**CS5590 APS**

**-**

**Python and Deep Learning Programming**

**LAB1**

**Team Members:**

**Dwarkamoye Mohanty (25)**

**Kruthika Basvankote Channappa (4)**

**Bhavaz Akula (2)**

**Introduction: -**

The assignment is targeted to cover python fundamentals and trying hands on machine learning models with different dataset.

Programming language python is purposely made such that its an interpreted and high level. The design strategy of python targets usage of indentation and white spaces such as readability increases. It follows an OOPS approach to manage clarity, coding logically and working without hesitation with large scale assignments. The two ML approaches we are covering is Supervised learning and Unsupervised Learning. Input factors or features are known in both cases while Output variable is known only in Supervised learning.

**Objectives: -**

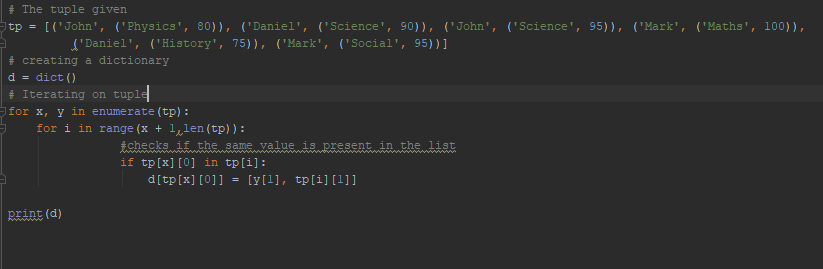
* Creating list, tuple, dictionaries to do operations.
* Training classifier or model with different dataset and check the accuracy score for each model.
* Creating multiple regression and evaluating the model with R2 and RMSE score.
* Using a clustering algorithm on a dataset to find the distribution of data across clusters i.e. Unsupervised approach.

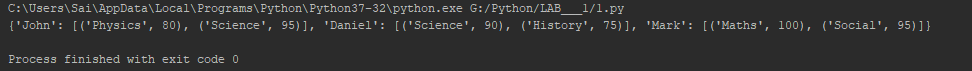
**Methods: -**

* Created Python file using PyCharm IDE.
* Downloaded the dataset from Kaggle.
* Trained model with that dataset.
* Plotted using matplotlib.

**Workflow: -**

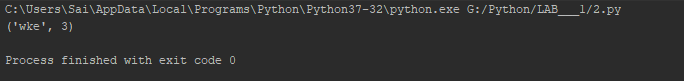




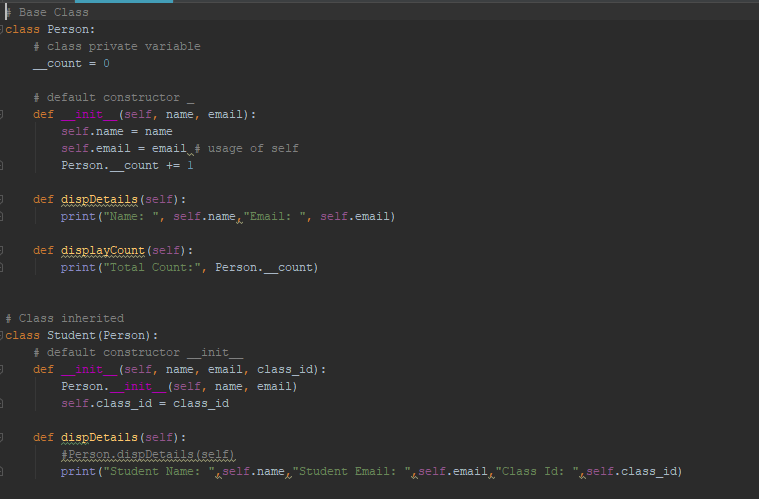


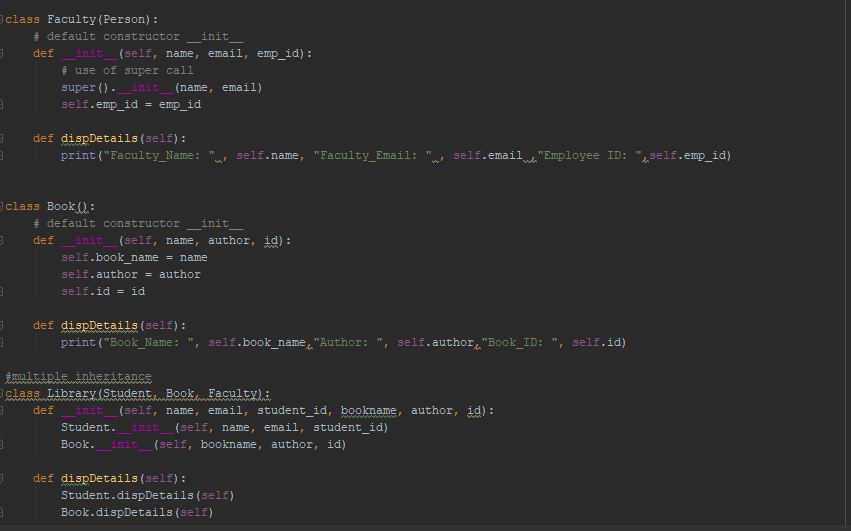
2.

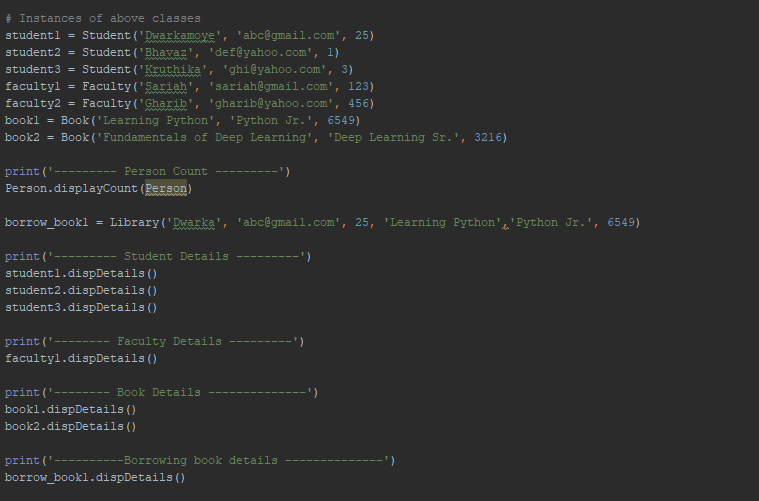


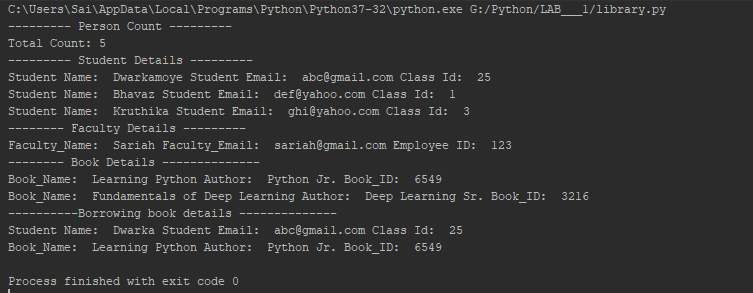


3.

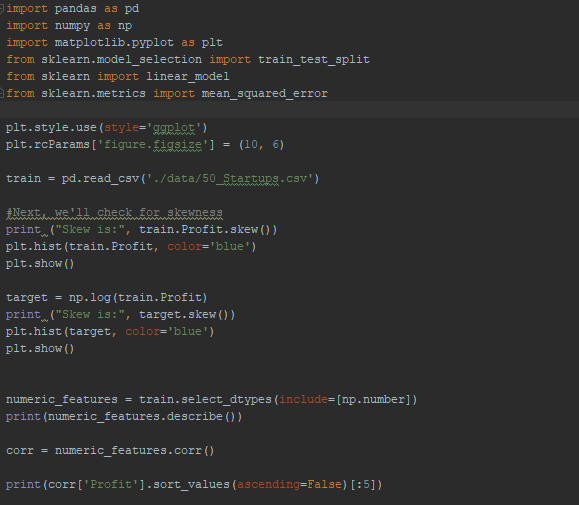


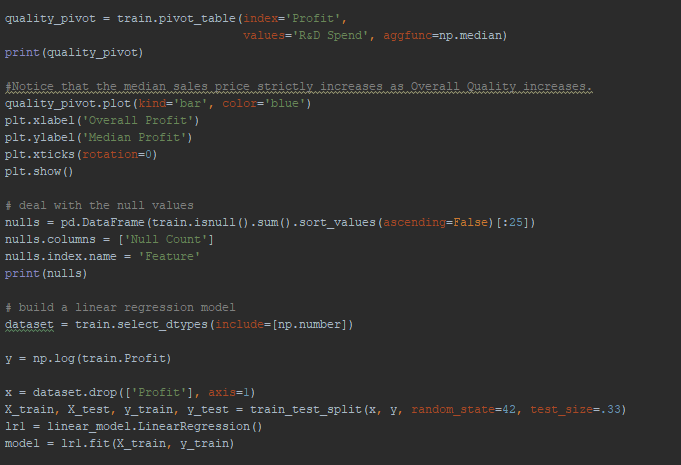


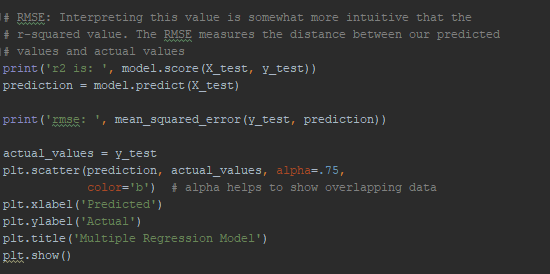


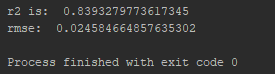


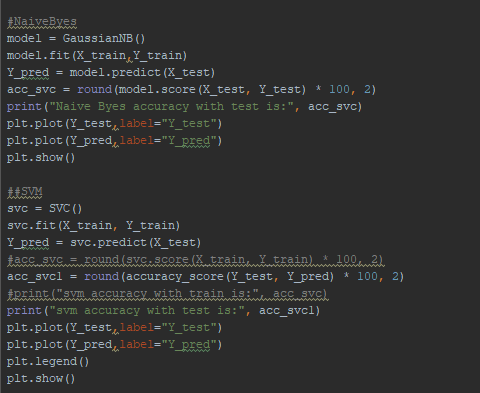
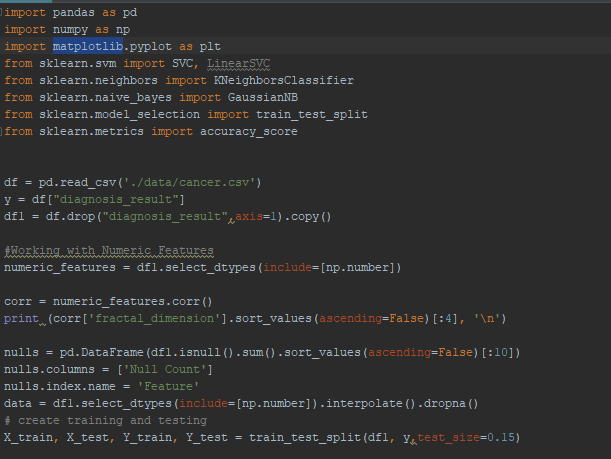
4.

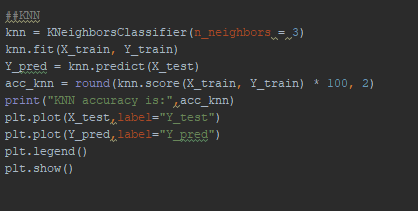


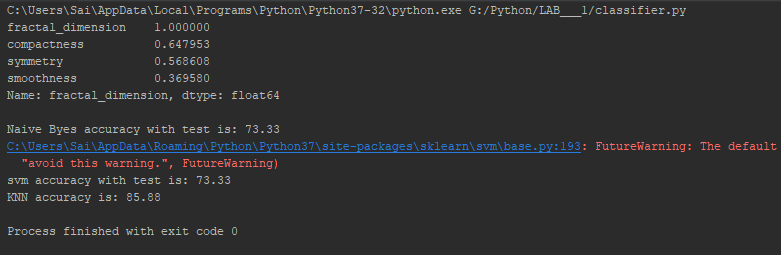




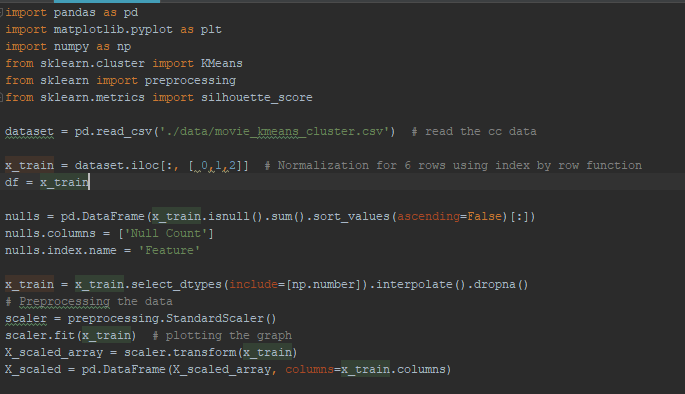


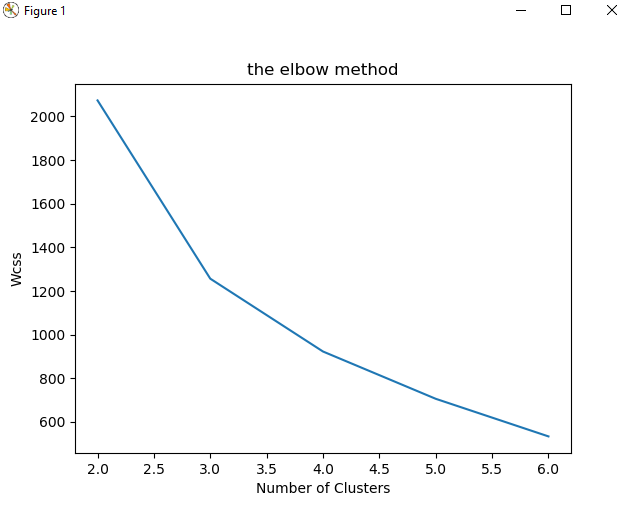
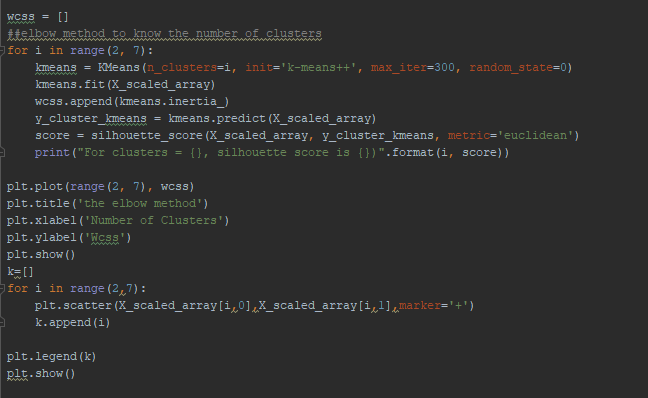
5.

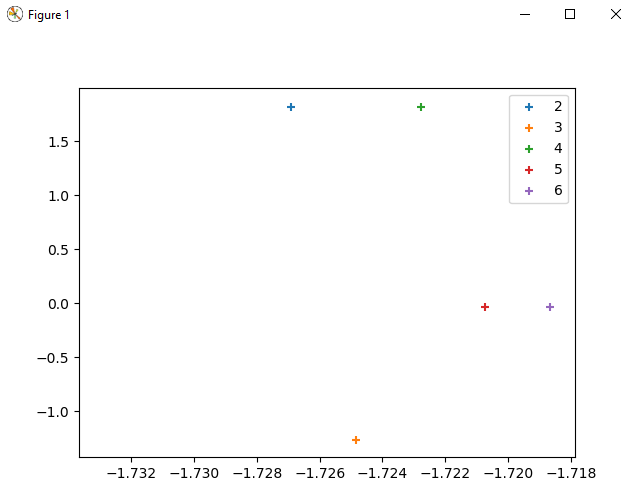


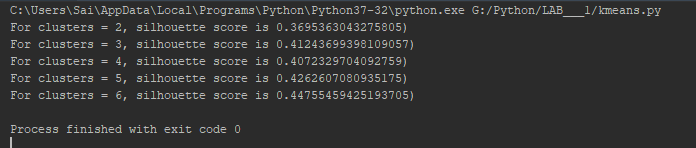


6.









**Datasets: -**

<https://www.kaggle.com/>

1. We have downloaded 3 datasets,

|  |  |
| --- | --- |
| Dataset | Usage |
| 50 Startups | Multiple Regression |
| Cancer | Classifiers |
| Movie Kmeans Cluster | K Means |

**Evaluation & Discussion: -**

A machine learning model's exhibition is viewed as great dependent on its expectation and how well it sums up on a free test dataset. Based on the presentation of various models we pick the model which positions most elevated in performance. When we have an input x and we apply a function f on the input x to anticipate an output y. Difference between the genuine yield and anticipated yield is the error. Our objective with ML algo is to produce a model which limits the mistake of the test dataset.

**Conclusion: -**

We have completed the assignments covering fundamentals of python and have trained classifier with models like SVM, KNN, Naïve Byes. We have also applied Kmeans clustering and plotted the clusters successfully. We have also created Multiple regression and evaluated R2, RMSE scores successfully.

**GITHUB LINK: -**

<https://github.com/Dwarkamoye/Python-CS5590-490-0001/tree/master/Lab-1>