10 WriteIn 'Hi AI';
20 Goto 10;

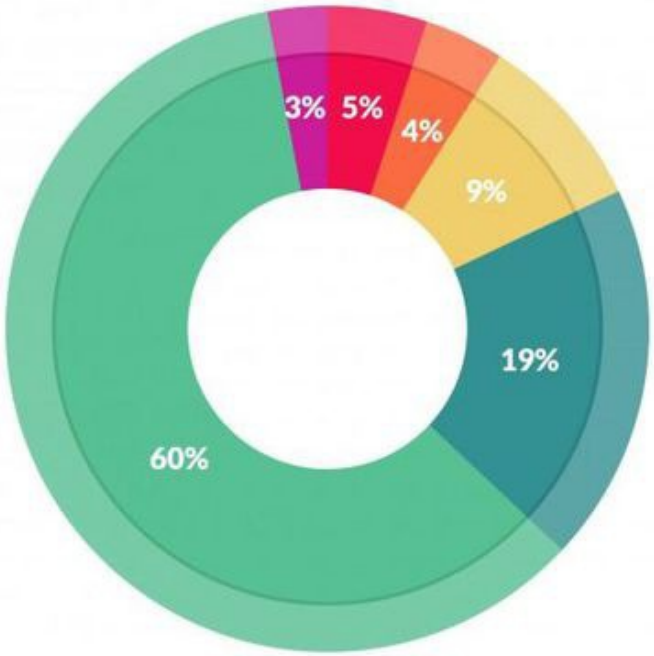
ML VSCode Data **Vision**

BASTA!

maXbox

Agenda BASTA 23/09/2020

- Data Preprocessing & Understanding
- Matplotlib Basics in Jupyter
- GraphViz (dot) language support, a C# wrapper for GraphViz graph generator for dotnet core.
- Integrate (VS Code/Jupyter Notebook/dotnet try)
- Installation, Optimisation & Sources

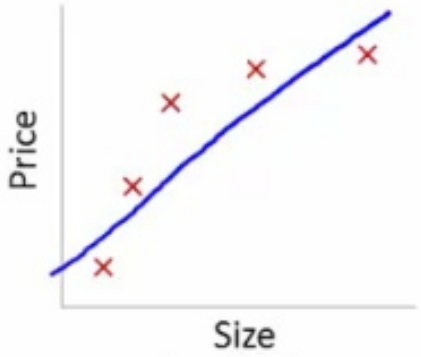


What data scientists spend the most time doing

- *Building training sets: 3%*
- *Cleaning and organizing data: 60%*
- *Collecting data sets; 19%*
- *Mining data for patterns: 9%*
- *Refining algorithms: 4%*
- *Other: 5%*

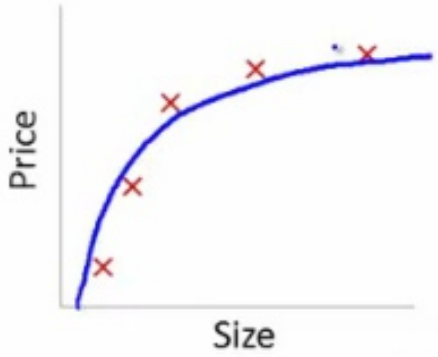
Now Demo: TensorFlowTest64.dproj

Machine Learning Fit



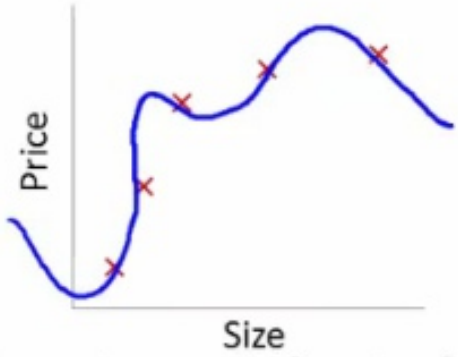
$\theta_0 + \theta_1 x$

High bias
(underfit)



$\theta_0 + \theta_1 x + \theta_2 x^2$

"Just right"

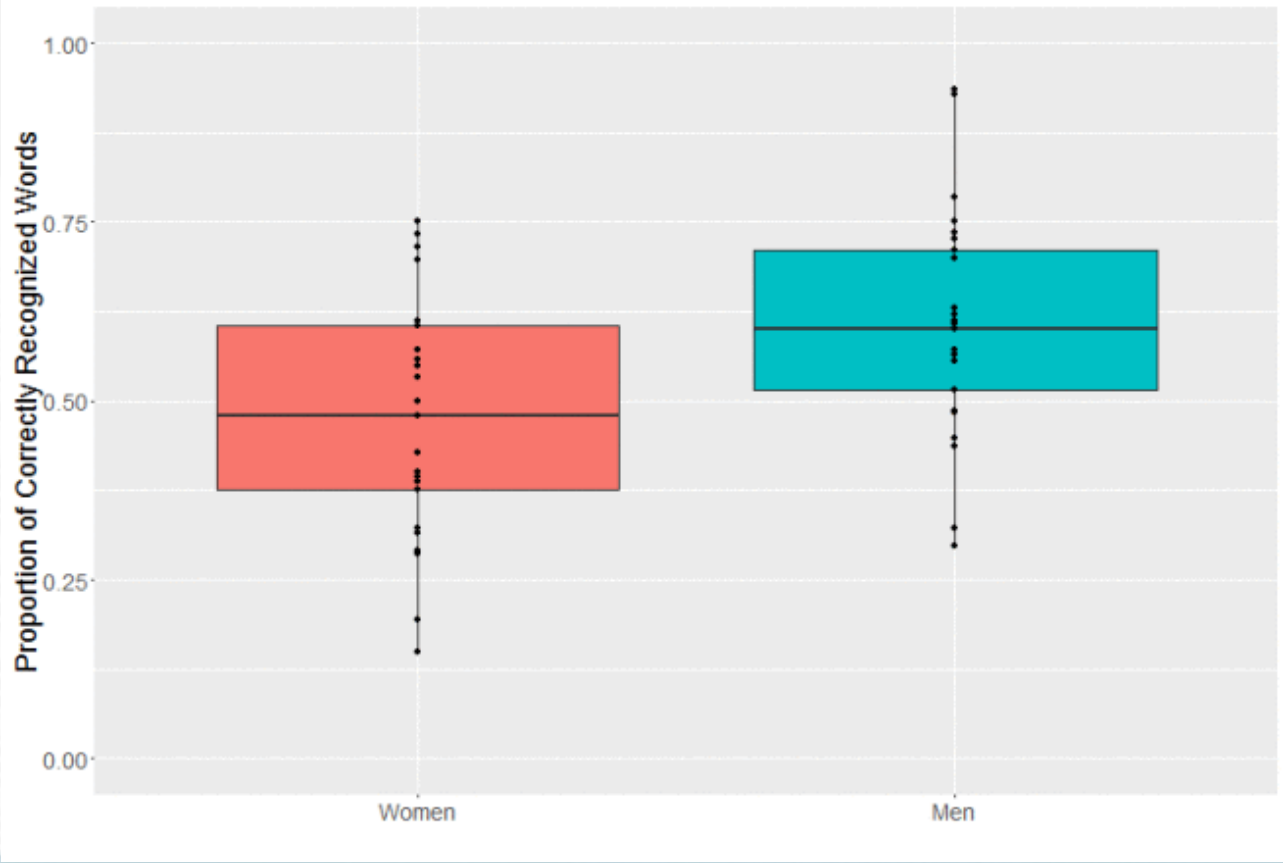


$\theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$

High variance
(overfit)

https://makingnoiseandhearingthings.com/2016/07/12/googles-speech-recognition-has-a-gender-bias/

Speech Recognition Boxplot

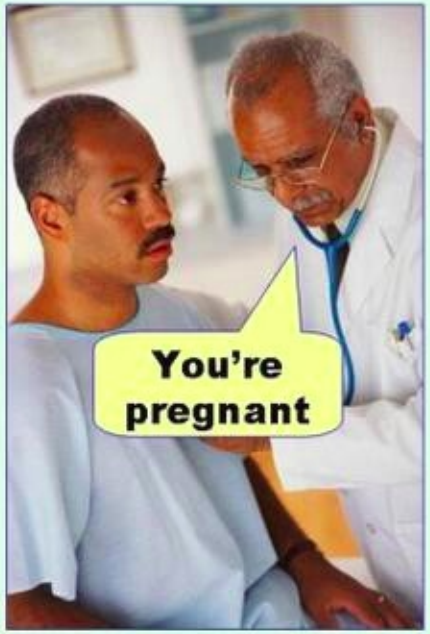


Confusion Matrix Data Truth

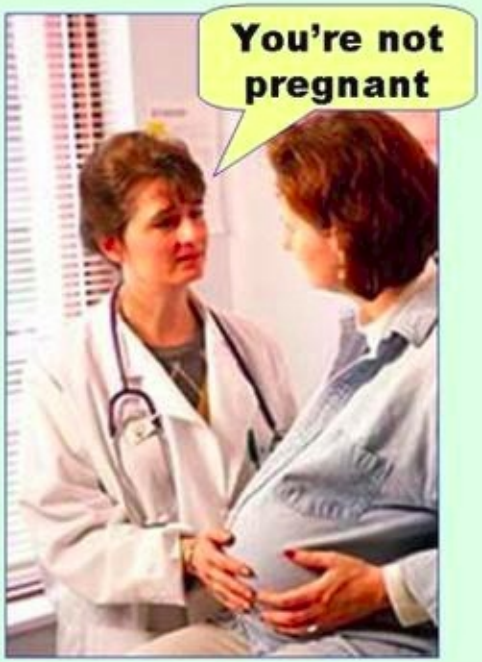
n=165	**Predicted: NO**	**Predicted: YES**	
Actual: NO	TN = 50	FP = 10	60
Actual: YES	FN = 5	TP = 100	105
	55	110	

Type I & II error

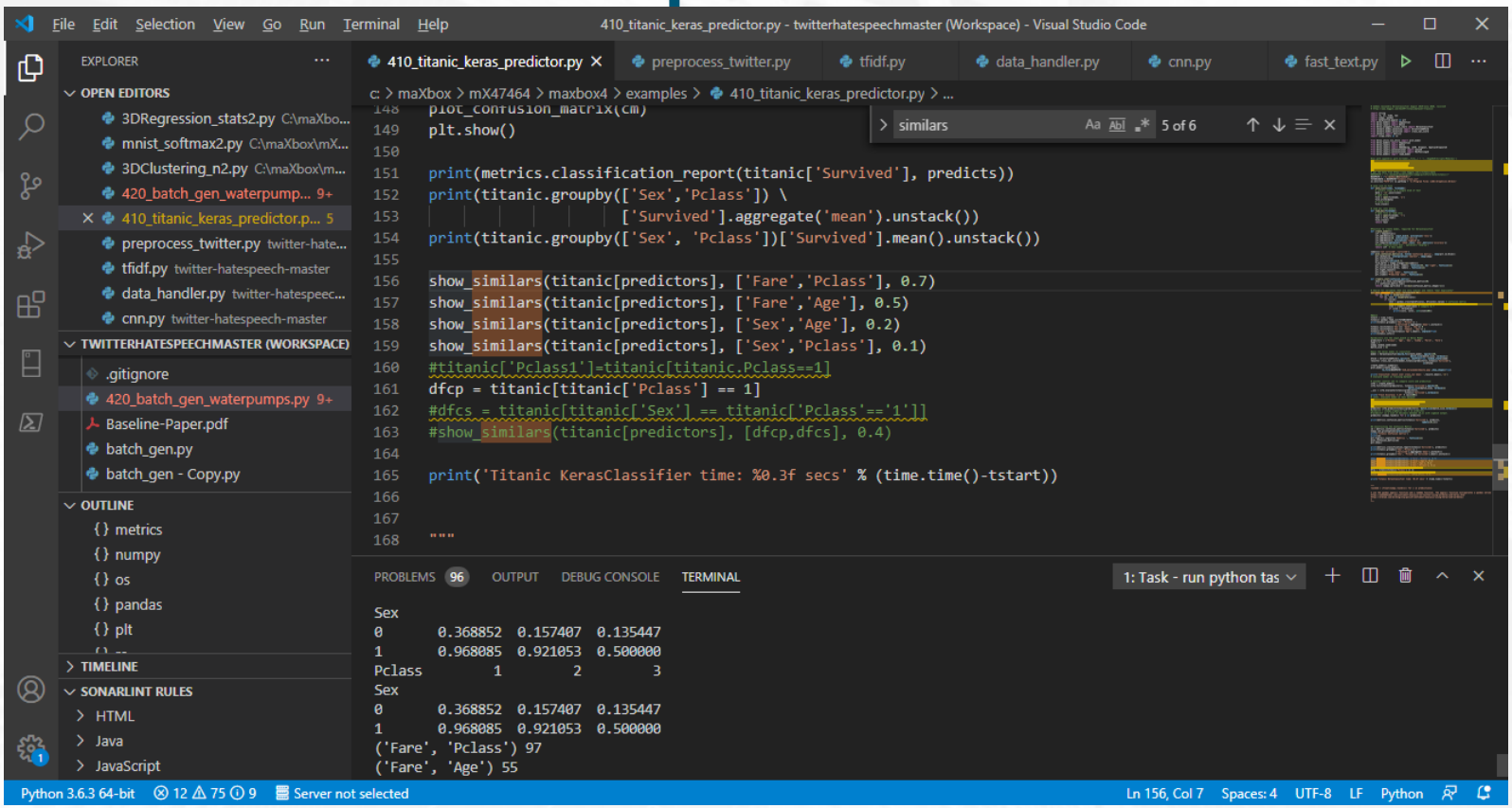
Type I error
(false positive)



Type II error
(false negative)



Demo matplotlib Plots



https://www.tensorflow.org/

Jupyter matplot notebook Demo

- Barchart (Histograms, Maps, Plot)
- Confusion-, Correlation Matrix
- PieChart, Density Plot, Heatmaps

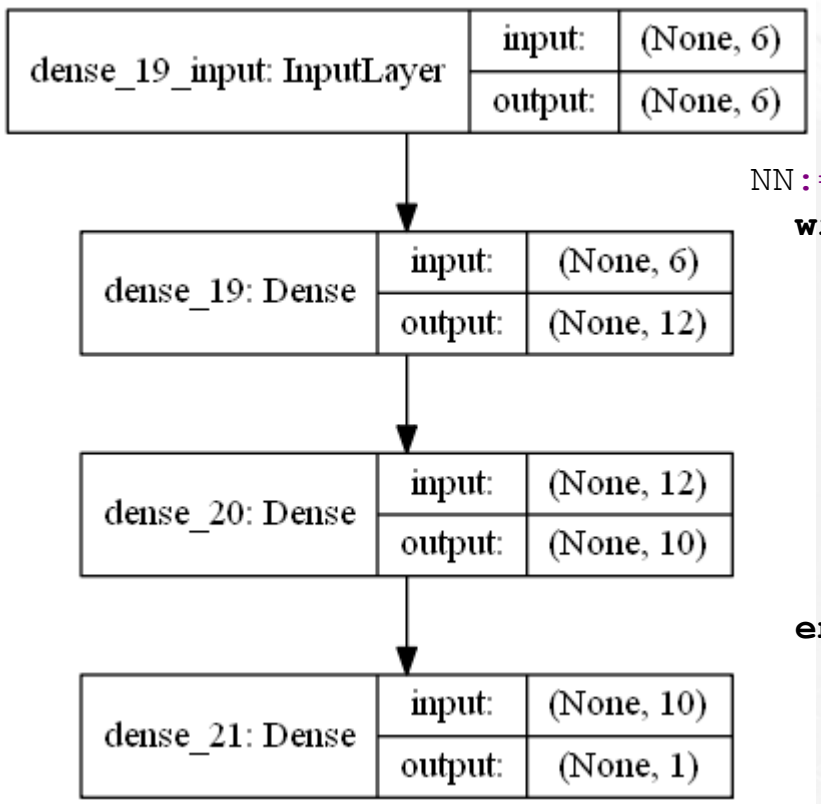
https://colab.research.google.com/drive/1pQPsMzru1eXYb7Q9WV5V3dqnki_KBJ2n

https://colab.research.google.com/github/maxkleiner/maXbox4/blob/master/BASTA_2020_matplotlib_presentation2.ipynb

https://github.com/maxkleiner/maXbox/blob/master/waterpumpspredictor2.ipynb

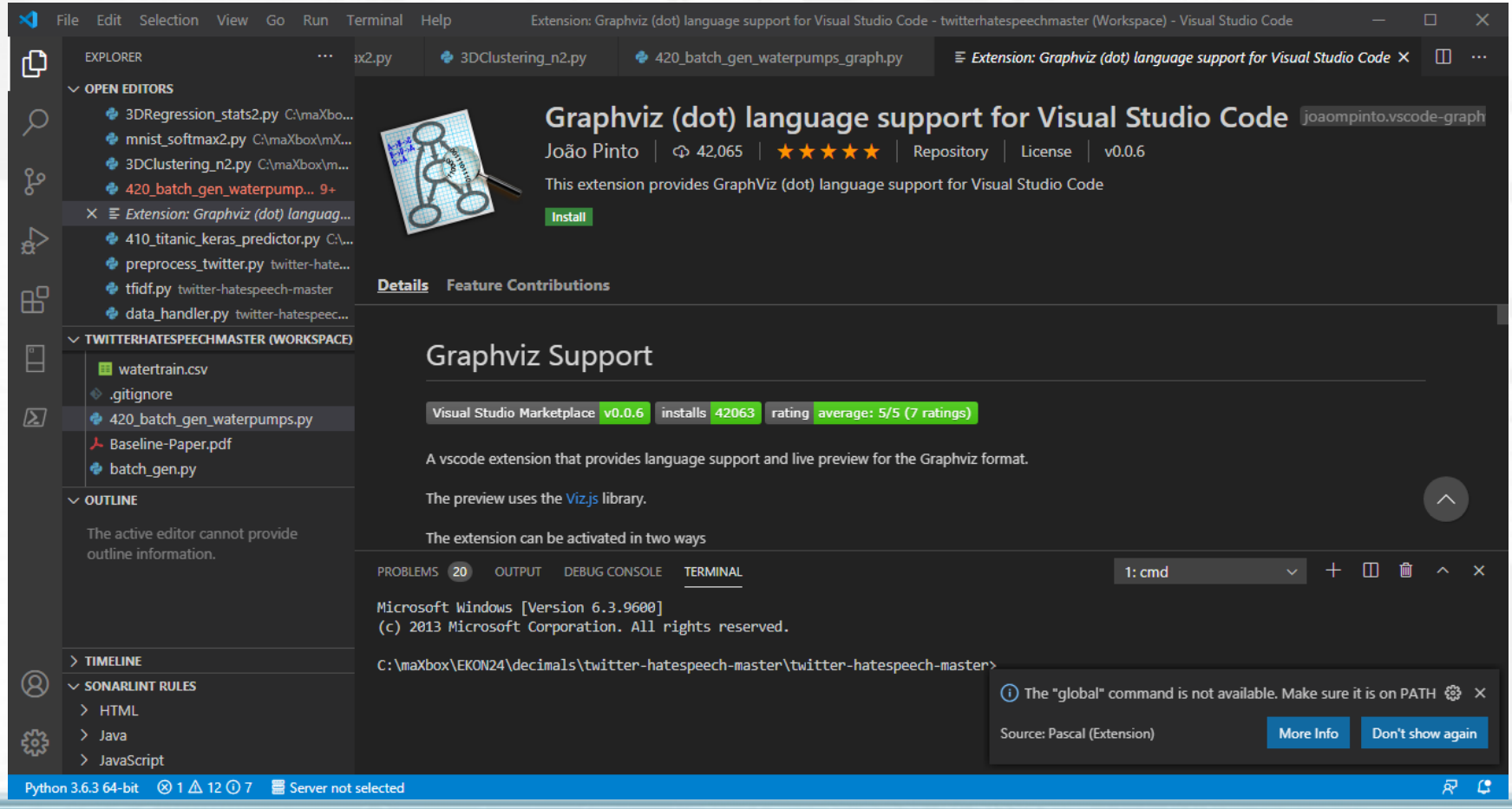
GraphViz-C-Sharp-Wrapper Demo:
https://sourceforge.net/projects/maxbox/files/Examples/EKON/BASTA2020/visout.zip/download
https://github.com/helgeu/GraphViz-C-Sharp-Wrapper

Visualise Model with Graph




```
NN:= TFannNetwork.create(self)
  with NN do begin
    Layers.add('2')
    Layers.add('3')
    Layers.add('1')
    LearningRate:= 0.6999999880790711OO
    ConnectionRate:= 1.000
    TrainingAlgorithm:= taFANN_TRAIN_RPROP
    ActivationFunctionHidden:= afFANN_SIGMOID
    ActivationFunctionOutput:= afFANN_SIGMOID
  end;
```

Graph Viz


File Edit Selection View Go Run Terminal Help
Extension: Graphviz (dot) language support for Visual Studio Code - twitterhatespeechmaster (Workspace) - Visual Studio Code
EXPLORER
OPEN EDITORS
3DRegression_stats2.py C:\maXbo...
mnist_softmax2.py C:\maXbox\mX...
3DClustering_n2.py C:\maXbox\m...
420_batch_gen_waterpump... 9+
Extension: Graphviz (dot) languag...
410_titanic_keras_predictor.py C:\...
preprocess_twitter.py twitter-hate...
tfidf.py twitter-hatespeech-master
data_handler.py twitter-hatespeec...
TWITTERHATESPEECHMASTER (WORKSPACE)
watertrain.csv
.gitignore
420_batch_gen_waterpumps.py
Baseline-Paper.pdf
batch_gen.py
OUTLINE
The active editor cannot provide outline information.
TIMELINE
SONARLINT RULES
HTML
Java
JavaScript
3DClustering_n2.py
420_batch_gen_waterpumps_graph.py
Extension: Graphviz (dot) language support for Visual Studio Code
Graphviz (dot) language support for Visual Studio Code
joaompinto.vscode-graph
João Pinto | 42,065 | ★★★★★ | Repository | License | v0.0.6
This extension provides GraphViz (dot) language support for Visual Studio Code
Install
Details
Feature Contributions
Graphviz Support
Visual Studio Marketplace v0.0.6
installs 42063
rating average: 5/5 (7 ratings)
A vscode extension that provides language support and live preview for the Graphviz format.
The preview uses the Viz.js library.
The extension can be activated in two ways
PROBLEMS 20
OUTPUT
DEBUG CONSOLE
TERMINAL
1: cmd
Microsoft Windows [Version 6.3.9600]
(c) 2013 Microsoft Corporation. All rights reserved.
C:\maXbox\EKON24\decimals\twitter-hatespeech-master\twitter-hatespeech-master>
The "global" command is not available. Make sure it is on PATH
Source: Pascal (Extension)
More Info
Don't show again
Python 3.6.3 64-bit
1 12 7
Server not selected

Graphviz

Graphviz is open source graph visualization software. Graph visualization is a way of representing structural information as diagrams of abstract graphs and networks.

It has important applications in networking, bioinformatics, software engineering, database and web design, machine learning, and in visual interfaces for other technical domains. Open In Colab

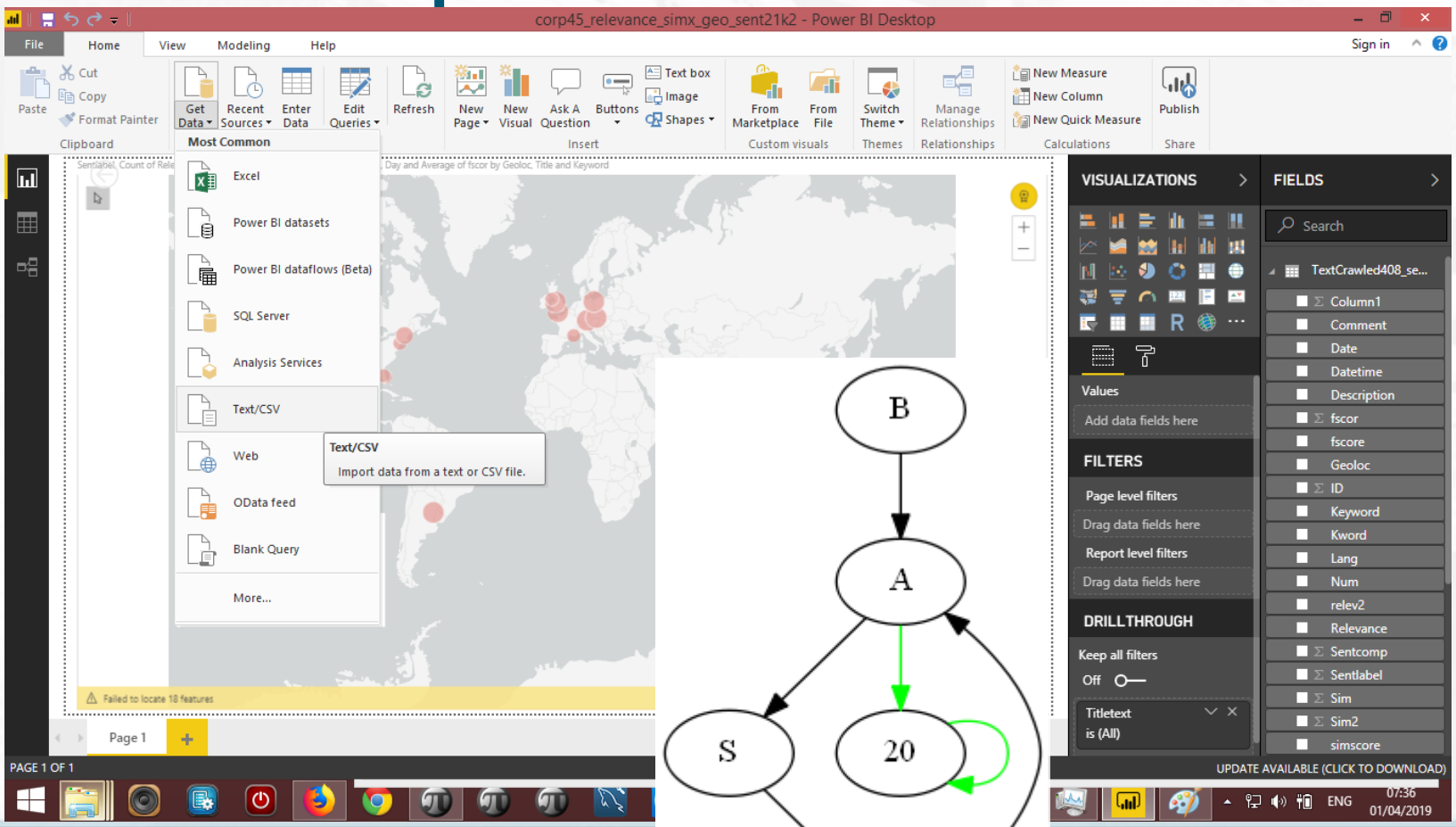
https://colab.research.google.com/github/maxkleiner/maXbox/blob/master/Copy_of_simple_image_classification_with_any_dataset.ipynb

https://colab.research.google.com/github/maxkleiner/maXbox/blob/master/Copy_of_simple_image_classification_with_any_dataset.ipynb

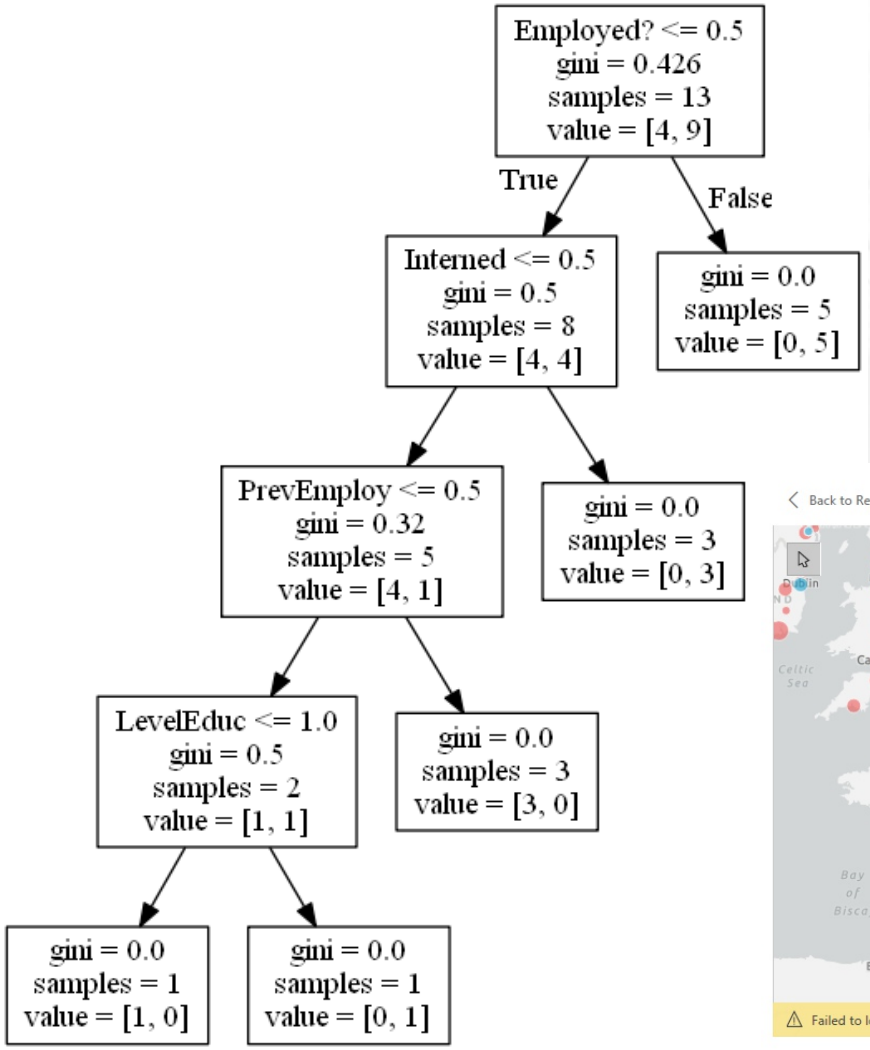
GraphViz-C-Sharp-Wrapper

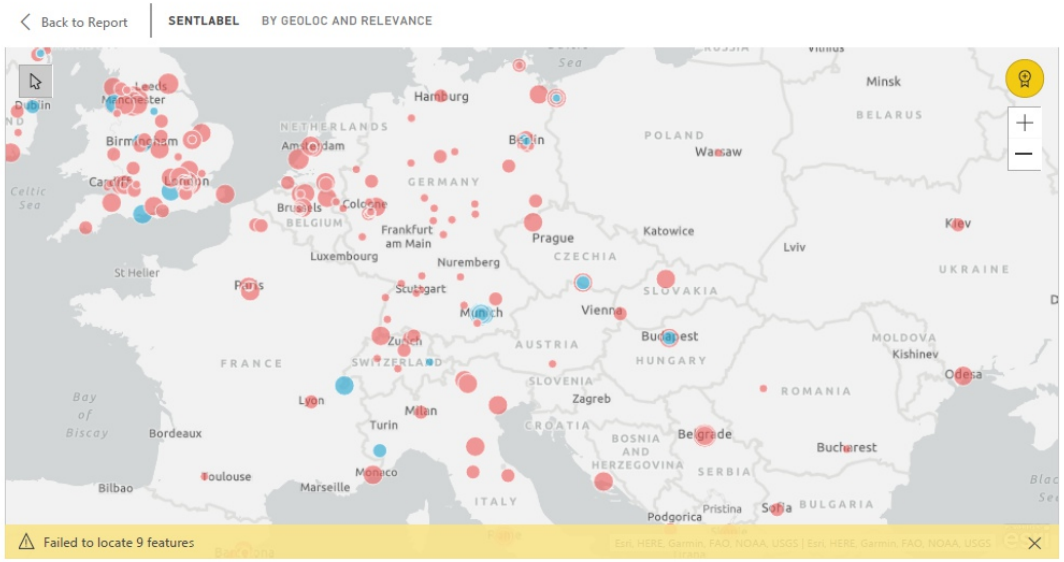
- using System.Collections;
- using System.Runtime.InteropServices;
- using GraphVizWrapper;
- using GraphVizWrapper.Commands;
- using GraphVizWrapper.Queries;
- using Graphviz4Net.Graphs;
- using System.Drawing.Imaging;
- using System.Drawing;
- using System.Drawing.Drawing2D;
- using System.IO; //memory stream

Graph Viz Demo



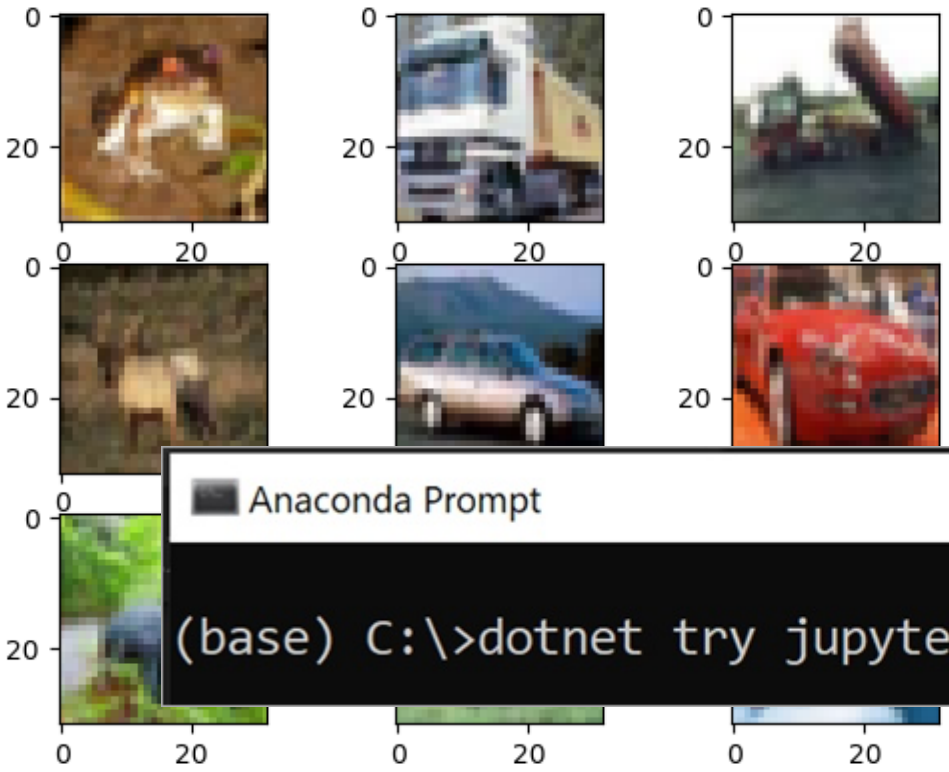
991_pyoma_svr22integrate_dotnetcore2.txt

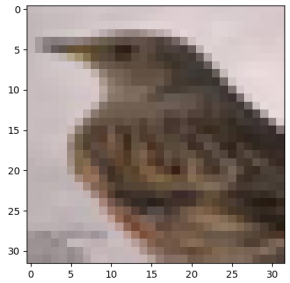

Employed? <= 0.5
gini = 0.426
samples = 13
value = [4, 9]
True
False
Interned <= 0.5
gini = 0.5
samples = 8
value = [4, 4]
gini = 0.0
samples = 5
value = [0, 5]
PrevEmploy <= 0.5
gini = 0.32
samples = 5
value = [4, 1]
gini = 0.0
samples = 3
value = [0, 3]
LevelEduc <= 1.0
gini = 0.5
samples = 2
value = [1, 1]
gini = 0.0
samples = 3
value = [3, 0]
gini = 0.0
samples = 1
value = [1, 0]
gini = 0.0
samples = 1
value = [0, 1]


Back to Report
SENTLABEL
BY GEOLOC AND RELEVANCE
Failed to locate 9 features

3D or 4D Jupyter in VS Code

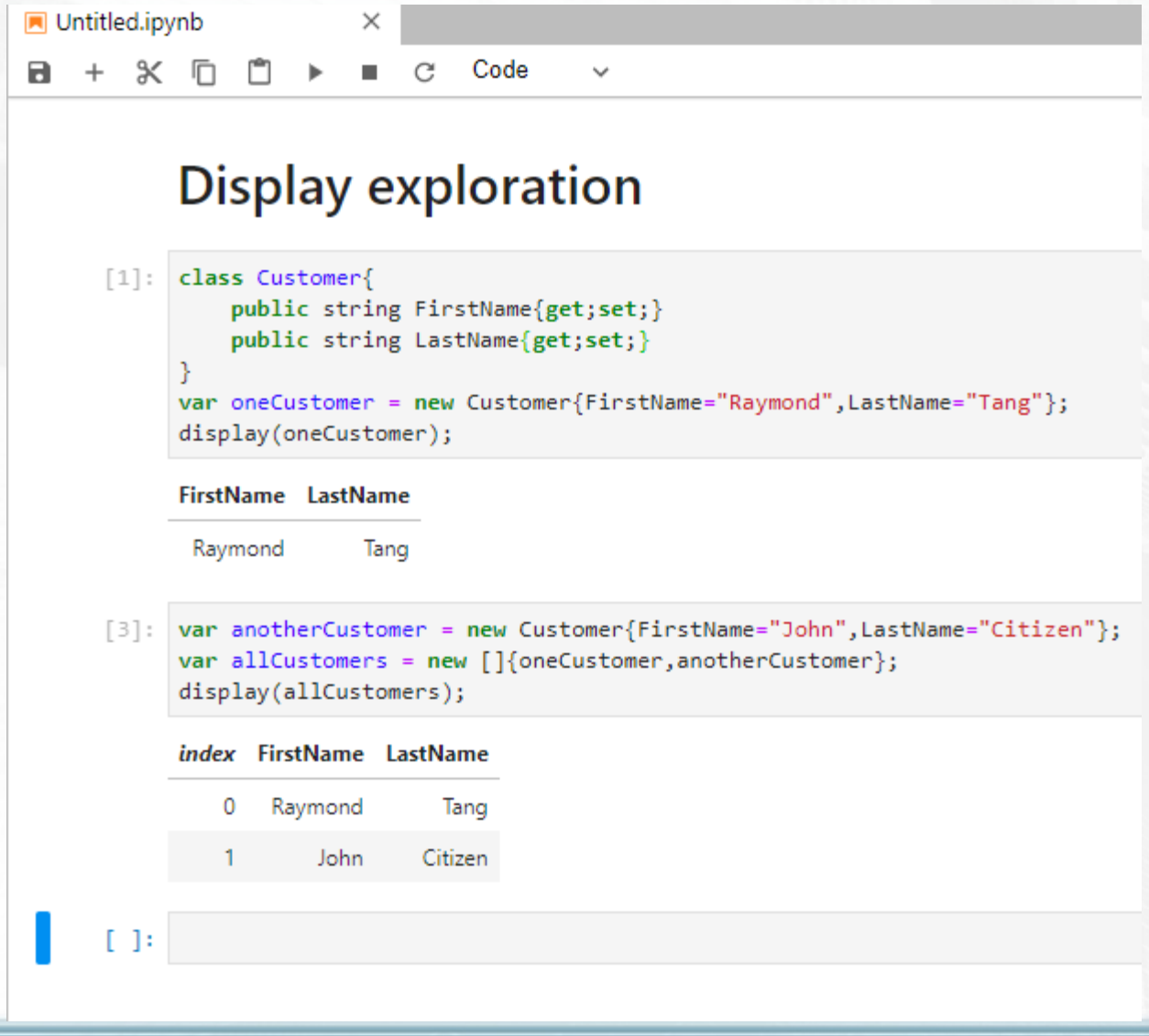
https://code.visualstudio.com/docs/python/jupyter-support



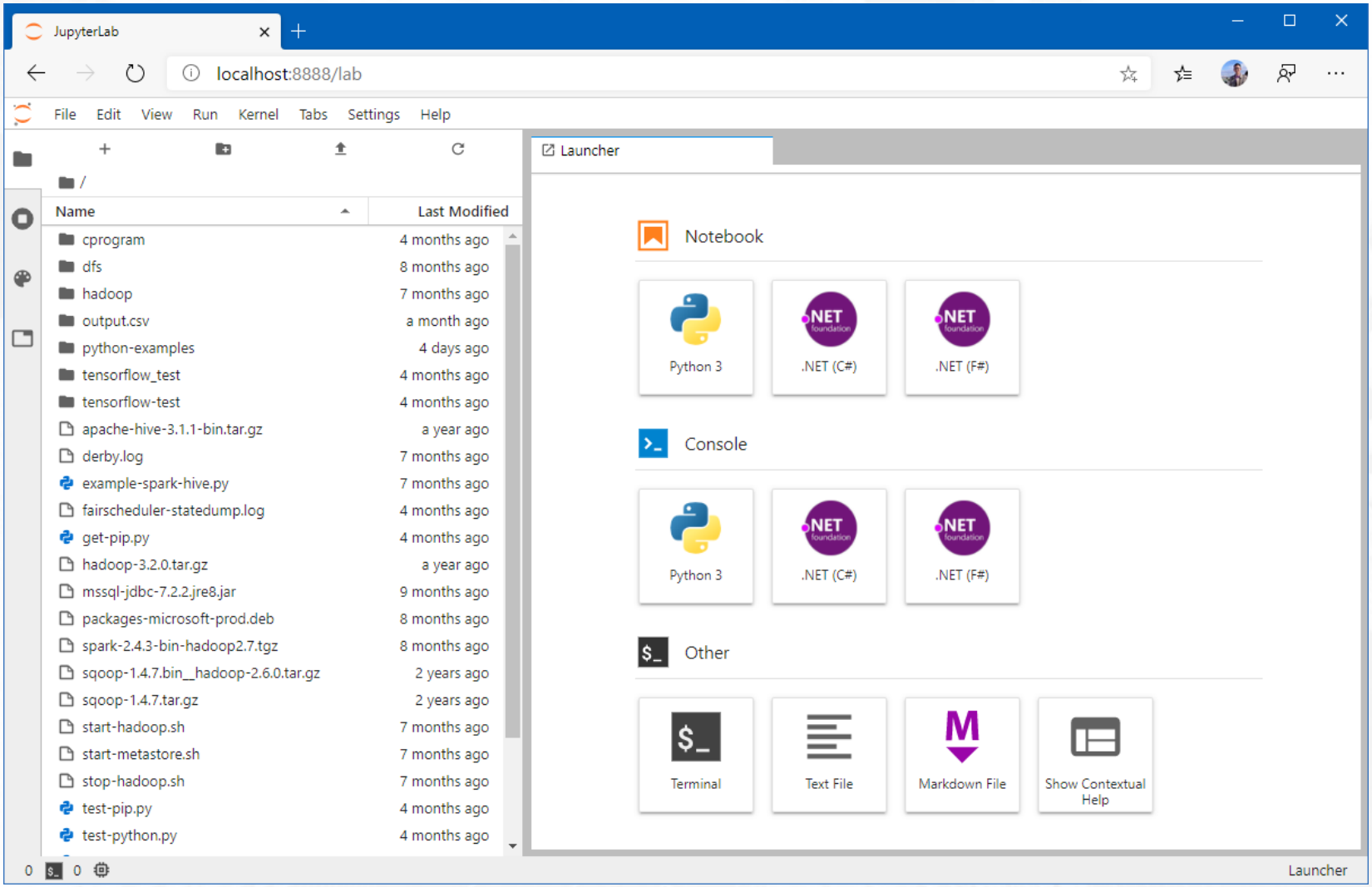



```
classes=('plane','car','bird','cat',
         'deer','dog','frog','horse','ship','truck')
```

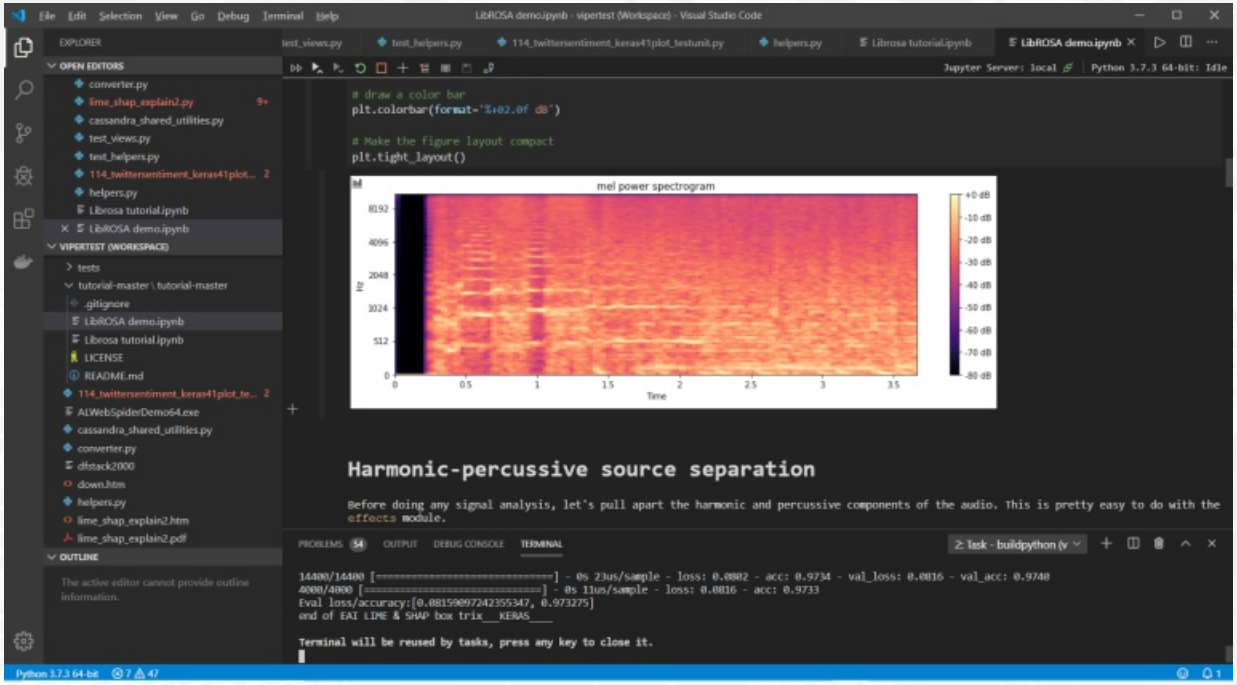
https://www.hanselman.com/blog/AnnouncingNETJupyterNotebooks.aspx



Data Class Display


JupyterLab
localhost:8888/lab
File Edit View Run Kernel Tabs Settings Help
/
Name
Last Modified
cprogram 4 months ago
dfs 8 months ago
hadoop 7 months ago
output.csv a month ago
python-examples 4 days ago
tensorflow_test 4 months ago
tensorflow-test 4 months ago
apache-hive-3.1.1-bin.tar.gz a year ago
derby.log 7 months ago
example-spark-hive.py 7 months ago
fairscheduler-statedump.log 4 months ago
get-pip.py 4 months ago
hadoop-3.2.0.tar.gz a year ago
mssql-jdbc-7.2.2.jre8.jar 9 months ago
packages-microsoft-prod.deb 8 months ago
spark-2.4.3-bin-hadoop2.7.tgz 8 months ago
sqoop-1.4.7.bin__hadoop-2.6.0.tar.gz 2 years ago
sqoop-1.4.7.tar.gz 2 years ago
start-hadoop.sh 7 months ago
start-metastore.sh 7 months ago
stop-hadoop.sh 7 months ago
test-pip.py 4 months ago
test-python.py 4 months ago
Launcher
Notebook
Python 3
.NET (C#)
.NET (F#)
Console
Python 3
.NET (C#)
.NET (F#)
Other
Terminal
Text File
Markdown File
Show Contextual Help
Launcher

Jupyter in VS Code



https://maxbox4.wordpress.com/2020/01/26/jupyter-in-visual-studio-code/

Links & Sources

- Almost all files:
- https://matplotlib.org/
- https://www.sonarlint.org/vscode/
- https://www.graphviz.org/resources/
-
- https://sourceforge.net/projects/maxbox/files/Examples/EKON/BASTA2020/
-
-
- https://maxbox4.wordpress.com/blog/
-
- https://github.com/maxkleiner/maXbox4/releases

max@kleiner.com
www.softwareschule.ch