**Universities Classification(private vs public)**

**MINI PROJECT**

Submitted in partial fulfillment of the requirement of University of Mumbai

For the Degree of

**(Electronics and Telecommunication Engineering)**

**By**

1. **Chinmay Gambhirrao “ID No: TU2F1920043”**

**Under the Guidance of**

**Prof. Nitesh Karmakar**

Knowledge Solutions India

**Department of Electronics and Telecommunication Engineering**

**TERNA ENGINEERING COLLEGE**

**Plot no.12, Sector-22, Opp. Nerul Railway station,**

**Phase-11, Nerul (w), Navi Mumbai 400706**

**UNIVERSITY OF MUMBAI**

**TERNA ENGINEERING COLLEGE, NERUL,**

**NAVI MUMBAI**

**Department of Electronics and Telecommunication Engineering**

Academic Year 2021-22

**CERTIFICATE**

This is to certify that

1. **Chinmay Gambhirrao “ID No: TU2F1920043”**

Has satisfactorily completed the requirements of the Mini Project of Subject

“Machine Learning”

As prescribed by the University of Mumbai under the guidance of

Prof. Nitesh Karmakar

**Subject In-charge HoD**

**Index**

|  |  |  |
| --- | --- | --- |
| **TABLE OF CONTENTS** | | |
| **Sr. No.** | **Title** | **Page No.** |
| Chapter 1 | Introduction   * 1. Sub chapter |  |
| Chapter 2 | Problem Statement |  |
| Chapter 3 | Implementation |  |
| Chapter 4 | Results |  |
| Chapter 5 | Conclusion |  |
|  | References |  |

**Chapter 1**

**Introduction**

K-Means Clustering is an unsupervised machine learning algorithm. In contrast to traditional supervised machine learning algorithms, K-Means attempts to classify data without having first been trained with labeled data. Once the algorithm has been run and the groups are defined, any new data can be easily assigned to the most relevant group.

For this project we will attempt to use KMeans Clustering to cluster Universities into to two groups, Private and Public.

**Chapter 2**

**Problem Statement**

* 1. **Problem Statement:**

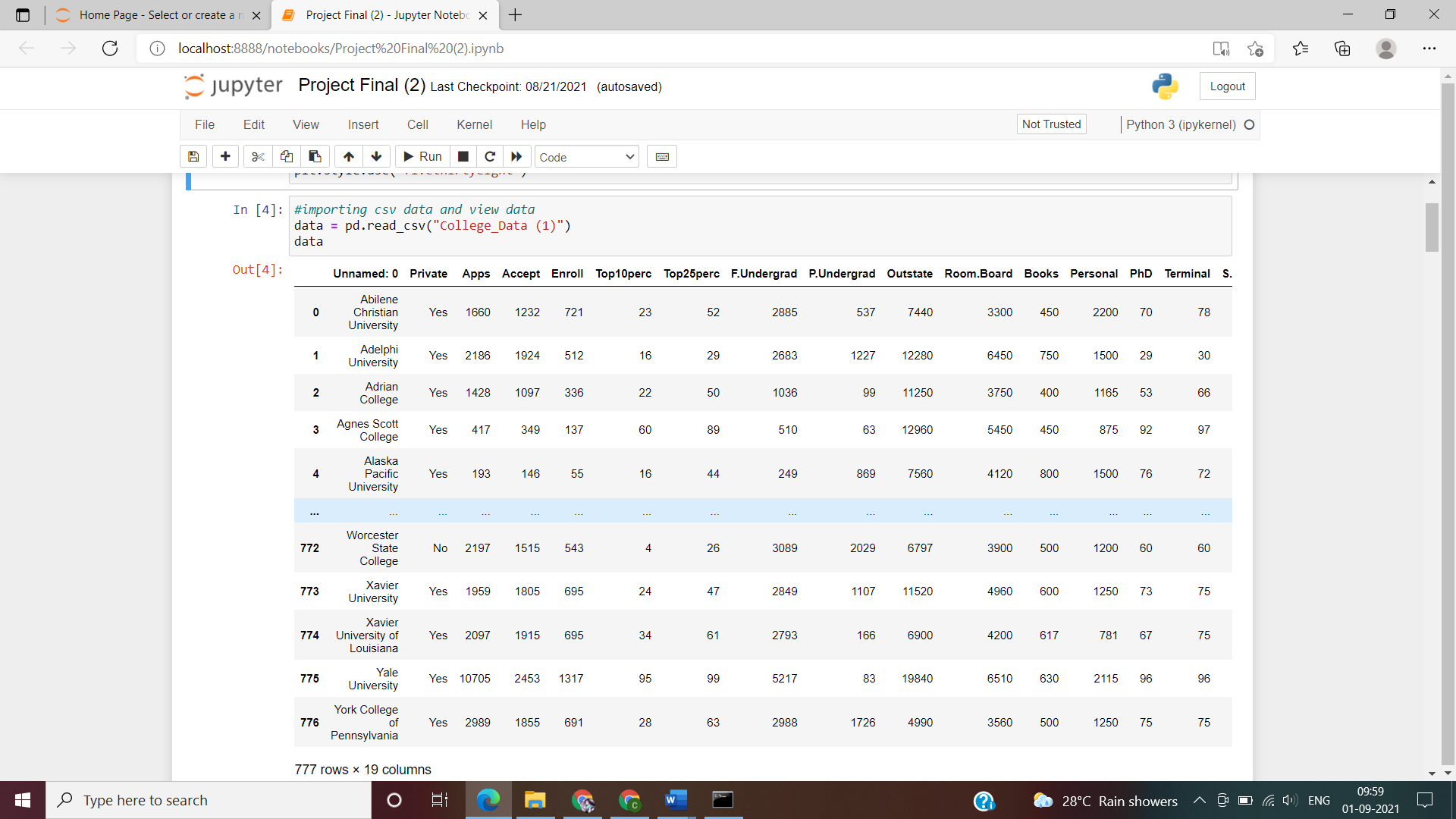
A student who does at least cursory research on a university website can find out in under a minute whether a school is private or public, but we’ll see if a model can correctly categorize schools based on features that may seem very relevant (out-of-state tuition, graduation rate) and others that might not seem relevant at all (percentage of alumni who donate).

**Chapter 3**

**Datasets**

The dataset provides patient information which includes over 700 records and 19 attributes. The attributes include: unamed, private, Apps, Accept, Enroll, Top10perc, Top25perc, F.Undergrad, Outstate, Room.Board, Books, Personal, Phd, Terminal, S.F.Ratio, perc.alumni, Expend, Grade.Rate. The data set is in csv(Comma Separated Value) format which is further prepared to data frame as supported by pandas library in python.

|  |  |  |  |
| --- | --- | --- | --- |



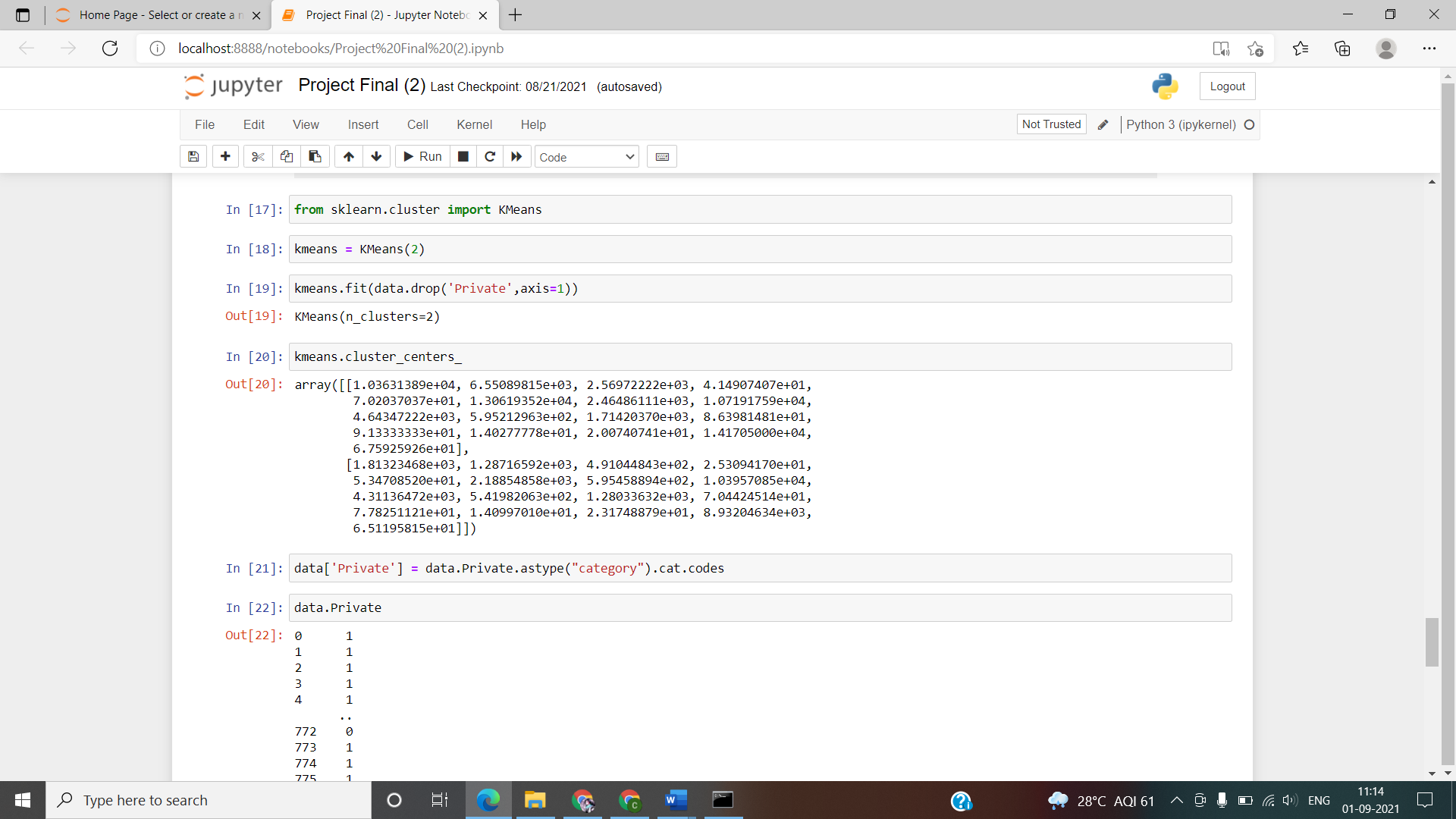
**Chapter 4**

**METHODS AND ALGORITHMS USED**

**4.1 KMeans Clustering**

K-means clustering is one of the simplest and popular unsupervised machine learning algorithms. K-means clustering is a method of vector quantization, originally from signal processing, that aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. This results in a partitioning of the data space into Voronoi cells k-means clustering minimizes within-cluster variances.

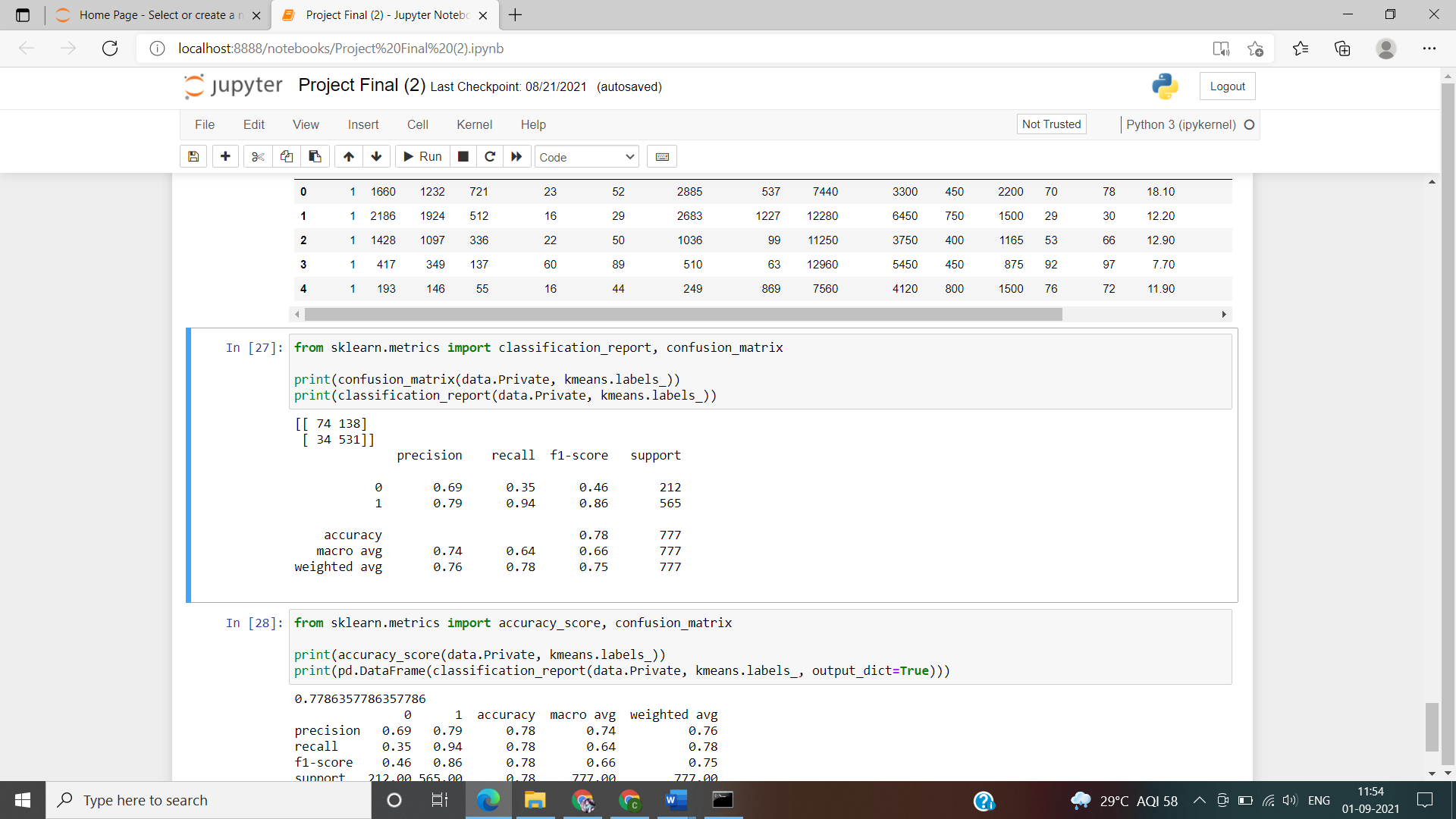
After importing K Means we will get array like

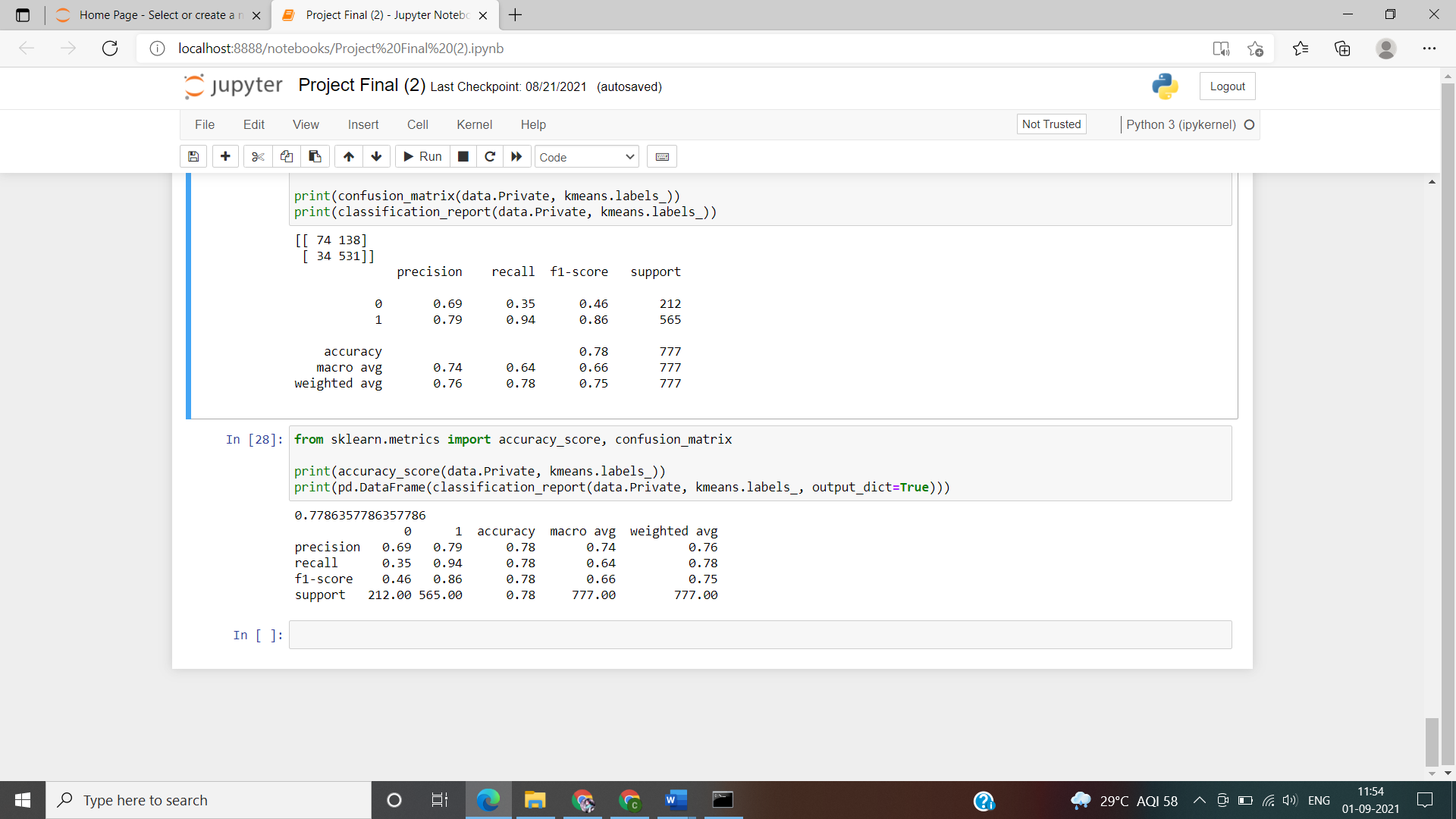


**Chapter 5**

**Results**

In this section we will see our results that how our model is predicting what is its accuracy and all.





From the above images we can say that our model is showing 0.7786 which is good accuracy.

**Chapter 6**

**LIBRARIES USED**

1. Numpy
2. Pandas
3. Seaborn
4. Matplotlib
5. Sklearn

**Chapter 7**

**Conclusion**

From this we can say that our model has successfully distinguish the public and government college in the various aspects. As our model is 77% accurate which we used K means Clustering.

**References**

[www.geeksforgeeks.org](http://www.geeksforgeeks.org)

[www.youtube.com](http://www.youtube.com)