

Configuration Manual

MSc Research Project
Data Analytics

Punith Kumar Nagaraje Gowda
Student ID: x18130771

School of Computing
National College of Ireland

Supervisor: Dr. Cristina Muntean

National College of Ireland
MSc Project Submission Sheet
School of Computing



Student Name:Punith Kumar Nagaraje Gowda.
Student ID:x18130771.....
Programme:M.Sc in Data Analytics..... **Year:**2019.....
Module:Research Project.....
Lecturer:Dr. Cristina Muntean.....
Submission Due Date:12/12/2019.....
Project Title:Performance of Motor Vehicle based on Driving and Vehicle Data using Machine Learning.....
Word Count:780..... **Page Count:**6.....

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:Punith Kumar Nagaraje Gowda.....
Date:12/12/2019.....

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	<input type="checkbox"/>
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	<input type="checkbox"/>
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	<input type="checkbox"/>

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Punith Kumar Nagaraje Gowda
Student ID: x18130771

1 Introduction

The manual is designed to explain the steps in detail to run the research project and also specifies the configuration of the machine used to build and run the models. The steps include installation of required software and the packages and also the minimum recommended configuration to run the project smoothly.

2 System Specification

2.1 Hardware

Hardware configuration of the machine used for the project:

Processor : 1.8 Ghz Dual-Core Intel Core i5

Memory : 8 GB 1600 MHz DDR3

Storage : 256 GB SSD

Operating System : macOS Catalina Version 10.15.1

Recommended hardware configuration:

Processor : Dual-Core Intel Core i5 or equivalent

Memory : 6 GB

Storage : 128 GB

The recommended configuration ensures that the code runs smoothly without any hurdles while the minimum configuration would be of RAM 4GB which will be sufficient to clinch the results but with more processing time.

2.2 Software

Microsoft excel : Used for data exploration and initial understanding.

Jupyter Notebook : For loading, cleaning and pre-processing the data. And for building the models and evaluating them. Python was used for the implementation.

PyCharm community edition : For quick debugging of Python code.

3 Download and installation

At first python needs to be installed¹ based on the operating system, the latest version is recommended for the installation. Fig. 1 shows the snapshot from where python can be downloaded. Python version 3.7.3 for the macOS was downloaded and installed, the then latest version.

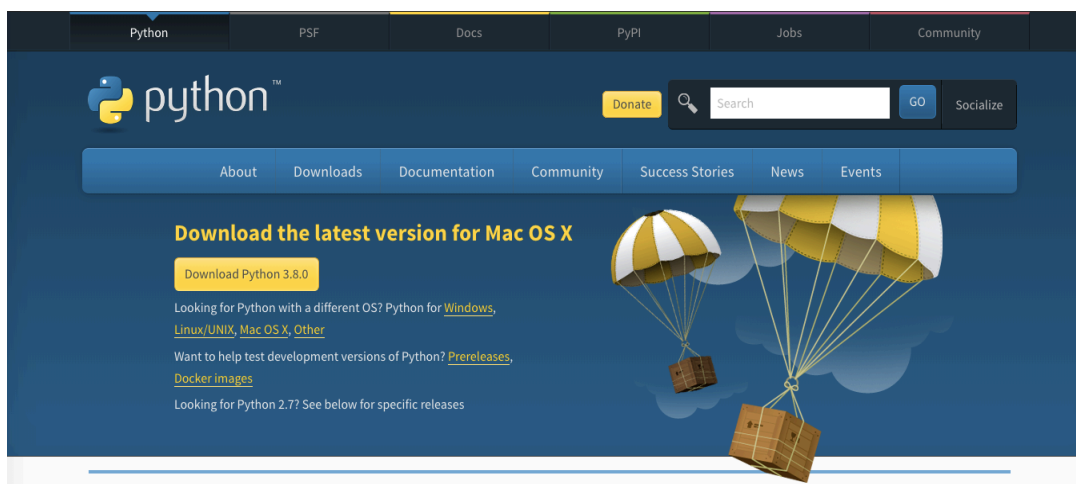


Figure 1. Snapshot from python.org

After python is installed, a development environment is required to write the code, execute it and to see the results. The most popular and user-friendly environments are Jupyter Notebook and Spyder which comes with the python distribution, anaconda². Fig. 2 is the snapshot of download screen of anaconda for macOS, based on the operating system corresponding installer can be downloaded and also has an option of graphical installer or the command line installer. Graphical installer is an easy way of installing as it does not demand for any knowledge of commands to install.

¹ <https://www.python.org/downloads/>

² <https://www.anaconda.com/distribution/>

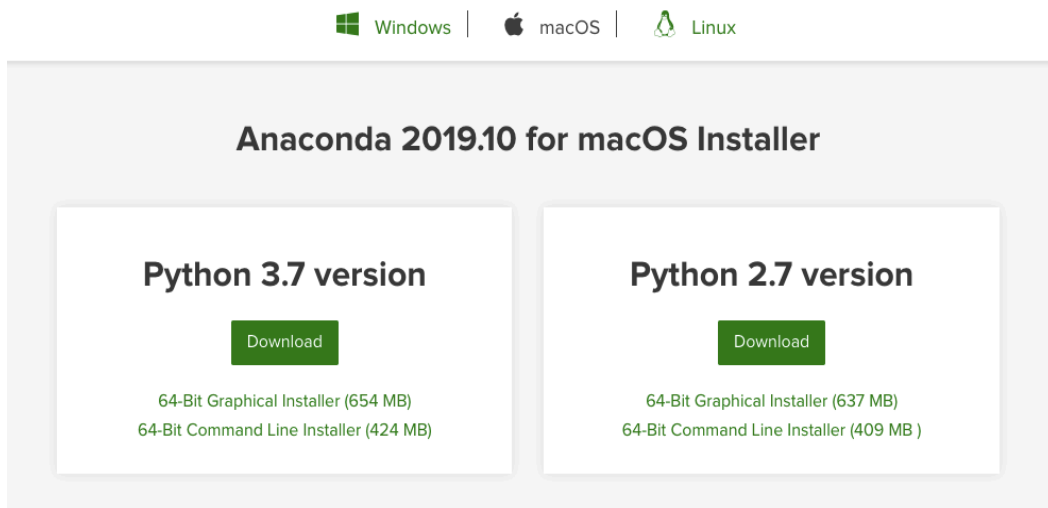


Figure 2. Snapshot from anaconda.com

Fig. 3 shows the anaconda navigator which is downloaded and installed. The anaconda comes with Jupyter Notebook and Spyder pre-installed. Launching the Jupyter Notebook or Spyder one can start writing code in python by creating a new python file.

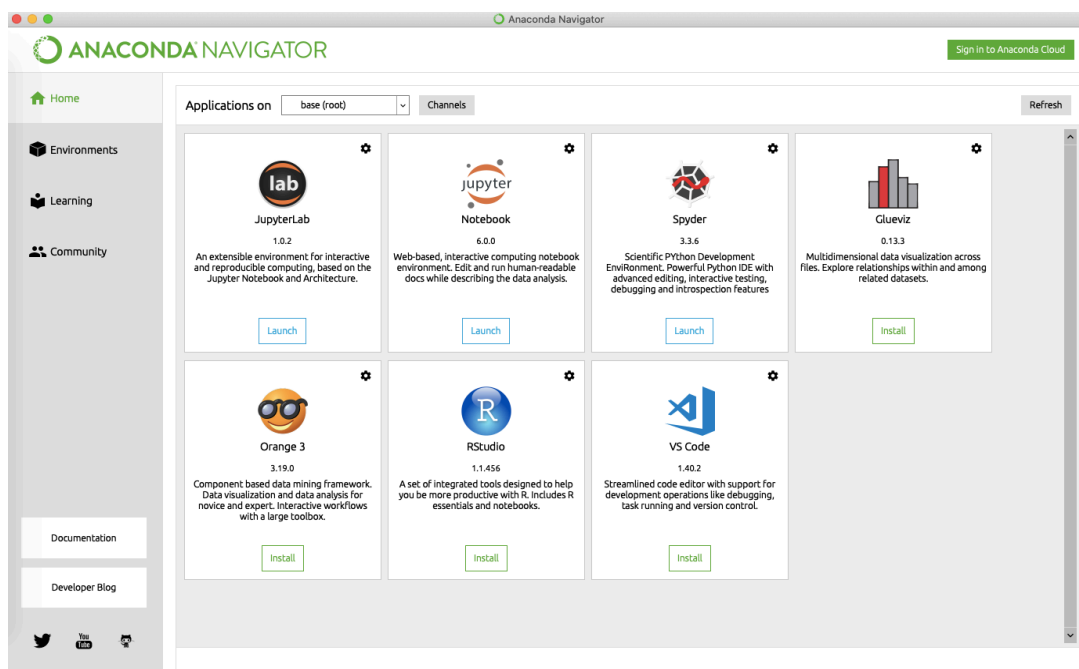


Figure 3. Snapshot of anaconda navigator

The other popular python IDE is PyCharm. Fig. 4 shows the download page of PyCharm³ for macOS. The community version of PyCharm was downloaded as it is free and open-source, professional version has more features compared to community version but is not

³ <https://www.jetbrains.com/pycharm/download/#section=mac>

free. The features available in the community version is adequate and PyCharm was used to debug the python code if there was any error.

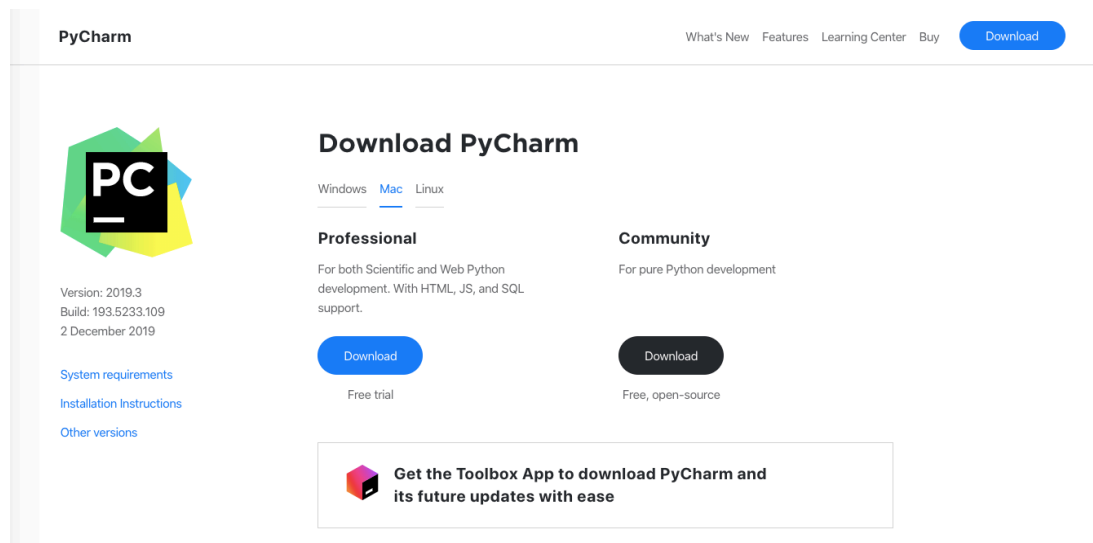


Figure 4. Snapshot from jetbrains pycharm download page

The system requirements for PyCharm is shown in the Fig. 5. It consists of both minimum and recommended requirements making it easier for download.

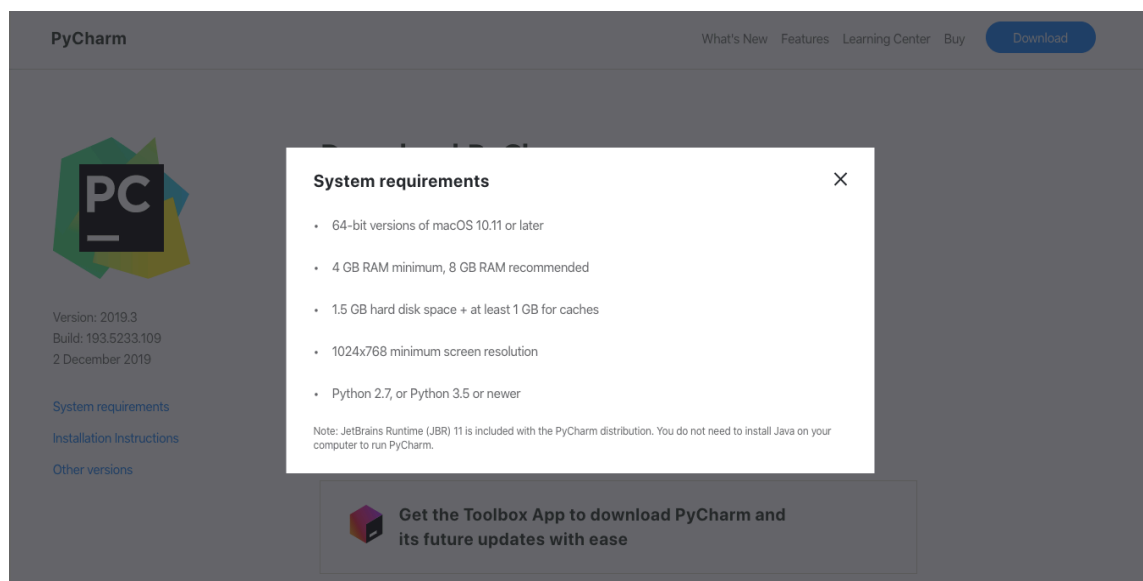


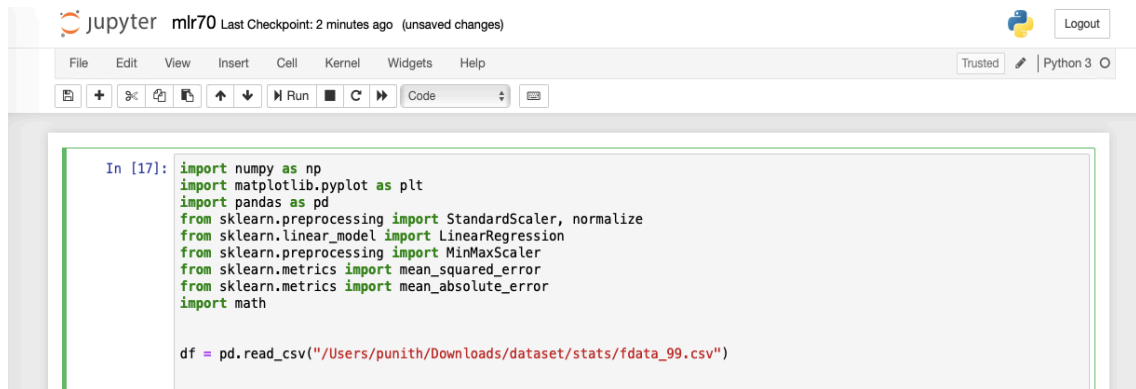
Figure 5. System requirements for PyCharm

4 Project Development

After the installation of the necessary software python coding can be done by launching the Jupyter Notebook from the anaconda navigator or the PyCharm. Detailed steps with respect to Jupyter Notebook is given, clicking on the new button on the right-hand side

will give a drop-down menu, click on the Python 3 section and this will take you to the new file. It consists of cells where the code or the script can be written, there is an option to add more cells or remove cells. The entire file can be executed at once or each cell can be executed individually.

The packages necessary for the project are installed using the command 'pip install package_name'.



```
In [17]: import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.preprocessing import StandardScaler, normalize
from sklearn.linear_model import LinearRegression
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error
from sklearn.metrics import mean_absolute_error
import math

df = pd.read_csv("/Users/punith/Downloads/dataset/stats/fdata_99.csv")
```

Figure 6. Snapshot of python script in Jupyter Notebook

Fig. 6 shows the python script in a cell in Jupyter Notebook. The '+' button helps in adding a cell, the cell will be added below the selected cell and the 'Run' button executes the selected cell.

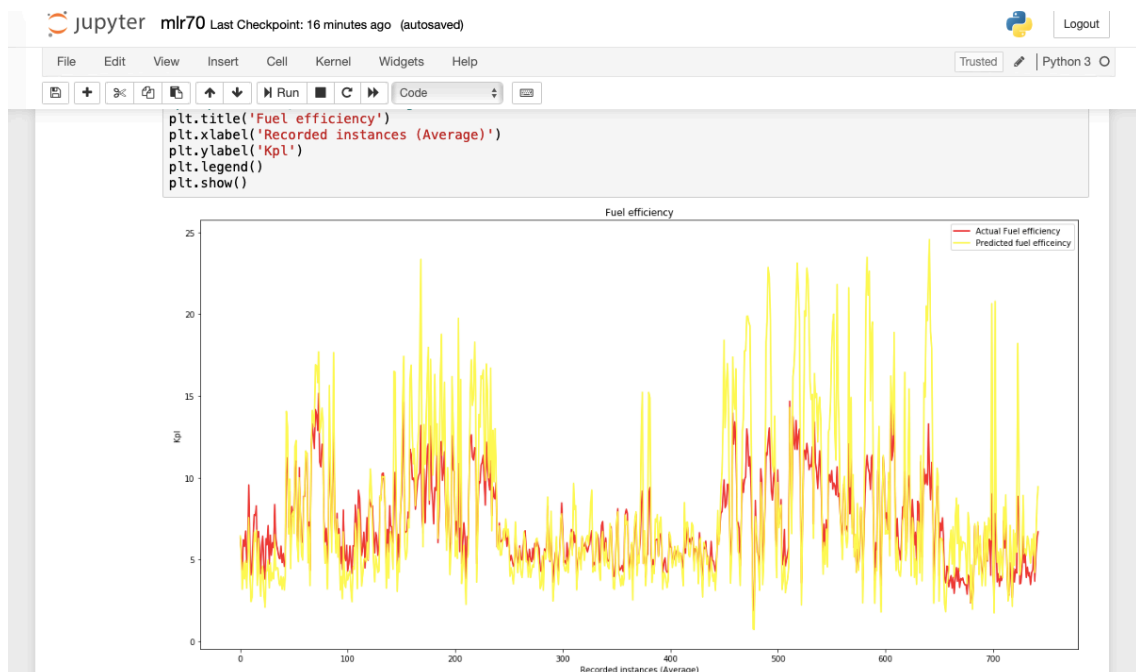


Figure 7. Snapshot of output in Jupyter Notebook

After the click on Run button, the script runs successfully, and the output is given below the cell as shown in the Fig. 7.

The snapshots give an overview of how a script can be added in Jupyter Notebook and how the scripts are run. The scripts can be saved as python notebook with an extension of `‘.ipynb’` and can be opened in Jupyter Notebook or can be saved as a python file with the extension `‘.py’`. All the scripts used for the project are submitted as part of the ICT solution along with the dataset and can be imported and executed quickly.