FutureFit

CSE 427 Group Presentation

Group no. 12

MEET THE TEAM

Habibun Nabi Hemel

22241042

Ishrak Hamim Mahi

22101792

AGENDA

INTRODUCTION

In order to create a college-level placement management system, one of our goals is to create a placement predictor that assists in improving students' skills prior to the commencement of the recruitment process and forecasts the likelihood that they will be placed. Machine learning is what we're using to anticipate placement. In order to place students into the proper clusters and assist them improve their profiles, we take into account K-nearest neighbor (KNN), Support Vector Machine (SVM), Logistic Regression, and Random Forest.The accuracy of reputable algorithms is recognized, and by comparing different machine learning techniques, recruiters and students can benefit from each other's assistance during placements and related activities

Dataset Description

1

The dataset is a synthetic dataset with 220000 rows and 19 columns.

2

No work has been done with this dataset before. The usability score of the dataset was more than 8.

3

There were enough room to preprocess the data. There were several rows with numerical data and categorical data.

4

The target variable of the dataset was Placement Package. This was a continuous data. We had converted it into a classification task later.

Literature Review

1. Upon going through other papers we have come to the conclusion that all the work were performed with a limited amount of data.

2. In the papers the researchers have used several ML and NN models. Among the models Random Forest and Logistic Regressioon worked the best in most of the cases .

3. None of the workings included any implementation of the model.
with a GUI platform.

Methodology

We have followed 4 steps to prepare the project.

1. EDA

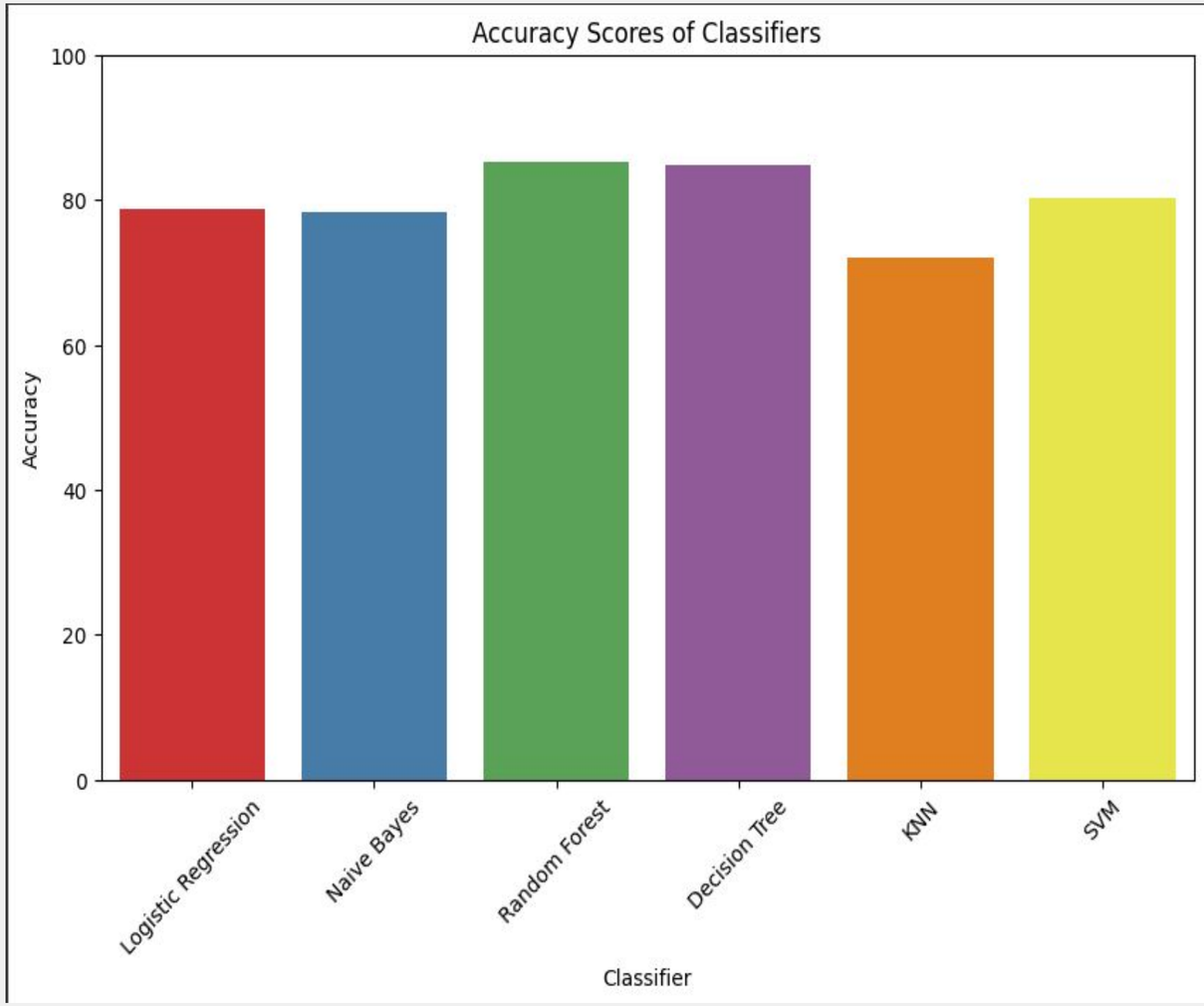
2. Data Pre Processing

3. Model Trainig

4. Cloud Hosting

Result Analysis

The table summarizes the accuracy scores of various machine learning models. Both Decision Tree and Random Forest models achieve high accuracy rates, with Random Forest slightly outperforming Decision Tree. Logistic Regression and Naive Bayes models exhibit lower but still respectable accuracy scores. However, the KNN model demonstrates the lowest accuracy among the models listed. Overall, Random Forest appears to be the most accurate model among those evaluated



CONCLUSION

Placement prediction system is a system which predicts the placement status of final year BSC students. For data analysis and prediction different machine learning algorithms are used in the python environment. We analyse the accuracy of different algorithms and it is shown in the above table. It is clear that Random Forest gives an accuracy of **85.13%** . We believe the room for improvement is still open. In our future work we will be trying some Neural Network based algorithms. We are also planning to try a Logistic Regression based Ensemble Algorithm. For our GUI Web Application, we will be adding the feature where the users can define the thresholds of different classes.

THANK YOU!

Useful Links

Click on the icons

kaggle

colab

Overleaf