',
    '168',
    '8',
    '10',
    '64',
    '16',
    '8',
    '17',
    '72',
    '4',
    '6',
    '10',
    '35',
    '35',
    '4',
    '102',
    '57',
    '22',
    '44',
    '45',
    '5',
    '9',
    '21',
    '32',
    '17',
    '16',
    '38',
    '21',
    '40',
    '13',
    '8',
    '17',
```

```
    '15',
    '22',
    '68',
    '40',
    '17',
    '89',
    '4',
    '12',
    '86',
    '72',
    '61',
    '12',
    '481',
    '42',
    '5',
    '90',
    '5',
    '14',
    '20',
    '13',
    '36',
    '19',
    '9',
    '18',
    '4',
    '60',
    '10',
    '10',
    '4',
    '3',
    '11',
    '11',
    '6',
    '13',
    '100000',
    '9',
    '101',
    '10000',
    '20000',
    '100000',
    '5000',
    '500',
    '73',
    '43',
    '23',
    '3',
    '5',
    '256',
    '591',
    '70',
    '91',
    '10',
    '128',
    '6',
    '12',
    '3231961',
    '5409',
    '26',
    '10',
    '23',
    '24',
    '13',
    '8',
    '27',
    '50',
    '90',
    '138672',
    '386',
    '12',
    '147',
    '6',
    '8',
    '27',
    '10000',
    '20000',
    '10000',
    '54877',
```

```
    '4702',
    '24',
    '29',
    '17',
    '3',
    '128',
    '10',
    '242',
    '120',
    '4',
    '13',
    '52',
    '47',
    '857',
    '9',
    '7',
    '200',
    '4',
    '561',
    '64',
    '8',
    '7',
    '10',
    '9',
    '4',
    '8',
    '77',
    '51',
    '18',
    '1950000',
    '18',
    '1300',
    '41',
    '6',
    '5625',
    '5',
    '7',
    '33',
    '5',
    '15',
    '9',
    '19',
    '5',
    '5',
    '1000000',
    '129',
    '20',
    '152',
    '24',
    '16',
    '35',
    '17',
    '5',
    '18',
    '28',
    '7',
    '309',
    '3',
    '6',
    '16',
    '13',
    '5',
    '6',
    '8',
    '2',
    '4',
    '148',
    '55',
    '17',
    '8',
    '42',
    '26',
    '50',
    '2',
    '281',
    '120',
    '6',
```

```
    '120432',
    '150000',
    '529',
    '16',
    '2500',
    '5',
    '68',
    '16',
    '100',
    '216',
    '23',
    '33',
    '140256',
    '19',
    '20',
    '20',
    '49',
    '12',
    '30',
    '5232',
    '20',
    '1',
    '61',
    '27',
    '53',
    '11',
    '25',
    '0',
    '20',
    '9',
    '3',
    '561',
    '82',
    '13',
    '16',
    '13',
    '28',
    '4',
    '24',
    '34',
    '8',
    '128',
    '15',
    '513',
    '7',
    '7',
    '5',
    '15',
    '480000',
    '11',
    '54',
    '561',
    '64',
    '6',
    '116',
    '19',
    '5812',
    '9',
    '32',
    '29',
    '67',
    '25',
    '6400',
    '10',
    '5',
    '13',
    '98',
    '36',
    '69',
    '2158859',
    '518',
    '15',
    '179',
    '12',
    '38',
    '65',
    '102',
```

```
    '86',
    '7',
    '53',
    '20',
    '1',
    '71',
    '29',
    '20531',
    '65',
    '3',
    '22',
    '38',
    '22',
    '4814',
    '698',
    '13',
    '10',
    '59',
    '56',
    '482',
    '171',
    '5',
    '500',
    '411',
    '8519',
    '21',
    '21',
    '171',
    '7',
    '49',
    '12',
    '3',
    '21',
    '9',
    '8',
    '7',
    '2',
    '11',
    '11',
    '173',
    '5',
    '15',
    '3',
    '105',
    '25000',
    '18',
    '21000',
    '115',
    '21',
    '206',
    '43680',
    '8',
    '10',
    '59',
    '10',
    '5',
    '5',
    '1000',
    '138',
    '16',
    '2',
    '10',
    '8',
    '6',
    '129',
    '81',
    '12',
    '6',
    '22',
    '18',
    '9',
    '754',
    '14',
    '5',
    '18',
    '7',
    '7',
```

```
 '7',
 '5',
 '6',
 '4',
 '18',
 '11',
 '25']
```

In [458]:

```
año = datos1[6:3310:7]
```

In [459]:

```
año
```

Out[459]:

```
[<p class="normal">1995 </p>,
 <p class="normal">1996 </p>,
 <p class="normal"> </p>,
 <p class="normal">1998 </p>,
 <p class="normal">1998 </p>,
 <p class="normal">1992 </p>,
 <p class="normal">1987 </p>,
 <p class="normal">1992 </p>,
 <p class="normal">1993 </p>,
 <p class="normal">1987 </p>,
 <p class="normal">1994 </p>,
 <p class="normal">1994 </p>,
 <p class="normal"> </p>,
 <p class="normal">1988 </p>,
 <p class="normal">1992 </p>,
 <p class="normal">1995 </p>,
 <p class="normal">1995 </p>,
 <p class="normal">1990 </p>,
 <p class="normal">1997 </p>,
 <p class="normal">1996 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">1989 </p>,
 <p class="normal">1994 </p>,
 <p class="normal"> </p>,
 <p class="normal"> </p>,
 <p class="normal">1995 </p>,
 <p class="normal">1992 </p>,
 <p class="normal">1987 </p>,
 <p class="normal">1997 </p>,
 <p class="normal">1998 </p>,
 <p class="normal">1995 </p>,
 <p class="normal">1998 </p>,
 <p class="normal">1994 </p>,
 <p class="normal">1989 </p>,
 <p class="normal">1996 </p>,
 <p class="normal">1990 </p>,
 <p class="normal">1990 </p>,
 <p class="normal">1987 </p>,
 <p class="normal">1999 </p>,
 <p class="normal">1989 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">1989 </p>,
 <p class="normal">1990 </p>,
 <p class="normal">1998 </p>,
 <p class="normal">1989 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">1994 </p>,
 <p class="normal">1990 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">1988 </p>,
 <p class="normal">1990 </p>,
```

```
<p class="normal">1991 </p>,
<p class="normal">1990 </p>,
<p class="normal"> </p>,
<p class="normal">1992 </p>,
<p class="normal">1988 </p>,
<p class="normal">1990 </p>,
<p class="normal">1996 </p>,
<p class="normal">1995 </p>,
<p class="normal">1990 </p>,
<p class="normal"> </p>,
<p class="normal">1992 </p>,
<p class="normal">1992 </p>,
<p class="normal">1994 </p>,
<p class="normal">1987 </p>,
<p class="normal">1994 </p>,
<p class="normal">1994 </p>,
<p class="normal">1997 </p>,
<p class="normal">1991 </p>,
<p class="normal">1995 </p>,
<p class="normal">1998 </p>,
<p class="normal">1998 </p>,
<p class="normal">1993 </p>,
<p class="normal">1988 </p>,
<p class="normal">1992 </p>,
<p class="normal">1993 </p>,
<p class="normal">1988 </p>,
<p class="normal">1989 </p>,
<p class="normal">1988 </p>,
<p class="normal">1987 </p>,
<p class="normal">1993 </p>,
<p class="normal">1988 </p>,
<p class="normal">1999 </p>,
<p class="normal">2001 </p>,
<p class="normal">2001 </p>,
<p class="normal">1992 </p>,
<p class="normal">1993 </p>,
<p class="normal">1997 </p>,
<p class="normal">1991 </p>,
<p class="normal">1987 </p>,
<p class="normal">1994 </p>,
<p class="normal">1988 </p>,
<p class="normal">1987 </p>,
<p class="normal">1993 </p>,
<p class="normal">1988 </p>,
<p class="normal">1988 </p>,
<p class="normal">1991 </p>,
<p class="normal">1996 </p>,
<p class="normal">1990 </p>,
<p class="normal">1999 </p>,
<p class="normal">1999 </p>,
<p class="normal">2002 </p>,
<p class="normal">2000 </p>,
<p class="normal">1999 </p>,
<p class="normal">1999 </p>,
<p class="normal">2001 </p>,
<p class="normal">1999 </p>,
<p class="normal">1999 </p>,
<p class="normal">2000 </p>,
<p class="normal">1999 </p>,
<p class="normal">2000 </p>,
<p class="normal">1999 </p>,
<p class="normal">1999 </p>,
<p class="normal">1998 </p>,
<p class="normal">1999 </p>,
<p class="normal">2001 </p>,
<p class="normal">1999 </p>,
<p class="normal">2003 </p>,
<p class="normal">1999 </p>,
```

```
<p class="normal">1997 </p>,
<p class="normal">1999 </p>,
<p class="normal">1999 </p>,
<p class="normal">1998 </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal">1994 </p>,
<p class="normal"> </p>,
<p class="normal">1993 </p>,
<p class="normal">1990 </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal"> </p>,
<p class="normal">1989 </p>,
<p class="normal">2006 </p>,
<p class="normal">2006 </p>,
<p class="normal">2007 </p>,
<p class="normal">2007 </p>,
<p class="normal">2007 </p>,
<p class="normal">2007 </p>,
<p class="normal">2008 </p>,
<p class="normal">2008 </p>,
<p class="normal">2008 </p>,
<p class="normal">2007 </p>,
<p class="normal">2008 </p>,
<p class="normal">2008 </p>,
<p class="normal">2008 </p>,
<p class="normal">2008 </p>,
<p class="normal">2008 </p>,
<p class="normal">2008 </p>,
<p class="normal">2008 </p>,
<p class="normal">2008 </p>,
<p class="normal">2008 </p>,
<p class="normal">2008 </p>,
<p class="normal">2008 </p>,
<p class="normal">2009 </p>,
<p class="normal">2008 </p>,
<p class="normal">2008 </p>,
<p class="normal">2008 </p>,
<p class="normal">2009 </p>,
<p class="normal">2009 </p>,
<p class="normal">2009 </p>,
<p class="normal">2009 </p>,
<p class="normal">2009 </p>,
<p class="normal">2009 </p>,
<p class="normal">2010 </p>,
<p class="normal">2009 </p>,
<p class="normal">2010 </p>,
<p class="normal">2010 </p>,
<p class="normal">2010 </p>,
<p class="normal">2010 </p>,
<p class="normal">2010 </p>,
<p class="normal">2010 </p>,
<p class="normal">2010 </p>,
<p class="normal">2010 </p>,
<p class="normal">2010 </p>,
<p class="normal">2010 </p>,
<p class="normal">2011 </p>,
<p class="normal">2011 </p>,
<p class="normal">2011 </p>,
<p class="normal">2011 </p>,
<p class="normal">2011 </p>,
<p class="normal">2011 </p>,
<p class="normal">2011 </p>,
<p class="normal">2011 </p>,
<p class="normal">2011 </p>,
<p class="normal">2011 </p>,
<p class="normal">2011 </p>,
<p class="normal">2011 </p>,
<p class="normal">2011 </p>,
<p class="normal">2011 </p>,
```

```
<p class="normal">2011 </p>,
<p class="normal">2011 </p>,
<p class="normal">2011 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2012 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2014 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2013 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2013 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
```

```
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2015 </p>,
<p class="normal">2014 </p>,
<p class="normal">2014 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2014 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2015 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2015 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2015 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2017 </p>,
<p class="normal">2016 </p>,
<p class="normal">2017 </p>,
<p class="normal">2016 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2017 </p>,
```

```
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2016 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2016 </p>,
<p class="normal">2017 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2016 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2016 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2017 </p>,
<p class="normal">2016 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2016 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
<p class="normal">2018 </p>,
```

```
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2019 </p>,
 <p class="normal">2019 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2018 </p>,
 <p class="normal">2019 </p>]
```

In [460]:

```
año_s=str(año)
re.findall(pat3,año_s)
```

Out[460]:

```
['1995',
 '1996',
 '1998',
 '1998',
 '1992',
 '1987',
 '1992',
 '1993',
 '1987',
 '1994',
 '1994',
 '1988',
 '1992',
 '1995',
 '1995',
 '1990',
 '1997',
 '1996',
 '1988',
 '1989',
 '1994',
 '1995',
 '1992',
 '1987',
 '1997',
 '1998',
 '1995',
 '1998',
 '1994',
 '1989',
 '1996',
 '1990',
 '1990',
 '1987',
 '1999',
 '1989',
 '1988',
 '1988',
 '1989',
 '1990',
```

```
    '1998',
    '1989',
    '1988',
    '1994',
    '1990',
    '1988',
    '1988',
    '1990',
    '1991',
    '1990',
    '1992',
    '1988',
    '1990',
    '1996',
    '1995',
    '1990',
    '1992',
    '1992',
    '1994',
    '1987',
    '1994',
    '1994',
    '1997',
    '1991',
    '1995',
    '1998',
    '1998',
    '1993',
    '1988',
    '1992',
    '1993',
    '1988',
    '1989',
    '1988',
    '1987',
    '1993',
    '1988',
    '1999',
    '2001',
    '2001',
    '1992',
    '1993',
    '1997',
    '1991',
    '1987',
    '1994',
    '1988',
    '1987',
    '1993',
    '1988',
    '1988',
    '1991',
    '1996',
    '1990',
    '1999',
    '1999',
    '2002',
    '2000',
    '1999',
    '1999',
    '2001',
    '1999',
    '1999',
    '2000',
    '1999',
    '2000',
    '1999',
    '1999',
    '1998',
    '1999',
    '2001',
    '1999',
    '2003',
    '1999',
    '1999',
    '1997',
    '1999',
```

```
    '1999',
    '1998',
    '1994',
    '1993',
    '1990',
    '1989',
    '2006',
    '2006',
    '2007',
    '2007',
    '2007',
    '2007',
    '2008',
    '2008',
    '2008',
    '2007',
    '2008',
    '2008',
    '2008',
    '2008',
    '2008',
    '2008',
    '2008',
    '2008',
    '2008',
    '2008',
    '2008',
    '2009',
    '2008',
    '2008',
    '2008',
    '2009',
    '2009',
    '2009',
    '2009',
    '2009',
    '2009',
    '2010',
    '2009',
    '2010',
    '2010',
    '2010',
    '2010',
    '2010',
    '2010',
    '2010',
    '2010',
    '2010',
    '2010',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2011',
    '2012',
    '2012',
    '2012',
    '2012',
    '2012',
    '2012',
    '2012',
    '2012',
    '2012',
    '2012',
```

```
    '2012',
    '2012',
    '2012',
    '2012',
    '2012',
    '2012',
    '2012',
    '2012',
    '2012',
    '2012',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2013',
    '2014',
    '2013',
    '2013',
    '2013',
    '2014',
    '2014',
    '2014',
    '2014',
    '2014',
    '2014',
    '2014',
    '2014',
    '2014',
    '2014',
    '2013',
    '2014',
    '2014',
    '2014',
    '2014',
    '2014',
    '2014',
    '2014',
    '2014',
    '2014',
    '2014
```

```
    '2014',
    '2014',
    '2014',
    '2014',
    '2015',
    '2014',
    '2014',
    '2015',
    '2015',
    '2015',
    '2015',
    '2015',
    '2015',
    '2015',
    '2015',
    '2014',
    '2015',
    '2015',
    '2015',
    '2015',
    '2015',
    '2015',
    '2015',
    '2015',
    '2015',
    '2015',
    '2015',
    '2015',
    '2015',
    '2015',
    '2015',
    '2016',
    '2016',
    '2015',
    '2016',
    '2016',
    '2015',
    '2016',
    '2016',
    '2016',
    '2016',
    '2016',
    '2016',
    '2016',
    '2016',
    '2016',
    '2016',
    '2016',
    '2016',
    '2016',
    '2016',
    '2016',
    '2016',
    '2017',
    '2016',
    '2017',
    '2016',
    '2017',
    '2017',
    '2016',
    '2016',
    '2017',
    '2017',
    '2016',
    '2017',
    '2017',
    '2017',
    '2017',
    '2016',
    '2017',
    '2016',
    '2016',
    '2016',
    '2017',
    '2017',
```

```
'2017',
'2017',
'2017',
'2017',
'2017',
'2016',
'2016',
'2016',
'2016',
'2017',
'2017',
'2017',
'2017',
'2017',
'2016',
'2017',
'2017',
'2017',
'2017',
'2017',
'2017',
'2017',
'2016',
'2016',
'2016',
'2017',
'2017',
'2017',
'2017',
'2017',
'2017',
'2017',
'2017',
'2016',
'2018',
'2018',
'2016',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
'2018',
```

```
 '2018',
 '2018',
 '2018',
 '2018',
 '2018',
 '2018',
 '2018',
 '2018',
 '2019',
 '2019',
 '2018',
 '2018',
 '2018',
 '2019']
```

In [461]:

```
from bs4 import Comment
```

In [462]:

```
comments = soup.findAll(text=lambda text:isinstance(text, Comment))
```

In [463]:

```
comments
```

Out[463]:

```
[' Stylesheet link ',
 ' SITE HEADER (INCLUDES LOGO AND SEARCH BOX) ',
 ' SITE HEADER (INCLUDES LOGO AND SEARCH BOX) ',
 ' Search Google ',
 ' Search Google ',
 ' <td><p class="normal, whitetext"><b>Abstract</b></p></td> ',
 ' <td><p class="normal, whitetext"><b>Area</b></p></td> ',
 ' <td><p class="normal">Predict the age of abalone from physical measurements </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Predict whether income exceeds $50K/yr based on census data.  Also known
as "Census Income" dataset. </p></td> ',
 ' <td><p class="normal">Social </p></td> ',
 ' <td><p class="normal">Steel annealing data </p></td> ',
 ' <td><p class="normal">Physical </p></td> ',
 ' <td><p class="normal">Log of anonymous users of www.microsoft.com; predict areas of the web
site a user visited based on data on other areas the user visited. </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Distinguish between the presence and absence of cardiac arrhythmia and cl
assify it in one of the 16 groups. </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Dataset artificially generated by using first order theory which
describes structure of ten capital letters of English alphabet </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Nominal audiology dataset from Baylor </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Standardized version of the original audiology database </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Revised from CMU StatLib library, data concerns city-cycle fuel
consumption </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">From 1985 Ward\'s Automotive Yearbook </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">Badges labeled with a "+" or "-" as a function of a person\'s name <
/p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">Balance scale weight & distance database </p></td> ',
 ' <td><p class="normal">Social </p></td> ',
 ' <td><p class="normal">Data previously used in cognitive psychology experiment; 4 data sets repr
esent different conditions of an experiment </p></td> ',
 ' <td><p class="normal">Social </p></td> ',
 ' <td><p class="normal">Breast Cancer Data (Restricted Access) </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Original Wisconsin Breast Cancer Database </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Prognostic Wisconsin Breast Cancer Database </p></td> ',
```

```
  <td><p class="normal">Prognostic Wisconsin Breast Cancer Database </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Diagnostic Wisconsin Breast Cancer Database </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Bridges database that has original and numeric-discretized datasets 
</p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">Derived from simple hierarchical decision model, this database may be use
ful for testing constructive induction and structure discovery methods. </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">Predict whether income exceeds $50K/yr based on census data.  Also known
as "Adult" dataset. </p></td> ',
 ' <td><p class="normal">Social </p></td> ',
 ' <td><p class="normal">Knight Pin Chess End-Game Database Creator </p></td> ',
 ' <td><p class="normal">Game </p></td> ',
 ' <td><p class="normal">King+Rook versus King+Pawn on a7 (usually abbreviated KRKPA7). </p><
/td> ',
 ' <td><p class="normal">Game </p></td> ',
 ' <td><p class="normal">Chess Endgame Database for White King and Rook against Black King
(KRK). </p></td> ',
 ' <td><p class="normal">Game </p></td> ',
 ' <td><p class="normal">6 different domain theories for generating legal moves of chess </p>
</td> ',
 ' <td><p class="normal">Game </p></td> ',
 ' <td><p class="normal">Time-series data based on chorales; challenge is to learn generative
grammar; data in Lisp </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">Contains connect-4 positions </p></td> ',
 ' <td><p class="normal">Game </p></td> ',
 ' <td><p class="normal">This data concerns credit card applications; good mix of attributes 
</p></td> ',
 ' <td><p class="normal">Financial </p></td> ',
 ' <td><p class="normal">Includes domain theory (generated by talking to Japanese domain experts);
data in Lisp </p></td> ',
 ' <td><p class="normal">Financial </p></td> ',
 ' <td><p class="normal">Relative CPU Performance Data, described in terms of its cycle time,
memory size, etc. </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Dataset is a subset of the 1987 National Indonesia Contraceptive
Prevalence Survey. </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Forest CoverType dataset </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Used in decision tree induction for mitigating process delays known as "c
ylinder bands" in rotogravure printing </p></td> ',
 ' <td><p class="normal">Physical </p></td> ',
 ' <td><p class="normal">Aim for this dataset is to determine the type of Eryhemato-Squamous
Disease. </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">This diabetes dataset is from AIM \'94 </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Generates application domains based on specific parameters, number of fea
tures, and proportion of positive to negative examples </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">Five concepts, expressed as predicates, to be learned </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">Assorted small-scale domain theories </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Data for classifying if patients will survive for at least one year after
a heart attack </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">This data contains protein localization sites </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">From Collins Gem Guide to Flags, 1986 </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">Cases collected mostly from investigations in physical science; intention
is to evaluate function-finding algorithms </p></td> ',
 ' <td><p class="normal">Physical </p></td> ',
 ' <td><p class="normal">From USA Forensic Science Service; 6 types of glass; defined in terms of
their oxide content (i.e. Na, Fe, K, etc) </p></td> ',
 ' <td><p class="normal">Physical </p></td> ',
 ' <td><p class="normal">Dataset contains cases from study conducted on the survival of patients w
ho had undergone surgery for breast cancer </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Topic: human subjects study </p></td> ',
 ' <td><p class="normal">Social </p></td> ',
 ' <td><p class="normal">4 databases: Cleveland, Hungary, Switzerland, and the VA Long Beach 
</p></td> '
```

```
</p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">From G.Gong: CMU; Mostly Boolean or numeric-valued attribute types; Inclu
des cost data (donated by Peter Turney) </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Well documented attributes; 368 instances with 28 attributes (continuous,
discrete, and nominal); 30% missing values </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Data set prepared for the use of participants for the 1994 AAAI Spring Sy
mposium on Artificial Intelligence in Medicine. </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Image data described by high-level numeric-valued attributes, 7
classes </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">This dataset represents a set of possible advertisements on Internet page
s. </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Classification of radar returns from the ionosphere </p></td> ',
 ' <td><p class="normal">Physical </p></td> ',
 ' <td><p class="normal">Famous database; from Fisher, 1936 </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal"> Goal: Predict which letter-name was spoken--a simple classification
task. </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Relational dataset </p></td> ',
 ' <td><p class="normal">Social </p></td> ',
 ' <td><p class="normal">From Collective Bargaining Review </p></td> ',
 ' <td><p class="normal">Social </p></td> ',
 ' <td><p class="normal">From Classification and Regression Trees book; We provide here 2 C
programs for generating sample databases </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Database for fitting contact lenses </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">Database of character image features; try to identify the letter </p
></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">BUPA Medical Research Ltd. database donated by Richard S. Forsyth </
p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">All code for Logic Theorist </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Lung cancer data; no attribute definitions </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">This lymphography domain was obtained from the University Medical Centre,
Institute of Oncology, Ljubljana, Yugoslavia.  (Restricted access) </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Fault diagnosis problem of electromechanical devices; also PUMPS DATA SET
is newer version with domain theory and results </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Meta-Data was used in order to give advice about which classification met
hod is appropriate for a particular dataset (taken from results of Statlog project). </p></td
> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">Learning concepts from sensor data of a mobile robot; set of data sets&nb
sp;</p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">E. Coli promoter gene sequences (DNA) with partial domain theory </p
></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">From CMU connectionist bench repository; Classifies secondary structure o
f certain globular proteins </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Primate splice-junction gene sequences (DNA) with associated imperfect do
main theory </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">A set of three artificial domains over the same attribute space; Used to
test a wide range of induction algorithms </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">Horn-clause model that qualitatively simulates moral reasoning; Theory
includes negated literals </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">This dataset consists of features of handwritten numerals (`0\'--`9\') ex
tracted from a collection of Dutch utility maps </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">From Audobon Society Field Guide; mushrooms described in terms of
physical characteristics; classification: poisonous or edible </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">The goal is to learn to predict whether new molecules will be musks or
```

' <td> The goal is to learn to predict whether new molecules will be musks or
non-musks </td> ',
' <td> Physical </td> ',
' <td> The goal is to learn to predict whether new molecules will be musks or
non-musks </td> ',
' <td> Physical </td> ',
' <td> Nursery Database was derived from a hierarchical decision model
originally developed to rank applications for nursery schools. </td> ',
' <td> Social </td> ',
' <td> Used in research to generate features for an inductive learning system&nb
sp; </td> ',
' <td> Game </td> ',
' <td> The problem consists of classifying all the blocks of the page layout of
a document that has been detected by a segmentation process. </td> ',
' <td> Computer </td> ',
' <td> Two versions of this database available; see folder </td> ',
' <td> Computer </td> ',
' <td> Digit database of 250 samples from 44 writers </td> ',
' <td> Computer </td> ',
' <td> Dataset of patient features </td> ',
' <td> Life </td> ',
' <td> From Ljubljana Oncology Institute </td> ',
' <td> Life </td> ',
' <td> Assorted domains like blocksworld, eightpuzzle, and schedworld.
</td> ',
' <td> Other </td> ',
' <td> Two sets of datasets are given: pyrimidines and triazines </td>
',
' <td> Physical </td> ',
' <td> The file animals.c is a data generator of structured instances
representing quadruped animals </td> ',
' <td> Life </td> ',
' <td> Data was from a simulation of a servo system </td> ',
' <td> Computer </td> ',
' <td> Tiny database; all nominal values </td> ',
' <td> Physical </td> ',
' <td> Each class attribute counts the number of solar flares of a certain class
that occur in a 24 hour period </td> ',
' <td> Physical </td> ',
' <td> Michalski\'s famous soybean disease database </td> ',
' <td> Life </td> ',
' <td> Michalski\'s famous soybean disease database </td> ',
' <td> Life </td> ',
' <td> Task: predict the number of O-rings that experience thermal distress on a
flight at 31 degrees F given data on the previous 23 shuttle flights </td> ',
' <td> Physical </td> ',
' <td> From IRAS data -- NASA Ames Research Center </td> ',
' <td> Physical </td> ',
' <td> Classifying Email as Spam or Non-Spam </td> ',
' <td> Computer </td> ',
' <td> Data on cardiac Single Proton Emission Computed Tomography (SPECT) images
. Each patient classified into two categories: normal and abnormal. </td> ',
' <td> Life </td> ',
' <td> Data on cardiac Single Proton Emission Computed Tomography (SPECT) images
. Each patient classified into two categories: normal and abnormal. </td> ',
' <td> Life </td> ',
' <td> Data on sponges; Attributes in spanish </td> ',
' <td> Life </td> ',
' <td> Various Databases: Vehicle silhouttes, Landsat Sattelite, Shuttle,
Australian Credit Approval, Heart Disease, Image Segmentation, German Credit </td> ',
' <td> Other </td> ',
' <td> Student Loan Relational Domain </td> ',
' <td> Social </td> ',
' <td> The data consist of evaluations of teaching performance; scores are "low"
, "medium", or "high" </td> ',
' <td> Other </td> ',
' <td> Binary classification task on possible configurations of tic-tac-toe game
  </td> ',
' <td> Game </td> ',
' <td> 10 separate databases from Garavan Institute </td> ',
' <td> Life </td> ',
' <td> 2 data formats (structured, one-instance-per-line) </td> ',
' <td> Other </td> ',
' <td> Data in original (LISP-readable) form </td> ',
' <td> Other </td> ',
' <td> 1984 United Stated Congressional Voting Records; Classify as Republican o
r Democrat </td> ',

' <td> Social </td> ',
' <td> Multiple classes predict plant state </td> ',
' <td> Physical </td> ',
' <td> CART book\'s waveform domains </td> ',
' <td> Physical </td> ',
' <td> CART book\'s waveform domains </td> ',
' <td> Physical </td> ',
' <td> Using chemical analysis determine the origin of wines </td> ',
' <td> Physical </td> ',
' <td> Predicting the Cellular Localization Sites of Proteins </td> ',
' <td> Life </td> ',
' <td> Artificial, 7 classes of animals </td> ',
' <td> Life </td> ',
' <td> Various datasets without documentation (feel free to explore!) </td> ',
' <td> Other </td> ',
' <td> This data set consists of 20000 messages taken from 20 newsgroups. </td> ',
' <td> Other </td> ',
' <td> This data consists of sample of Auslan (Australian Sign Language) signs. Examples of 95 signs were collected from five signers with a total of 6650 sign samples. </td> ',
' <td> Other </td> ',
' <td> This data consists of sample of Auslan (Australian Sign Language) signs. 27 examples of each of 95 Auslan signs were captured from a native signer using high-quality position trackers </td> ',
' <td> Other </td> ',
' <td> The USCensus1990raw data set contains a one percent sample of the Public Use Microdata Samples (PUMS) person records drawn from the full 1990 census sample. </td> ',
' <td> Social </td> ',
' <td> This data set contains weighted census data extracted from the 1994 and 1995 current population surveys conducted by the U.S. Census Bureau. </td> ',
' <td> Social </td> ',
' <td> This data set is from the 1999 Computational Intelligence and Learning (COIL) competition. The data contains measurements of river chemical concentrations and algae densities. </td> ',
' <td> Physical </td> ',
' <td> This dataset contains image features extracted from a Corel image collection. Four sets of features are available based on the color histogram, color histogram layout, color moments, and co-occurence </td> ',
' <td> Other </td> ',
' <td> Data giving characteristics of each ORF (potential gene) in the E. coli genome. Sequence, homology (similarity to other genes) and structural information, and function (if known) are provided. </td> ',
' <td> Life </td> ',
' <td> This data arises from a large study to examine EEG correlates of genetic predisposition to alcoholism. It contains measurements from 64 electrodes placed on the scalp sampled at 256 Hz </td> ',
' <td> Life </td> ',
' <td> The data set contains oceanographic and surface meteorological readings taken from a series of buoys positioned throughout the equatorial Pacific. </td> ',
' <td> Physical </td> ',
' <td> This data contains a record of user interactions with the Entree Chicago restaurant recommendation system. </td> ',
' <td> Other </td> ',
' <td> This data consists of 640 black and white face images of people taken with varying pose (straight, left, right, up), expression (neutral, happy, sad, angry), eyes (wearing sunglasses or not), and size </td> ',
' <td> Other </td> ',
' <td> This data set used in the CoIL 2000 Challenge contains information on customers of an insurance company. The data consists of 86 variables and includes product usage data and socio-demographic data </td> ',
' <td> Social </td> ',
' <td> This data contains general demographic information on internet users in 1997. </td> ',
' <td> Computer </td> ',
' <td> This data set contains unweighted PUMS census data from the Los Angeles and Long Beach areas for the years 1970, 1980, and 1990. </td> ',
' <td> Social </td> ',
' <td> This dataset records 640 time series of 12 LPC cepstrum coefficients taken from nine male speakers. </td> ',
' <td> Other </td> ',
' <td> This is the data set used for The Second International Knowledge Discovery and Data Mining Tools Competition, which was held in conjunction with KDD-98 </td> ',
' <td> Other </td> ',
' <td> This is the data set used for The Third International Knowledge Discovery

and Data Mining Tools Competition, which was held in conjunction with KDD-99 </td> ',
' <td> Computer </td> ',
' <td> Data giving characteristics of each ORF (potential gene) in the M. tuberculosis bacterium. Sequence, homology (similarity to other genes) and structural information, and function (if known) are provided </td> ',
' <td> Life </td> ',
' <td> This data set contains a list of over 10000 films including many older, odd, and cult films. There is information on actors, casts, directors, producers, studios, etc. </td> ',
' <td> Other </td> ',
' <td> This data describes the page visits of users who visited msnbc.com on September 28, 1999. Visits are recorded at the level of URL category (see description) and are recorded in time order. </td> ',
' <td> Computer </td> ',
' <td> This data set consists of (a) 129,000 abstracts describing NSF awards for basic research, (b) bag-of-word data files extracted from the abstracts, (c) a list of words used for indexing the bag-of-word </td> ',
' <td> Other </td> ',
' <td> This dataset contains time series sensor readings of the Pioneer-1 mobile robot. The data is broken into "experiences" in which the robot takes action for some period of time and experiences a control </td> ',
' <td> Computer </td> ',
' <td> This data set is designed for testing indexing schemes in time series databases. The data appears highly periodic, but never exactly repeats itself. </td> ',
' <td> Other </td> ',
' <td> This is a collection of documents that appeared on Reuters newswire in 1987. The documents were assembled and indexed with categories. </td> ',
' <td> Other </td> ',
' <td> This dataset contains force and torque measurements on a robot after failure detection. Each failure is characterized by 15 force/torque samples collected at regular time intervals </td> ',
' <td> Physical </td> ',
' <td> This data consists of synthetically generated control charts. </td> ',
' <td> Other </td> ',
' <td> This database contains HTML source of web pages plus the ratings of a single user on these web pages. Web pages are on four seperate subjects (Bands- recording artists; Goats; Sheep; and BioMedical) </td> ',
' <td> Computer </td> ',
' <td> This file contains 9 sets of sanitized user data drawn from the command histories of 8 UNIX computer users at Purdue over the course of up to 2 years. </td> ',
' <td> Computer </td> ',
' <td> The JARtool project was a pioneering effort to develop an automatic system for cataloging small volcanoes in the large set of Venus images returned by the Magellan spacecraft. </td> ',
' <td> Physical </td> ',
' <td> This file concerns credit card applications. This database exists elsewhere in the repository (Credit Screening Database) in a slightly different form </td> ',
' <td> Financial </td> ',
' <td> This dataset classifies people described by a set of attributes as good or bad credit risks. Comes in two formats (one all numeric). Also comes with a cost matrix </td> ',
' <td> Financial </td> ',
' <td> This dataset is a heart disease database similar to a database already present in the repository (Heart Disease databases) but in a slightly different form </td> ',
' <td> Life </td> ',
' <td> Multi-spectral values of pixels in 3x3 neighbourhoods in a satellite image, and the classification associated with the central pixel in each neighbourhood </td> ',
' <td> Physical </td> ',
' <td> This dataset is an image segmentation database similar to a database already present in the repository (Image segmentation database) but in a slightly different form. </td> ',
' <td> Other </td> ',
' <td> The shuttle dataset contains 9 attributes all of which are numerical. Approximately 80% of the data belongs to class 1 </td> ',
' <td> Physical </td> ',
' <td> 3D objects within a 2D image by application of an ensemble of shape feature extractors to the 2D silhouettes of the objects. </td> ',
' <td> Other </td> ',
' <td> The file "nettalk.data" contains a list of 20,008 English words, along with a phonetic transcription for each word. The task is to train a network to produce the proper phonemes </td> ',
' <td> Other </td> ',
' <td> The task is to train a network to discriminate between sonar signals bounced off a metal cylinder and those bounced off a roughly cylindrical rock. </td> ',

' <td> Physical </td> ',
' <td> Speaker independent recognition of the eleven steady state vowels of Brit
ish English using a specified training set of lpc derived log area ratios. </td> ',
' <td> Other </td> ',
' <td> Domain Theory on Economic Sanctions; Undocumented </td> ',
' <td> Financial </td> ',
' <td> Undocumented </td> ',
' <td> Life </td> ',
' <td> Little Documentation </td> ',
' <td> Physical </td> ',
' <td> This data comes from the main door of the CalIt2 building at UCI. </
p></td> ',
' <td> Other </td> ',
' <td> Loop sensor data was collected for the Glendale on ramp for the 101 North
freeway in Los Angeles </td> ',
' <td> Other </td> ',
' <td> Purpose is to predict poker hands </td> ',
' <td> Game </td> ',
' <td> Data are MC generated to simulate registration of high energy gamma parti
cles in an atmospheric Cherenkov telescope </td> ',
' <td> Physical </td> ',
' <td> Data consists of written characters in a UNIPEN-like format </td
> ',
' <td> Computer </td> ',
' <td> Discrimination of benign and malignant mammographic masses based on BI-RA
DS attributes and the patient\'s age. </td> ',
' <td> Life </td> ',
' <td> This is a difficult regression task, where the aim is to predict the burn
ed area of forest fires, in the northeast region of Portugal, by using meteorological and other da
ta (see details at: http://www.dsi.uminho.pt/~pcortez/forestfires). </td> ',
' <td> Physical </td> ',
' <td> This dataset is created by reading out 200 files from the 10 largest Reut
ers \r\nclasses and using an Automatic Speech Recognition system to create \r\ncorresponding trans
criptions. </td> ',
' <td> Business </td> ',
' <td> This data set contains five text collections in the form of bags-of-words
. </td> ',
' <td> Other </td> ',
' <td> Concrete is the most important material in civil engineering. The
concrete compressive strength is a highly nonlinear function of age and ingredients.   </t
d> ',
' <td> Physical </td> ',
' <td> Each record represents 100 points on a two-dimensional graph. When plotte
d in order (from 1 through 100) as the Y co-ordinate, the points will create either a Hill (a ⌧bum
p⌧ in the terrain) or a Valley (a ⌧dip⌧ in the terrain). </td> ',
' <td> Other </td> ',
' <td> ARCENE\'s task is to distinguish cancer versus normal patterns from mass-
spectrometric data. This is a two-class classification problem with continuous input variables.
This dataset is one of 5 datasets of the NIPS 2003 feature selection challenge. </td> ',
' <td> Life </td> ',
' <td> DEXTER is a text classification problem in a bag-of-word representation.
This is a two-class classification problem with sparse continuous input variables. This dataset is
one of five datasets of the NIPS 2003 feature selection challenge.\r\n </td> ',
' <td> Other </td> ',
' <td> DOROTHEA is a drug discovery dataset. Chemical compounds represented by s
tructural molecular features must be classified as active (binding to thrombin) or inactive. This
is one of 5 datasets of the NIPS 2003 feature selection challenge. </td> ',
' <td> Life </td> ',
' <td> GISETTE is a handwritten digit recognition problem. The problem is to sep
arate the highly confusible digits \'4\' and \'9\'. This dataset is one of five datasets of the NI
PS 2003 feature selection challenge.\r\n </td> ',
' <td> Computer </td> ',
' <td> MADELON is an artificial dataset, which was part of the NIPS 2003 feature
selection challenge. This is a two-class classification problem with continuous input variables. T
he difficulty is that the problem is multivariate and highly non-linear.   </td> ',
' <td> Other </td> ',
' <td> Two ground ozone level data sets are included in this collection. One is
the eight hour peak set (eighthr.data), the other is the one hour peak set (onehr.data). Those dat
a were collected from 1998 to 2004 at the Houston, Galveston and Brazoria area. </td> ',
' <td> Physical </td> ',
' <td> The objective is to determine the set of boolean rules that describe the
interactions of the nodes within this plant signaling network. The dataset includes 300 separate
boolean pseudodynamic simulations using an asynchronous update scheme.   </td> ',
' <td> Life </td> ',
' <td> Oxford Parkinson\'s Disease Detection Dataset </td> ',
' <td> Life </td> ',
' <td> Multiple, labelled samples of pen tip trajectories recorded whilst writin

g individual characters. All samples are from the same writer, for the purposes of primitive extra
ction. Only characters with a single pen-down segment were considered. </td> ',
' <td> Computer </td> ',
' <td> Data taken from the Blood Transfusion Service Center in Hsin-Chu City in
Taiwan -- this is a classification problem.   </td> ',
' <td> Business </td> ',
' <td> A pen-based database with more than 11k isolated handwritten
characters </td> ',
' <td> Computer </td> ',
' <td> 1593 handwritten digits from around 80 persons were scanned, stretched in
a rectangular box 16x16 in a gray scale of 256 values. </td> ',
' <td> Computer </td> ',
' <td> Data from a semi-conductor manufacturing process </td> ',
' <td> Computer </td> ',
' <td> Data has been extracted from the USDA plants database. It contains all pl
ants (species and genera) in the database and the states of USA and Canada where they occur.
</td> ',
' <td> Life </td> ',
' <td> The data set contains 15 classes of 24 instances each. Each class
references to a hand movement type in LIBRAS (Portuguese\r\nname \'L
ngua BRAsileira de Sinais\',
oficial brazilian signal language). </td> ',
' <td> Other </td> ',
' <td> Concrete is a highly complex material. The slump flow of concrete is not
only determined by the water content, but that is also influenced by other concrete
ingredients. </td> ',
' <td> Computer </td> ',
' <td> Communities within the United States. The data combines socio-economic da
ta from the 1990 US Census, law enforcement data from the 1990 US LEMAS survey, and crime data fro
m the 1995 FBI UCR. </td> ',
' <td> Social </td> ',
' <td> The data was created by a medical expert as a data set to test the expert
system, \r\nwhich will perform the presumptive diagnosis of two diseases of the urinary
system.\r\n </td> ',
' <td> Life </td> ',
' <td> Two datasets are included, related to red and white vinho verde wine samp
les, from the north of Portugal. The goal is to model wine quality based on physicochemical tests
(see [Cortez et al., 2009], http://www3.dsi.uminho.pt/pcortez/wine/). </td> ',
' <td> Business </td> ',
' <td> Anonymized 120-day subset of the ICML-09 URL data containing 2.4 million
examples and 3.2 million features. </td> ',
' <td> Computer </td> ',
' <td> The goal is to model mutant p53 transcriptional activity (active vs
inactive) based on data extracted from biophysical simulations.\r\n </td> ',
' <td> Life </td> ',
' <td> Oxford Parkinson\'s Disease Telemonitoring Dataset </td> ',
' <td> Life </td> ',
' <td> Marine sponges of the Demospongiae class classification domain.
</td> ',
' <td> Life </td> ',
' <td> This dataset contains sentences extracted from user reviews on a given to
pic. Example topics are “performance of Toyota Camry” and “sound quality of ipod nano”.
</td> ',
' <td> Computer </td> ',
' <td> Dataset with electrical impedance measurements of freshly excised tissue
samples from the breast. </td> ',
' <td> Life </td> ',
' <td> The dataset consists of measurements of fetal heart rate (FHR) and
uterine contraction (UC) features on cardiotocograms classified by expert obstetricians.
</td> ',
' <td> Life </td> ',
' <td> The data were collected as the SCITOS G5 robot navigates through the room
following the wall in a clockwise direction, for 4 rounds, using 24 ultrasound sensors arranged ci
rcularly around its \'waist\'. </td> ',
' <td> Computer </td> ',
' <td> This dataset contains timeseries of mel-frequency cepstrum coefficients (
MFCCs) corresponding to spoken Arabic digits. Includes data from 44 male and 44 female native Ara
bic speakers. </td> ',
' <td> Other </td> ',
' <td> Data contains recordings of five people performing different activities.
Each person wore four sensors (tags) while performing the same scenario five times.   </td
> ',
' <td> Life </td> ',
' <td> AutoUniv is an advanced data generator for classifications tasks. The aim
is to reflect the nuances and heterogeneity of real data. Data can be generated in .csv, ARFF or
C4.5 formats. </td> ',
' <td> Other </td> ',
' <td> A dataset of steel plates’ faults, classified into 7 different types. \r\

nThe goal was to train machine learning for automatic pattern recognition.\r\n </td> ',
' <td> Physical </td> ',
' <td> This dataset is taken from the MiniBooNE experiment and is used to distinguish electron neutrinos (signal) from muon neutrinos (background). </td> ',
' <td> Physical </td> ',
' <td> Prediction of the release year of a song from audio features. Songs are mostly western, commercial tracks ranging from 1922 to 2011, with a peak in the year 2000s. </td> ',
' <td> Other </td> ',
' <td> 15 months worth of daily data (440 daily records) that describes the occupancy rate, between 0 and 1, of different car lanes of the San Francisco bay area freeways across time. </td> ',
' <td> Computer </td> ',
' <td> This data set contains user reviews of cars and and hotels collected from Tripadvisor (~259,000 \r\nreviews) and Edmunds (~42,230 reviews).   </td> ',
' <td> Computer </td> ',
' <td> The dataset consists of 384 features extracted from CT images. The class variable is numeric and denotes the relative location of the CT slice on the axial axis of the human body. </td> ',
' <td> Computer </td> ',
' <td> This is a dataset of 8235 online handwritten assamese characters. The “online” process involves capturing of data as text is written on a digitizing tablet with an electronic pen. </td> ',
' <td> Computer </td> ',
' <td> These highly imbalanced bioassay datasets are from the differing types of screening that can be performed using HTS technology. 21 datasets were created from 12 bioassays. </td> ',
' <td> Life </td> ',
' <td> Element-wise comparison of records with personal data from a record linkage setting. The task is to decide from a comparison pattern whether the underlying records belong to one person. </td> ',
' <td> Other </td> ',
' <td> Communities in the US. Data combines socio-economic data from the \'90 Census, law enforcement data from the 1990 Law Enforcement Management and Admin Stats survey, and crime data from the 1995 FBI UCR </td> ',
' <td> Social </td> ',
' <td> Data set containing values for six biomechanical features used to classify orthopaedic patients into 3 classes (normal, disk hernia or spondilolysthesis) or 2 classes (normal or abnormal). </td> ',
' <td>   </td> ',
' <td> The Physical Action Data Set includes 10 normal and 10 aggressive physical actions that measure the human activity. The data have been collected by 4 subjects using the Delsys EMG wireless apparatus. </td> ',
' <td> Physical </td> ',
' <td> The Physical Action Data Set includes 10 normal and 10 aggressive physical actions that measure the human activity. The data have been collected by 10 subjects using the Vicon 3D tracker. </td> ',
' <td> Physical </td> ',
' <td> The dataset is used for authorship identification in online Writeprint which is a new research field of pattern recognition.   </td> ',
' <td> Physical </td> ',
' <td> Amazon\'s InfoSec is getting smarter about the way Access data is leveraged. This is an anonymized sample of access provisioned within the company. </td> ',
' <td> Business </td> ',
' <td> The dataset is used for authorship identification in online Writeprint which is a new research field of pattern recognition.   </td> ',
' <td> Computer </td> ',
' <td> This data was collected from text ads found on twelve websites that deal with various farm animal related topics. The binary labels are based on whether or not the content owner approves of the ad. </td> ',
' <td> Business </td> ',
' <td> It contains 64 e-mails which I have manually collected from DBWorld mailing list. They are classified in: \'announces of conferences\' and \'everything else\'. </td> ',
' <td> Computer </td> ',
' <td> KEGG Metabolic pathways modeled as directed relation network. Variety of graphical features presented. </td> ',
' <td> Life </td> ',
' <td> KEGG Metabolic pathways modeled as un-directed reaction network. Variety of graphical features presented. </td> ',
' <td> Life </td> ',
' <td> The data is related with direct marketing campaigns (phone calls) of a Portuguese banking institution. The classification goal is to predict if the client will subscribe a term deposit (variable y). </td> ',
' <td> Business </td> ',
' <td> This dataset provides user vote data on which video from a pair of videos

is funnier collected on YouTube Comedy Slam. The task is to automatically predict this preference based on video metadata. </td> ',
' <td> Computer </td> ',
' <td> This archive contains 13910 measurements from 16 chemical sensors utilized in simulations for drift compensation in a discrimination task of 6 gases at various levels of concentrations. </td> ',
' <td> Computer </td> ',
' <td> This data set contains 10 variables that are age, gender, total Bilirubin, direct Bilirubin, total proteins, albumin, A/G ratio, SGPT, SGOT and Alkphos. </td> ',
' <td> Life </td> ',
' <td> The OPPORTUNITY Dataset for Human Activity Recognition from Wearable, Object, and Ambient Sensors is a dataset devised to benchmark human activity recognition algorithms (classification, automatic data segmentation, sensor fusion, feature extraction, etc). </td> ',
' <td> Computer </td> ',
' <td> Nomao collects data about places (name, phone, localization...) from many sources.\r\nDeduplication consists in detecting what data refer to the same place.\r\nInstances in the dataset compare 2 spots. </td> ',
' <td> Computer </td> ',
' <td> The SMS Spam Collection is a public set of SMS labeled messages that have been collected for mobile phone spam research. </td> ',
' <td> Computer </td> ',
' <td> The Skin Segmentation dataset is constructed over B, G, R color space. Skin and Nonskin dataset is generated using skin textures from face images of diversity of age, gender, and race people. </td> ',
' <td> Computer </td> ',
' <td> The dataset concerns with the classification of two mental stages from recorded EEG signals: Planning (during imagination of motor act) and Relax state.   </td> ',
' <td> Computer </td> ',
' <td> The PAMAP2 Physical Activity Monitoring dataset contains data of 18 different physical activities, performed by 9 subjects wearing 3 inertial measurement units and a heart rate monitor. </td> ',
' <td> Computer </td> ',
' <td> The dataset was obtained from a recommender system prototype. The task was to generate a top-n list of restaurants according to the consumer preferences.   </td> ',
' <td> Computer </td> ',
' <td> This is a data set containing 1080 documents of free text business descriptions of Brazilian companies categorized into a\r\nsubset of 9 categories </td> ',
' <td> Business </td> ',
' <td> Measurements of electric power consumption in one household with a one-minute sampling rate over a period of almost 4 years. Different electrical quantities and some sub-metering values are available. </td> ',
' <td> Physical </td> ',
' <td> Measurements of geometrical properties of kernels belonging to three different varieties of wheat. A soft X-ray technique and GRAINS package were used to construct all seven, real-valued attributes. </td> ',
' <td> Life </td> ',
' <td> Northix is designed to be a schema matching benchmark problem for data integration of two entity relationship databases.   </td> ',
' <td> Computer </td> ',
' <td> Since there is no numerical sequential data stream available in standard data sets, this data set is generated from the original T40I10D100K data set </td> ',
' <td>   </td> ',
' <td> A textual corpus of 4000 legal cases for automatic summarization and citation analysis. For each document we collect catchphrases, citations sentences, citation catchphrases and citation classes. </td> ',
' <td> Other </td> ',
' <td> Human Activity Recognition database built from the recordings of 30 subjects performing activities of daily living (ADL) while carrying a waist-mounted smartphone with embedded inertial sensors. </td> ',
' <td> Computer </td> ',
' <td> Sixteen samples of leaf each of one-hundred plant species. For each sample, a shape descriptor, fine scale margin and texture histogram are given. </td> ',
' <td> Life </td> ',
' <td> This study looked into assessing the heating load and cooling load requirements of buildings (that is, energy efficiency) as a function of building parameters. </td> ',
' <td> Computer </td> ',
' <td> Delft data set, used to predict the hydodynamic performance of sailing yachts from dimensions and velocity. </td> ',
' <td> Physical </td> ',
' <td> >100 volunteers provide a semen sample analyzed according to the WHO 2010 criteria. Sperm concentration are related to socio-demographic data, environmental factors, health status, and life habits </td> ',
' <td> Life </td> ',

' <td> This dataset contains the annotated readings of 3 acceleration sensors at the hip and leg of Parkinson\'s disease patients that experience freezing of gait (FoG) during walking tasks.\r\n </td> ',
' <td> Life </td> ',
' <td> 3D road network with highly accurate elevation information (+-20cm) from Denmark used in eco-routing and fuel/Co2-estimation routing algorithms. </td> ',
' <td> Computer </td> ',
' <td> Data sets includes returns of Istanbul Stock Exchange with seven other international index; SP, DAX, FTSE, NIKKEI, BOVESPA, MSCE_EU, MSCI_EM from Jun 5, 2009 to Feb 22, 2011. </td> ',
' <td> Business </td> ',
' <td> This data-set contains examples of buzz events from two different social networks: Twitter, and Tom\'s Hardware, a forum network focusing on new technology with more conservative dynamics. </td> ',
' <td> Computer </td> ',
' <td> Given a theorem, predict which of five heuristics will give the fastest proof when used by a first-order prover. A sixth prediction declines to attempt a proof, should the theorem be too difficult. </td> ',
' <td> Computer </td> ',
' <td> A dataset with 5 classes (sitting-down, standing-up, standing, walking, and sitting) collected on 8 hours of activities of 4 healthy subjects. We also established a baseline performance index. </td> ',
' <td> Computer </td> ',
' <td> The dataset contains 18000 time-series recordings from a chemical detection platform at six different locations in a wind tunnel facility in response to ten high-priority chemical gaseous substances </td> ',
' <td> Computer </td> ',
' <td> Given Latin hypercube samples of 18 climate model input parameter values, predict climate model simulation crashes and determine the parameter value combinations that cause the failures. </td> ',
' <td> Physical </td> ',
' <td> A dataset to explore machine learning approaches for the identification of microorganisms from mass-spectrometry data. </td> ',
' <td> Life </td> ',
' <td> Data set containing values for 41 attributes (molecular descriptors) used to classify 1055 chemicals into 2 classes (ready and not ready biodegradable). </td> ',
' <td> Other </td> ',
' <td> In this paper, we look for to recognize the causes of users tend\r\nto cyber space in Kohkiloye and Boyer Ahmad Province in\r\nIran </td> ',
' <td> Computer </td> ',
' <td> The dataset comprises motion sensor data of 19 daily and sports activities each performed by 8 subjects in their own style for 5 minutes. Five Xsens MTx units are used on the torso, arms, and legs.\r\n </td> ',
' <td> Computer </td> ',
' <td> It is the real dataset about the students\' knowledge status about the subject of Electrical DC Machines. The dataset had been obtained from Ph.D. Thesis. </td> ',
' <td> Computer </td> ',
' <td> This test collection contains feature characteristics of documents originally written in five different languages and their translations, over a common set of 6 categories.   </td> ',
' <td> Life </td> ',
' <td> NYSK (New York v. Strauss-Kahn) is a collection of English news articles about the case relating to allegations of sexual assault against the former IMF director Dominique Strauss-Kahn (May 2011). </td> ',
' <td> Social </td> ',
' <td> This data set contains a total 5820 evaluation scores provided by students from Gazi University in Ankara (Turkey). There is a total of 28 course specific questions and additional 5 attributes. </td> ',
' <td> Other </td> ',
' <td> The dataset is about the users\' learning activities and knowledge levels on subjects of DC Electrical Machines. The dataset had been obtained from online web-courses and reported in my Ph.D. Thesis. </td> ',
' <td> Computer </td> ',
' <td> The data set consists of 14 EEG values and a value indicating the eye state. </td> ',
' <td> Life </td> ',
' <td> This is a data set of Physicochemical Properties of Protein Tertiary Structure. The data set is taken from CASP 5-9. There are 45730 decoys and size varying from 0 to 21 armstrong. </td> ',
' <td> Life </td> ',
' <td> The data describe the problem of high energy (higher than 10^4 J) seismic bumps forecasting in a coal \r\nmine. Data come from two of longwalls located in a Polish coal mine. </td> ',
' <td> Other </td> ',
' <td> Data were extracted from images that were taken for the evaluation of an authentication procedure for banknotes. </td> ',

' <td> Computer </td> ',
' <td> Data used for USPTO Algorithm Competition. Contains drawing pages from US patents with manually labeled figure and part labels. </td> ',
' <td> Other </td> ',
' <td> This dataset contains about 120k instances, each described by 13 feature types, with class information, specially useful for exploring multiview topics (cotraining, ensembles, clustering,..). </td> ',
' <td> Computer </td> ',
' <td> This archive contains 13910 measurements from 16 chemical sensors exposed to 6 different gases at various concentration levels. </td> ',
' <td> Computer </td> ',
' <td> This dataset comprises information regarding the ADLs performed by two us ers on a daily basis in their \r\nown homes.   </td> ',
' <td> Computer </td> ',
' <td> This data was used in Thompson et al. (2013). A list of possible game act ions is discussed in Thompson, Blair, Chen, & Henrey (2013). </td> ',
' <td> Game </td> ',
' <td> Six young health subjects were asked to perform 5 variations of the biceps curl weight lifting exercise. One of the variations is the one predicted by the health prof essional. </td> ',
' <td> Physical </td> ',
' <td> This dataset is collected from a monitor system mounted in a domotic hous e. It corresponds to approximately 40 days of monitoring data. </td> ',
' <td> Computer </td> ',
' <td> This dataset contains the hourly and daily count of rental bikes between years 2011 and 2012 in Capital bikeshare system with the corresponding weather and seasonal inform ation. </td> ',
' <td> Social </td> ',
' <td> The data from Twitter was collected during 360 consecutive days. It was d one by querying 1497 English keywords sampled from Wikipedia. This dataset is proposed in a Learning to rank setting. </td> ',
' <td> Computer </td> ',
' <td> The data is dedicated to classification problem related to the post-operative life expectancy in the lung cancer patients: class 1 - death within one year after surge ry, class 2 - survival. </td> ',
' <td> Life </td> ',
' <td> 3 different exercises: sitting, standing and walking in the muscles: bice ps femoris, vastus medialis, rectus femoris and semitendinosus addition to goniometry in the exerc ises. </td> ',
' <td> Computer </td> ',
' <td> This is a classification problem to distinguish between a signal process which produces supersymmetric particles and a background process which does not. </td> ',
' <td> Physical </td> ',
' <td> This is a classification problem to distinguish between a signal process which produces Higgs bosons and a background process which does not.   </td> ',
' <td> Physical </td> ',
' <td> Predict the Bankruptcy from Qualitative parameters from experts. </td> ',
' <td> Computer </td> ',
' <td> 126 samples from 14 participants, 309 features. Aim: assess whether voice rehabilitation treatment lead to phonations considered \'acceptable\' or \'unacceptable\' (binary class classification problem). </td> ',
' <td> Life </td> ',
' <td> Recordings of 16 volunteers performing 14 Activities of Daily Living (ADL) while carrying a single wrist-worn tri-axial accelerometer. </td> ',
' <td> Computer </td> ',
' <td> High-resolution Remote Sensing data set (Quickbird). Small number of trai ning samples of diseased trees, large number for other land cover. Testing data set from stratified random sample of image. </td> ',
' <td> Life </td> ',
' <td> The dataset collects data from an Android smartphone positioned in the ch est pocket from 22 participants walking in the wild over a predefined path. \r\n </td> ',
' <td> Other </td> ',
' <td> The dataset collects data from a wearable accelerometer mounted on the ch est. The dataset is intended for Activity Recognition research purposes. </td> ',
' <td> Other </td> ',
' <td> This dataset consists in a collection of shape and texture features extra cted from digital images of leaf specimens originating from a total of 40 different plant species. </td> ',
' <td> Computer </td> ',
' <td> This dataset contain Attributes of dresses and their recommendations according to their sales.Sales are monitor on the basis of alternate days.   </td> ',
' <td> Computer </td> ',
' <td> This data can be effectively produced the result to fewer parameter of th e Load profile can be reduced in the Database   </td> ',
' <td> Life </td> ',
' <td> NASA data set, obtained from a series of aerodynamic and acoustic tests o

f two and three-dimensional airfoil blade sections conducted in an anechoic wind tunnel.
</td> ',
' <td> Physical </td> ',
' <td> The data set refers to clients of a wholesale distributor. It includes th
e annual spending in monetary units (m.u.) on diverse product categories </td> ',
' <td> Business </td> ',
' <td> This problem of Sentiment Analysis (SA) has been studied well on the Engl
ish language but not Arabic one. Two main approaches have been devised: corpus-based and lexicon-b
ased.   </td> ',
' <td> Social </td> ',
' <td> The dataset contains 9568 data points collected from a Combined Cycle Pow
er Plant over 6 years (2006-2011), when the plant was set to work with full load.   </td>
',
' <td> Computer </td> ',
' <td> Classification of urban land cover using high resolution aerial imagery.
Intended to assist sustainable urban planning efforts. </td> ',
' <td> Physical </td> ',
' <td> This data has been prepared to analyze factors related to readmission as
well as other \r\n\r\noutcomes pertaining to patients with diabetes. </td> ',
' <td> Life </td> ',
' <td> The data set is composed of 60 chorales (5665 events) by J.S. Bach (1675-
1750).\r\nEach event of each chorale is labelled using 1 among 101 chord labels and
described\r\nthrough 14 features. </td> ',
' <td> Other </td> ',
' <td> Stone flakes are waste products of the stone tool production in\r\nthe pr
ehistoric era. The variables are means of geometric and\r\nstylistic features of the flakes contai
ned in different inventories. </td> ',
' <td> Other </td> ',
' <td> This is a collection of 8 files containing the match statistics for both
women and men at the four major tennis tournaments of the year 2013. Each file has 42 columns and
a minimum of 76 rows. </td> ',
' <td> Other </td> ',
' <td> The training data belongs to 20 Parkinson\'s Disease (PD) patients and 20
healthy subjects. From all subjects, multiple types of sound recordings (26) are taken. <
/td> ',
' <td> Life </td> ',
' <td> The dataset is composed by features extracted from 7 videos with people g
esticulating, aiming at studying Gesture Phase Segmentation. It contains 50 attributes divided
into two files for each video. </td> ',
' <td> Other </td> ',
' <td> This data consists of odors of 20 different perfumes. Data was obtained b
y using a handheld odor meter (OMX-GR sensor) per second for 28 seconds period. </td> ',
' <td> Computer </td> ',
' <td> Instances in this dataset contain features extracted from blog posts. The
task associated with the data is to predict how many comments the post will receive. </td
> ',
' <td> Social </td> ',
' <td> The REALDISP dataset is devised to evaluate techniques dealing with the e
ffects of sensor displacement in wearable activity recognition as well as to benchmark general act
ivity recognition algorithms   </td> ',
' <td> Computer </td> ',
' <td> Dataset is well suited for segmentation tasks. It contains 101 scanned pa
ges from different newspapers and magazines in Russian with ground truth pixel-based masks. <
/p></td> ',
' <td> Computer </td> ',
' <td> This data set compromises the metadata for the 2014 AAAI conference\'s ac
cepted papers, including paper titles, authors, abstracts, and keywords of varying
granularity. </td> ',
' <td> Computer </td> ',
' <td> The data set contains 58 time series acquired from 16 chemical sensors un
der gas flow modulation conditions. The sensors were exposed to different gaseous binary mixtures
of acetone and ethanol. </td> ',
' <td> Computer </td> ',
' <td> A chemical detection platform composed of 8 chemoresistive gas sensors wa
s exposed to turbulent gas mixtures generated naturally in a wind tunnel. The acquired time series
of the sensors are provided. </td> ',
' <td> Computer </td> ',
' <td> The UJIIndoorLoc is a Multi-Building Multi-Floor indoor localization data
base to test Indoor Positioning System that rely on WLAN/WiFi fingerprint. </td> ',
' <td> Computer </td> ',
' <td> Contains sentences from the abstract and introduction of 30 articles anno
tated with a modified Argumentative Zones annotation scheme. These articles come from biology, mac
hine learning and psychology. </td> ',
' <td> Other </td> ',
' <td> This dataset contains weekly data for the Dow Jones Industrial Index. It
has been used in computational investing research. </td> ',
' <td> Business </td> ',

' <td> The “sEMG for Basic Hand movements” includes 2 databases of surface electromyographic signals of 6 hand movements using Delsys\' EMG System. Healthy subjects conducted six daily life grasps. </td> ',
' <td> Life </td> ',
' <td> This data set compromises the metadata for the 2013 AAAI conference\'s accepted papers (main track only), including paper titles, abstracts, and keywords of varying granularity. </td> ',
' <td> Computer </td> ',
' <td> Instances in this dataset contain audio features extracted from 1059 wave files. The task associated with the data is to predict the geographical origin of music.\r\n </td> ',
' <td> Other </td> ',
' <td> Data have been generated from a sophisticated simulator of a Gas Turbines (GT), mounted on a Frigate characterized by a COmbined Diesel eLectric And Gas (CODLAG) propulsion plant type. </td> ',
' <td> Computer </td> ',
' <td> This dataset supports the development of models that make possible to interpret Grammatical Facial Expressions from Brazilian Sign Language (Libras). </td> ',
' <td> Computer </td> ',
' <td> Corpus intended to do cleaning (or binarization) and enhancement of noisy grayscale printed text images using supervised learning methods. Noisy images and their corresponding ground truth provided. </td> ',
' <td> Computer </td> ',
' <td> The MHEALTH (Mobile Health) dataset is devised to benchmark techniques dealing with human behavior analysis based on multimodal body sensing. </td> ',
' <td> Computer </td> ',
' <td> Predict student performance in secondary education (high school).   </td> ',
' <td> Social </td> ',
' <td> This data set contains electricity consumption of 370 points/clients.\r\n </td> ',
' <td> Computer </td> ',
' <td> The data set contains the recordings of 16 chemical sensors exposed to two dynamic gas mixtures at varying concentrations. For each mixture, signals were acquired continuously during 12 hours. </td> ',
' <td> Computer </td> ',
' <td> MicroblogPCU data is crawled from sina weibo microblog[http://weibo.com/]. This data can be used to study machine learning methods as well as do some social network research.   </td> ',
' <td> Computer </td> ',
' <td> The data are binary attack-point vectors and their clave-direction class(es) according to the partido-alto-based paradigm. </td> ',
' <td> Other </td> ',
' <td> Features are extracted from motor current. The motor has intact and defective components. This results in 11 different classes with different conditions.   </td> ',
' <td> Computer </td> ',
' <td> TV Commercials data set consists of standard audio-visual features of video shots extracted from 150 hours of TV news broadcast of 3 Indian and 2 international news channels (30 Hours each).   </td> ',
' <td> Computer </td> ',
' <td> This dataset collected mainly from: PhishTank archive, MillerSmiles archive, Google’s searching operators. </td> ',
' <td> Computer Security </td> ',
' <td> Design an observing network to monitor emissions of a greenhouse gas (GHG) in California given time series of synthetic observations and tracers from weather model simulations.\r\n </td> ',
' <td> Physical </td> ',
' <td> This dataset contains features extracted from the Messidor image set to predict whether an image contains signs of diabetic retinopathy or not.   </td> ',
' <td> Life </td> ',
' <td> The data contains lists of octamers (8 amino acids) and a flag (-1 or 1) depending on whether HIV-1 protease will cleave in the central position (between amino acids 4 and 5). </td> ',
' <td> Life </td> ',
' <td> The dataset contains sentences labelled with positive or negative sentiment. </td> ',
' <td> Other </td> ',
' <td> This dataset summarizes a heterogeneous set of features about articles published by Mashable in a period of two years. The goal is to predict the number of shares in social networks (popularity). </td> ',
' <td> Business </td> ',
' <td> Multi-temporal remote sensing data of a forested area in Japan. The goal is to map different forest types using spectral data. </td> ',
' <td> Life </td> ',
' <td> Survey of faculty members from two Spanish universities on teaching uses of Wikipedia </td> ',

' <td> Social </td> ',
' <td> The dataset contains a million randomly sampled video instances listing 1
0 fundamental video characteristics along with the YouTube video ID.   </td> ',
' <td> Computer </td> ',
' <td> This dataset can be used to predict the chronic kidney disease and it can
be collected from the hospital nearly 2 months of period. </td> ',
' <td> Other </td> ',
' <td> The data here are the ZZAlpha® machine learning recommendations made for
various US traded stock portfolios the morning of each day during the 3 year period Jan 1, 2012 -
Dec 31, 2014.   </td> ',
' <td> Business </td> ',
' <td> 20 photos of leaves for each of 32 different species. </td> ',
' <td> Other </td> ',
' <td> An accurate dataset describing trajectories performed by all the 442 taxi
s running in the city of Porto, in Portugal.\r\n </td> ',
' <td> Computer </td> ',
' <td> This Data set provides preprocessed and cleaned vital signals which can b
e used in designing algorithms for cuff-less estimation of the blood pressure. </td> ',
' <td> Life </td> ',
' <td> Activity recognition data set built from the recordings of 30 subjects pe
rforming basic activities and postural transitions while carrying a waist-mounted smartphone with
embedded inertial sensors.\r\n </td> ',
' <td> Life </td> ',
' <td> Expression levels of 77 proteins measured in the cerebral cortex of 8 cla
sses of control and Down syndrome mice exposed to context fear conditioning, a task used to assess
associative learning. </td> ',
' <td> Life </td> ',
' <td> The UJIIndoorLoc-Mag is an indoor localization database to test Indoor Po
sitioning System that rely on Earth\'s magnetic field variations. </td> ',
' <td> Computer </td> ',
' <td> The Heterogeneity Human Activity Recognition (HHAR) dataset from
Smartphones and Smartwatches is a dataset devised to benchmark human activity recognition
algorithms (classification, automatic data segmentation, sensor fusion, feature extraction, etc.)
in real-world contexts; specifically, the dataset is gathered with a variety of different device m
odels and use-scenarios, in order to reflect sensing heterogeneities to be expected in real
deployments. </td> ',
' <td> Computer </td> ',
' <td> Educational Process Mining data set is built from the recordings of 115 s
ubjects\' activities through a logging application while learning with an educational
simulator. </td> ',
' <td> Computer </td> ',
' <td> The search for exotic particles requires sorting through a large number o
f collisions to find the events of interest. This data set challenges one to detect a new particle
of unknown mass. </td> ',
' <td> Physical </td> ',
' <td> This dataset contains temporal data from a Wireless Sensor Network deploy
ed in real-world office environments. The task is intended as real-life benchmark in the area of A
mbient Assisted Living. </td> ',
' <td> Computer </td> ',
' <td> Open University Learning Analytics Dataset contains data about courses, s
tudents and their interactions with Virtual Learning Environment for seven selected courses and mo
re than 30000 students. </td> ',
' <td> Computer </td> ',
' <td> This research aimed at the case of customers' default payments in Taiwan
and compares the predictive accuracy of probability of default among six data mining
methods. </td> ',
' <td> Business </td> ',
' <td> Mesothelioma's disease data set were prepared at Dicle University Faculty
of Medicine in Turkey.\r\nThree hundred and twenty-four Mesothelioma patient data. In the dataset,
all samples have 34 features. </td> ',
' <td> Computer </td> ',
' <td> This is a transnational data set which contains all the transactions occu
rring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store online
retail. </td> ',
' <td> Business </td> ',
' <td> In SIFT10M, each data point is a SIFT feature which is extracted from Cal
tech-256 by the open source VLFeat library. The corresponding patches of the SIFT features are pro
vided. </td> ',
' <td> Computer </td> ',
' <td> The dataset has been feed by Android app called Go!Track. It is available
at Goolge Play Store(https://play.google.com/store/apps/details?id=com.go.router).   </td>
',
' <td> Computer </td> ',
' <td> I extract features from malacious and non-malacious and create and traini
ng dataset to teach svm classifier.Dataset made of unknown executable to detect if it is virus or
normal safe executable. </td> ',
' <td> Computer </td> ',

' <td> Computer </td> ',
' <td> Experimental data used for binary classification (room occupancy) from Temperature,Humidity,Light and CO2. Ground-truth occupancy was obtained from time stamped pictures that were taken every minute. </td> ',
' <td> Computer </td> ',
' <td> Handwriting database consists of 25 PWP(People with Parkinson) and 15 healthy individuals.Three types of recordings (Static Spiral Test, Dynamic Spiral Test and Stability Test) are taken. </td> ',
' <td> Computer </td> ',
' <td> References to news pages collected from an web aggregator in the period from 10-March-2014 to 10-August-2014. The resources are grouped into clusters that represent pages discussing the same story. </td> ',
' <td> Other </td> ',
' <td> Contains the responses of a gas multisensor device deployed on the field in an Italian city. Hourly responses averages are recorded along with gas concentrations references from a certified analyzer.   </td> ',
' <td> Computer </td> ',
' <td> >5 replicates of an 8-MOX gas sensor array were exposed to different gas conditions (4 volatiles at 10 concentration levels each). </td> ',
' <td> Computer </td> ',
' <td> >100 recordings of a sensor array under different conditions in a home setting: background, wine and banana presentations. The array includes 8 MOX gas sensors, and humidity and temperature sensors.\r\n </td> ',
' <td> Computer </td> ',
' <td> Instances in this dataset contain features extracted from facebook posts. The task associated with the data is to predict how many comments the post will receive. </td> ',
' <td> Other </td> ',
' <td> This data is an addition to an existing dataset on UCI. We collected more data to improve the accuracy of our human activity recognition algorithms applied in the domain of Ambient Assisted Living.   </td> ',
' <td> Computer </td> ',
' <td> The dataset is about bankruptcy prediction of Polish companies.The bankrupt companies were analyzed in the period 2000-2012, while the still operating companies were evaluated from 2007 to 2013. </td> ',
' <td> Business </td> ',
' <td> This dataset contains temporal data from a Wireless Sensor Network worn by an actor performing the activities: bending, cycling, lying down, sitting, standing, walking. </td> ',
' <td> Computer </td> ',
' <td> Dota 2 is a popular computer game with two teams of 5 players. At the start of the game each player chooses a unique hero with different strengths and weaknesses. </td> ',
' <td> Game </td> ',
' <td> Facebook performance metrics of a renowned cosmetic\'s brand Facebook page. </td> ',
' <td> Business </td> ',
' <td> UbiqLog is the smartphone lifelogging tool that runs on the smartphone of 35 users for about 2 months. \r\n </td> ',
' <td> Computer </td> ',
' <td> This data set contains the distribution of words in the full text of the NIPS conference papers published from 1987 to 2015. </td> ',
' <td> Computer </td> ',
' <td> Pulsar candidates collected during the HTRU survey. Pulsars are a type of star, of considerable scientific interest. Candidates must be classified in to pulsar and non-pulsar classes to aid discovery. </td> ',
' <td> Physical </td> ',
' <td> Classify type of drug consumer by personality data </td> ',
' <td> Social </td> ',
' <td> Experimental data used to create regression models of appliances energy use in a low energy building. </td> ',
' <td> Computer </td> ',
' <td> The dataset was created for the comparison and evaluation of hybrid indoor positioning methods. The dataset presented contains data from W-LAN and Bluetooth interfaces, and Magnetometer.   </td> ',
' <td> Computer </td> ',
' <td> KDC-4007 dataset Collection is the Kurdish Documents Classification text used in categories regarding Kurdish Sorani news and articles. </td> ',
' <td> Computer </td> ',
' <td> >A multisource and multivariate dataset for indoor localisation methods based on WLAN and Geo-Magnetic field fingerprinting </td> ',
' <td> Computer </td> ',
' <td> The DrivFace contains images sequences of subjects while driving in real scenarios. It is composed of 606 samples of 640×480, acquired over different days from 4 drivers with several facial features. </td> ',
' <td> Computer </td> ',
' <td> >\r\n\r\n </td> ',
' <td> Computer </td> ',

' <td> Computer </td> ',
' <td> It is a public set of comments collected for spam research. It has five d
atasets composed by 1,956 real messages extracted from five videos that were among the 10 most vie
wed on the collection period. </td> ',
' <td> Computer </td> ',
' <td> This hourly data set contains the PM2.5 data of US Embassy in Beijing. Me
anwhile, meteorological data from Beijing Capital International Airport are also included. <
/p></td> ',
' <td> Physical </td> ',
' <td> Sanitized and anonymized Cargo 2000 (C2K) airfreight tracking and tracing
events, covering five months of business execution (3,942 process instances, 7,932 transport legs,
56,082 activities).   </td> ',
' <td> Business </td> ',
' <td> This dataset focuses on the prediction of indicators/diagnosis of
cervical cancer. The features cover demographic information, habits, and historic medical
records. </td> ',
' <td> Life </td> ',
' <td> This dataset explores the subjective quality assessment of digital
colposcopies. </td> ',
' <td> Life </td> ',
' <td> KASANDR is a novel, publicly available collection for recommendation syst
ems that records the behavior of customers of the European leader in e-Commerce advertising, Kelko
o.   </td> ',
' <td> Life </td> ',
' <td> FMA features 106,574 tracks and includes song title, album, artist, genre
s; play counts, favorites, comments; description, biography, tags; together with audio (343 days,
917 GiB) and features. </td> ',
' <td> Computer </td> ',
' <td> Contains the responses of a gas multisensor device deployed on the field
in an Italian city.   </td> ',
' <td> Other </td> ',
' <td> This dataset is a pre-processed and re-structured/reshaped version of a v
ery commonly used dataset featuring epileptic seizure detection.   </td> ',
' <td> Life </td> ',
' <td> This is an image database of Handwritten Devanagari characters. There are
46 classes of characters with 2000 examples each. The dataset is split into training set(85%) and
testing set(15%).   </td> ',
' <td> Computer </td> ',
' <td> The data set of performances of weighted scoring stock portfolios are obt
ained with mixture design from the US stock market historical database. </td> ',
' <td> Business </td> ',
' <td> 5 types of hand postures from 12 users were recorded using unlabeled mark
ers attached to fingers of a glove in a motion capture environment. Due to resolution and
occlusion, missing values are common. </td> ',
' <td> Computer </td> ',
' <td> Predict a pattern of neurodegeneration in the dataset of speech features
obtained from patients with early untreated Parkinson’s disease and patients at high risk
developing Parkinson’s disease. </td> ',
' <td> Life </td> ',
' <td> Data include over 100 Team Activity Measures and outcomes (ML classes) ob
tained from activities of 74 student teams during the creation of final class project in SW Eng.
classes at SFSU, Fulda, FAU </td> ',
' <td> Computer </td> ',
' <td> This hourly data set contains the PM2.5 data in Beijing, Shanghai, Guangz
hou, Chengdu and Shenyang. Meanwhile, meteorological data for each city are also included. <
/p></td> ',
' <td> Physical </td> ',
' <td> Handwriting database consists of 62 PWP(People with Parkinson) and 15 hea
lthy individuals. Three types of recordings (Static Spiral Test, Dynamic Spiral Test and Stability
Test) are taken. </td> ',
' <td> Computer </td> ',
' <td> Contains weekly purchased quantities of 800 over products over 52 weeks.
Normalised values are provided too. </td> ',
' <td>   </td> ',
' <td> This dataset includes quantitative and categorical features from online r
eviews from 21 hotels located in Las Vegas Strip, extracted from TripAdvisor
(http://www.tripadvisor.com). </td> ',
' <td> Business </td> ',
' <td> This dataset includes Online Textual Reviews from both online (e.g., Trip
Advisor) and offline (e.g., Guests\' book) sources from the Areias do Seixo Eco-Resort. <
/td> ',
' <td> Business </td> ',
' <td> This dataset contains keystroke dynamics data collected on a touch mobile
device (Nexus 7). The dataset contains 2856 records, 51 records per subject for 56 subjects.
; </td> ',
' <td> Computer </td> ',
' <td> Crowdsourced data from OpenStreetMap is used to automate the
classification of satellite images into different land cover classes (impervious, farm, forest

classification of satellite images into different land cover classes (impervious, farm, forest, grass, orchard, water).   </td> ',
' <td> Physical </td> ',
' <td> This collection of data is part of the RNA-Seq (HiSeq) PANCAN data set, it is a random extraction of gene expressions of patients having different types of tumor: BRCA, KIRC, COAD, LUAD and PRAD. </td> ',
' <td> Life </td> ',
' <td> The dataset was created for the comparison and evaluation of hybrid indoor positioning methods. The dataset presented contains data from W-LAN and Bluetooth interfaces, and Magnetometer.   </td> ',
' <td> Computer </td> ',
' <td> The research project presents this database, shows the images of chestnuts that will be processed to determine the presence or absence of defects </td> ',
' <td> Computer </td> ',
' <td> One of the primary challenges in identifying the risks of the Burst Header Packet (BHP) flood attacks in Optical Burst Switching networks (OBS) is the scarcity of reliable historical data.   </td> ',
' <td> Computer </td> ',
' <td> 5 types of hand postures from 12 users were recorded using unlabeled markers on fingers of a glove in a motion capture environment. Due to resolution and occlusion, missing values are common. </td> ',
' <td> Computer </td> ',
' <td> Acoustic features extracted from syllables of anuran (frogs) calls, including the family, the genus, and the species labels (multilabel).   </td> ',
' <td> Life </td> ',
' <td> The TTC-3600 data set is a collection of Turkish news and articles including categorized 3,600 documents from 6 well-known portals in Turkey. It has 4 different forms in ARFF Weka format. </td> ',
' <td> Computer </td> ',
' <td> This dataset contains features extracted from colonoscopy videos used to detect gastrointestinal lesions. It contains 76 lesions: 15 serrated adenomas, 21 hyperplastic lesions and 40 adenoma.   </td> ',
' <td> Computer </td> ',
' <td> The dataset was collected during 60 days, this is a real database of a brazilian logistics company. </td> ',
' <td> Business </td> ',
' <td> This sentiment analysis data set contains scientific paper reviews from an international conference on computing and informatics. The task is to predict the orientation or the evaluation of a review. </td> ',
' <td> Computer </td> ',
' <td> It was collected for CAD diagnosis. </td> ',
' <td> Life </td> ',
' <td> It was collected for CAD diagnosis. </td> ',
' <td> Life </td> ',
' <td> This dataset contains the dynamic features of 107,888 executables, collected by VirusShare from Nov/2010 to Jul/2014. </td> ',
' <td> Computer </td> ',
' <td> The dataset consists of data collected from heavy Scania trucks in everyday usage.   </td> ',
' <td> Computer </td> ',
' <td> This set Provides data regarding wireless communications between vehicles and road side units. two separate data sets are provided (normal scenario) and in the presence of attacker (jammer). </td> ',
' <td> Computer </td> ',
' <td> This dataset was collected by Shan-Hung Wu and DataLab members at NTHU, Taiwan. There\'re 325 user-perceived clusters from 100 users and their corresponding descriptions. </td> ',
' <td> Computer </td> ',
' <td> Character images from scanned and computer generated fonts. </td> ',
' <td> Computer </td> ',
' <td> This dataset includes 1) 12234 documents (8251 training, 3983 test) extracted from DeliciousT140 dataset, 2) class labels for all documents, 3) labels for a subset of sentences of the test documents. </td> ',
' <td> Computer </td> ',
' <td> Children screening data for autism suitable for classification and predictive tasks   </td> ',
' <td> Life </td> ',
' <td> Autistic Spectrum Disorder Screening Data for Adolescent. This dataset is related to classification and predictive tasks. </td> ',
' <td> Life </td> ',
' <td> The datasets\' positive class consists of component failures for a specific component of the APS system. The negative class consists of trucks with failures for components not related to the APS. </td> ',
' <td> Computer </td> ',
' <td> Collected in indoor space by observing signal strengths of seven WiFi signals visible on a smartphone. The decision variable is one of the four rooms.   </td> ',
' <td> Computer </td> '

```
  <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Hepatocellular Carcinoma dataset (HCC dataset) was collected at a Univers
ity Hospital in Portugal. It contains real clinical data of 165 patients diagnosed with HCC. 
</p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">12 features categorized as conventional and social media features. Both c
onventional features, collected from movies databases on Web as well as social media
features(YouTube,Twitter). </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Persian questions gathered from a jeopardy game broadcasted on Iranian na
tional television.  </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">Autistic Spectrum Disorder Screening Data for Adult. This dataset is rela
ted to classification and predictive tasks. </p></td> ',
 ' <td><p class="normal">Social </p></td> ',
 ' <td><p class="normal">Sequential motion data from 14 healthy older people aged 66 to 86 years o
ld using a batteryless, wearable sensor on top of their clothing for the recognition of activities
in clinical environments. </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">This dataset contains information about wart treatment results of 90 pati
ents using immunotherapy. </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">This dataset contains information about wart treatment results of 90 pati
ents using cryotherapy. </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">This dataset contains OCT data (in mat format) and color fundus data (in
jpg format) of left & right eyes of 50 healthy persons. </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Discrete Tone Images(DTI)are available which needs to be analyzed in
detail. Here, we created this dataset for those who do research in DTI. \r\n </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Large data set of news items and their respective social feedback on mult
iple platforms: Facebook, Google+ and LinkedIn. </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Fault diagnosis of four liquid ultrasonic flowmeters </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">This data set compromises the metadata for the 2014 ICMLA conference\'s a
ccepted papers, including ID, paper titles, author\'s keywords, abstracts and sessions in which th
ey were exposed. </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">This dataset contains RSSI readings gathered from an array of Bluetooth L
ow Energy (BLE) iBeacons in a real-world and operational indoor environment for localization and n
avigation purposes. </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">A container crane has the function of transporting containers from one po
int to another point. </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Data set includes construction cost, sale prices, project variables, and
economic variables corresponding to real estate single-family residential apartments in Tehran,
Iran.  </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">The data was collected in 2015 using Twitter API. This dataset contains h
ealth news from more than 15 major health news agencies such as BBC, CNN, and NYT.  </p></td>
',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">ChIP-seq experiments characterize protein modifications or binding
at\r\nspecific genomic locations in specific samples. The machine learning\r\nproblem in these dat
a is structured binary classification. </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Running times for multiplying two 2048 x 2048 matrices using a GPU OpenCL
SGEMM kernel with varying parameters (using the library \'CLTune\'). </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">The dataset contains 7 datasets of User - Item matrices, where each entry
represents how many times a user consumed an item. Item is used as an umbrella term for various ca
tegories. </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">This dataset addresses the lack of public botnet datasets, especially for
the IoT. It suggests *real* traffic data, gathered from 9 commercial IoT devices authentically inf
ected by Mirai and BASHLITE. </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">The database was created with records of absenteeism at work from July 20
07 to July 2010 at a courier company in Brazil. </p></td> ',
 ' <td><p class="normal">Business </p></td> ',
 ' <td><p class="normal">First self-care activities dataset based on ICF-CY. </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">The data set addresses the condition assessment of a hydraulic test rig b
ased on multi sensor data. Four fault types are superimposed with several severity grades impeding
selective quantification. </p></td> ',
```

selective quantification. </td> ',
' <td> Computer </td> ',
' <td> This dataset contains 10721 initial and calculated atomic coordinates of carbon nanotubes. </td> ',
' <td> Computer </td> ',
' <td> This dataset contains 640 performance measurements from a simulation of 2-Dimensional Multiprocessor Optical Interconnection Network.   </td> ',
' <td> Computer </td> ',
' <td> 1000 sports articles were labeled using Amazon Mechanical Turk as objective or subjective. The raw texts, extracted features, and the URLs from which the articles were retrieved are provided. </td> ',
' <td> Social </td> ',
' <td> Clinical features were observed or measured for 64 patients with breast cancer and 52 healthy controls.   </td> ',
' <td> Life </td> ',
' <td> The data-set contains four (4) sets of mobile sensor readings data (humidity, temperature) corresponding to a swarm of four (4) Unmanned Surface Vehicles (USVs) in a test-bed in Athens (Greece).   </td> ',
' <td> Computer </td> ',
' <td> The dataset was used to test an architecture based on a trust model capable to cope with the evaluation of the trustworthiness of users interacting in pervasive environments. </td> ',
' <td> Computer </td> ',
' <td> To create the largest authorship attribution dataset, we extracted works of 50 well-known authors. To have a non-exhaustive learning, in training there are 45 authors whereas, in the testing, it\'s 50 </td> ',
' <td> Computer </td> ',
' <td> 20 falls and 16 daily living activities were performed by 17 volunteers with 5 repetitions while wearing 6 sensors (3.060 instances) that attached to their head, chest, waist, wrist, thigh and ankle. </td> ',
' <td> Life </td> ',
' <td> 5879 captioned images (image and text) from social media related to damage during natural disasters/wars, and belong to 6 classes: Fires, Floods, Natural landscape, Infrastructural, Human, Non-damage. </td> ',
' <td> Social </td> ',
' <td> This database consists on 30 subjects performing Brain Computer Interface for Steady State Visual Evoked Potentials (BCI-SSVEP).   </td> ',
' <td> Life </td> ',
' <td> Roman Urdu (the scripting style for Urdu language) is one of the limited resource languages.A data corpus comprising of more than 20000 records was collected. </td> ',
' <td> Computer </td> ',
' <td> The Avila data set has been extracted from 800 images of the \'Avila Bible\', an XII century giant Latin copy of the Bible. The prediction task consists in associating each pattern to a copyist. </td> ',
' <td> Computer </td> ',
' <td> PANDOR is a novel and publicly available dataset for online recommendation provided by Purch (http://www.purch.com/).   </td> ',
' <td> Life </td> ',
' <td> The dataset provides patient reviews on specific drugs along with related conditions. Reviews and ratings are grouped into reports on the three aspects benefits, side effects and overall comment. </td> ',
' <td>   </td> ',
' <td> The dataset provides patient reviews on specific drugs along with related conditions and a 10 star patient rating reflecting overall patient satisfaction. </td> ',
' <td> Life </td> ',
' <td> The dataset is generated from Physical Unclonable Functions (PUFs) simulation, specifically XOR Arbiter PUFs. PUFs are used for authentication purposes. For more info, refer to our paper below. </td> ',
' <td> Computer </td> ',
' <td> Two file s contain data on 21263 superconductors and their relevant features. </td> ',
' <td> Physical </td> ',
' <td> WESAD (Wearable Stress and Affect Detection) contains data of 15 subjects during a stress-affect lab study, while wearing physiological and motion sensors. </td> ',
' <td> Computer </td> ',
' <td> The data-set contains eight (2x4) data-sets of mobile sensor readings data (humidity, temperature) corresponding to a swarm of four Unmanned Surface Vehicles (USVs) in a test-bed, Athens, Greece. </td> ',
' <td> Computer </td> ',
' <td> The dataset tried to find the end semester percentage prediction based on different social, economic and academic attributes.   </td> ',
' <td> Computer </td> ',
' <td> Of the 12,330 sessions in the dataset,\r\n84.5% (10,422) were negative class samples that did not\r\nend with shopping, and the rest (1908) were positive class\r\nsamples ending with shopping. </td> ',
' <td> Business </td> ',

```
 ' <td><p class="normal">Business </p></td> ',
 ' <td><p class="normal">The handwritten dataset was collected from 170 participants with a total
of 5,180 numeral patterns. The dataset is named Prince Mohammad Bin Fahd University - Urdu/Arabic
Database (PMU-UD).  </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">The data used in this study were gathered from 188 patients with PD (107
men and 81 women) with ages ranging from 33 to 87 (65.1±10.9). </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">The local stability analysis of the 4-node star system (electricity produ
cer is in the center) implementing Decentral Smart Grid Control concept.  </p></td> ',
 ' <td><p class="normal">Physical </p></td> ',
 ' <td><p class="normal">This dataset contains information about caesarian section results of  80
pregnant women with the most important characteristics of delivery problems in the medical
field. </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">BAUM-1 dataset contains 1184 multimodal facial video clips collected from
31 subjects. The 1184 video clips contain spontaneous facial expressions and speech of 13
emotional and mental states. </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">A multilingual audio-visual affective face database consisting of 1047 vi
deo clips of 286 subjects.  </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Exhaustive one year non-confidential data in the year 2015 to 2016 of fir
ms is collected from the Auditor Office of India to build a predictor for classifying suspicious f
irms. </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">User interest information extracted from user reviews published in holida
yiq.com about various types of point of interests in South India </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">The “real estate valuation” is a regression problem. The market
historical data set of real estate valuation are collected from Sindian Dist., New Taipei City, Ta
iwan.  </p></td> ',
 ' <td><p class="normal">Business </p></td> ',
 ' <td><p class="normal">. </p></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">A data extract of a non-federal dataset posted here
https://catalog.data.gov/dataset/somerville-happiness-survey-responses-2011-2013-2015 </p></t
d> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Measurement of the S21,consists of 10 sweeps, each sweep contains 601 fre
quency points with spacing of 0.167MHz to cover a 100MHz band centered at 2.4GHz. </p></td> '
,
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">These are files of raw EMG data recorded by MYO Thalmic bracelet </p
></td> ',
 ' <td><p class="normal">Life </p></td> ',
 ' <td><p class="normal">Data collected from car parks in Birmingham that are operated by NCP from
\r\nBirmingham City Council. UK Open Government Licence
(OGL).\r\nhttps://data.birmingham.gov.uk/dataset/birmingham-parking </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">The database was created with records of behavior of the urban traffic of
the city of Sao Paulo in Brazil. </p></td> ',
 ' <td><p class="normal">Computer </p></td> ',
 ' <td><p class="normal">Reviews on destinations in 10 categories mentioned across East Asia. Each
traveler rating is mapped as Excellent(4), Very Good(3), Average(2), Poor(1), and Terrible(0) and
average rating is used. </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">Google reviews on attractions from 24 categories across Europe are consid
ered. Google user rating ranges from 1 to 5 and average user rating per category is
calculated. </p></td> ',
 ' <td><p class="normal">Other </p></td> ',
 ' <td><p class="normal">There are three classes/diseases: Bacterial leaf blight, Brown spot, and
Leaf smut, each having 40 images. The format of all images is jpg.   </p></td> ',
 ' <td><p class="normal">Computer </p></td> ']
```

In [464]:

```
descripcion_base=comments[7:1000:2]
```

In [465]:

```
descripcion_base
```

Out[465]:

[' <td> Predict the age of abalone from physical measurements </td> ',
 ' <td> Predict whether income exceeds $50K/yr based on census data. Also known as "Census Income" dataset. </td> ',
 ' <td> Steel annealing data </td> ',
 ' <td> Log of anonymous users of www.microsoft.com; predict areas of the web site a user visited based on data on other areas the user visited. </td> ',
 ' <td> Distinguish between the presence and absence of cardiac arrhythmia and classify it in one of the 16 groups. </td> ',
 ' <td> Dataset artificially generated by using first order theory which describes structure of ten capital letters of English alphabet </td> ',
 ' <td> Nominal audiology dataset from Baylor </td> ',
 ' <td> Standardized version of the original audiology database </td> ',
 ' <td> Revised from CMU StatLib library, data concerns city-cycle fuel consumption </td> ',
 ' <td> From 1985 Ward\'s Automotive Yearbook </td> ',
 ' <td> Badges labeled with a "+" or "-" as a function of a person\'s name </td> ',
 ' <td> Balance scale weight & distance database </td> ',
 ' <td> Data previously used in cognitive psychology experiment; 4 data sets represent different conditions of an experiment </td> ',
 ' <td> Breast Cancer Data (Restricted Access) </td> ',
 ' <td> Original Wisconsin Breast Cancer Database </td> ',
 ' <td> Prognostic Wisconsin Breast Cancer Database </td> ',
 ' <td> Diagnostic Wisconsin Breast Cancer Database </td> ',
 ' <td> Bridges database that has original and numeric-discretized datasets </td> ',
 ' <td> Derived from simple hierarchical decision model, this database may be useful for testing constructive induction and structure discovery methods. </td> ',
 ' <td> Predict whether income exceeds $50K/yr based on census data. Also known as "Adult" dataset. </td> ',
 ' <td> Knight Pin Chess End-Game Database Creator </td> ',
 ' <td> King+Rook versus King+Pawn on a7 (usually abbreviated KRKPA7). </td> ',
 ' <td> Chess Endgame Database for White King and Rook against Black King (KRK). </td> ',
 ' <td> 6 different domain theories for generating legal moves of chess </td> ',
 ' <td> Time-series data based on chorales; challenge is to learn generative grammar; data in Lisp </td> ',
 ' <td> Contains connect-4 positions </td> ',
 ' <td> This data concerns credit card applications; good mix of attributes </td> ',
 ' <td> Includes domain theory (generated by talking to Japanese domain experts); data in Lisp </td> ',
 ' <td> Relative CPU Performance Data, described in terms of its cycle time, memory size, etc. </td> ',
 ' <td> Dataset is a subset of the 1987 National Indonesia Contraceptive Prevalence Survey. </td> ',
 ' <td> Forest CoverType dataset </td> ',
 ' <td> Used in decision tree induction for mitigating process delays known as "cylinder bands" in rotogravure printing </td> ',
 ' <td> Aim for this dataset is to determine the type of Eryhemato-Squamous Disease. </td> ',
 ' <td> This diabetes dataset is from AIM \'94 </td> ',
 ' <td> Generates application domains based on specific parameters, number of features, and proportion of positive to negative examples </td> ',
 ' <td> Five concepts, expressed as predicates, to be learned </td> ',
 ' <td> Assorted small-scale domain theories </td> ',
 ' <td> Data for classifying if patients will survive for at least one year after a heart attack </td> ',
 ' <td> This data contains protein localization sites </td> ',
 ' <td> From Collins Gem Guide to Flags, 1986 </td> ',
 ' <td> Cases collected mostly from investigations in physical science; intention is to evaluate function-finding algorithms </td> ',
 ' <td> From USA Forensic Science Service; 6 types of glass; defined in terms of their oxide content (i.e. Na, Fe, K, etc) </td> ',
 ' <td> Dataset contains cases from study conducted on the survival of patients who had undergone surgery for breast cancer </td> ',
 ' <td> Topic: human subjects study </td> ',
 ' <td> 4 databases: Cleveland, Hungary, Switzerland, and the VA Long Beach </td> ',
 ' <td> From G.Gong: CMU; Mostly Boolean or numeric-valued attribute types; Includes cost data (donated by Peter Turney) </td> ',
 ' <td> Well documented attributes; 368 instances with 28 attributes (continuous, discrete, and nominal); 30% missing values </td> ',
 ' <td> Data set prepared for the use of participants for the 1994 AAAI Spring Symposium on Artificial Intelligence in Medicine. </td> ',

' <td> Image data described by high-level numeric-valued attributes, 7 classes </td> ',
' <td> This dataset represents a set of possible advertisements on Internet page s. </td> ',
' <td> Classification of radar returns from the ionosphere </td> ',
' <td> Famous database; from Fisher, 1936 </td> ',
' <td> Goal: Predict which letter-name was spoken--a simple classification task. </td> ',
' <td> Relational dataset </td> ',
' <td> From Collective Bargaining Review </td> ',
' <td> From Classification and Regression Trees book; We provide here 2 C programs for generating sample databases </td> ',
' <td> Database for fitting contact lenses </td> ',
' <td> Database of character image features; try to identify the letter </td> ',
' <td> BUPA Medical Research Ltd. database donated by Richard S. Forsyth </td> ',
' <td> All code for Logic Theorist </td> ',
' <td> Lung cancer data; no attribute definitions </td> ',
' <td> This lymphography domain was obtained from the University Medical Centre, Institute of Oncology, Ljubljana, Yugoslavia. (Restricted access) </td> ',
' <td> Fault diagnosis problem of electromechanical devices; also PUMPS DATA SET is newer version with domain theory and results </td> ',
' <td> Meta-Data was used in order to give advice about which classification method is appropriate for a particular dataset (taken from results of Statlog project). </td> ',
' <td> Learning concepts from sensor data of a mobile robot; set of data sets </td> ',
' <td> E. Coli promoter gene sequences (DNA) with partial domain theory </td> ',
' <td> From CMU connectionist bench repository; Classifies secondary structure of certain globular proteins </td> ',
' <td> Primate splice-junction gene sequences (DNA) with associated imperfect domain theory </td> ',
' <td> A set of three artificial domains over the same attribute space; Used to test a wide range of induction algorithms </td> ',
' <td> Horn-clause model that qualitatively simulates moral reasoning; Theory includes negated literals </td> ',
' <td> This dataset consists of features of handwritten numerals (`0\'--`9\') extracted from a collection of Dutch utility maps </td> ',
' <td> From Audobon Society Field Guide; mushrooms described in terms of physical characteristics; classification: poisonous or edible </td> ',
' <td> The goal is to learn to predict whether new molecules will be musks or non-musks </td> ',
' <td> The goal is to learn to predict whether new molecules will be musks or non-musks </td> ',
' <td> Nursery Database was derived from a hierarchical decision model originally developed to rank applications for nursery schools. </td> ',
' <td> Used in research to generate features for an inductive learning system </td> ',
' <td> The problem consists of classifying all the blocks of the page layout of a document that has been detected by a segmentation process. </td> ',
' <td> Two versions of this database available; see folder </td> ',
' <td> Digit database of 250 samples from 44 writers </td> ',
' <td> Dataset of patient features </td> ',
' <td> From Ljubljana Oncology Institute </td> ',
' <td> Assorted domains like blocksworld, eightpuzzle, and schedworld. </td> ',
' <td> Two sets of datasets are given: pyrimidines and triazines </td> ',
' <td> The file animals.c is a data generator of structured instances representing quadruped animals </td> ',
' <td> Data was from a simulation of a servo system </td> ',
' <td> Tiny database; all nominal values </td> ',
' <td> Each class attribute counts the number of solar flares of a certain class that occur in a 24 hour period </td> ',
' <td> Michalski\'s famous soybean disease database </td> ',
' <td> Michalski\'s famous soybean disease database </td> ',
' <td> Task: predict the number of O-rings that experience thermal distress on a flight at 31 degrees F given data on the previous 23 shuttle flights </td> ',
' <td> From IRAS data -- NASA Ames Research Center </td> ',
' <td> Classifying Email as Spam or Non-Spam </td> ',
' <td> Data on cardiac Single Proton Emission Computed Tomography (SPECT) images. Each patient classified into two categories: normal and abnormal. </td> ',
' <td> Data on cardiac Single Proton Emission Computed Tomography (SPECT) images. Each patient classified into two categories: normal and abnormal. </td> ',
' <td> Data on sponges; Attributes in spanish </td> ',

' <td> Various Databases: Vehicle silhouttes, Landsat Sattelite, Shuttle, Australian Credit Approval, Heart Disease, Image Segmentation, German Credit </td> ',
' <td> Student Loan Relational Domain </td> ',
' <td> The data consist of evaluations of teaching performance; scores are "low", "medium", or "high" </td> ',
' <td> Binary classification task on possible configurations of tic-tac-toe game </td> ',
' <td> 10 separate databases from Garavan Institute </td> ',
' <td> 2 data formats (structured, one-instance-per-line) </td> ',
' <td> Data in original (LISP-readable) form </td> ',
' <td> 1984 United Stated Congressional Voting Records; Classify as Republican or Democrat </td> ',
' <td> Multiple classes predict plant state </td> ',
' <td> CART book\'s waveform domains </td> ',
' <td> CART book\'s waveform domains </td> ',
' <td> Using chemical analysis determine the origin of wines </td> ',
' <td> Predicting the Cellular Localization Sites of Proteins </td> ',
' <td> Artificial, 7 classes of animals </td> ',
' <td> Various datasets without documentation (feel free to explore!) </td> ',
' <td> This data set consists of 20000 messages taken from 20 newsgroups. </td> ',
' <td> This data consists of sample of Auslan (Australian Sign Language) signs. Examples of 95 signs were collected from five signers with a total of 6650 sign samples. </td> ',
' <td> This data consists of sample of Auslan (Australian Sign Language) signs. 27 examples of each of 95 Auslan signs were captured from a native signer using high-quality position trackers </td> ',
' <td> The USCensus1990raw data set contains a one percent sample of the Public Use Microdata Samples (PUMS) person records drawn from the full 1990 census sample. </td> ',
' <td> This data set contains weighted census data extracted from the 1994 and 1995 current population surveys conducted by the U.S. Census Bureau. </td> ',
' <td> This data set is from the 1999 Computational Intelligence and Learning (COIL) competition. The data contains measurements of river chemical concentrations and algae densities. </td> ',
' <td> This dataset contains image features extracted from a Corel image collection. Four sets of features are available based on the color histogram, color histogram layout, color moments, and co-occurence </td> ',
' <td> Data giving characteristics of each ORF (potential gene) in the E. coli genome. Sequence, homology (similarity to other genes) and structural information, and function (if known) are provided. </td> ',
' <td> This data arises from a large study to examine EEG correlates of genetic predisposition to alcoholism. It contains measurements from 64 electrodes placed on the scalp sampled at 256 Hz </td> ',
' <td> The data set contains oceanographic and surface meteorological readings taken from a series of buoys positioned throughout the equatorial Pacific. </td> ',
' <td> This data contains a record of user interactions with the Entree Chicago restaurant recommendation system. </td> ',
' <td> This data consists of 640 black and white face images of people taken with varying pose (straight, left, right, up), expression (neutral, happy, sad, angry), eyes (wearing sunglasses or not), and size </td> ',
' <td> This data set used in the CoIL 2000 Challenge contains information on customers of an insurance company. The data consists of 86 variables and includes product usage data and socio-demographic data </td> ',
' <td> This data contains general demographic information on internet users in 1997. </td> ',
' <td> This data set contains unweighted PUMS census data from the Los Angeles and Long Beach areas for the years 1970, 1980, and 1990. </td> ',
' <td> This dataset records 640 time series of 12 LPC cepstrum coefficients taken from nine male speakers. </td> ',
' <td> This is the data set used for The Second International Knowledge Discovery and Data Mining Tools Competition, which was held in conjunction with KDD-98 </td> ',
' <td> This is the data set used for The Third International Knowledge Discovery and Data Mining Tools Competition, which was held in conjunction with KDD-99 </td> ',
' <td> Data giving characteristics of each ORF (potential gene) in the M. tuberculosis bacterium. Sequence, homology (similarity to other genes) and structural information, and function (if known) are provided </td> ',
' <td> This data set contains a list of over 10000 films including many older, odd, and cult films. There is information on actors, casts, directors, producers, studios, etc. </td> ',
' <td> This data describes the page visits of users who visited msnbc.com on September 28, 1999. Visits are recorded at the level of URL category (see description) and are recorded in time order. </td> ',
' <td> This data set consists of (a) 129,000 abstracts describing NSF awards for basic research, (b) bag-of-word data files extracted from the abstracts, (c) a list of words used for indexing the bag-of-word </td> ',

' <td> This dataset contains time series sensor readings of the Pioneer-1 mobile robot. The data is broken into "experiences" in which the robot takes action for some period of time and experiences a control </td> ',
' <td> This data set is designed for testing indexing schemes in time series databases. The data appears highly periodic, but never exactly repeats itself. </td> ',
' <td> This is a collection of documents that appeared on Reuters newswire in 1987. The documents were assembled and indexed with categories. </td> ',
' <td> This dataset contains force and torque measurements on a robot after failure detection. Each failure is characterized by 15 force/torque samples collected at regular time intervals </td> ',
' <td> This data consists of synthetically generated control charts. </td> ',
' <td> This database contains HTML source of web pages plus the ratings of a single user on these web pages. Web pages are on four seperate subjects (Bands- recording artists; Goats; Sheep; and BioMedical) </td> ',
' <td> This file contains 9 sets of sanitized user data drawn from the command histories of 8 UNIX computer users at Purdue over the course of up to 2 years. </td> ',
' <td> The JARtool project was a pioneering effort to develop an automatic system for cataloging small volcanoes in the large set of Venus images returned by the Magellan spacecraft. </td> ',
' <td> This file concerns credit card applications. This database exists elsewhere in the repository (Credit Screening Database) in a slightly different form </td> ',
' <td> This dataset classifies people described by a set of attributes as good or bad credit risks. Comes in two formats (one all numeric). Also comes with a cost matrix </td> ',
' <td> This dataset is a heart disease database similar to a database already present in the repository (Heart Disease databases) but in a slightly different form </td> ',
' <td> Multi-spectral values of pixels in 3x3 neighbourhoods in a satellite image, and the classification associated with the central pixel in each neighbourhood </td> ',
' <td> This dataset is an image segmentation database similar to a database already present in the repository (Image segmentation database) but in a slightly different form. </td> ',
' <td> The shuttle dataset contains 9 attributes all of which are numerical. Approximately 80% of the data belongs to class 1 </td> ',
' <td> 3D objects within a 2D image by application of an ensemble of shape feature extractors to the 2D silhouettes of the objects. </td> ',
' <td> The file "nettalk.data" contains a list of 20,008 English words, along with a phonetic transcription for each word. The task is to train a network to produce the proper phonemes </td> ',
' <td> The task is to train a network to discriminate between sonar signals bounced off a metal cylinder and those bounced off a roughly cylindrical rock. </td> ',
' <td> Speaker independent recognition of the eleven steady state vowels of British English using a specified training set of lpc derived log area ratios. </td> ',
' <td> Domain Theory on Economic Sanctions; Undocumented </td> ',
' <td> Undocumented </td> ',
' <td> Little Documentation </td> ',
' <td> This data comes from the main door of the CalIt2 building at UCI. </td> ',
' <td> Loop sensor data was collected for the Glendale on ramp for the 101 North freeway in Los Angeles </td> ',
' <td> Purpose is to predict poker hands </td> ',
' <td> Data are MC generated to simulate registration of high energy gamma particles in an atmospheric Cherenkov telescope </td> ',
' <td> Data consists of written characters in a UNIPEN-like format </td> ',
' <td> Discrimination of benign and malignant mammographic masses based on BI-RADS attributes and the patient\'s age. </td> ',
' <td> This is a difficult regression task, where the aim is to predict the burned area of forest fires, in the northeast region of Portugal, by using meteorological and other data (see details at: http://www.dsi.uminho.pt/~pcortez/forestfires). </td> ',
' <td> This dataset is created by reading out 200 files from the 10 largest Reuters \r\nclasses and using an Automatic Speech Recognition system to create \r\ncorresponding transcriptions. </td> ',
' <td> This data set contains five text collections in the form of bags-of-words. </td> ',
' <td> Concrete is the most important material in civil engineering. The concrete compressive strength is a highly nonlinear function of age and ingredients.   </td> ',
' <td> Each record represents 100 points on a two-dimensional graph. When plotted in order (from 1 through 100) as the Y co-ordinate, the points will create either a Hill (a ⬚bump⬚ in the terrain) or a Valley (a ⬚dip⬚ in the terrain). </td> ',
' <td> ARCENE\'s task is to distinguish cancer versus normal patterns from mass-spectrometric data. This is a two-class classification problem with continuous input variables. This dataset is one of 5 datasets of the NIPS 2003 feature selection challenge. </td> ',
' <td> DEXTER is a text classification problem in a bag-of-word representation.

This is a two-class classification problem with sparse continuous input variables. This dataset is one of five datasets of the NIPS 2003 feature selection challenge.\r\n </td> ',
' <td> DOROTHEA is a drug discovery dataset. Chemical compounds represented by s tructural molecular features must be classified as active (binding to thrombin) or inactive. This is one of 5 datasets of the NIPS 2003 feature selection challenge. </td> ',
' <td> GISETTE is a handwritten digit recognition problem. The problem is to sep arate the highly confusible digits \'4\' and \'9\'. This dataset is one of five datasets of the NI PS 2003 feature selection challenge.\r\n </td> ',
' <td> MADELON is an artificial dataset, which was part of the NIPS 2003 feature selection challenge. This is a two-class classification problem with continuous input variables. T he difficulty is that the problem is multivariate and highly non-linear.   </td> ',
' <td> Two ground ozone level data sets are included in this collection. One is the eight hour peak set (eighthr.data), the other is the one hour peak set (onehr.data). Those dat a were collected from 1998 to 2004 at the Houston, Galveston and Brazoria area. </td> ',
' <td> The objective is to determine the set of boolean rules that describe the interactions of the nodes within this plant signaling network. The dataset includes 300 separate boolean pseudodynamic simulations using an asynchronous update scheme.   </td> ',
' <td> Oxford Parkinson\'s Disease Detection Dataset </td> ',
' <td> Multiple, labelled samples of pen tip trajectories recorded whilst writin g individual characters. All samples are from the same writer, for the purposes of primitive extra ction. Only characters with a single pen-down segment were considered. </td> ',
' <td> Data taken from the Blood Transfusion Service Center in Hsin-Chu City in Taiwan -- this is a classification problem.   </td> ',
' <td> A pen-based database with more than 11k isolated handwritten characters </td> ',
' <td> 1593 handwritten digits from around 80 persons were scanned, stretched in a rectangular box 16x16 in a gray scale of 256 values. </td> ',
' <td> Data from a semi-conductor manufacturing process </td> ',
' <td> Data has been extracted from the USDA plants database. It contains all pl ants (species and genera) in the database and the states of USA and Canada where they occur. </td> ',
' <td> The data set contains 15 classes of 24 instances each. Each class references to a hand movement type in LIBRAS (Portuguese\r\nname \'L?ngua BRAsileira de Sinais\', oficial brazilian signal language). </td> ',
' <td> Concrete is a highly complex material. The slump flow of concrete is not only determined by the water content, but that is also influenced by other concrete ingredients. </td> ',
' <td> Communities within the United States. The data combines socio-economic da ta from the 1990 US Census, law enforcement data from the 1990 US LEMAS survey, and crime data fro m the 1995 FBI UCR. </td> ',
' <td> The data was created by a medical expert as a data set to test the expert system, \r\nwhich will perform the presumptive diagnosis of two diseases of the urinary system.\r\n </td> ',
' <td> Two datasets are included, related to red and white vinho verde wine samp les, from the north of Portugal. The goal is to model wine quality based on physicochemical tests (see [Cortez et al., 2009], http://www3.dsi.uminho.pt/pcortez/wine/). </td> ',
' <td> Anonymized 120-day subset of the ICML-09 URL data containing 2.4 million examples and 3.2 million features. </td> ',
' <td> The goal is to model mutant p53 transcriptional activity (active vs inactive) based on data extracted from biophysical simulations.\r\n </td> ',
' <td> Oxford Parkinson\'s Disease Telemonitoring Dataset </td> ',
' <td> Marine sponges of the Demospongiae class classification domain. </td> ',
' <td> This dataset contains sentences extracted from user reviews on a given to pic. Example topics are “performance of Toyota Camry” and “sound quality of ipod nano”.   </td> ',
' <td> Dataset with electrical impedance measurements of freshly excised tissue samples from the breast. </td> ',
' <td> The dataset consists of measurements of fetal heart rate (FHR) and uterine contraction (UC) features on cardiotocograms classified by expert obstetricians. </td> ',
' <td> The data were collected as the SCITOS G5 robot navigates through the room following the wall in a clockwise direction, for 4 rounds, using 24 ultrasound sensors arranged ci rcularly around its \'waist\'. </td> ',
' <td> This dataset contains timeseries of mel-frequency cepstrum coefficients (MFCCs) corresponding to spoken Arabic digits. Includes data from 44 male and 44 female native Ara bic speakers. </td> ',
' <td> Data contains recordings of five people performing different activities. Each person wore four sensors (tags) while performing the same scenario five times.   </td > ',
' <td> AutoUniv is an advanced data generator for classifications tasks. The aim is to reflect the nuances and heterogeneity of real data. Data can be generated in .csv, ARFF or C4.5 formats. </td> ',
' <td> A dataset of steel plates’ faults, classified into 7 different types. \r\ nThe goal was to train machine learning for automatic pattern recognition.\r\n </td> ',
' <td> This dataset is taken from the MiniBooNE experiment and is used to distinguish electron neutrinos (signal) from muon neutrinos (background). </td> ',

' <td> Prediction of the release year of a song from audio features. Songs are m
ostly western, commercial tracks ranging from 1922 to 2011, with a peak in the year 2000s. </
p></td> ',
' <td> 15 months worth of daily data (440 daily records) that describes the occu
pancy rate, between 0 and 1, of different car lanes of the San Francisco bay area freeways across
time. </td> ',
' <td> This data set contains user reviews of cars and and hotels collected from
Tripadvisor (~259,000 \r\nreviews) and Edmunds (~42,230 reviews).   </td> ',
' <td> The dataset consists of 384 features extracted from CT images. The class
variable is numeric and denotes the relative location of the CT slice on the axial axis of the hum
an body. </td> ',
' <td> This is a dataset of 8235 online handwritten assamese characters. The “on
line” process involves capturing of data as text is written on a digitizing tablet with an
electronic pen. </td> ',
' <td> These highly imbalanced bioassay datasets are from the differing types of
screening that can be performed using HTS technology. 21 datasets were created from 12
bioassays. </td> ',
' <td> Element-wise comparison of records with personal data from a record linka
ge setting. The task is to decide from a comparison pattern whether the underlying records belong
to one person. </td> ',
' <td> Communities in the US. Data combines socio-economic data from the \'90 Ce
nsus, law enforcement data from the 1990 Law Enforcement Management and Admin Stats survey, and cr
ime data from the 1995 FBI UCR </td> ',
' <td> Data set containing values for six biomechanical features used to classif
y orthopaedic patients into 3 classes (normal, disk hernia or spondilolysthesis) or 2 classes (nor
mal or abnormal). </td> ',
' <td> The Physical Action Data Set includes 10 normal and 10 aggressive physica
l actions that measure the human activity. The data have been collected by 4 subjects using the De
lsys EMG wireless apparatus. </td> ',
' <td> The Physical Action Data Set includes 10 normal and 10 aggressive physica
l actions that measure the human activity. The data have been collected by 10 subjects using the V
icon 3D tracker. </td> ',
' <td> The dataset is used for authorship identification in online Writeprint wh
ich is a new research field of pattern recognition.   </td> ',
' <td> Amazon\'s InfoSec is getting smarter about the way Access data is
leveraged. This is an anonymized sample of access provisioned within the company. </td> '
,
' <td> The dataset is used for authorship identification in online Writeprint wh
ich is a new research field of pattern recognition.   </td> ',
' <td> This data was collected from text ads found on twelve websites that deal
with various farm animal related topics. The binary labels are based on whether or not the conten
t owner approves of the ad. </td> ',
' <td> It contains 64 e-mails which I have manually collected from DBWorld maili
ng list. They are classified in: \'announces of conferences\' and \'everything else\'. </
td> ',
' <td> KEGG Metabolic pathways modeled as directed relation network. Variety of
graphical features presented. </td> ',
' <td> KEGG Metabolic pathways modeled as un-directed reaction network. Variety
of graphical features presented. </td> ',
' <td> The data is related with direct marketing campaigns (phone calls) of a Po
rtuguese banking institution. The classification goal is to predict if the client will subscribe a
term deposit (variable y). </td> ',
' <td> This dataset provides user vote data on which video from a pair of videos
is funnier collected on YouTube Comedy Slam. The task is to automatically predict this preference
based on video metadata. </td> ',
' <td> This archive contains 13910 measurements from 16 chemical sensors
utilized in simulations for drift compensation in a discrimination task of 6 gases at various
levels of concentrations. </td> ',
' <td> This data set contains 10 variables that are age, gender, total
Bilirubin, direct Bilirubin, total proteins, albumin, A/G ratio, SGPT, SGOT and Alkphos.
</td> ',
' <td> The OPPORTUNITY Dataset for Human Activity Recognition from Wearable, Obj
ect, and Ambient Sensors is a dataset devised to benchmark human activity recognition algorithms (
classification, automatic data segmentation, sensor fusion, feature extraction, etc). </t
d> ',
' <td> Nomao collects data about places (name, phone, localization...) from many
sources.\r\nDeduplication consists in detecting what data refer to the same place.\r\nInstances in
the dataset compare 2 spots. </td> ',
' <td> The SMS Spam Collection is a public set of SMS labeled messages that have
been collected for mobile phone spam research. </td> ',
' <td> The Skin Segmentation dataset is constructed over B, G, R color space. Sk
in and Nonskin dataset is generated using skin textures from face images of diversity of age, gend
er, and race people. </td> ',
' <td> The dataset concerns with the classification of two mental stages from re
corded EEG signals: Planning (during imagination of motor act) and Relax state.   </td> ',
' <td> The PAMAP2 Physical Activity Monitoring dataset contains data of 18 diffe
rent physical activities, performed by 9 subjects wearing 3 inertial measurement units and a heart

rate monitor. </td> ',
' <td> The dataset was obtained from a recommender system prototype. The task wa
s to generate a top-n list of restaurants according to the consumer preferences.   </td> '
,
' <td> This is a data set containing 1080 documents of free text business descri
ptions of Brazilian companies categorized into a\r\nsubset of 9 categories </td> ',
' <td> Measurements of electric power consumption in one household with a one-mi
nute sampling rate over a period of almost 4 years. Different electrical quantities and some sub-m
etering values are available. </td> ',
' <td> Measurements of geometrical properties of kernels belonging to three
different varieties of wheat. A soft X-ray technique and GRAINS package were used to construct all
seven, real-valued attributes. </td> ',
' <td> Northix is designed to be a schema matching benchmark problem for data in
tegration of two entity relationship databases.   </td> ',
' <td> Since there is no numerical sequential data stream available in standard
data sets, this data set is generated from the original T40I10D100K data set </td> ',
' <td> A textual corpus of 4000 legal cases for automatic summarization and cita
tion analysis. For each document we collect catchphrases, citations sentences, citation
catchphrases and citation classes. </td> ',
' <td> Human Activity Recognition database built from the recordings of 30 subje
cts performing activities of daily living (ADL) while carrying a waist-mounted smartphone with
embedded inertial sensors. </td> ',
' <td> Sixteen samples of leaf each of one-hundred plant species. For each sampl
e, a shape descriptor, fine scale margin and texture histogram are given. </td> ',
' <td> This study looked into assessing the heating load and cooling load requir
ements of buildings (that is, energy efficiency) as a function of building parameters. </
td> ',
' <td> Delft data set, used to predict the hydodynamic performance of sailing ya
chts from dimensions and velocity. </td> ',
' <td> 100 volunteers provide a semen sample analyzed according to the WHO 2010
criteria. Sperm concentration are related to socio-demographic data, environmental factors, health
status, and life habits </td> ',
' <td> This dataset contains the annotated readings of 3 acceleration sensors at
the hip and leg of Parkinson\'s disease patients that experience freezing of gait (FoG) during wal
king tasks.\r\n </td> ',
' <td> 3D road network with highly accurate elevation information (+-20cm) from
Denmark used in eco-routing and fuel/Co2-estimation routing algorithms. </td> ',
' <td> Data sets includes returns of Istanbul Stock Exchange with seven other in
ternational index; SP, DAX, FTSE, NIKKEI, BOVESPA, MSCE_EU, MSCI_EM from Jun 5, 2009 to Feb 22, 20
11. </td> ',
' <td> This data-set contains examples of buzz events from two different social
networks: Twitter, and Tom\'s Hardware, a forum network focusing on new technology with more conse
rvative dynamics. </td> ',
' <td> Given a theorem, predict which of five heuristics will give the fastest p
roof when used by a first-order prover. A sixth prediction declines to attempt a proof, should the
theorem be too difficult. </td> ',
' <td> A dataset with 5 classes (sitting-down, standing-up, standing, walking, a
nd sitting) collected on 8 hours of activities of 4 healthy subjects. We also established a
baseline performance index. </td> ',
' <td> The dataset contains 18000 time-series recordings from a chemical
detection platform at six different locations in a wind tunnel facility in response to ten high-pr
iority chemical gaseous substances </td> ',
' <td> Given Latin hypercube samples of 18 climate model input parameter values,
predict climate model simulation crashes and determine the parameter value combinations that cause
the failures. </td> ',
' <td> A dataset to explore machine learning approaches for the identification o
f microorganisms from mass-spectrometry data. </td> ',
' <td> Data set containing values for 41 attributes (molecular descriptors) used
to classify 1055 chemicals into 2 classes (ready and not ready biodegradable). </td> ',
' <td> In this paper, we look for to recognize the causes of users tend\r\nto cy
ber space in Kohkiloye and Boyer Ahmad Province in\r\nIran </td> ',
' <td> The dataset comprises motion sensor data of 19 daily and sports
activities each performed by 8 subjects in their own style for 5 minutes. Five Xsens MTx units are
used on the torso, arms, and legs.\r\n </td> ',
' <td> It is the real dataset about the students\' knowledge status about the su
bject of Electrical DC Machines. The dataset had been obtained from Ph.D. Thesis. </td> '
,
' <td> This test collection contains feature characteristics of documents
originally written in five different languages and their translations, over a common set of 6 cate
gories.   </td> ',
' <td> NYSK (New York v. Strauss-Kahn) is a collection of English news articles
about the case relating to allegations of sexual assault against the former IMF director Dominique
Strauss-Kahn (May 2011). </td> ',
' <td> This data set contains a total 5820 evaluation scores provided by student
s from Gazi University in Ankara (Turkey). There is a total of 28 course specific questions and ad
ditional 5 attributes. </td> ',
' <td> The dataset is about the users\' learning activities and knowledge levels

' <td> The dataset is about the users' learning activities and knowledge levels
on subjects of DC Electrical Machines. The dataset had been obtained from online web-courses and
reported in my Ph.D. Thesis. </td> ',
' <td> The data set consists of 14 EEG values and a value indicating the eye sta
te. </td> ',
' <td> This is a data set of Physicochemical Properties of Protein Tertiary Stru
cture. The data set is taken from CASP 5-9. There are 45730 decoys and size varying from 0 to 21 a
rmstrong. </td> ',
' <td> The data describe the problem of high energy (higher than 10^4 J) seismic
bumps forecasting in a coal \r\nmine. Data come from two of longwalls located in a Polish coal min
e. </td> ',
' <td> Data were extracted from images that were taken for the evaluation of an
authentication procedure for banknotes. </td> ',
' <td> Data used for USPTO Algorithm Competition. Contains drawing pages from US
patents with manually labeled figure and part labels. </td> ',
' <td> This dataset contains about 120k instances, each described by 13 feature
types, with class information, specially useful for exploring multiview topics (cotraining,
ensembles, clustering,..). </td> ',
' <td> This archive contains 13910 measurements from 16 chemical sensors exposed
to 6 different gases at various concentration levels. </td> ',
' <td> This dataset comprises information regarding the ADLs performed by two us
ers on a daily basis in their \r\nown homes.   </td> ',
' <td> This data was used in Thompson et al. (2013). A list of possible game act
ions is discussed in Thompson, Blair, Chen, & Henrey (2013). </td> ',
' <td> Six young health subjects were asked to perform 5 variations of the
biceps curl weight lifting exercise. One of the variations is the one predicted by the health prof
essional. </td> ',
' <td> This dataset is collected from a monitor system mounted in a domotic hous
e. It corresponds to approximately 40 days of monitoring data. </td> ',
' <td> This dataset contains the hourly and daily count of rental bikes between
years 2011 and 2012 in Capital bikeshare system with the corresponding weather and seasonal inform
ation. </td> ',
' <td> The data from Twitter was collected during 360 consecutive days. It was d
one by querying 1497 English keywords sampled from Wikipedia. This dataset is proposed in a
Learning to rank setting. </td> ',
' <td> The data is dedicated to classification problem related to the post-
operative life expectancy in the lung cancer patients: class 1 - death within one year after surge
ry, class 2 - survival. </td> ',
' <td> >3 different exercises: sitting, standing and walking in the muscles: bice
ps femoris, vastus medialis, rectus femoris and semitendinosus addition to goniometry in the exerc
ises. </td> ',
' <td> This is a classification problem to distinguish between a signal process
which produces supersymmetric particles and a background process which does not. </td> ',
' <td> This is a classification problem to distinguish between a signal process
which produces Higgs bosons and a background process which does not.   </td> ',
' <td> Predict the Bankruptcy from Qualitative parameters from experts. </p
></td> ',
' <td> 126 samples from 14 participants, 309 features. Aim: assess whether voice
rehabilitation treatment lead to phonations considered \'acceptable\' or \'unacceptable\' (binary
class classification problem). </td> ',
' <td> Recordings of 16 volunteers performing 14 Activities of Daily Living
(ADL) while carrying a single wrist-worn tri-axial accelerometer. </td> ',
' <td> High-resolution Remote Sensing data set (Quickbird). Small number of trai
ning samples of diseased trees, large number for other land cover. Testing data set from
stratified random sample of image. </td> ',
' <td> The dataset collects data from an Android smartphone positioned in the ch
est pocket from 22 participants walking in the wild over a predefined path. \r\n </td> ',
' <td> The dataset collects data from a wearable accelerometer mounted on the ch
est. The dataset is intended for Activity Recognition research purposes. </td> ',
' <td> This dataset consists in a collection of shape and texture features extra
cted from digital images of leaf specimens originating from a total of 40 different plant
species. </td> ',
' <td> This dataset contain Attributes of dresses and their recommendations
according to their sales.Sales are monitor on the basis of alternate days.   </td> ',
' <td> This data can be effectively produced the result to fewer parameter of th
e Load profile can be reduced in the Database   </td> ',
' <td> NASA data set, obtained from a series of aerodynamic and acoustic tests o
f two and three-dimensional airfoil blade sections conducted in an anechoic wind tunnel.
</td> ',
' <td> The data set refers to clients of a wholesale distributor. It includes th
e annual spending in monetary units (m.u.) on diverse product categories </td> ',
' <td> This problem of Sentiment Analysis (SA) has been studied well on the Engl
ish language but not Arabic one. Two main approaches have been devised: corpus-based and lexicon-b
ased.   </td> ',
' <td> The dataset contains 9568 data points collected from a Combined Cycle Pow
er Plant over 6 years (2006-2011), when the plant was set to work with full load.   </td>
',
' <td> Classification of urban land cover using high resolution aerial imagery.

' <td> Classification of urban land cover using high resolution aerial imagery. Intended to assist sustainable urban planning efforts. </td> ',
' <td> This data has been prepared to analyze factors related to readmission as well as other \r\n\r\noutcomes pertaining to patients with diabetes. </td> ',
' <td> The data set is composed of 60 chorales (5665 events) by J.S. Bach (1675-1750).\r\nEach event of each chorale is labelled using 1 among 101 chord labels and described\r\nthrough 14 features. </td> ',
' <td> Stone flakes are waste products of the stone tool production in\r\nthe prehistoric era. The variables are means of geometric and\r\nstylistic features of the flakes contained in different inventories. </td> ',
' <td> This is a collection of 8 files containing the match statistics for both women and men at the four major tennis tournaments of the year 2013. Each file has 42 columns and a minimum of 76 rows. </td> ',
' <td> The training data belongs to 20 Parkinson\'s Disease (PD) patients and 20 healthy subjects. From all subjects, multiple types of sound recordings (26) are taken. </td> ',
' <td> The dataset is composed by features extracted from 7 videos with people gesticulating, aiming at studying Gesture Phase Segmentation. It contains 50 attributes divided into two files for each video. </td> ',
' <td> This data consists of odors of 20 different perfumes. Data was obtained by using a handheld odor meter (OMX-GR sensor) per second for 28 seconds period. </td> ',
' <td> Instances in this dataset contain features extracted from blog posts. The task associated with the data is to predict how many comments the post will receive. </td> ',
' <td> The REALDISP dataset is devised to evaluate techniques dealing with the effects of sensor displacement in wearable activity recognition as well as to benchmark general activity recognition algorithms   </td> ',
' <td> Dataset is well suited for segmentation tasks. It contains 101 scanned pages from different newspapers and magazines in Russian with ground truth pixel-based masks. </td> ',
' <td> This data set compromises the metadata for the 2014 AAAI conference\'s accepted papers, including paper titles, authors, abstracts, and keywords of varying granularity. </td> ',
' <td> The data set contains 58 time series acquired from 16 chemical sensors under gas flow modulation conditions. The sensors were exposed to different gaseous binary mixtures of acetone and ethanol. </td> ',
' <td> A chemical detection platform composed of 8 chemoresistive gas sensors was exposed to turbulent gas mixtures generated naturally in a wind tunnel. The acquired time series of the sensors are provided. </td> ',
' <td> The UJIIndoorLoc is a Multi-Building Multi-Floor indoor localization database to test Indoor Positioning System that rely on WLAN/WiFi fingerprint. </td> ',
' <td> Contains sentences from the abstract and introduction of 30 articles annotated with a modified Argumentative Zones annotation scheme. These articles come from biology, machine learning and psychology. </td> ',
' <td> This dataset contains weekly data for the Dow Jones Industrial Index. It has been used in computational investing research. </td> ',
' <td> The “sEMG for Basic Hand movements” includes 2 databases of surface electromyographic signals of 6 hand movements using Delsys\' EMG System. Healthy subjects conducted six daily life grasps. </td> ',
' <td> This data set compromises the metadata for the 2013 AAAI conference\'s accepted papers (main track only), including paper titles, abstracts, and keywords of varying granularity. </td> ',
' <td> Instances in this dataset contain audio features extracted from 1059 wave files. The task associated with the data is to predict the geographical origin of music.\r\n </td> ',
' <td> Data have been generated from a sophisticated simulator of a Gas Turbines (GT), mounted on a Frigate characterized by a COmbined Diesel eLectric And Gas (CODLAG) propulsion plant type. </td> ',
' <td> This dataset supports the development of models that make possible to interpret Grammatical Facial Expressions from Brazilian Sign Language (Libras). </td> ',
' <td> Corpus intended to do cleaning (or binarization) and enhancement of noisy grayscale printed text images using supervised learning methods. Noisy images and their corresponding ground truth provided. </td> ',
' <td> The MHEALTH (Mobile Health) dataset is devised to benchmark techniques dealing with human behavior analysis based on multimodal body sensing. </td> ',
' <td> Predict student performance in secondary education (high school).   </td> ',
' <td> This data set contains electricity consumption of 370 points/clients.\r\n </td> ',
' <td> The data set contains the recordings of 16 chemical sensors exposed to two dynamic gas mixtures at varying concentrations. For each mixture, signals were acquired continuously during 12 hours. </td> ',
' <td> MicroblogPCU data is crawled from sina weibo microblog[http://weibo.com/]. This data can be used to study machine learning methods as well as do some social network research.   </td> ',
' <td> The data are binary attack-point vectors and their clave-direction class(es) according to the partido-alto-based paradigm. </td> ',
' <td> Features are extracted from motor current. The motor has intact and defec

' <td> Features are extracted from motor current. The motor has intact and defec
tive components. This results in 11 different classes with different conditions.   </td> '
,
' <td> TV Commercials data set consists of standard audio-visual features of vi
deo shots extracted from 150 hours of TV news broadcast of 3 Indian and 2 international news
channels (30 Hours each).   </td> ',
' <td> This dataset collected mainly from: PhishTank archive, MillerSmiles archi
ve, Google’s searching operators. </td> ',
' <td> Design an observing network to monitor emissions of a greenhouse gas (GHG
) in California given time series of synthetic observations and tracers from weather model
simulations.\r\n </td> ',
' <td> This dataset contains features extracted from the Messidor image set to p
redict whether an image contains signs of diabetic retinopathy or not.   </td> ',
' <td> The data contains lists of octamers (8 amino acids) and a flag (-1 or 1)
depending on whether HIV-1 protease will cleave in the central position (between amino acids 4 and
5). </td> ',
' <td> The dataset contains sentences labelled with positive or negative sentime
nt. </td> ',
' <td> This dataset summarizes a heterogeneous set of features about articles pu
blished by Mashable in a period of two years. The goal is to predict the number of shares in socia
l networks (popularity). </td> ',
' <td> Multi-temporal remote sensing data of a forested area in Japan. The goal
is to map different forest types using spectral data. </td> ',
' <td> Survey of faculty members from two Spanish universities on teaching uses
of Wikipedia </td> ',
' <td> The dataset contains a million randomly sampled video instances listing 1
0 fundamental video characteristics along with the YouTube video ID.   </td> ',
' <td> This dataset can be used to predict the chronic kidney disease and it can
be collected from the hospital nearly 2 months of period. </td> ',
' <td> The data here are the ZZAlpha® machine learning recommendations made for
various US traded stock portfolios the morning of each day during the 3 year period Jan 1, 2012 -
Dec 31, 2014.   </td> ',
' <td> 20 photos of leaves for each of 32 different species. </td> ',
' <td> An accurate dataset describing trajectories performed by all the 442 taxi
s running in the city of Porto, in Portugal.\r\n </td> ',
' <td> This Data set provides preprocessed and cleaned vital signals which can b
e used in designing algorithms for cuff-less estimation of the blood pressure. </td> ',
' <td> Activity recognition data set built from the recordings of 30 subjects pe
rforming basic activities and postural transitions while carrying a waist-mounted smartphone with
embedded inertial sensors.\r\n </td> ',
' <td> Expression levels of 77 proteins measured in the cerebral cortex of 8 cla
sses of control and Down syndrome mice exposed to context fear conditioning, a task used to assess
associative learning. </td> ',
' <td> The UJIIndoorLoc-Mag is an indoor localization database to test Indoor Po
sitioning System that rely on Earth\'s magnetic field variations. </td> ',
' <td> The Heterogeneity Human Activity Recognition (HHAR) dataset from
Smartphones and Smartwatches is a dataset devised to benchmark human activity recognition
algorithms (classification, automatic data segmentation, sensor fusion, feature extraction, etc.)
in real-world contexts; specifically, the dataset is gathered with a variety of different device m
odels and use-scenarios, in order to reflect sensing heterogeneities to be expected in real
deployments. </td> ',
' <td> Educational Process Mining data set is built from the recordings of 115 s
ubjects\' activities through a logging application while learning with an educational
simulator. </td> ',
' <td> The search for exotic particles requires sorting through a large number o
f collisions to find the events of interest. This data set challenges one to detect a new particle
of unknown mass. </td> ',
' <td> This dataset contains temporal data from a Wireless Sensor Network deploy
ed in real-world office environments. The task is intended as real-life benchmark in the area of A
mbient Assisted Living. </td> ',
' <td> Open University Learning Analytics Dataset contains data about courses, s
tudents and their interactions with Virtual Learning Environment for seven selected courses and mo
re than 30000 students. </td> ',
' <td> This research aimed at the case of customers’ default payments in Taiwan
and compares the predictive accuracy of probability of default among six data mining
methods. </td> ',
' <td> Mesothelioma’s disease data set were prepared at Dicle University Faculty
of Medicine in Turkey.\r\nThree hundred and twenty-four Mesothelioma patient data. In the dataset,
all samples have 34 features. </td> ',
' <td> This is a transnational data set which contains all the transactions occu
rring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store online
retail. </td> ',
' <td> In SIFT10M, each data point is a SIFT feature which is extracted from Cal
tech-256 by the open source VLFeat library. The corresponding patches of the SIFT features are pro
vided. </td> ',
' <td> The dataset has been feed by Android app called Go!Track. It is available
at Goolge Play Store(https://play.google.com/store/apps/details?id=com.go.router).   </td>
'

,
' <td> I extract features from malacious and non-malacious and create and training dataset to teach svm classifier.Dataset made of unknown executable to detect if it is virus or normal safe executable. </td> ',
' <td> Experimental data used for binary classification (room occupancy) from Temperature,Humidity,Light and CO2. Ground-truth occupancy was obtained from time stamped pictures that were taken every minute. </td> ',
' <td> Handwriting database consists of 25 PWP(People with Parkinson) and 15 healthy individuals.Three types of recordings (Static Spiral Test, Dynamic Spiral Test and Stability Test) are taken. </td> ',
' <td> References to news pages collected from an web aggregator in the period from 10-March-2014 to 10-August-2014. The resources are grouped into clusters that represent pages discussing the same story. </td> ',
' <td> Contains the responses of a gas multisensor device deployed on the field in an Italian city. Hourly responses averages are recorded along with gas concentrations references from a certified analyzer.   </td> ',
' <td> 5 replicates of an 8-MOX gas sensor array were exposed to different gas conditions (4 volatiles at 10 concentration levels each). </td> ',
' <td> 100 recordings of a sensor array under different conditions in a home setting: background, wine and banana presentations. The array includes 8 MOX gas sensors, and humidity and temperature sensors.\r\n </td> ',
' <td> Instances in this dataset contain features extracted from facebook posts. The task associated with the data is to predict how many comments the post will receive. </td> ',
' <td> This data is an addition to an existing dataset on UCI. We collected more data to improve the accuracy of our human activity recognition algorithms applied in the domain of Ambient Assisted Living.   </td> ',
' <td> The dataset is about bankruptcy prediction of Polish companies.The bankrupt companies were analyzed in the period 2000-2012, while the still operating companies were evaluated from 2007 to 2013. </td> ',
' <td> This dataset contains temporal data from a Wireless Sensor Network worn by an actor performing the activities: bending, cycling, lying down, sitting, standing, walking. </td> ',
' <td> Dota 2 is a popular computer game with two teams of 5 players. At the start of the game each player chooses a unique hero with different strengths and weaknesses. </td> ',
' <td> Facebook performance metrics of a renowned cosmetic\'s brand Facebook page. </td> ',
' <td> UbiqLog is the smartphone lifelogging tool that runs on the smartphone of 35 users for about 2 months. \r\n </td> ',
' <td> This data set contains the distribution of words in the full text of the NIPS conference papers published from 1987 to 2015. </td> ',
' <td> Pulsar candidates collected during the HTRU survey. Pulsars are a type of star, of considerable scientific interest. Candidates must be classified in to pulsar and non-pulsar classes to aid discovery. </td> ',
' <td> Classify type of drug consumer by personality data </td> ',
' <td> Experimental data used to create regression models of appliances energy use in a low energy building. </td> ',
' <td> The dataset was created for the comparison and evaluation of hybrid indoor positioning methods. The dataset presented contains data from W-LAN and Bluetooth interfaces, and Magnetometer.   </td> ',
' <td> KDC-4007 dataset Collection is the Kurdish Documents Classification text used in categories regarding Kurdish Sorani news and articles. </td> ',
' <td> A multisource and multivariate dataset for indoor localisation methods based on WLAN and Geo-Magnetic field fingerprinting </td> ',
' <td> The DrivFace contains images sequences of subjects while driving in real scenarios. It is composed of 606 samples of 640×480, acquired over different days from 4 drivers with several facial features. </td> ',
' <td> \r\n\r\n </td> ',
' <td> It is a public set of comments collected for spam research. It has five datasets composed by 1,956 real messages extracted from five videos that were among the 10 most viewed on the collection period. </td> ',
' <td> This hourly data set contains the PM2.5 data of US Embassy in Beijing. Meanwhile, meteorological data from Beijing Capital International Airport are also included.   </td> ',
' <td> Sanitized and anonymized Cargo 2000 (C2K) airfreight tracking and tracing events, covering five months of business execution (3,942 process instances, 7,932 transport legs, 56,082 activities).   </td> ',
' <td> This dataset focuses on the prediction of indicators/diagnosis of cervical cancer. The features cover demographic information, habits, and historic medical records. </td> ',
' <td> This dataset explores the subjective quality assessment of digital colposcopies. </td> ',
' <td> KASANDR is a novel, publicly available collection for recommendation systems that records the behavior of customers of the European leader in e-Commerce advertising, Kelkoo.   </td> ',
' <td> FMA features 106,574 tracks and includes song title, album, artist, genres, play counts, favorites, comments, description, biography, tags, together with audio (343 days

s; play counts, favorites, comments; description, biography, tags; together with audio (345 days, 917 GiB) and features. </td> ',
' <td> Contains the responses of a gas multisensor device deployed on the field in an Italian city.   </td> ',
' <td> This dataset is a pre-processed and re-structured/reshaped version of a very commonly used dataset featuring epileptic seizure detection.   </td> ',
' <td> This is an image database of Handwritten Devanagari characters. There are 46 classes of characters with 2000 examples each. The dataset is split into training set(85%) and testing set(15%).   </td> ',
' <td> The data set of performances of weighted scoring stock portfolios are obtained with mixture design from the US stock market historical database. </td> ',
' <td> 5 types of hand postures from 12 users were recorded using unlabeled markers attached to fingers of a glove in a motion capture environment. Due to resolution and occlusion, missing values are common. </td> ',
' <td> Predict a pattern of neurodegeneration in the dataset of speech features obtained from patients with early untreated Parkinson’s disease and patients at high risk developing Parkinson’s disease. </td> ',
' <td> Data include over 100 Team Activity Measures and outcomes (ML classes) obtained from activities of 74 student teams during the creation of final class project in SW Eng. classes at SFSU, Fulda, FAU </td> ',
' <td> This hourly data set contains the PM2.5 data in Beijing, Shanghai, Guangzhou, Chengdu and Shenyang. Meanwhile, meteorological data for each city are also included. </td> ',
' <td> Handwriting database consists of 62 PWP(People with Parkinson) and 15 healthy individuals. Three types of recordings (Static Spiral Test, Dynamic Spiral Test and Stability Test) are taken. </td> ',
' <td> Contains weekly purchased quantities of 800 over products over 52 weeks. Normalised values are provided too. </td> ',
' <td> This dataset includes quantitative and categorical features from online reviews from 21 hotels located in Las Vegas Strip, extracted from TripAdvisor (http://www.tripadvisor.com). </td> ',
' <td> This dataset includes Online Textual Reviews from both online (e.g., Trip Advisor) and offline (e.g., Guests\' book) sources from the Areias do Seixo Eco-Resort. </td> ',
' <td> This dataset contains keystroke dynamics data collected on a touch mobile device (Nexus 7). The dataset contains 2856 records, 51 records per subject for 56 subjects.   </td> ',
' <td> Crowdsourced data from OpenStreetMap is used to automate the classification of satellite images into different land cover classes (impervious, farm, forest, grass, orchard, water).   </td> ',
' <td> This collection of data is part of the RNA-Seq (HiSeq) PANCAN data set, it is a random extraction of gene expressions of patients having different types of tumor: BRCA, KIRC, COAD, LUAD and PRAD. </td> ',
' <td> The dataset was created for the comparison and evaluation of hybrid indoor positioning methods. The dataset presented contains data from W-LAN and Bluetooth interfaces, and Magnetometer.   </td> ',
' <td> The research project presents this database, shows the images of chestnuts that will be processed to determine the presence or absence of defects </td> ',
' <td> One of the primary challenges in identifying the risks of the Burst Header Packet (BHP) flood attacks in Optical Burst Switching networks (OBS) is the scarcity of reliable historical data.   </td> ',
' <td> 5 types of hand postures from 12 users were recorded using unlabeled markers on fingers of a glove in a motion capture environment. Due to resolution and occlusion, missing values are common. </td> ',
' <td> Acoustic features extracted from syllables of anuran (frogs) calls, including the family, the genus, and the species labels (multilabel).   </td> ',
' <td> The TTC-3600 data set is a collection of Turkish news and articles including categorized 3,600 documents from 6 well-known portals in Turkey. It has 4 different forms in ARFF Weka format. </td> ',
' <td> This dataset contains features extracted from colonoscopy videos used to detect gastrointestinal lesions. It contains 76 lesions: 15 serrated adenomas, 21 hyperplastic lesions and 40 adenoma.   </td> ',
' <td> The dataset was collected during 60 days, this is a real database of a brazilian logistics company. </td> ',
' <td> This sentiment analysis data set contains scientific paper reviews from an international conference on computing and informatics. The task is to predict the orientation or the evaluation of a review. </td> ',
' <td> It was collected for CAD diagnosis. </td> ',
' <td> It was collected for CAD diagnosis. </td> ',
' <td> This dataset contains the dynamic features of 107,888 executables, collected by VirusShare from Nov/2010 to Jul/2014. </td> ',
' <td> The dataset consists of data collected from heavy Scania trucks in everyday usage.   </td> ',
' <td> This set Provides data regarding wireless communications between vehicles and road side units. two separate data sets are provided (normal scenario) and in the presence of attacker (jammer). </td> ',
' <td> This dataset was collected by Shan-Hung Wu and DataLab members at NTHU, Taiwan. There\'re 225 user perceived clusters from 100 users and their corresponding

aiwan. There\'re 325 user-perceived clusters from 100 users and their corresponding descriptions. </td> ',
' <td> Character images from scanned and computer generated fonts. </td> ',
' <td> This dataset includes 1) 12234 documents (8251 training, 3983 test) extracted from DeliciousT140 dataset, 2) class labels for all documents, 3) labels for a subset of sentences of the test documents. </td> ',
' <td> Children screening data for autism suitable for classification and predictive tasks   </td> ',
' <td> Autistic Spectrum Disorder Screening Data for Adolescent. This dataset is related to classification and predictive tasks. </td> ',
' <td> The datasets\' positive class consists of component failures for a specific component of the APS system. The negative class consists of trucks with failures for components not related to the APS. </td> ',
' <td> Collected in indoor space by observing signal strengths of seven WiFi signals visible on a smartphone. The decision variable is one of the four rooms.   </td> ',
' <td> Hepatocellular Carcinoma dataset (HCC dataset) was collected at a University Hospital in Portugal. It contains real clinical data of 165 patients diagnosed with HCC. </td> ',
' <td> 12 features categorized as conventional and social media features. Both conventional features, collected from movies databases on Web as well as social media features(YouTube,Twitter). </td> ',
' <td> Persian questions gathered from a jeopardy game broadcasted on Iranian national television.   </td> ',
' <td> Autistic Spectrum Disorder Screening Data for Adult. This dataset is related to classification and predictive tasks. </td> ',
' <td> Sequential motion data from 14 healthy older people aged 66 to 86 years old using a batteryless, wearable sensor on top of their clothing for the recognition of activities in clinical environments. </td> ',
' <td> This dataset contains information about wart treatment results of 90 patients using immunotherapy. </td> ',
' <td> This dataset contains information about wart treatment results of 90 patients using cryotherapy. </td> ',
' <td> This dataset contains OCT data (in mat format) and color fundus data (in jpg format) of left & right eyes of 50 healthy persons. </td> ',
' <td> Discrete Tone Images(DTI)are available which needs to be analyzed in detail. Here, we created this dataset for those who do research in DTI. \r\n </td> ',
' <td> Large data set of news items and their respective social feedback on multiple platforms: Facebook, Google+ and LinkedIn. </td> ',
' <td> Fault diagnosis of four liquid ultrasonic flowmeters </td> ',
' <td> This data set compromises the metadata for the 2014 ICMLA conference\'s accepted papers, including ID, paper titles, author\'s keywords, abstracts and sessions in which they were exposed. </td> ',
' <td> This dataset contains RSSI readings gathered from an array of Bluetooth Low Energy (BLE) iBeacons in a real-world and operational indoor environment for localization and navigation purposes. </td> ',
' <td> A container crane has the function of transporting containers from one point to another point. </td> ',
' <td> Data set includes construction cost, sale prices, project variables, and economic variables corresponding to real estate single-family residential apartments in Tehran, Iran.   </td> ',
' <td> The data was collected in 2015 using Twitter API. This dataset contains health news from more than 15 major health news agencies such as BBC, CNN, and NYT.   </td> ',
' <td> ChIP-seq experiments characterize protein modifications or binding at\r\nspecific genomic locations in specific samples. The machine learning\r\nproblem in these data is structured binary classification. </td> ',
' <td> Running times for multiplying two 2048 x 2048 matrices using a GPU OpenCL SGEMM kernel with varying parameters (using the library \'CLTune\'). </td> ',
' <td> The dataset contains 7 datasets of User - Item matrices, where each entry represents how many times a user consumed an item. Item is used as an umbrella term for various categories. </td> ',
' <td> This dataset addresses the lack of public botnet datasets, especially for the IoT. It suggests *real* traffic data, gathered from 9 commercial IoT devices authentically infected by Mirai and BASHLITE. </td> ',
' <td> The database was created with records of absenteeism at work from July 2007 to July 2010 at a courier company in Brazil. </td> ',
' <td> First self-care activities dataset based on ICF-CY. </td> ',
' <td> The data set addresses the condition assessment of a hydraulic test rig based on multi sensor data. Four fault types are superimposed with several severity grades impeding selective quantification. </td> ',
' <td> This dataset contains 10721 initial and calculated atomic coordinates of carbon nanotubes. </td> ',
' <td> This dataset contains 640 performance measurements from a simulation of 2-Dimensional Multiprocessor Optical Interconnection Network.   </td> ',
' <td> 1000 sports articles were labeled using Amazon Mechanical Turk as objective or subjective. The raw texts, extracted features, and the URLs from which the articles w

ere retrieved are provided. </td> ',
' <td> Clinical features were observed or measured for 64 patients with breast c
ancer and 52 healthy controls.   </td> ',
' <td> The data-set contains four (4) sets of mobile sensor readings data (humid
ity, temperature) corresponding to a swarm of four (4) Unmanned Surface Vehicles (USVs) in a test-
bed in Athens (Greece).   </td> ',
' <td> The dataset was used to test an architecture based on a trust model
capable to cope with the evaluation of the trustworthiness of users interacting in pervasive
environments. </td> ',
' <td> To create the largest authorship attribution dataset, we extracted works
of 50 well-known authors. To have a non-exhaustive learning, in training there are 45 authors wher
eas, in the testing, it\'s 50 </td> ',
' <td> 20 falls and 16 daily living activities were performed by 17 volunteers w
ith 5 repetitions while wearing 6 sensors (3.060 instances) that attached to their head, chest, wa
ist, wrist, thigh and ankle. </td> ',
' <td> 5879 captioned images (image and text) from social media related to damag
e during natural disasters/wars, and belong to 6 classes: Fires, Floods, Natural landscape,
Infrastructural, Human, Non-damage. </td> ',
' <td> This database consists on 30 subjects performing Brain Computer Interface
for Steady State Visual Evoked Potentials (BCI-SSVEP).   </td> ',
' <td> Roman Urdu (the scripting style for Urdu language) is one of the limited
resource languages.A data corpus comprising of more than 20000 records was collected. </t
d> ',
' <td> The Avila data set has been extracted from 800 images of the \'Avila Bibl
e\', an XII century giant Latin copy of the Bible. The prediction task consists in associating ea
ch pattern to a copyist. </td> ',
' <td> PANDOR is a novel and publicly available dataset for online
recommendation provided by Purch (http://www.purch.com/).   </td> ',
' <td> The dataset provides patient reviews on specific drugs along with related
conditions. Reviews and ratings are grouped into reports on the three aspects benefits, side
effects and overall comment. </td> ',
' <td> The dataset provides patient reviews on specific drugs along with related
conditions and a 10 star patient rating reflecting overall patient satisfaction. </td> ',
' <td> The dataset is generated from Physical Unclonable Functions (PUFs)
simulation, specifically XOR Arbiter PUFs. PUFs are used for authentication purposes. For more inf
o, refer to our paper below. </td> ',
' <td> Two file s contain data on 21263 superconductors and their relevant
features. </td> ',
' <td> WESAD (Wearable Stress and Affect Detection) contains data of 15 subjects
during a stress-affect lab study, while wearing physiological and motion sensors. </td> '
,
' <td> The data-set contains eight (2x4) data-sets of mobile sensor readings dat
a (humidity, temperature) corresponding to a swarm of four Unmanned Surface Vehicles (USVs) in a t
est-bed, Athens, Greece. </td> ',
' <td> The dataset tried to find the end semester percentage prediction based on
different social, economic and academic attributes.   </td> ',
' <td> Of the 12,330 sessions in the dataset,\r\n84.5% (10,422) were negative cl
ass samples that did not\r\nend with shopping, and the rest (1908) were positive class\r\nsamples
ending with shopping. </td> ',
' <td> The handwritten dataset was collected from 170 participants with a total
of 5,180 numeral patterns. The dataset is named Prince Mohammad Bin Fahd University - Urdu/Arabic
Database (PMU-UD).   </td> ',
' <td> The data used in this study were gathered from 188 patients with PD (107
men and 81 women) with ages ranging from 33 to 87 (65.1±10.9). </td> ',
' <td> The local stability analysis of the 4-node star system (electricity produ
cer is in the center) implementing Decentral Smart Grid Control concept.   </td> ',
' <td> This dataset contains information about caesarian section results of 80
pregnant women with the most important characteristics of delivery problems in the medical
field. </td> ',
' <td> BAUM-1 dataset contains 1184 multimodal facial video clips collected from
31 subjects. The 1184 video clips contain spontaneous facial expressions and speech of 13
emotional and mental states. </td> ',
' <td> A multilingual audio-visual affective face database consisting of 1047 vi
deo clips of 286 subjects.   </td> ',
' <td> Exhaustive one year non-confidential data in the year 2015 to 2016 of fir
ms is collected from the Auditor Office of India to build a predictor for classifying suspicious f
irms. </td> ',
' <td> User interest information extracted from user reviews published in holida
yiq.com about various types of point of interests in South India </td> ',
' <td> The “real estate valuation” is a regression problem. The market
historical data set of real estate valuation are collected from Sindian Dist., New Taipei City, Ta
iwan.   </td> ',
' <td> . </td> ',
' <td> A data extract of a non-federal dataset posted here
https://catalog.data.gov/dataset/somerville-happiness-survey-responses-2011-2013-2015 </t
d> ',
' <td> Measurement of the S21,consists of 10 sweeps, each sweep contains 601 fre

```
quency points with spacing of 0.167MHz to cover a 100MHz band centered at 2.4GHz. </p></td> '
,
 ' <td><p class="normal">These are files of raw EMG data recorded by MYO Thalmic bracelet </p
></td> ',
 ' <td><p class="normal">Data collected from car parks in Birmingham that are operated by NCP from
\r\nBirmingham City Council. UK Open Government Licence
(OGL).\r\nhttps://data.birmingham.gov.uk/dataset/birmingham-parking </p></td> ',
 ' <td><p class="normal">The database was created with records of behavior of the urban traffic of
the city of Sao Paulo in Brazil. </p></td> ',
 ' <td><p class="normal">Reviews on destinations in 10 categories mentioned across East Asia. Each
traveler rating is mapped as Excellent(4), Very Good(3), Average(2), Poor(1), and Terrible(0) and
average rating is used. </p></td> ',
 ' <td><p class="normal">Google reviews on attractions from 24 categories across Europe are consid
ered. Google user rating ranges from 1 to 5 and average user rating per category is
calculated. </p></td> ',
 ' <td><p class="normal">There are three classes/diseases: Bacterial leaf blight, Brown spot, and
Leaf smut, each having 40 images. The format of all images is jpg.   </p></td> ']
```

In [466]:

```
len(descripcion_base)
```

Out[466]:

```
469
```

In [467]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

In [468]:

```
pd_links=pd.DataFrame(links1)
```

In [497]:

```
pd_links.columns=['Links']
```

In [498]:

```
pd_links
```

Out[498]:

	Links
0	http://archive.ics.uci.edu/ml/datasets/Abalone
1	http://archive.ics.uci.edu/ml/datasets/Adult
2	http://archive.ics.uci.edu/ml/datasets/Annealing
3	http://archive.ics.uci.edu/ml/datasets/Anonymo...
4	http://archive.ics.uci.edu/ml/datasets/Arrhythmia
5	http://archive.ics.uci.edu/ml/datasets/Artific...
6	http://archive.ics.uci.edu/ml/datasets/Audiolo...
7	http://archive.ics.uci.edu/ml/datasets/Audiolo...
8	http://archive.ics.uci.edu/ml/datasets/Auto+MPG
9	http://archive.ics.uci.edu/ml/datasets/Automobile
10	http://archive.ics.uci.edu/ml/datasets/Badges
11	http://archive.ics.uci.edu/ml/datasets/Balance...
12	http://archive.ics.uci.edu/ml/datasets/Balloons
13	http://archive.ics.uci.edu/ml/datasets/Breast+...
14	http://archive.ics.uci.edu/ml/datasets/Breast+...

	Links
15	http://archive.ics.uci.edu/ml/datasets/Breast+...
16	http://archive.ics.uci.edu/ml/datasets/Breast+...
17	http://archive.ics.uci.edu/ml/datasets/Pittsbu...
18	http://archive.ics.uci.edu/ml/datasets/Car+Eva...
19	http://archive.ics.uci.edu/ml/datasets/Census+...
20	http://archive.ics.uci.edu/ml/datasets/Chess+%...
21	http://archive.ics.uci.edu/ml/datasets/Chess+%...
22	http://archive.ics.uci.edu/ml/datasets/Chess+%...
23	http://archive.ics.uci.edu/ml/datasets/Chess+%...
24	http://archive.ics.uci.edu/ml/datasets/Bach+Ch...
25	http://archive.ics.uci.edu/ml/datasets/Connect-4
26	http://archive.ics.uci.edu/ml/datasets/Credit+...
27	http://archive.ics.uci.edu/ml/datasets/Japanes...
28	http://archive.ics.uci.edu/ml/datasets/Compute...
29	http://archive.ics.uci.edu/ml/datasets/Contrac...
...	...
439	http://archive.ics.uci.edu/ml/datasets/EEG+Ste...
440	http://archive.ics.uci.edu/ml/datasets/Roman+U...
441	http://archive.ics.uci.edu/ml/datasets/Avila
442	http://archive.ics.uci.edu/ml/datasets/PANDOR
443	http://archive.ics.uci.edu/ml/datasets/Drug+Re...
444	http://archive.ics.uci.edu/ml/datasets/Drug+Re...
445	http://archive.ics.uci.edu/ml/datasets/Physica...
446	http://archive.ics.uci.edu/ml/datasets/Superco...
447	http://archive.ics.uci.edu/ml/datasets/WESAD+%...
448	http://archive.ics.uci.edu/ml/datasets/GNFUV+U...
449	http://archive.ics.uci.edu/ml/datasets/Student...
450	http://archive.ics.uci.edu/ml/datasets/Online+...
451	http://archive.ics.uci.edu/ml/datasets/PMU-UD
452	http://archive.ics.uci.edu/ml/datasets/Parkins...
453	http://archive.ics.uci.edu/ml/datasets/Electri...
454	http://archive.ics.uci.edu/ml/datasets/Caesari...
455	http://archive.ics.uci.edu/ml/datasets/BAUM-1
456	http://archive.ics.uci.edu/ml/datasets/BAUM-2
457	http://archive.ics.uci.edu/ml/datasets/Audit+Data
458	http://archive.ics.uci.edu/ml/datasets/BuddyMo...
459	http://archive.ics.uci.edu/ml/datasets/Real+es...
460	http://archive.ics.uci.edu/ml/datasets/Early+b...
461	http://archive.ics.uci.edu/ml/datasets/Somervi...
462	http://archive.ics.uci.edu/ml/datasets/2.4+GHZ...
463	http://archive.ics.uci.edu/ml/datasets/EMG+dat...
464	http://archive.ics.uci.edu/ml/datasets/Parking...
465	http://archive.ics.uci.edu/ml/datasets/Behavio...
466	http://archive.ics.uci.edu/ml/datasets/Travel+...
467	http://archive.ics.uci.edu/ml/datasets/Tarvel+...
468	http://archive.ics.uci.edu/ml/datasets/Rice+Le...

469 rows × 1 columns

In [499]:

```
pd_nombre=pd.DataFrame(lista_links1)
```

In [500]:

```
pd_nombre.columns=['Nombre base de datos']
```

In [501]:

```
pd_nombre
```

Out[501]:

	Nombre base de datos
0	datasets/Abalone
1	datasets/Adult
2	datasets/Annealing
3	datasets/Anonymous+Microsoft+Web+Data
4	datasets/Arrhythmia
5	datasets/Artificial+Characters
6	datasets/Audiology+%28Original%29
7	datasets/Audiology+%28Standardized%29
8	datasets/Auto+MPG
9	datasets/Automobile
10	datasets/Badges
11	datasets/Balance+Scale
12	datasets/Balloons
13	datasets/Breast+Cancer
14	datasets/Breast+Cancer+Wisconsin+%28Original%29
15	datasets/Breast+Cancer+Wisconsin+%28Prognostic%29
16	datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29
17	datasets/Pittsburgh+Bridges
18	datasets/Car+Evaluation
19	datasets/Census+Income
20	datasets/Chess+%28King-Rook+vs.+King-Knight%29
21	datasets/Chess+%28King-Rook+vs.+King-Pawn%29
22	datasets/Chess+%28King-Rook+vs.+King%29
23	datasets/Chess+%28Domain+Theories%29
24	datasets/Bach+Chorales
25	datasets/Connect-4
26	datasets/Credit+Approval
27	datasets/Japanese+Credit+Screening
28	datasets/Computer+Hardware
29	datasets/Contraceptive+Method+Choice
...	...
439	datasets/EEG+Steady-State+Visual+Evoked+Potent...
440	datasets/Roman+Urdu+Data+Set
441	datasets/Avila
442	datasets/PANDOR
443	datasets/Drug+Review+Dataset+%28Druglib.com%29
444	datasets/Drug+Review+Dataset+%28Drugs.com%29
445	datasets/Physical+Unclonable+Functions
446	datasets/Superconductivty+Data
447	datasets/WESAD+%28Wearable+Stress+and+Affect+D...

	Nombre base de datos
448	datasets/GNFUV+Unmanned+Surface+Vehicles+Senso...
449	datasets/Student+Academics+Performance
450	datasets/Online+Shoppers+Purchasing+Intention+...
451	datasets/PMU-UD
452	datasets/Parkinson%27s+Disease+Classification
453	datasets/Electrical+Grid+Stability+Simulated+D...
454	datasets/Caesarian+Section+Classification+Dataset
455	datasets/BAUM-1
456	datasets/BAUM-2
457	datasets/Audit+Data
458	datasets/BuddyMove+Data+Set
459	datasets/Real+estate+valuation+data+set
460	datasets/Early+biomarkers+of+Parkinson%E2%80%9...
461	datasets/Somerville+Happiness+Survey
462	datasets/2.4+GHZ+Indoor+Channel+Measurements
463	datasets/EMG+data+for+gestures
464	datasets/Parking+Birmingham
465	datasets/Behavior+of+the+urban+traffic+of+the+...
466	datasets/Travel+Reviews
467	datasets/Tarvel+Review+Ratings
468	datasets/Rice+Leaf+Diseases

469 rows × 1 columns

In [502]:

```
pd_tipo=pd.DataFrame(tipo_datos)
```

In [503]:

```
pd_tipo.columns=['Tipo de Datos']
```

In [504]:

```
pd_tipo
```

Out[504]:

	Tipo de Datos
0	Multivariate
1	Multivariate
2	Multivariate
3	
4	Multivariate
5	Multivariate
6	Multivariate
7	Multivariate
8	Multivariate
9	Multivariate
10	Univariate, Text
11	Multivariate
12	Multivariate
13	Multivariate
14	Multivariate

	Tipo de Datos
15	Multivariate
16	Multivariate
17	Multivariate
18	Multivariate
19	Multivariate
20	Multivariate, Data-Generator...
21	Multivariate
22	Multivariate
23	Domain-Theory
24	Univariate, Time-Series
25	Multivariate, Spatial
26	Multivariate
27	Multivariate, Domain-Theory ...
28	Multivariate
29	Multivariate
...	...
440	Text
441	Multivariate
442	Multivariate
443	Multivariate, Text
444	Multivariate, Text
445	Multivariate
446	Multivariate
447	Multivariate, Time-Series
448	Multivariate, Sequential, Ti...
449	Multivariate
450	Multivariate
451	Univariate
452	Multivariate
453	Multivariate
454	Univariate
455	Time-Series
456	Time-Series
457	Multivariate
458	Multivariate, Text
459	Multivariate
460	Multivariate
461	
462	Multivariate
463	Time-Series
464	Multivariate, Univariate, Se...
465	Multivariate, Time-Series
466	Multivariate, Text
467	Multivariate, Text
468	Multivariate
469	In Collaboration With:

470 rows × 1 columns

In [505]:

```
pd_tipo_tarea=pd.DataFrame(tipo_tarea)
```

In [506]:

```
pd_tipo_tarea.columns=['Tipo de Tarea']
```

In [507]:

```
pd_tipo_tarea
```

Out[507]:

	Tipo de Tarea
0	Classification
1	Classification
2	Classification
3	Recommender-Systems
4	Classification
5	Classification
6	Classification
7	Classification
8	Regression
9	Regression
10	Classification
11	Classification
12	Classification
13	Classification
14	Classification
15	Classification, Regression
16	Classification
17	Classification
18	Classification
19	Classification
20	Classification
21	Classification
22	Classification
23	
24	
25	Classification
26	Classification
27	Classification
28	Regression
29	Classification
...	...
439	Classification, Regression
440	Classification
441	Classification
442	Recommendation
443	Classification, Regression, ...
444	Classification, Regression, ...
445	Classification
446	Regression

	Tipo de Tarea
446	Regression
447	Classification, Regression
448	Regression
449	Classification
450	Classification, Clustering
451	Classification
452	Classification
453	Classification, Regression
454	Classification
455	Classification
456	Classification
457	Classification
458	Classification, Clustering
459	Regression
460	Classification
461	Classification
462	Classification
463	Classification
464	Classification, Regression, ...
465	Classification, Regression
466	Classification, Clustering
467	Classification, Clustering
468	Classification

469 rows × 1 columns

In [508]:

```
pd_num_obs=pd.DataFrame(numero_obs)
```

In [509]:

```
pd_num_obs.columns=['Numero de Observaciones']
```

In [510]:

```
pd_num_obs
```

Out[510]:

	Numero de Observaciones
0	4177
1	48842
2	798
3	37711
4	452
5	6000
6	226
7	226
8	398
9	205
10	294
11	625

	Numero de Observaciones
11	625
12	16
13	286
14	699
15	198
16	569
17	108
18	1728
19	48842
20	
21	3196
22	28056
23	
24	100
25	67557
26	690
27	125
28	209
29	1473
...	...
439	9200
440	20000
441	20867
442	
443	4143
444	215063
445	6000000
446	21263
447	63000000
448	10190
449	300
450	12330
451	5180
452	756
453	10000
454	80
455	1184
456	1047
457	777
458	249
459	414
460	
461	143
462	7840
463	30000
464	35717
465	135
466	980
467	5456
468	120

469 rows × 1 columns

In [511]:

```
pd_num_var=pd.DataFrame(numero_var)
```

In [512]:

```
pd_num_var.columns=['Numero de Variables']
```

In [513]:

```
pd_num_var
```

Out[513]:

	Numero de Variables
0	8
1	14
2	38
3	294
4	279
5	7
6	
7	69
8	8
9	26
10	1
11	4
12	4
13	9
14	10
15	34
16	32
17	13
18	6
19	14
20	22
21	36
22	6
23	
24	6
25	42
26	15
27	
28	9
29	9
...	...
439	16
440	2
441	10
442	
443	8

	Numero de Variables
444	6
445	129
446	81
447	12
448	6
449	22
450	18
451	9
452	754
453	14
454	5
455	
456	
457	18
458	7
459	7
460	
461	7
462	5
463	6
464	4
465	18
466	11
467	25
468	

469 rows × 1 columns

In [514]:

```
pd_año = pd.DataFrame(año)
```

In [515]:

```
pd_año.columns=['Año']
```

In [516]:

```
pd_año
```

Out[516]:

	Año
0	1995
1	1996
2	
3	1998
4	1998
5	1992
6	1987

	Año
7	1992
8	1993
9	1987
10	1994
11	1994
12	
13	1988
14	1992
15	1995
16	1995
17	1990
18	1997
19	1996
20	1988
21	1989
22	1994
23	
24	
25	1995
26	
27	1992
28	1987
29	1997
...	...
439	2018
440	2018
441	2018
442	2018
443	2018
444	2018
445	2018
446	2018
447	2018
448	2018
449	2018

	Año
450	2018
451	2018
452	2018
453	2018
454	2018
455	2018
456	2018
457	2018
458	2018
459	2018
460	2018
461	2018
462	2018
463	2019
464	2019
465	2018
466	2018
467	2018
468	2019

469 rows × 1 columns

In [517]:

```
pd_descrip=pd.DataFrame(descripcion_base)
```

In [518]:

```
pd_descrip.columns=['Descripción base de datos']
```

In [519]:

```
pd_descrip
```

Out[519]:

	Descripción base de datos
0	<td> Predict the age of abal...
1	<td> Predict whether income ...
2	<td> Steel annealing data&nb...
3	<td> Log of anonymous users ...
4	<td> Distinguish between the...
5	<td> Dataset artificially ge...

	Descripción base de datos
6	<td> Nominal audiology data...
7	<td> Standardized version of...
8	<td> Revised from CMU StatLi...
9	<td> From 1985 Ward's Automo...
10	<td> Badges labeled with a "...
11	<td> Balance scale weight & ...
12	<td> Data previously used in...
13	<td> Breast Cancer Data (Res...
14	<td> Original Wisconsin Brea...
15	<td> Prognostic Wisconsin Br...
16	<td> Diagnostic Wisconsin Br...
17	<td> Bridges database that h...
18	<td> Derived from simple hie...
19	<td> Predict whether income ...
20	<td> Knight Pin Chess End-Ga...
21	<td> King+Rook versus King+P...
22	<td> Chess Endgame Database ...
23	<td> 6 different domain theo...
24	<td> Time-series data based ...
25	<td> Contains connect-4 posi...
26	<td> This data concerns cred...
27	<td> Includes domain theory ...
28	<td> Relative CPU Performanc...
29	<td> Dataset is a subset of ...
...	...
439	<td> This database consists ...
440	<td> Roman Urdu (the scripti...
441	<td> The Avila data set has ...
442	<td> PANDOR is a novel and p...
443	<td> The dataset provides pa...
444	<td> The dataset provides pa...
445	<td> The dataset is generate...
446	<td> Two file s contain data...
447	<td> WESAD (Wearable Stress ...
448	<td> The data-set contains e...
449	<td> The dataset tried to fi...
450	<td> Of the 12,330 sessions ...
451	<td> The handwritten dataset...
452	<td> The data used in this s...
453	<td> The local stability ana...
454	<td> This dataset contains i...
455	<td> BAUM-1 dataset contains...
456	<td> A multilingual audio-vi...
457	<td> Exhaustive one year non...
458	<td> User interest informati...
459	<td> The “real estate valuat...
460	<td> . </td>
461	<td> A data extract of a non...
462	<td> Measurement of the S21,...
463	<td> These are files of raw ...

	Descripción base de datos
464	<td> Data collected from car...
465	<td> The database was create...
466	<td> Reviews on destinations...
467	<td> Google reviews on attra...
468	<td> There are three classes...

469 rows × 1 columns

In [520]:

```
pd_tipo_var=pd.DataFrame(tipo_var)
```

In [521]:

```
pd_tipo_var.columns=['Tipo de Variables']
```

In [522]:

```
pd_tipo_var
```

Out[522]:

	Tipo de Variables
0	Categorical, Integer, Real
1	Categorical, Integer
2	Categorical, Integer, Real
3	Categorical
4	Categorical, Integer, Real
5	Categorical, Integer, Real
6	Categorical
7	Categorical
8	Categorical, Real
9	Categorical, Integer, Real
10	
11	Categorical
12	Categorical
13	Categorical
14	Integer
15	Real
16	Real
17	Categorical, Integer
18	Categorical
19	Categorical, Integer
20	Categorical, Integer
21	Categorical
22	Categorical, Integer
23	
24	Categorical, Integer
25	Categorical
26	Categorical, Integer, Real

	Tipo de Variables
27	Categorical, Real, Integer
28	Integer
29	Categorical, Integer
...	...
439	Integer
440	
441	Real
442	Categorical
443	Integer
444	Integer
445	Integer
446	Real
447	Real
448	Real
449	
450	Integer, Real
451	
452	Integer, Real
453	Real
454	Integer
455	
456	
457	Real
458	Real
459	Integer, Real
460	Real
461	Integer
462	Real
463	Real
464	Real
465	Integer, Real
466	Real
467	Real
468	Integer

469 rows × 1 columns

In [525]:

```
Final_pd = pd.concat([pd_links,pd_nombre,pd_tipo, pd_tipo_tarea, pd_tipo_var,pd_num_obs, pd_num_var
, pd_año,pd_descrip], axis=1)
```

In [526]:

```
Final_pd
```

Out[526]:

	Links	Nombre base de datos	Tipo de Datos
0	http://archive.ics.uci.edu/ml/datasets/Abalone	datasets/Abalone	Multivariate
1	http://archive.ics.uci.edu/ml/datasets/Adult	datasets/Adult	Multivariate

	Links	Nombre base de datos	Tipo de Datos
2	http://archive.ics.uci.edu/ml/datasets/Annealing	datasets/Annealing	Multivariate
3	http://archive.ics.uci.edu/ml/datasets/Anonymo...	datasets/Anonymous+Microsoft+Web+Data	
4	http://archive.ics.uci.edu/ml/datasets/Arrhythmia	datasets/Arrhythmia	Multivariate
5	http://archive.ics.uci.edu/ml/datasets/Artific...	datasets/Artificial+Characters	Multivariate
6	http://archive.ics.uci.edu/ml/datasets/Audiolo...	datasets/Audiology+%28Original%29	Multivariate
7	http://archive.ics.uci.edu/ml/datasets/Audiolo...	datasets/Audiology+%28Standardized%29	Multivariate
8	http://archive.ics.uci.edu/ml/datasets/Auto+MPG	datasets/Auto+MPG	Multivariate
9	http://archive.ics.uci.edu/ml/datasets/Automobile	datasets/Automobile	Multivariate
10	http://archive.ics.uci.edu/ml/datasets/Badges	datasets/Badges	Univariate, Text
11	http://archive.ics.uci.edu/ml/datasets/Balance...	datasets/Balance+Scale	Multivariate
12	http://archive.ics.uci.edu/ml/datasets/Balloons	datasets/Balloons	Multivariate
13	http://archive.ics.uci.edu/ml/datasets/Breast+...	datasets/Breast+Cancer	Multivariate
14	http://archive.ics.uci.edu/ml/datasets/Breast+...	datasets/Breast+Cancer+Wisconsin+%28Original%29	Multivariate
15	http://archive.ics.uci.edu/ml/datasets/Breast+...	datasets/Breast+Cancer+Wisconsin+%28Prognostic%29	Multivariate
16	http://archive.ics.uci.edu/ml/datasets/Breast+...	datasets/Breast+Cancer+Wisconsin+%28Diagnostic%29	Multivariate
17	http://archive.ics.uci.edu/ml/datasets/Pittsbu...	datasets/Pittsburgh+Bridges	Multivariate
18	http://archive.ics.uci.edu/ml/datasets/Car+Eva...	datasets/Car+Evaluation	Multivariate
19	http://archive.ics.uci.edu/ml/datasets/Census+...	datasets/Census+Income	Multivariate
20	http://archive.ics.uci.edu/ml/datasets/Chess+%...	datasets/Chess+%28King-Rook+vs.+King-Knight%29	Multivariate, Data-Generator...
21	http://archive.ics.uci.edu/ml/datasets/Chess+%...	datasets/Chess+%28King-Rook+vs.+King-Pawn%29	Multivariate
22	http://archive.ics.uci.edu/ml/datasets/Chess+%...	datasets/Chess+%28King-Rook+vs.+King%29	Multivariate
23	http://archive.ics.uci.edu/ml/datasets/Chess+%...	datasets/Chess+%28Domain+Theories%29	Domain-Theory
24	http://archive.ics.uci.edu/ml/datasets/Bach+Ch...	datasets/Bach+Chorales	Univariate, Time-Series
25	http://archive.ics.uci.edu/ml/datasets/Connect-4	datasets/Connect-4	Multivariate, Spatial
26	http://archive.ics.uci.edu/ml/datasets/Credit+...	datasets/Credit+Approval	Multivariate
27	http://archive.ics.uci.edu/ml/datasets/Japanes...	datasets/Japanese+Credit+Screening	Multivariate, Domain-Theory ...

	Links	Nombre base de datos	Tipo de Datos
28	http://archive.ics.uci.edu/ml/datasets/Compute...	datasets/Computer+Hardware	Multivariate
29	http://archive.ics.uci.edu/ml/datasets/Contrac...	datasets/Contraceptive+Method+Choice	Multivariate
...	...	...	...
440	http://archive.ics.uci.edu/ml/datasets/Roman+U...	datasets/Roman+Urdu+Data+Set	Text
441	http://archive.ics.uci.edu/ml/datasets/Avila	datasets/Avila	Multivariate
442	http://archive.ics.uci.edu/ml/datasets/PANDOR	datasets/PANDOR	Multivariate
443	http://archive.ics.uci.edu/ml/datasets/Drug+Re...	datasets/Drug+Review+Dataset+%28Druglib.com%29	Multivariate, Text
444	http://archive.ics.uci.edu/ml/datasets/Drug+Re...	datasets/Drug+Review+Dataset+%28Drugs.com%29	Multivariate, Text
445	http://archive.ics.uci.edu/ml/datasets/Physica...	datasets/Physical+Unclonable+Functions	Multivariate
446	http://archive.ics.uci.edu/ml/datasets/Superco...	datasets/Superconductivty+Data	Multivariate
447	http://archive.ics.uci.edu/ml/datasets/WESAD+%...	datasets/WESAD+%28Wearable+Stress+and+Affect+D...	Multivariate, Time-Series
448	http://archive.ics.uci.edu/ml/datasets/GNFUV+U...	datasets/GNFUV+Unmanned+Surface+Vehicles+Senso...	Multivariate, Sequential, Ti...
449	http://archive.ics.uci.edu/ml/datasets/Student...	datasets/Student+Academics+Performance	Multivariate
450	http://archive.ics.uci.edu/ml/datasets/Online+...	datasets/Online+Shoppers+Purchasing+Intention+...	Multivariate
451	http://archive.ics.uci.edu/ml/datasets/PMU-UD	datasets/PMU-UD	Univariate
452	http://archive.ics.uci.edu/ml/datasets/Parkins...	datasets/Parkinson%27s+Disease+Classification	Multivariate
453	http://archive.ics.uci.edu/ml/datasets/Electri...	datasets/Electrical+Grid+Stability+Simulated+D...	Multivariate
454	http://archive.ics.uci.edu/ml/datasets/Caesari...	datasets/Caesarian+Section+Classification+Dataset	Univariate
455	http://archive.ics.uci.edu/ml/datasets/BAUM-1	datasets/BAUM-1	Time-Series
456	http://archive.ics.uci.edu/ml/datasets/BAUM-2	datasets/BAUM-2	Time-Series
457	http://archive.ics.uci.edu/ml/datasets/Audit+Data	datasets/Audit+Data	Multivariate
458	http://archive.ics.uci.edu/ml/datasets/BuddyMo...	datasets/BuddyMove+Data+Set	Multivariate, Text
459	http://archive.ics.uci.edu/ml/datasets/Real+es...	datasets/Real+estate+valuation+data+set	Multivariate
460	http://archive.ics.uci.edu/ml/datasets/Early+b...	datasets/Early+biomarkers+of+Parkinson%E2%80%9...	Multivariate
461	http://archive.ics.uci.edu/ml/datasets/Somervi...	datasets/Somerville+Happiness+Survey	
462	http://archive.ics.uci.edu/ml/datasets/2.4+GHZ...	datasets/2.4+GHZ+Indoor+Channel+Measurements	Multivariate
463	http://archive.ics.uci.edu/ml/datasets/EMG+dat...	datasets/EMG+data+for+gestures	Time-Series
464	http://archive.ics.uci.edu/ml/datasets/Parking...	datasets/Parking+Birmingham	Multivariate, Univariate, Se...
465	http://archive.ics.uci.edu/ml/datasets/Behavio...	datasets/Behavior+of+the+urban+traffic+of+the+...	Multivariate, Time-Series

	Links	Nombre base de datos	Tipo de Datos
466	http://archive.ics.uci.edu/ml/datasets/Travel+...	datasets/Travel+Reviews	Multivariate, Text
467	http://archive.ics.uci.edu/ml/datasets/Tarvel+...	datasets/Tarvel+Review+Ratings	Multivariate, Text
468	http://archive.ics.uci.edu/ml/datasets/Rice+Le...	datasets/Rice+Leaf+Diseases	Multivariate
469	NaN	NaN	In Collaboration With:

470 rows × 9 columns

Diviértase.