PREDICTION OF GRADUATE ADMISSION

UNDER THE GUIDANCE OF

Dr. SHIVANJALI KHARE

DATA MINING

CSCI-6674-01

Team Miners

TEAM MEMBERS

PRASANTH REDDY THUMMA (pthum2@unh.newhaven.edu)

SUMA MODUGU (smodu1@unh.newhaven.edu)

MOHAMMAD NAZEER (mashail4@unh.newhaven.edu)

Table of contents

1. **Abstract**
2. **Introduction**
3. **Related work**
4. **The proposed method**
5. **The experimental results**
6. **Discussion - Conclusion and future work**
7. **Appendix for link to the GitHub repository**
8. **References - Proofreading with an email from Writing Center**

1. ABSTRACT

In academic institutions, the issue of student admission is particularly essential. This research looks at machine learning models for predicting a student's chances of being accepted into a master's degree. Students will be able to see ahead of time if they have a chance of being admitted. Multiple linear regression, k-nearest neighbor, random forest, and Multilayer Perceptron are the machine learning models. Experiments have shown that the Multilayer Perceptron model outperforms the competition.

2. INTRODUCTION

The world marketplaces are fast evolving, and people are constantly seeking the greatest information and experience.
Young individuals who want to make a name for themselves in the workplace are always on the lookout for higher degrees that can help them advance their abilities and knowledge. As a result, in the recent decade, the number of students applying for graduate school has increased. This reality has prompted us to research student grades and the likelihood of admission to master's programs in order to assist universities in forecasting the likelihood of accepting master's students who apply each year and providing the necessary resources.

Education is extremely important in one's life. The scope of the higher educational model is more challenging now a days and it is the topic of discussion across many of the students who are willing to get a good and reputed university or a college in order to complete their higher studies. Several problems and misguidances by third party consultants have improved a lot in which students are missing right place and right degree or course and enrolling with many dissatisfactions have identified in many cases. Many students are trying to complete their higher studies at U.S.A and majority of the immigrants are flying from India and China, every decade the number of students pursuing master at U.S.A gradually increases from year to year.

In general there is lack of knowledge in understanding the requirements to get the right admission is challenging to most of the students. The categories of the universities, eligibility and cost of the fee are different to different universities and colleges makes students very ambiguity in pertaining the right admission at right college. The admission is highly dependent on various factors like whether a college is public or govt undertaken, private universities along with few universities are offering funding and scholarships based on the skills of an individual student.

According to a study in the year 2018-2019 there was more than 1Million students have enrolled and got admission across U.S.A based various state and private universities by the immigrants. IIE reported to several education departments of the united stated, there should be a web application providing all the details of the universities and their regulations along with several requirements which needed to be addressed in order to get an admission and which should help student in order to get the right place for study in feasible and also helps student not to spend more cost on admission based counselling to the third party services.

This research mainly focus on the prediction of the chance of admission by using various data mining algorithms in order to undertake the provided dataset with preprocessing and avoiding the non key attributes of the dataset using data munging. Decision making techniques are used integrating with well defined data mining algorithms will provide applicants decision regarding the chance of admission.

3. RELATED WORK

[1] Abdul Fatah S. "Hybrid Recommender System for Predicting College Admission" 2012. In this paper for tackling college admissions prediction challenges, this work introduces a new hybrid recommender based on data mining approaches and knowledge discovery principles.

[2] Bibodi, J., Vadodaria, A., Rawat, A.. and Patel, J. (n.d.). "Admission Prediction System Using Data mining". The purpose of this research is to identify the elements that influence and guide students while choosing an engineering institution for their first year of study.

[3] Mane, R. V "Predicting Student Admission decisions byAssociation Rule Mining with Pattern Growth Approach" 2016.

[4] (Yuling Lin et al., 2020) used SVM for Stock Market Trend Prediction in two phases, one being feature selection and other prediction. They showed that the proposed SVM-based system can find out a good subset and evaluate stock indicators which provide useful information for investors.

[5] (Manisha Mudgal, 2020) has used the sentimental and technical analysis for predicting the stocks. He collected the dataset from twitter and other social media platforms to analyze the mood of market for giving more accurate and better prediction.

4. PROPOSED METHOD

The complete implementation is based on python using jupyter notebook which implicitly provides kernel connector with python interpreter. This is highly applicable in reading the datasets of various formats using ipynb files.

Several data structures are used in order apply various operations on data sets like slicing, zipping, merging, filtering which are used in categorizing the data based on the required parameters. Jupyter notebook can be accessed with any web browser like chrome, firefox and safari, this environment is provided with rich set of deployment techniques based on the cells used in the notebook and able to get the results in the form of tabular structure and graphs for data visualization. Kernel connector.py is provided to interpret the code provided in the jupyter notebook, this kernel connector executes and evaluates the code cells.

Python provides vast set of libraries in order which are need for classification of data, applying various techniques, generated results in the form of graphs, mean square error, accuracy calculation are highly used for prediction based results provided in the data mining techniques.

This complete project is evaluated and executed using python notebook and with local data. If the data set size is gradually increased it can also be deployed using google collab and using big data. Based on the size of the data set we can use different environments in order to predict the results of the dataset provided.

In this project, Dataset from Kaggle is used. The dataset is composed up of marks from 400 records and 9 columns. The data consists of student's GRE score, TOEFL score, CGPA, University Rating, SOP, LOR, Research Experience and Chance of Admit. The data's main purpose is to predict the chance of admission of a student to pursue their master's degree.

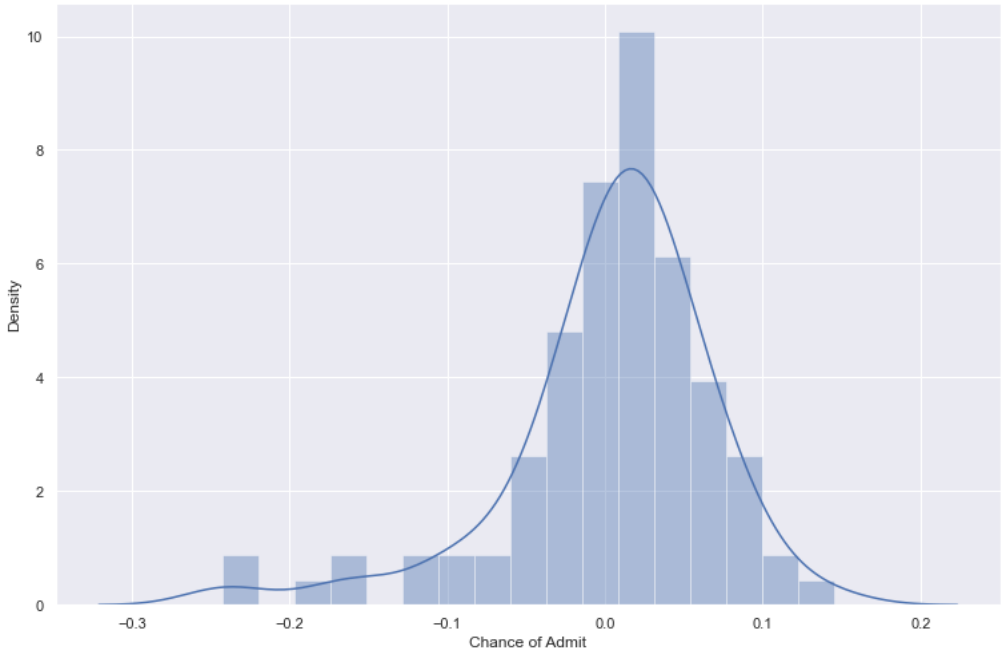
The key objective of the proposed system is to provide end to end guidance to the students aspiring and pursue their masters and higher education across united states.

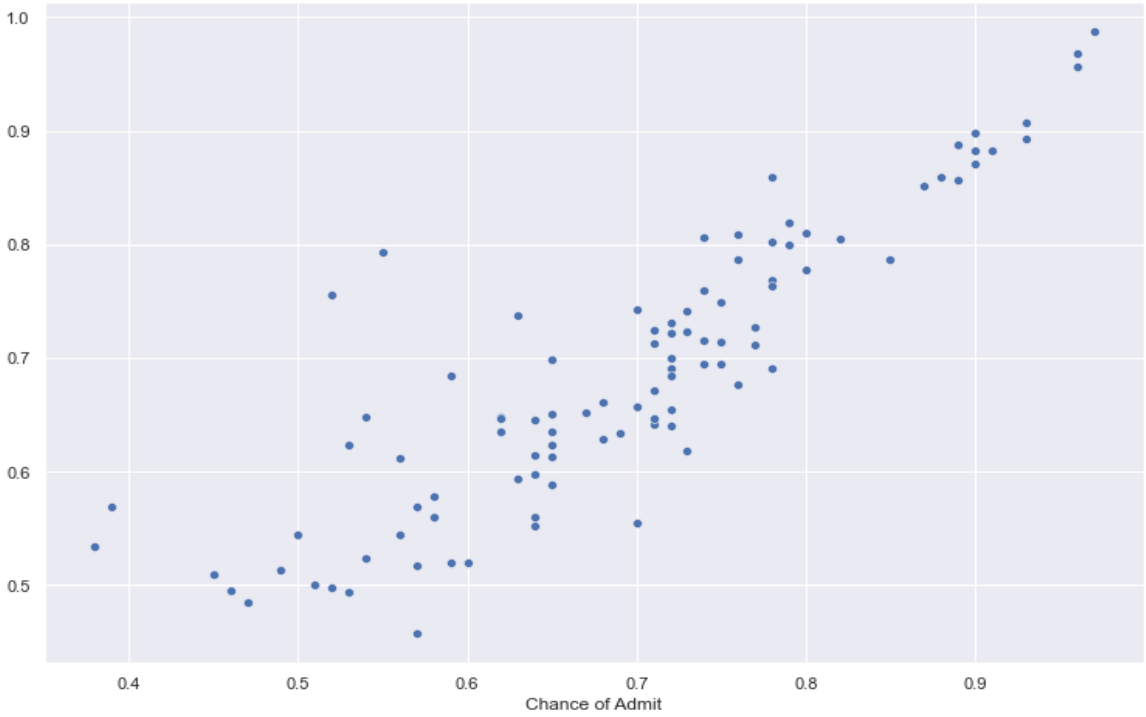
Without depending on the educational consultancy how the students can opt this model based web sites are hosted by considering the 85% of the universities across the country. It provides the easiest steps in selecting the type of study and requirements to be provided in order to get the chance of getting admission.

The chance criteria is based on the GRE, TOEFL and CGPA scores. Other parameters are considered in order to predict the chance of admission based on research experience, SOP and LOR of the student.

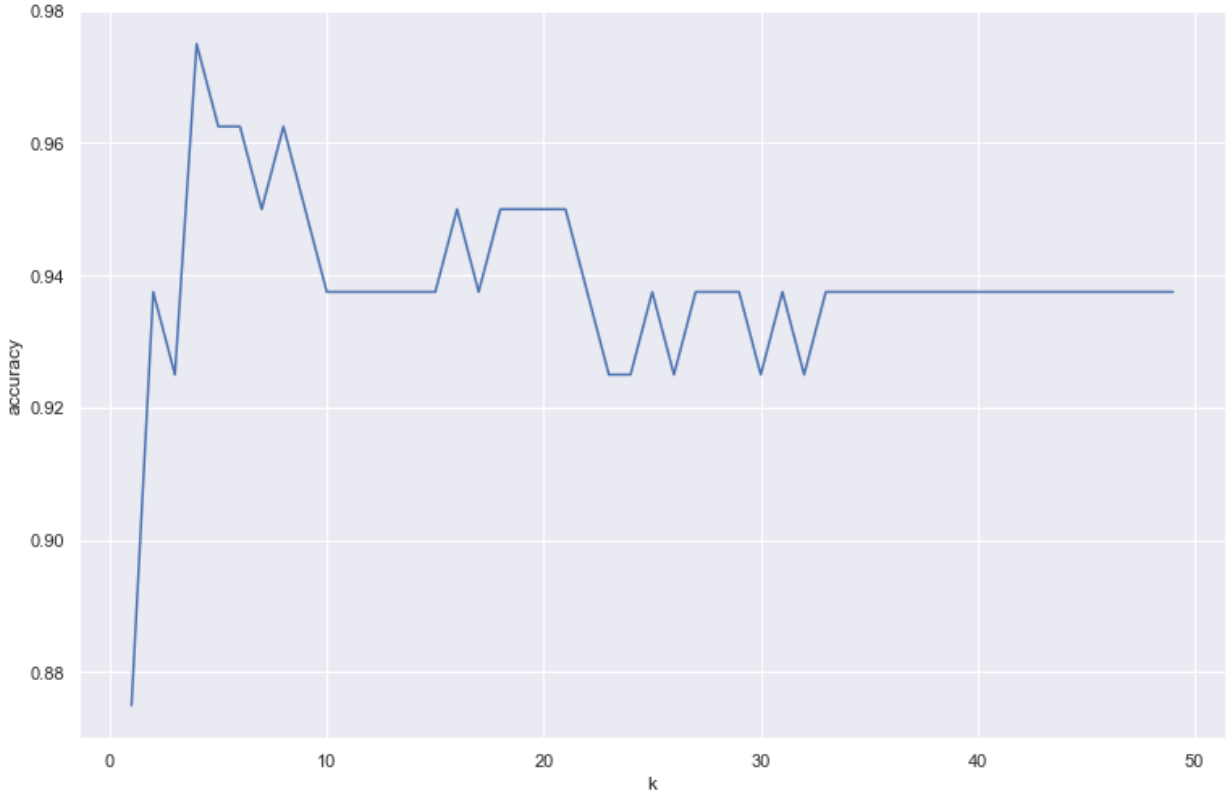
5. EXPERIMENTAL RESULTS

The following are the results of the work:

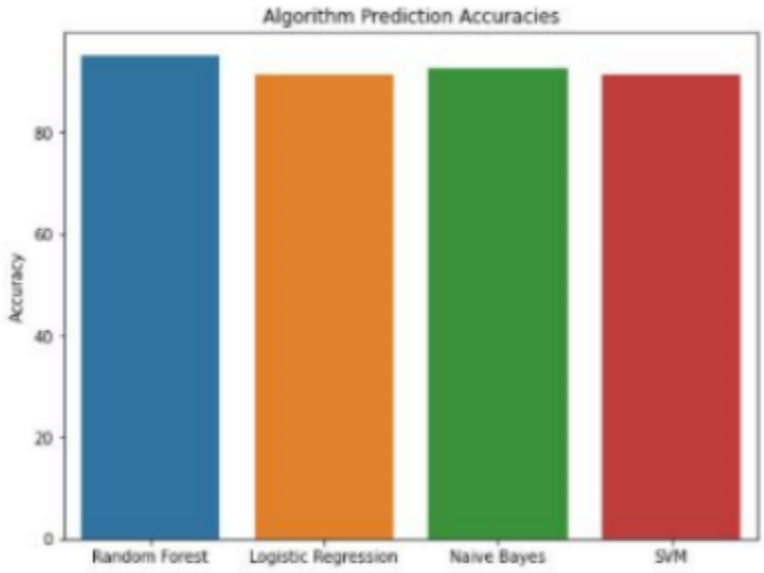




Accuracy plot:



Comparison Graph of algorithms



6. Discussion

In this project, Dataset from Kaggle is used. The dataset is composed up of marks from 400 records and 9 columns. The data consists of student's GRE score, TOEFL score, CGPA, University Rating, SOP, LOR, Research Experience and Chance of Admit. The data's main purpose is to predict the chance of admission of a student to pursue their master's degree.

The key objective of the proposed system is to provide end to end guidance to the students aspiring and pursue their masters and higher education across united states.

Without depending on the educational consultancy how the students can opt this model based web sites are hosted by considering the 85% of the universities across the country. It provides the easiest steps in selecting the type of study and requirements to be provided in order to get the chance of getting admission.

The chance criteria is based on the GRE, TOEFL and CGPA scores. Other parameters are considered in order to predict the chance of admission based on research experience, SOP and LOR of the student.

7. CONCLUSION AND FUTURE WORK

We can see that the Random Forest has the highest accuracy with 95% from the obtained results which is very impressive.

Automating the manual scoring parameters would save a lot of resources i.e training the model to give grades to LOR and SOP. Optimizing the model to have more accuracy which should increase the precision of the outcome.

This research project gives an overview of the predicting the chance of admission of a student willing to pursue master's degree across U.S.A based on various requirements by implementing different data mining algorithms. The techniques used in project are Logistic Regression, Naive Bayes, Random Forest and Support Vector Machine. Among all the models used, Random Forest gives the best accuracy to predict the chance of a student getting admitted.

8. Appendix for link to the GitHub repository

https://github.com/prasanth50/Phase6

9.REFERENCES

[1] Abdul Fatah S. “Hybrid Recommender System for Predicting College Admission” 2012. In this paper for tackling college admissions prediction challenges, this work introduces a new hybrid recommender based on data mining approaches and knowledge discovery principles.

[2] Bibodi, J., Vadodaria, A., Rawat, A.. and Patel, J. (n.d.). “Admission Prediction System Using Data mining”. The purpose of this research is to identify the elements that influence and guide students while choosing an engineering institution for their first year of study.

[3] Mane, R. V “Predicting Student Admission decisions byAssociation Rule Mining with Pattern Growth Approach” 2016.

[4] (Yuling Lin et al., 2020) used SVM for Stock Market Trend Prediction in two phases, one being feature selection and other prediction. They showed that the proposed SVM-based system can find out a good subset and evaluate stock indicators which provide useful information for investors.

[5] (Manisha Mudgal, 2020) has used the sentimental and technical analysis for predicting the stocks. He collected the dataset from twitter and other social media platforms to analyze the mood of market for giving more accurate and better prediction.

10. Proofreading with an email from Writing Center

RE: Proofreading with an email from Writing Center

TR Thumma, Prasanth Kumar Reddy <pthum2@unh.newhaven.edu>
2:55 PM

To: Modugu, Suma Cc: Nassar, Mohamad

final datamining document.docx
246.23 KB

Sent from Mail for Windows

From: Thumma, Prasanth Kumar Reddy
Sent: Wednesday, May 4, 2022 2:53 PM
To: Modugu, Suma
Cc: Nassar, Mohamad
Subject: Proofreading with an email from Writing Center

Hello Suma & Nazzer, Please find the following attachment for the final report and let me know for any changes and suggestions.

Thanks & Regards,
Prasanth.

Sent from Mail for Windows