**B.TECH. (2020-24)**

**Artificial Intelligence**

**Open Ended**

**LAB File**

on

**Fundamentals of Machine Learning**

**[CSE313]**

**Logo

Description automatically generated**

Submitted To

**Dr Monika Arora**

Submitted By

|  |  |  |
| --- | --- | --- |
| **Hitesh** | **A023119820027** | **5AI 1** |
| **Gauri Tyagi** | **A023119820028** | **5AI 1** |
| **Kushagra Dubey** | **A023119820029** | **5AI 1** |
| **BHOOMIKA SHARMA** | **A023119820030** | **5AI 1** |
| **SUNIDHI SINGH** | **A023119820032** | **5AI 1** |

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

AMITY SCHOOL OF ENGINEERING AND TECHNOLOGY

AMITY UNIVERSITY UTTAR PRADESH

NOIDA (U.P)

**OPEN ENDED EXPERIMENT**

**Aim**

To implement k means clustering algorithm over a dataset

**Software Used**

Google Colab

**Program Code and Output**





Table

Description automatically generated

Table

Description automatically generated

Graphical user interface, table

Description automatically generated with medium confidence

Table

Description automatically generated

Table

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated



Table

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

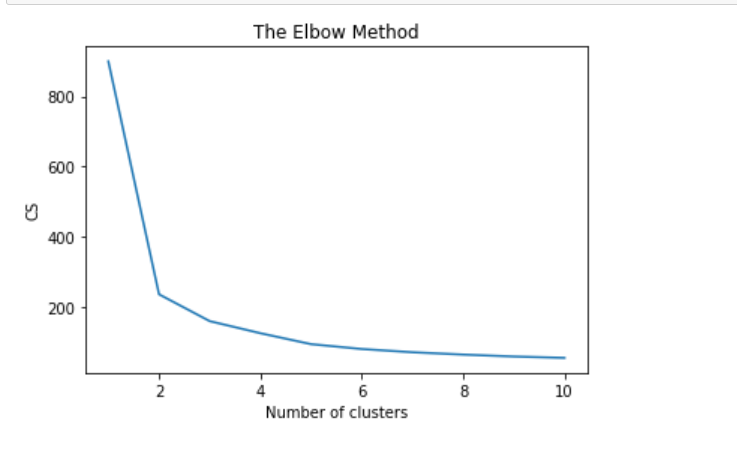
Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

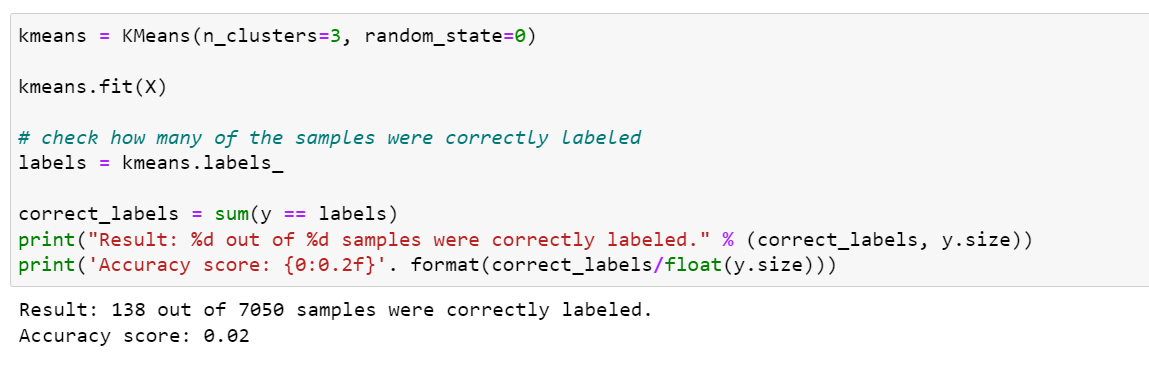
Text

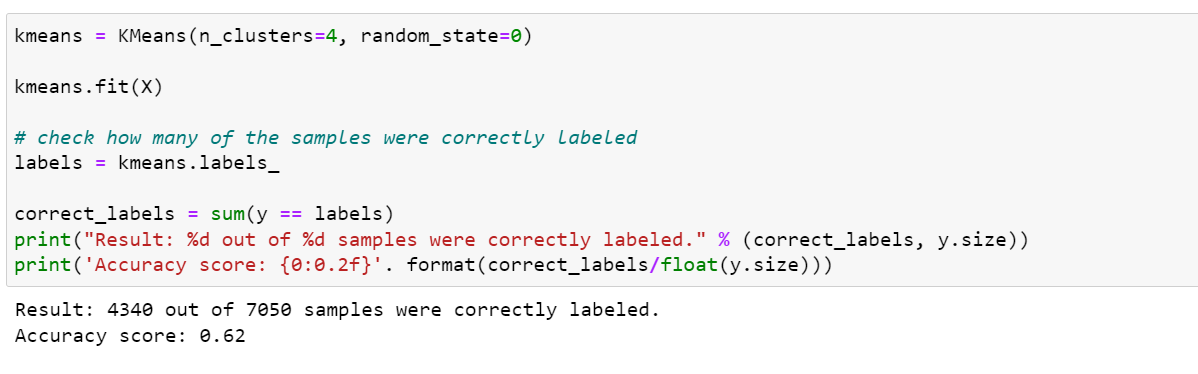
Description automatically generated with medium confidence



Graphical user interface, text, application, email

Description automatically generated





Graphical user interface, text, application, email

Description automatically generated

**Discussion and Conclusion**

The lesser the model inertia, the better the model fit. So, we use the Use elbow method to find optimal number of clusters. There is a kink at k=2.Hence k=2 can be considered a good number of the cluster to cluster this data. So, we have changed the value of k and found relatively higher classification accuracy of 62% with k=4. Hence, we can conclude that k=4 being the optimal number of clusters.

