UNIVERSITY RECOMMENDATION SYSTYM

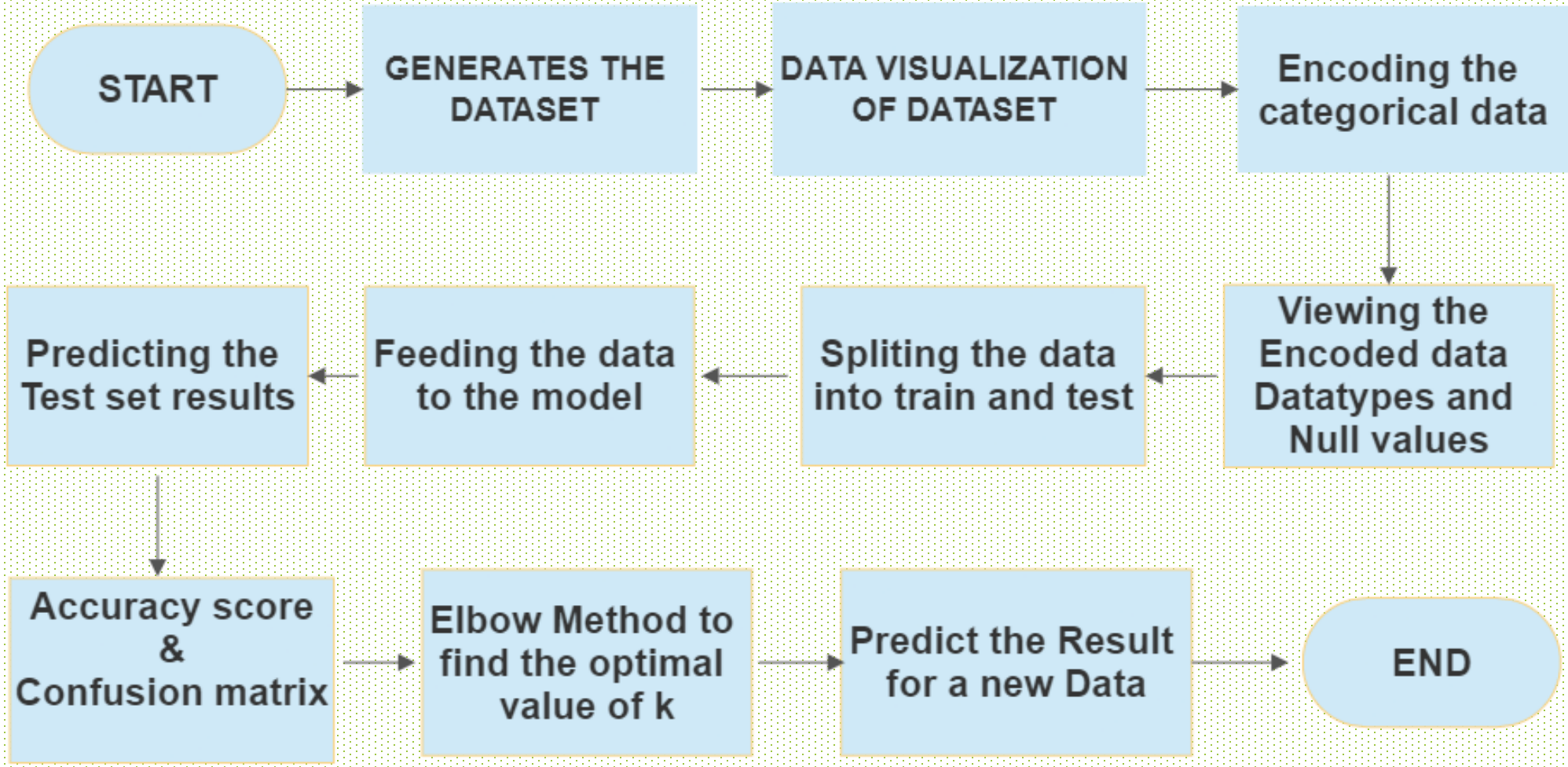
USING K NEAREST NEIGHBOR ALGORITHM

INDEX

PROPOSED SYSTEM

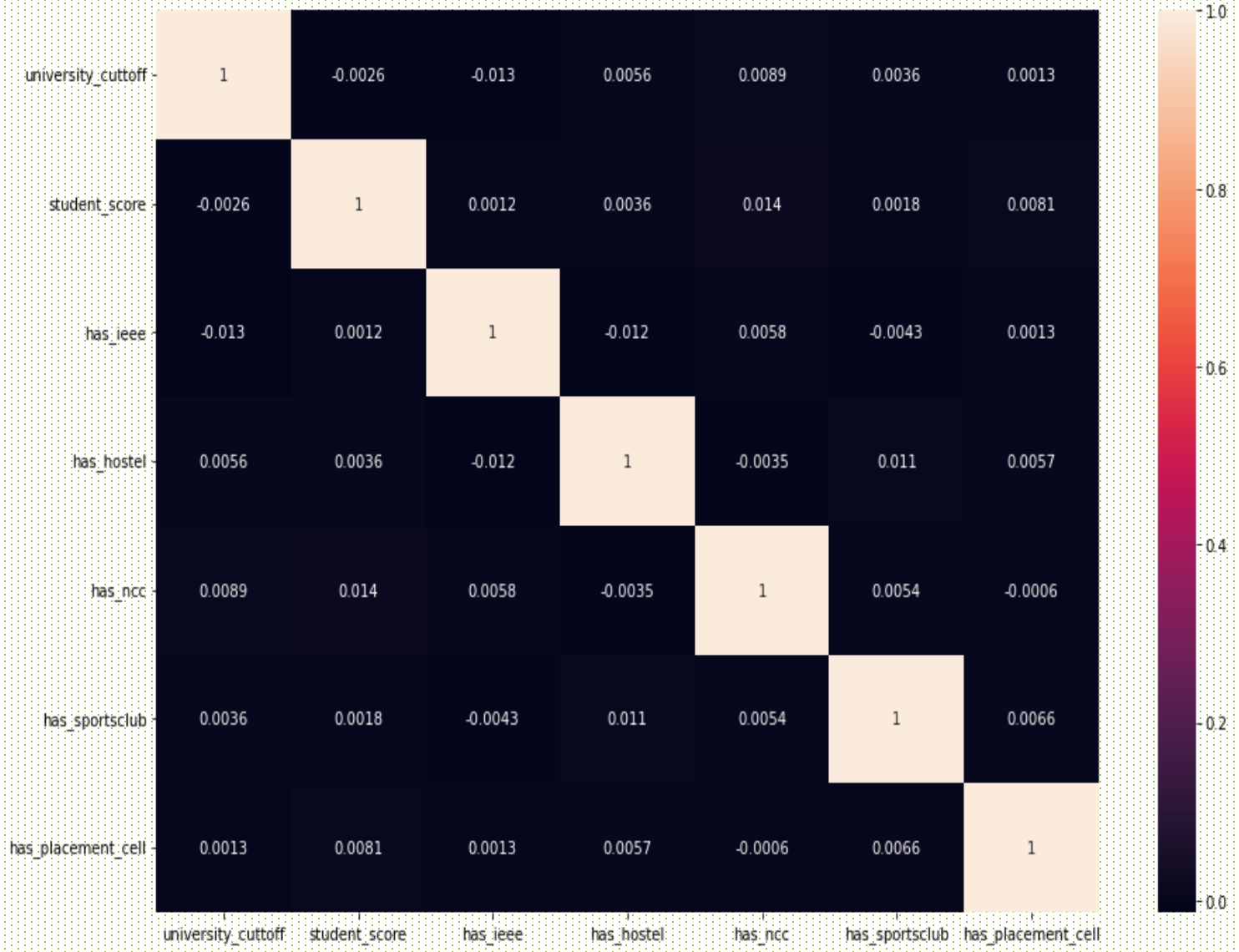
- The task is to create a model that recommends a list of universities based on the new data.
- Model trains on the dataset of universities.
- Model works on k nearest neighbour algorithm of supervised machine learning.

SYSTEM FLOWCHART



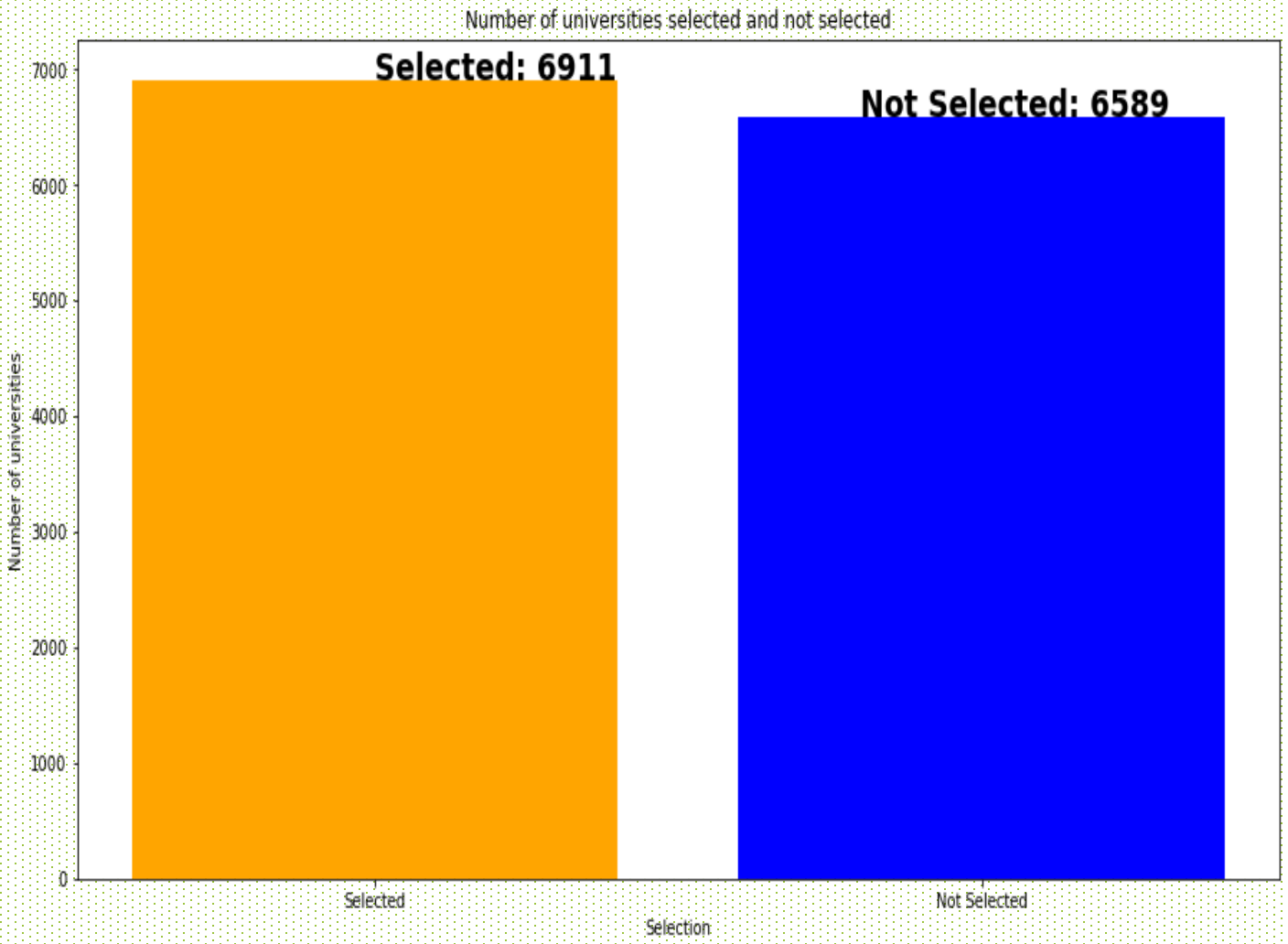
ANALYTICAL INFORMATION

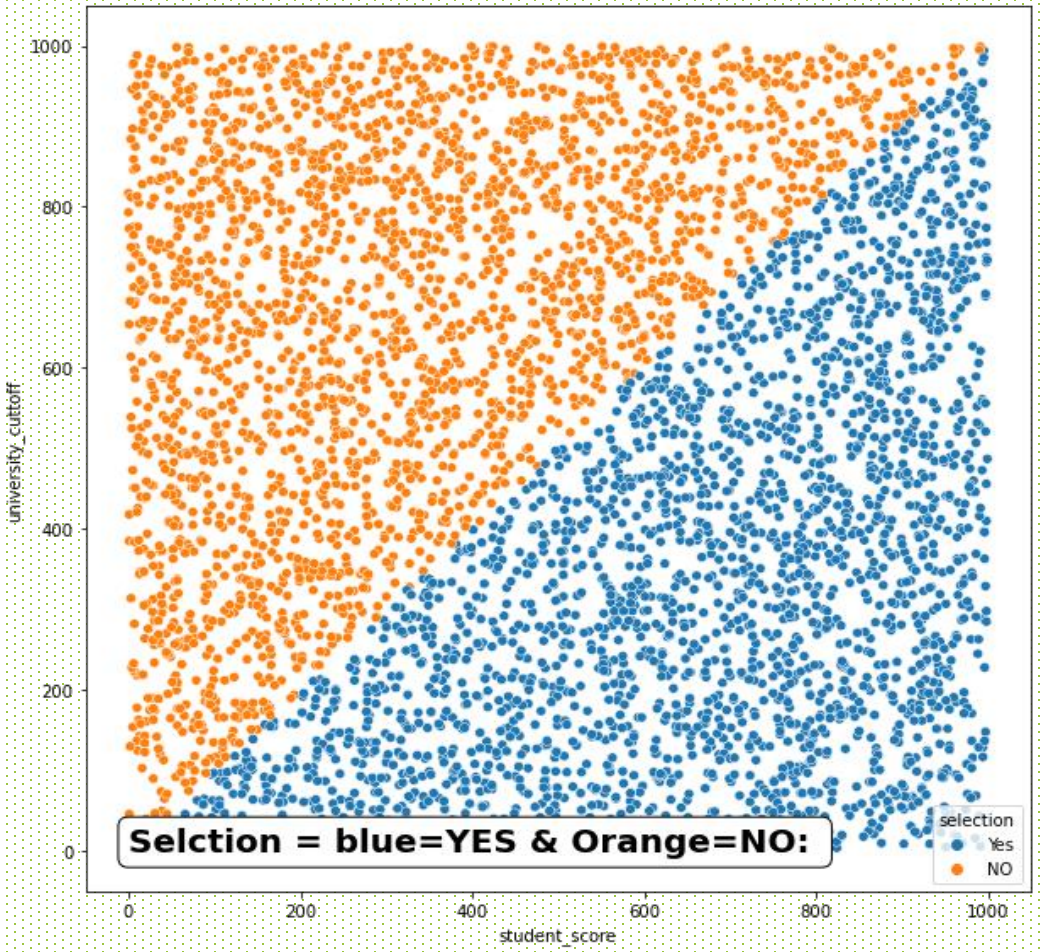
Correlation of Attributes in the Dataset



VISUALIZATION

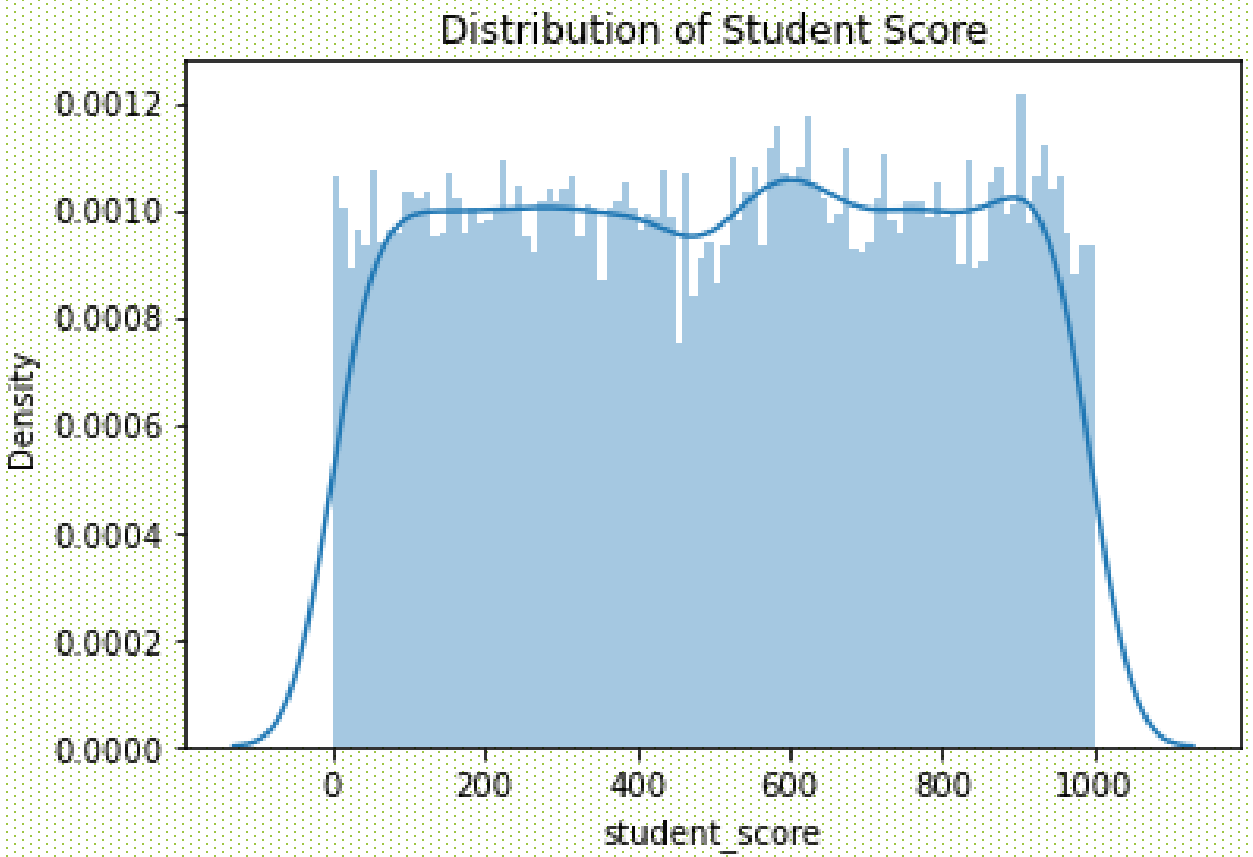
NUMBER OF UNIVERSITY SELECTED AND NOT SELECTED

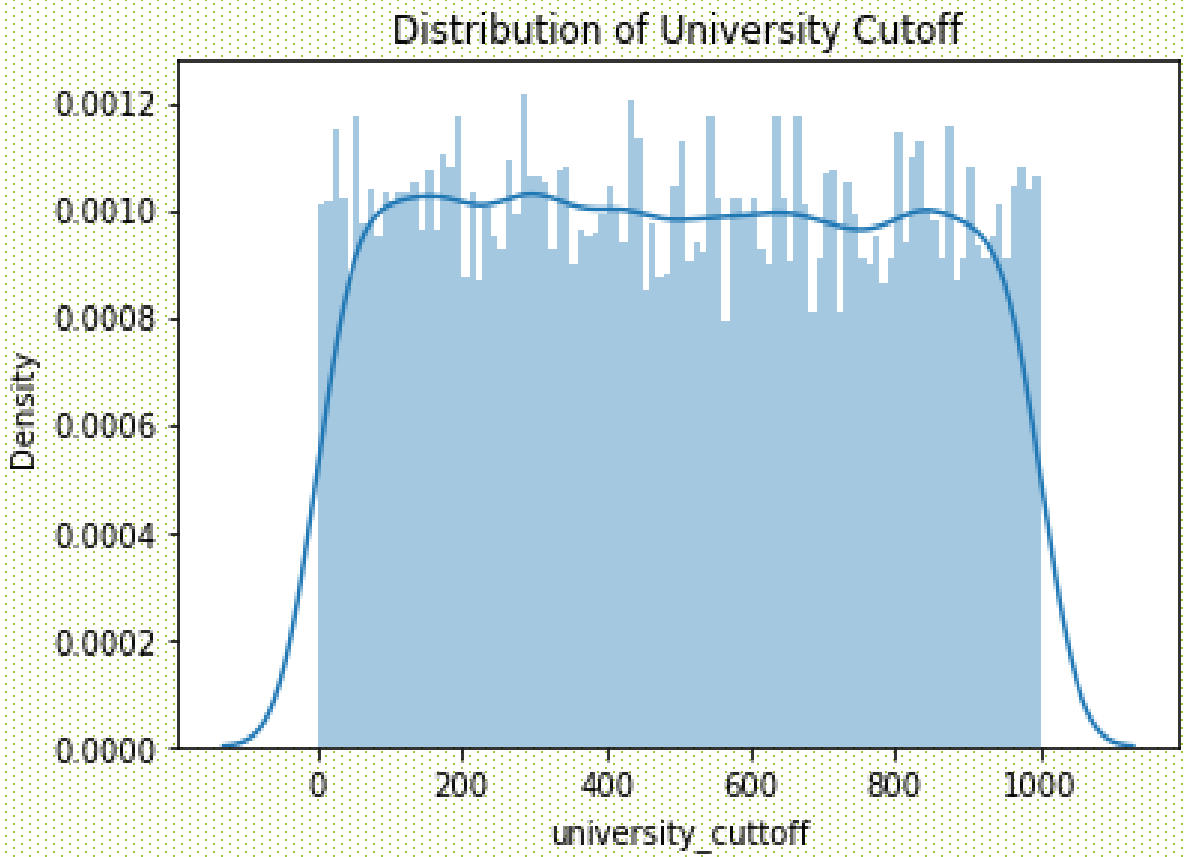




VISUALIZATION

DISTPLOT USING SEABORN LIBRARY



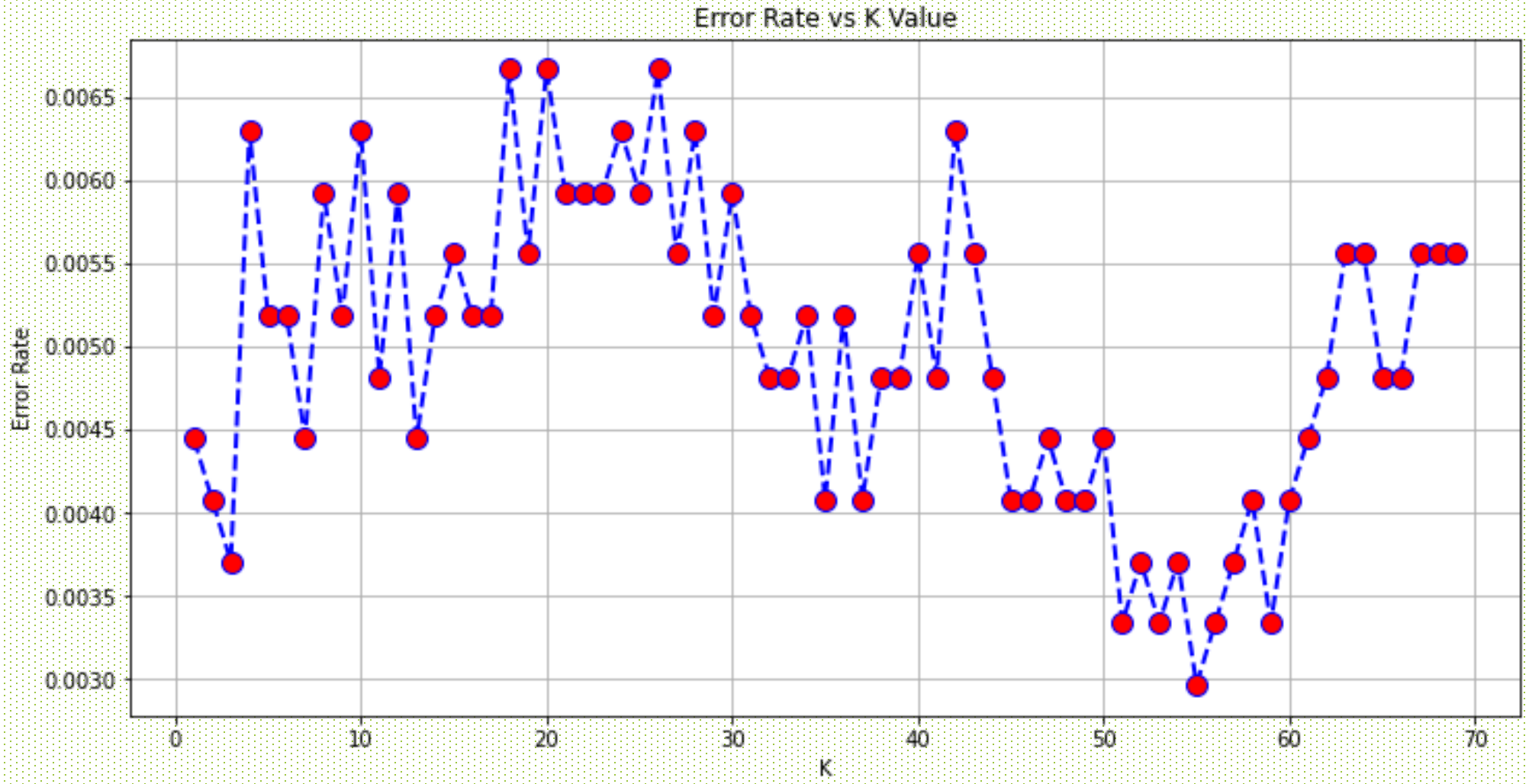


INFERENCE & INSIGHTS

- Model works on KNN classification algorithm.
- The First step is to encode the categorical data into integers
- The Second step is to split the data into train and test and have to select the features and label from the dataset
- Mostly the splitting is done in the ratio of 70:30 % or 80:20%(i.e train and test data)
- Then feed the splitted data into the model with the help of classifier and neighbors
- The main job is to choose the number of k values (i.e. nearest neighbors), and the Elbow method is used to obtain the optimal K value.
- Now predict test set results from the model and find the Accuracy Score, F1 Score, Precision & Recall Score and confusion matrix.
- The final step is to enter the new data into the model, after which the model will recommend the best universities based on the new data.

ELBOW METHOD FOR FINDING THE BEST K VALUE(NEAREST NEIGHBORS)

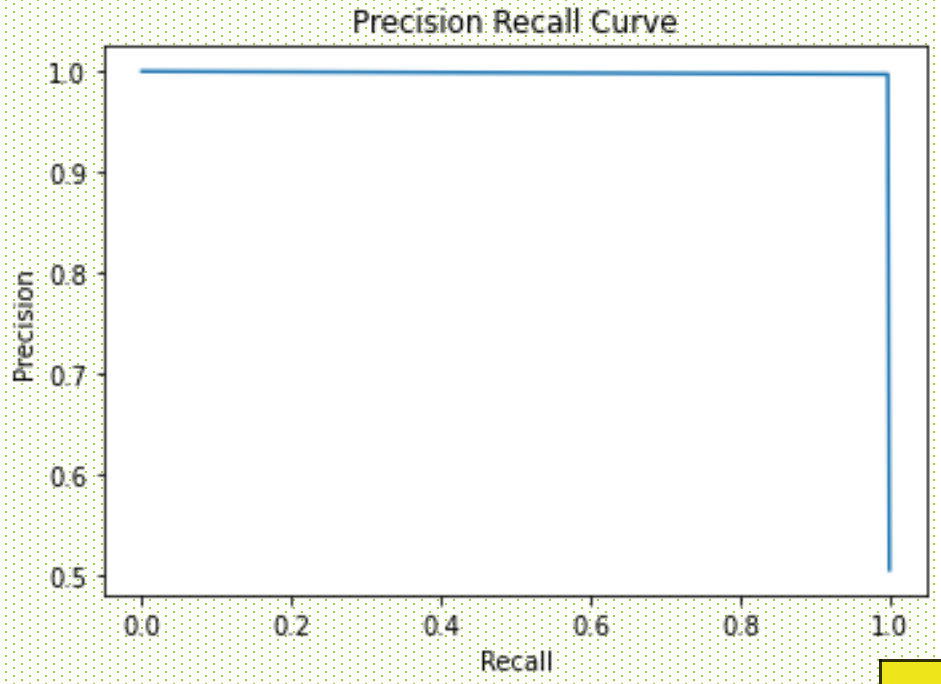
THE MARKER CLOSEST TO THE X AXIS HAS LEAST ERROR RATE
SO THAT MARKER IS BEST SUITABLE VALUE FOR K

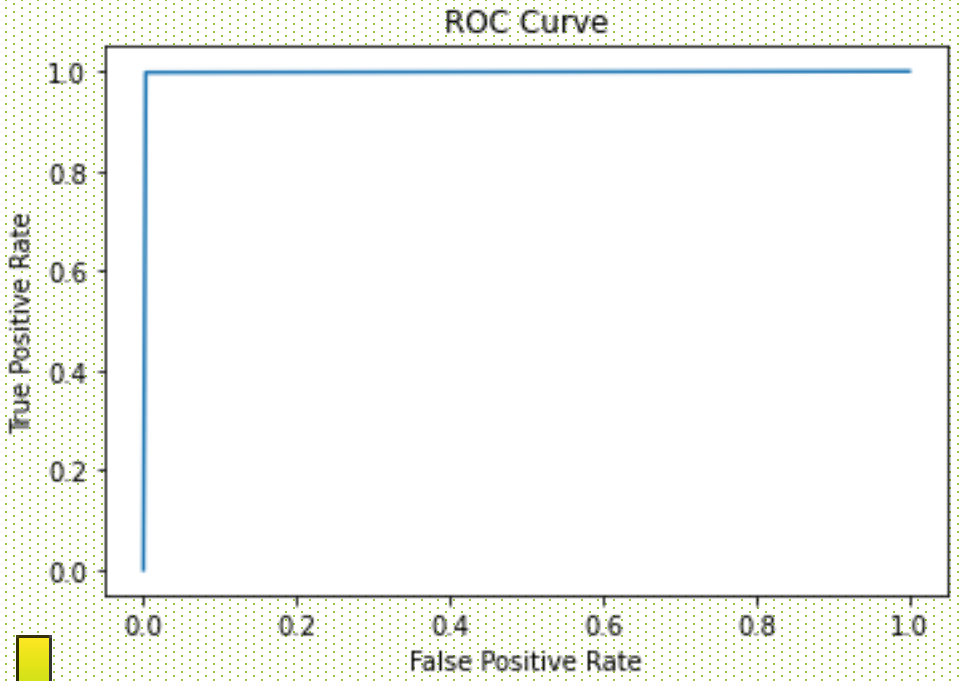


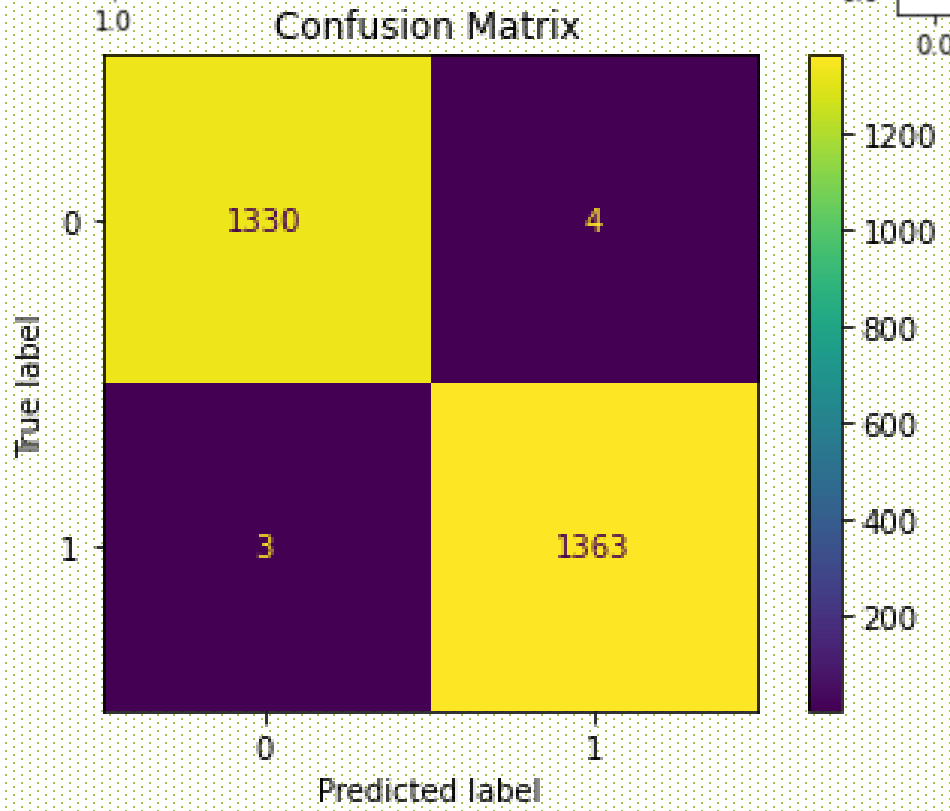
Accuracy Score of the model is 99.74074075%

Classification Report of the model

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1334
1	1.00	1.00	1.00	1366
accuracy			1.00	2700
macro avg	1.00	1.00	1.00	2700
weighted avg	1.00	1.00	1.00	2700


Precision Recall Curve
Precision
Recall
1.0
0.9
0.8
0.7
0.6
0.5
0.0
0.2
0.4
0.6
0.8
1.0


ROC Curve
True Positive Rate
False Positive Rate
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0


Confusion Matrix
True label
Predicted label
0
1
1330
4
3
1363
1200
1000
800
600
400
200

THANK YOU!

PROJECT BY
RUSHI HITESH PRAJAPATI