

COMP4434 Big Data Analytics

Lab 1 Introduction to Jupyter Notebook

HUANG Xiao

xiaohuang@comp.polyu.edu.hk

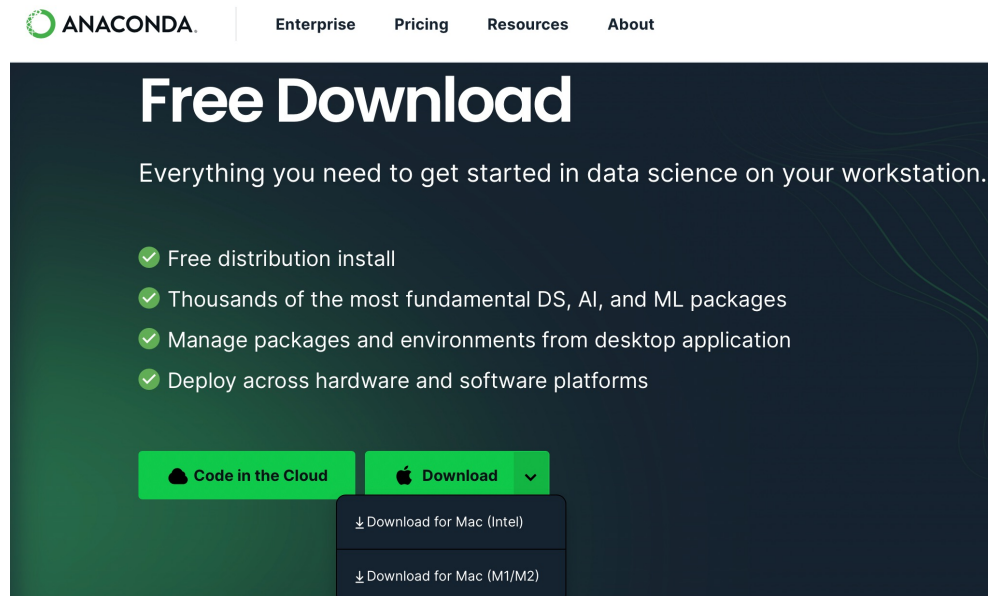
Jupyter Notebook



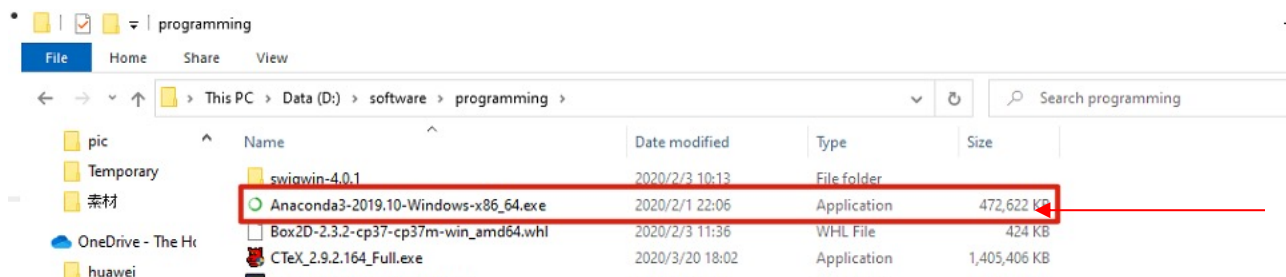
- Project Jupyter exists to develop open-source software, open-standards, and services for interactive computing across dozens of programming languages.
- Method 1: Anaconda
- Method 2: Google's Colab: <https://colab.research.google.com>
 - Tutorial: https://www.tutorialspoint.com/google_colab/google_colab_tutorial.pdf
- Method 3: PyCharm

Anaconda Installation

- Download Anaconda from <https://www.anaconda.com/download>



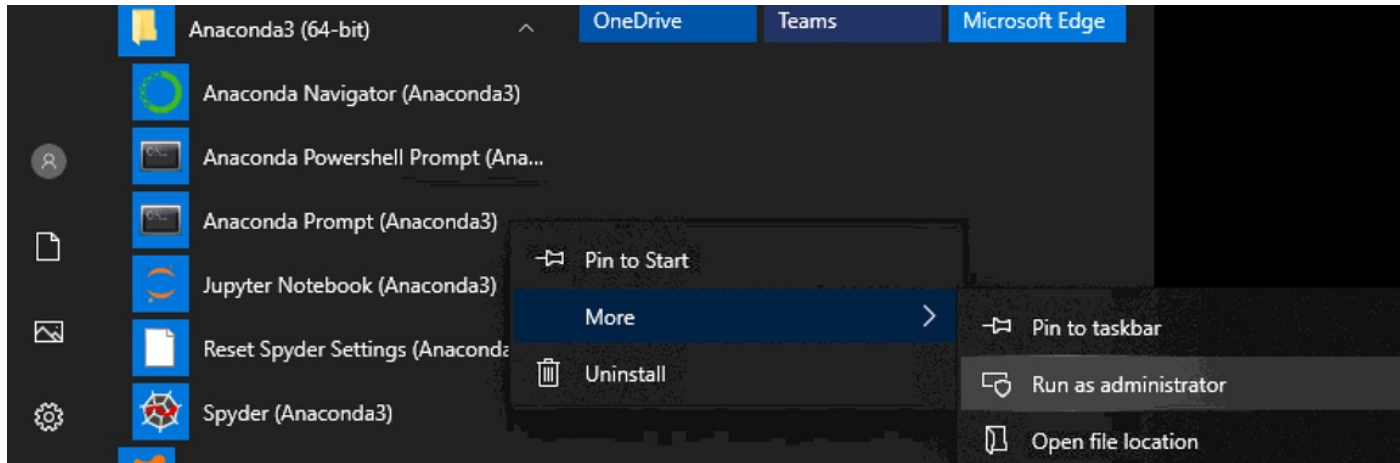
- Install Anaconda



Double-Click it to run

Upgrade Anaconda

- Open Anaconda Prompt



- Input “Conda list” and “conda upgrade --all”

```
Anaconda Prompt (Anaconda3)

(base) C:\Users\18104473r>conda list
# packages in environment at C:\ProgramData\Anaconda3:
#
# Name                   Version           Build    Channel
_ipyw_jlab_nb_ext_conf  0.1.0             py37_0
alabaster                 0.7.12            py37_0
anaconda                  2019.10           py37_0
anaconda-client           1.7.2             py37_0
anaconda-navigator        1.9.7             py37_0
anaconda-project          0.8.3             py_0
asn1crypto                1.0.1             py37_0
astroid                   2.3.1             py37_0
astropy                   3.2.1             py37he774522_0
atomicwrites              1.3.0             py37_1
attrs                     19.2.0            py_0
babel                     2.7.0             py_0
backcall                  0.1.0             py37_0
backports                 1.0               py_2
backports.functools_lru_cache 1.5.1            py_2
```

```
(base) C:\Windows\system32>conda upgrade --all
Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##

environment location: C:\ProgramData\Anaconda3
```

```
Proceed ([y]/n)? y

Preparing transaction: done
Verifying transaction: \
```

Jupyter Installation via Anaconda

- Check the version of Python

```
(base) C:\Windows\system32>python
Python 3.7.6 (default, Jan 8 2020, 20:23:39) [MSC v.1916 64 bit (AMD64)] :: Anaconda, Inc. on win32
Type "help", "copyright", "credits" or "license" for more information.
>>>
```

```
(base) C:\Windows\system32>python
Python 3.7.6 (default, Jan 8 2020, 20:23:39) [MSC v.1916 64 bit (AMD64)] :: Anaconda, Inc. on win32
Type "help", "copyright", "credits" or "license" for more information.
>>> exit()

(base) C:\Windows\system32>_
```

- Install Jupyter Notebook——“conda install jupyter notebook”

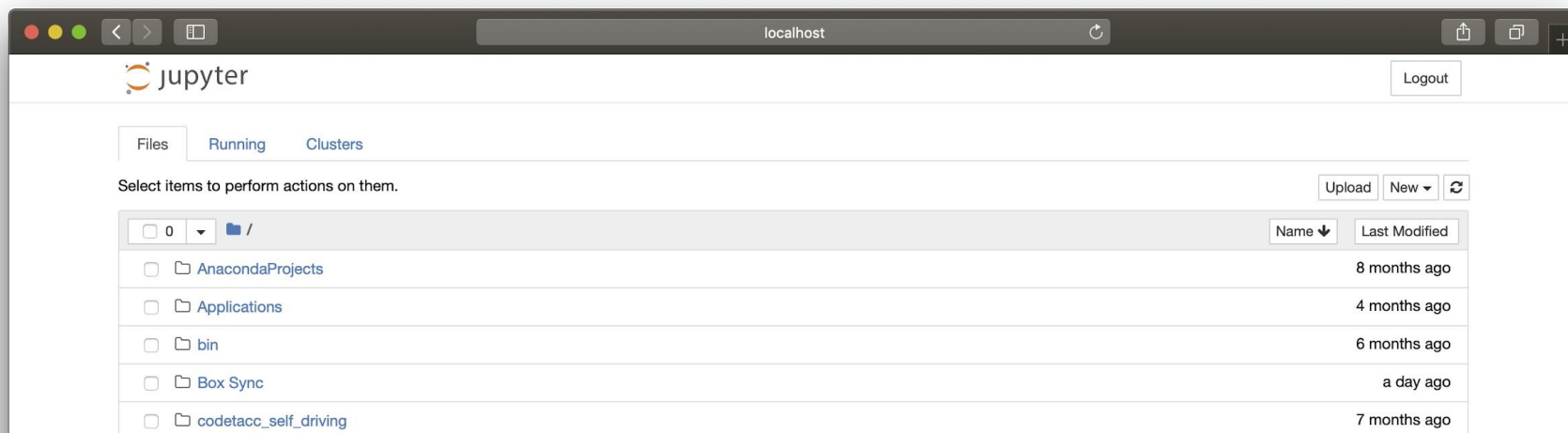
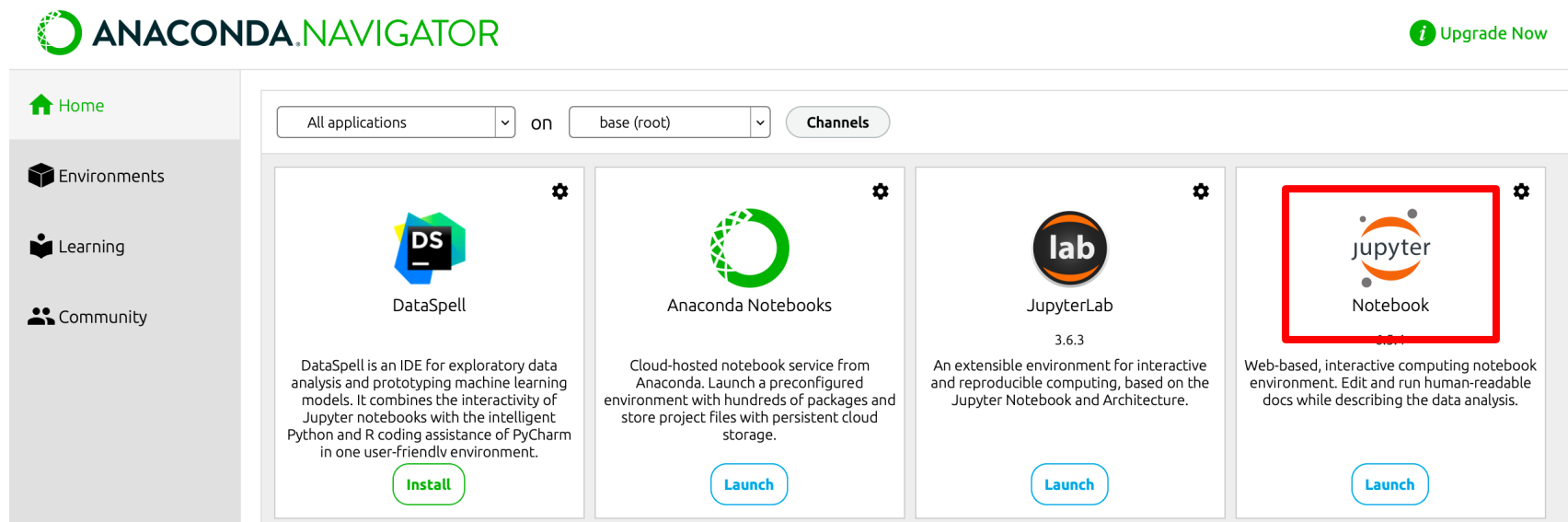
```
(base) C:\Windows\system32>conda install jupyter notebook
Collecting package metadata (current_repodata.json): done
Solving environment: done

# All requested packages already installed.

(base) C:\Windows\system32>_
```

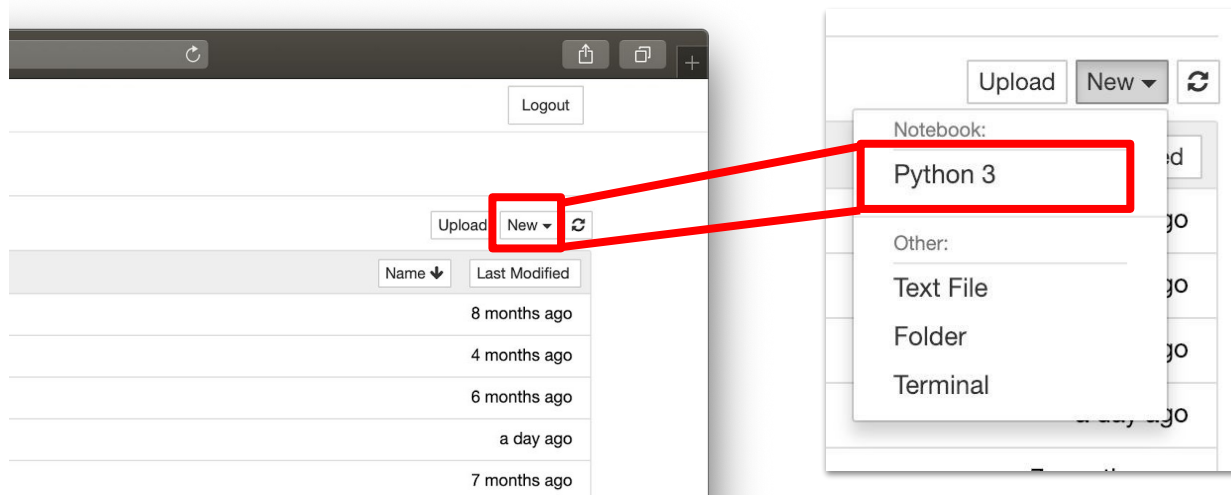
Jupyter Notebooks Terminology

■ Open Jupyter Notebook

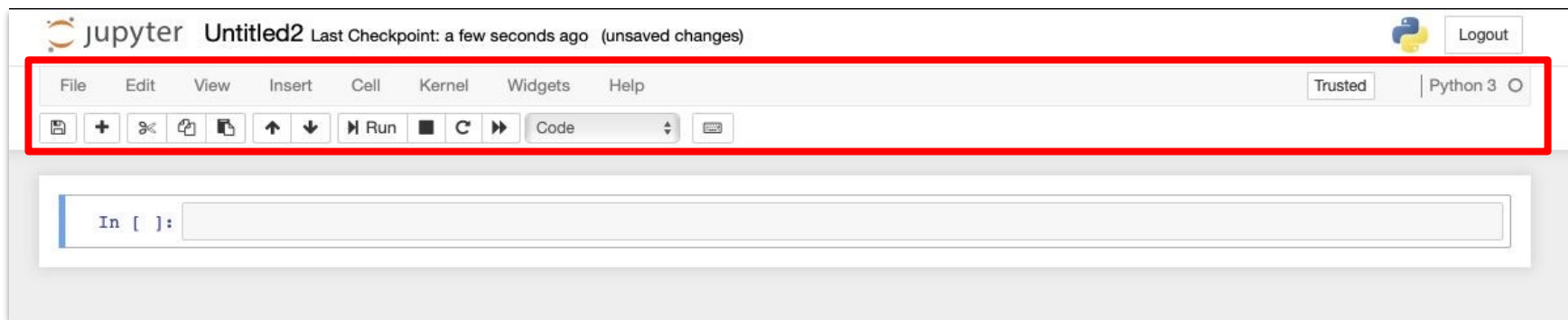


Jupyter Notebooks Kernels and Toolbar

Kernels



Toolbar

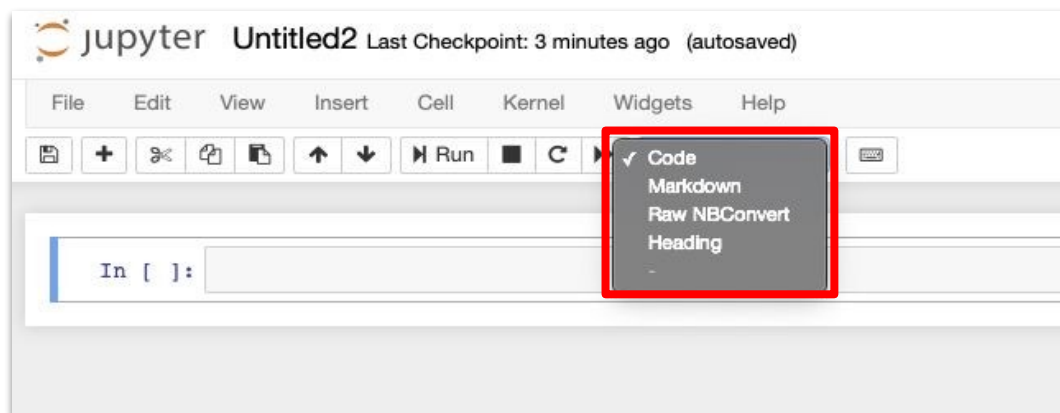


Jupyter Notebooks Cell

Cell

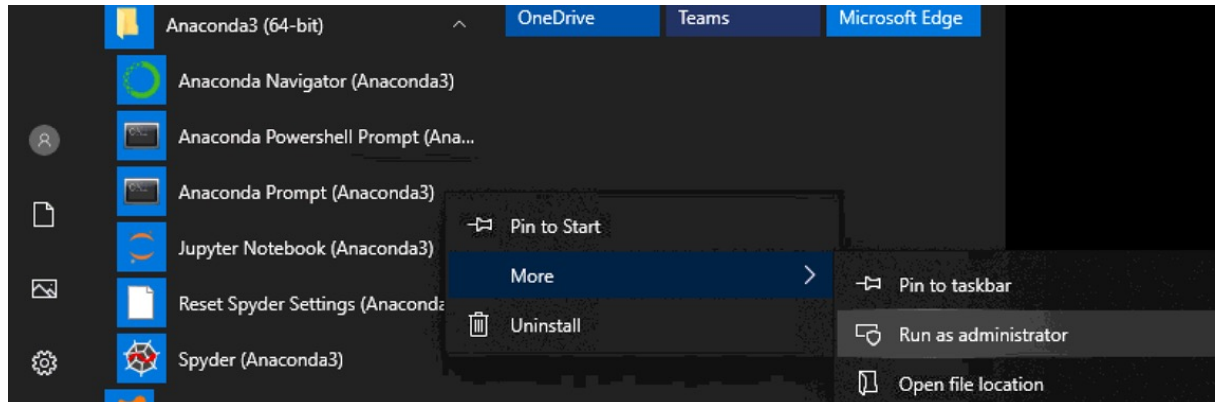


Cell Type Dropdown



Python packages Installation

- Open Anaconda Prompt



- Type 'conda install XX' or 'pip install XX'

```
(base) C:\Users\Yiliu>conda install numpy
Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##

  environment location: C:\Users\Yiliu\Anaconda3

added / updated specs:
- numpy
```

```
(base) C:\Windows\system32>conda install scikit-learn
Collecting package metadata (current_repodata.json): done
Solving environment: done

# All requested packages already installed.

(base) C:\Windows\system32>
```

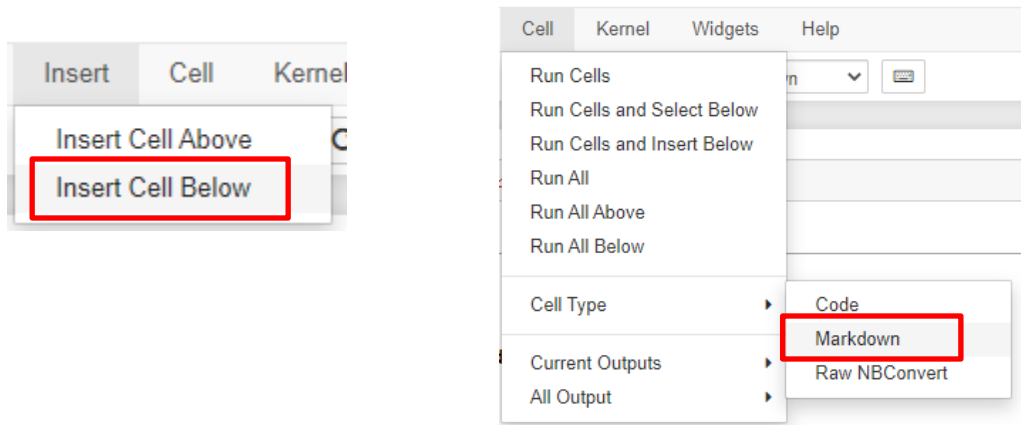
Simple Practice

■ Python -- “Hello world”

- Now, we may try to input a simple statement, which prints the “Hello World”. Enter “print(“Hello World”)” in the cell, and then click the [run] button on menu bar or press [Ctrl]+[Enter] keys.

```
In [1]: print("Hello world")  
Hello world
```

■ Markdown Document



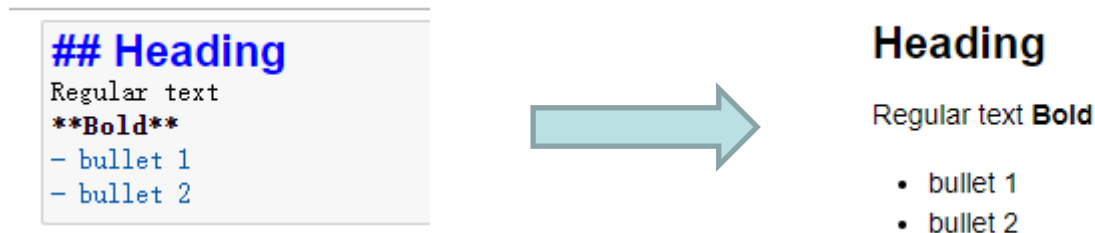
Simple Practice

■ Markdown Document

- Type the following lines in the modified cell. These are some simple commands to create a markdown document.

```
## Heading  
Regular text  
**Bold**  
- bullet 1  
- bullet 2
```

Press[Ctrl]+[Enter] keys.



<https://www.markdownguide.org/basic-syntax/>

Play with it on Jupyter Notebook

- Code link: https://github.com/cugzj/Simple-Linear-Regression/blob/master/LR_example.py

```
jupyter Untitled2 最后检查 1 天前 (自动保存) Logout
File Edit View Insert Cell Kernel Widgets Help 可信的 Python 3
In [11]: # imports
import numpy as np
import matplotlib.pyplot as plt
from sklearn.metrics import mean_squared_error, r2_score

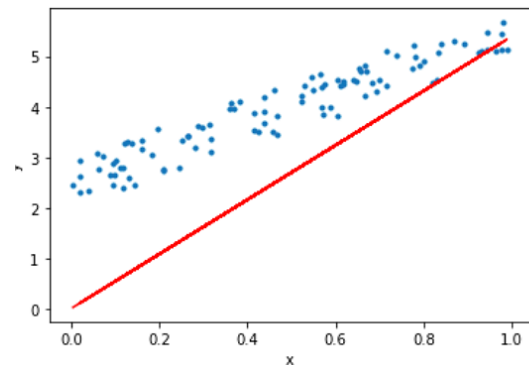
class LinearRegressionUsingGD:
    """Linear Regression Using Gradient Descent.
    Parameters
    eta : float
        Learning rate
    n_iterations : int
        No of passes over the training set
    Attributes
    w_ : weights/ after fitting the model
    cost_ : total error of the model after each iteration
    """

    def __init__(self, eta=0.005, n_iterations=1000):
        self.eta = eta
        self.n_iterations = n_iterations

    def fit(self, x, y):
        """Fit the training data
        Parameters
        x : array-like, shape = [n_samples, n_features]
            Training samples
        y : array-like, shape = [n_samples, n_target_values]
            Target values
        Returns
        self : object
        """

        self.cost_ = []
        self.w_ = np.zeros((x.shape[1], 1))
        m = x.shape[0]

        for _ in range(self.n_iterations):
            y_pred = np.dot(x, self.w_)
            residuals = y_pred - y
            gradient_vector = np.dot(x.T, residuals)
            self.w_ -= (self.eta / m) * gradient_vector
            cost = np.sum((residuals ** 2)) / (2 * m)
            self.cost_.append(cost)
```



More references

- Python:
 - Basic:https://www.tutorialspoint.com/python/python_basic_syntax.htm
 - Detail:https://bugs.python.org/file47781/Tutorial_EDIT.pdf
- Markdown basic syntax:
 - <https://www.markdownguide.org/basic-syntax/>
- Jupyter Notebook's history:
 - <https://ep2020.europython.eu/media/conference/slides/7UBMYed-a-brief-history-of-jupyter-notebooks.pdf>