[Sports articles for objectivity analysis](https://archive.ics.uci.edu/dataset/450/sports+articles+for+objectivity+analysis)

[**Sports articles for objectivity analysis**](https://archive.ics.uci.edu/dataset/450/sports+articles+for+objectivity+analysis)

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.

Classification

Multivariate, Text

1K Instances

59 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Social Science | Integer | 4/8/2018 | 6.39K |

[South German Credit (UPDATE)](https://archive.ics.uci.edu/dataset/573/south+german+credit+update)

[**South German Credit (UPDATE)**](https://archive.ics.uci.edu/dataset/573/south+german+credit+update)

700 good and 300 bad credits with 20 predictor variables. Data from 1973 to 1975. Stratified sample from actual credits with bad credits heavily oversampled. A cost matrix can be used.

Classification, Regression, Clustering

Multivariate

1K Instances

21 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Business | Integer, Real | 6/19/2020 | 14.39K |

[South German Credit](https://archive.ics.uci.edu/dataset/522/south+german+credit)

[**South German Credit**](https://archive.ics.uci.edu/dataset/522/south+german+credit)

700 good and 300 bad credits with 20 predictor variables. Data from 1973 to 1975. Stratified sample from actual credits with bad credits heavily oversampled. A cost matrix can be used.

Classification, Regression, Clustering

Multivariate

1K Instances

21 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Business | Integer, Real | 11/28/2019 | 10.94K |

[Statlog (German Credit Data)](https://archive.ics.uci.edu/dataset/144/statlog+german+credit+data)

[**Statlog (German Credit Data)**](https://archive.ics.uci.edu/dataset/144/statlog+german+credit+data)

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

Classification

Multivariate

1K Instances

20 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Social Science | Categorical, Integer | 11/16/1994 | 119.35K |

[QSAR biodegradation](https://archive.ics.uci.edu/dataset/254/qsar+biodegradation)

[**QSAR biodegradation**](https://archive.ics.uci.edu/dataset/254/qsar+biodegradation)

Data set containing values for 41 attributes (molecular descriptors) used to classify 1055 chemicals into 2 classes (ready and not ready biodegradable).

Classification

Multivariate

1.06K Instances

41 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Other | Integer, Real | 6/20/2013 | 5.24K |

[Geographical Original of Music](https://archive.ics.uci.edu/dataset/315/geographical+original+of+music)

[**Geographical Original of Music**](https://archive.ics.uci.edu/dataset/315/geographical+original+of+music)

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.

Classification, Regression

Multivariate

1.06K Instances

68 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Other | Real | 10/17/2014 | 6.07K |

[Burst Header Packet (BHP) flooding attack on Optical Burst Switching (OBS) Network](https://archive.ics.uci.edu/dataset/404/burst+header+packet+bhp+flooding+attack+on+optical+burst+switching+obs+network)

[**Burst Header Packet (BHP) flooding attack on Optical Burst Switching (OBS) Network**](https://archive.ics.uci.edu/dataset/404/burst+header+packet+bhp+flooding+attack+on+optical+burst+switching+obs+network)

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.

Classification

Text

1.08K Instances

22 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Computer Science | Integer | 8/27/2017 | 1.8K |

[Mice Protein Expression](https://archive.ics.uci.edu/dataset/342/mice+protein+expression)

[**Mice Protein Expression**](https://archive.ics.uci.edu/dataset/342/mice+protein+expression)

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 to assess associative learning.

Classification, Clustering

Multivariate

1.08K Instances

82 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Biology | Real | 8/3/2015 | 9.32K |

[Diabetic Retinopathy Debrecen](https://archive.ics.uci.edu/dataset/329/diabetic+retinopathy+debrecen)

[**Diabetic Retinopathy Debrecen**](https://archive.ics.uci.edu/dataset/329/diabetic+retinopathy+debrecen)

This dataset contains features extracted from the Messidor image set to predict whether an image contains signs of diabetic retinopathy or not.

Classification

Multivariate

1.15K Instances

19 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Health and Medicine | Integer, Real | 11/2/2014 | 8.75K |

[Productivity Prediction of Garment Employees](https://archive.ics.uci.edu/dataset/597/productivity+prediction+of+garment+employees)

[**Productivity Prediction of Garment Employees**](https://archive.ics.uci.edu/dataset/597/productivity+prediction+of+garment+employees)

This dataset includes important attributes of the garment manufacturing process and the productivity of the employees which had been collected manually and also been validated by the industry experts.

Classification, Regression

Multivariate, Time-Series

1.2K Instances

15 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Business | Integer, Real | 8/2/2020 | 25.97K |

[Website Phishing](https://archive.ics.uci.edu/dataset/379/website+phishing)

[**Website Phishing**](https://archive.ics.uci.edu/dataset/379/website+phishing)

Classification

Multivariate

1.35K Instances

10 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Computer Science | Integer | 11/1/2016 | 13.64K |

[Hepatitis C Virus (HCV) for Egyptian patients](https://archive.ics.uci.edu/dataset/503/hepatitis+c+virus+hcv+for+egyptian+patients)

[**Hepatitis C Virus (HCV) for Egyptian patients**](https://archive.ics.uci.edu/dataset/503/hepatitis+c+virus+hcv+for+egyptian+patients)

Egyptian patients who underwent treatment dosages for HCV about 18 months. Discretization should be applied based on expert recommendations; there is an attached file shows how.

Classification

Multivariate

1.39K Instances

28 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Health and Medicine | Real | 9/29/2019 | 9.95K |

[Hybrid Indoor Positioning Dataset from WiFi RSSI, Bluetooth and magnetometer](https://archive.ics.uci.edu/dataset/402/hybrid+indoor+positioning+dataset+from+wifi+rssi+bluetooth+and+magnetometer)

[**Hybrid Indoor Positioning Dataset from WiFi RSSI, Bluetooth and magnetometer**](https://archive.ics.uci.edu/dataset/402/hybrid+indoor+positioning+dataset+from+wifi+rssi+bluetooth+and+magnetometer)

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.

Classification

Multivariate, Sequential, Time-Series

1.54K Instances

65 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Computer Science | Real | 12/17/2016 | 2.4K |

[Miskolc IIS Hybrid IPS](https://archive.ics.uci.edu/dataset/375/miskolc+iis+hybrid+ips)

[**Miskolc IIS Hybrid IPS**](https://archive.ics.uci.edu/dataset/375/miskolc+iis+hybrid+ips)

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.

Classification, Clustering, Causal-Discovery

Text

1.54K Instances

67 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Computer Science | Integer | 7/3/2016 | 1.7K |

[One-hundred plant species leaves data set](https://archive.ics.uci.edu/dataset/241/one+hundred+plant+species+leaves+data+set)

[**One-hundred plant species leaves data set**](https://archive.ics.uci.edu/dataset/241/one+hundred+plant+species+leaves+data+set)

Sixteen samples of leaf each of one-hundred plant species. For each sample, a shape descriptor, fine scale margin and texture histogram are given.

Classification

null

1.6K Instances

64 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Biology | Real | 12/2/2012 | 3.96K |

[Drug consumption (quantified)](https://archive.ics.uci.edu/dataset/373/drug+consumption+quantified)

[**Drug consumption (quantified)**](https://archive.ics.uci.edu/dataset/373/drug+consumption+quantified)

Classify type of drug consumer by personality data

Classification

Multivariate

1.89K Instances

32 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Social Science | Real | 10/16/2016 | 22.04K |

[Steel Plates Faults](https://archive.ics.uci.edu/dataset/198/steel+plates+faults)

[**Steel Plates Faults**](https://archive.ics.uci.edu/dataset/198/steel+plates+faults)

A dataset of steel platesâ€™ faults, classified into 7 different types. The goal was to train machine learning for automatic pattern recognition.

Classification

Multivariate

1.94K Instances

27 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Physics and Chemistry | Integer, Real | 10/25/2010 | 9.23K |

[CNNpred: CNN-based stock market prediction using a diverse set of variables](https://archive.ics.uci.edu/dataset/554/cnnpred+cnn+based+stock+market+prediction+using+a+diverse+set+of+variables)

[**CNNpred: CNN-based stock market prediction using a diverse set of variables**](https://archive.ics.uci.edu/dataset/554/cnnpred+cnn+based+stock+market+prediction+using+a+diverse+set+of+variables)

This dataset contains several daily features of S&P 500, NASDAQ Composite, Dow Jones Industrial Average, RUSSELL 2000, and NYSE Composite from 2010 to 2017.

Classification, Regression

Sequential, Time-Series

1.99K Instances

84 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Computer Science | Real | 12/25/2019 | 6.68K |

[Estimation of Obesity Levels Based On Eating Habits and Physical Condition ](https://archive.ics.uci.edu/dataset/544/estimation+of+obesity+levels+based+on+eating+habits+and+physical+condition)

[**Estimation of Obesity Levels Based On Eating Habits and Physical Condition**](https://archive.ics.uci.edu/dataset/544/estimation+of+obesity+levels+based+on+eating+habits+and+physical+condition)

This dataset include data for the estimation of obesity levels in individuals from the countries of Mexico, Peru and Colombia, based on their eating habits and physical condition.

Classification, Regression, Clustering

Multivariate

2.11K Instances

16 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Health and Medicine | Integer | 8/26/2019 | 57.38K |

[AIDS Clinical Trials Group Study 175](https://archive.ics.uci.edu/dataset/890/aids+clinical+trials+group+study+175)

[**AIDS Clinical Trials Group Study 175**](https://archive.ics.uci.edu/dataset/890/aids+clinical+trials+group+study+175)

The AIDS Clinical Trials Group Study 175 Dataset contains healthcare statistics and categorical information about patients who have been diagnosed with AIDS. This dataset was initially published in 1996. The prediction task is to predict whether or not each patient died within a certain window of time or not.

Classification, Regression

Tabular, Multivariate

2.14K Instances

23 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Health and Medicine | Categorical, Integer | 9/25/2023 | 15.15K |

[Image Segmentation](https://archive.ics.uci.edu/dataset/50/image+segmentation)

[**Image Segmentation**](https://archive.ics.uci.edu/dataset/50/image+segmentation)

Image data described by high-level numeric-valued attributes, 7 classes

Classification

Multivariate

2.31K Instances

19 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Other | Real | 10/31/1990 | 41.89K |

[Statlog (Image Segmentation)](https://archive.ics.uci.edu/dataset/147/statlog+image+segmentation)

[**Statlog (Image Segmentation)**](https://archive.ics.uci.edu/dataset/147/statlog+image+segmentation)

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.

Classification

Multivariate

2.31K Instances

19 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Other | Real | 10/31/1990 | 13.04K |

[Australian Sign Language signs (High Quality)](https://archive.ics.uci.edu/dataset/115/australian+sign+language+signs+high+quality)

[**Australian Sign Language signs (High Quality)**](https://archive.ics.uci.edu/dataset/115/australian+sign+language+signs+high+quality)

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

Classification

Multivariate, Time-Series

2.57K Instances

22 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Other | Real | 2/25/2002 | 12.11K |

[seismic-bumps](https://archive.ics.uci.edu/dataset/266/seismic+bumps)

[**seismic-bumps**](https://archive.ics.uci.edu/dataset/266/seismic+bumps)

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.

Classification

Multivariate

2.58K Instances

19 Features

| **Subject Area** | **Feature Type** | **Date Donated** | **Views** |
| --- | --- | --- | --- |
| Other | Real | 4/2/2013 | 4.31K |

[MEU-Mobile KSD](https://archive.ics.uci.edu/dataset/399/meu+mobile+ksd)

[**MEU-Mobile KSD**](https://archive.ics.uci.edu/dataset/399/meu+mobile+ksd)

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.

Classification

Multivariate

2.86K