HINDUSTHAN INSTITUTE OF TECHNOLOGY, COIMBATORE

DEPARTMENT OF INFORMATION TECHNOLOGY

Project Based Experiential Learning Program

(Nalaiya Thiran)

UNIVERSITY ADMIT ELIGIBILITY PREDICTOR

BATCH　　**:** 2019-2023

YEAR　　　: FINAL YEAR

SEMESTER : VII

TEAM MENTOR　　: MRS.G.KAVIPRIYA

TEAM LEADER　　**:** MS.RAJALAKSHMI.H　　　(720819205040)

TEAM MEMBERS　**:** MS.AKSHARA NARAYANAN　(720819205003)

　　　　　　　　　　MR.KAVIYARASAN.G　　　(720819205020)

　　　　　　　　　　MR.RAGUL.E　　　　　(720819205038)

TEAM ID　　　　**:** B4-4M6E

INTRODUCTION

Project Overview

We know after 12th board results, the main problem of the students is to find an appropriate college for their future education. It is a tough decision to make for many students as to which college they should apply to. We have built a system that compares the students data with past admission data and suggests college in a sequence of their preference. The purpose of this analysis is to demonstrate the top contributing scores which helps the student to get the admission into the Master's degree program. What factors contributes to successful admission to a Master's degree program? We have used XGBOOST classifier, Decision tree classifier, etc. as our statistical model to predict the probability of getting admission to college. It was observed that the performance of XGBOOST was achieved highest among all. The Education Based Prediction System helps a person decide what colleges they can apply to with their scores. The dataset that is used for processing consists of the following parameters: University name, Quants and Verbal Scores (GRE) TOEFL and AWA Scores. The GRE Test (Graduate Record Examinations) is a standardized test used by many universities and graduate schools around the world as part of the graduate admissions process. Majority of universities in the USA follow similar guidelines for providing admission to students.

Purpose

A persons education plays a vital role in their life. While planning for education students often have several questions regarding the courses, universities, job opportunities, expenses involved, etc. Securing admission in their dream university is one of their main concerns. It is seen that often students prefer to pursue their education from universities which have global recognition. And when it comes to international students the first choice of the majority of them is the United States of America. With the majority of worlds highly reputed universities, wide range of courses offered in every sector, highly accredited education system and teaching, scholarships provided to students, best job market and many more advantages make it the dream destination for the international students. According to research, there are above 8 Million international students studying in more than 1700 public and 2500 private universities and colleges. The primary objective of this research is to develop a system to solve the problems the international students are facing while applying for universities. We will be developing a Student Admission Predictor (SAP) system which will help the students to predict the chances of their application being selected for a particular university for which they wish to apply based on their profile. Also, the system will provide a recommendation of universities to the student to which the student has a high possibility of getting admission.

LITERATURE SURVEY

Existing problem

1. (Mishra and Sahoo (2016)) conducted a research from a university point of view to predict the likelihood of a student enrolling in the university. They used K-means algorithm.

Disadvantage

The objective of the model was to increase the enrolment of the students in the university and not for helping students.

2. GRADE system was developed by (Waters and Miikkulainen (2013)) to support the admission process for the graduate students in the University of Texas Austin Department of Computer Science.

Disadvantage

The model was created using the Multiple Logistic regression algorithm, it was able to achieve accuracy rate of 67% only.

3. (Eberle et al. (n.d.)) used machine learning and predictive modelling to develop a model that to evaluate the admission policies and standards in the Tennessee Tech University.

Disadvantage

The model worked well in predicting the true positive scenarios where the student was had good profile to secure the admission, but it failed in efficiently identifying the true negatives because of which student who does not satisfy the defined criteria.

4. Bayesian Networks were used by (Thi et al. (2007)) to create a decision support system for evaluating the application submitted by international students in the university.

Disadvantage

The comparisons were made only with the students who were already admitted in the university and the data of the students who were denied admission were not included in the research this model proved to be less efficient due to the problem of class imbalance.

5. (Bibodi et al. (n.d.)) used multiple machine learning models to create a system that would help the students to shortlist the universities suitable for them also a second model was created to help the colleges to decide on enrolment of the student.

Disadvantage

It did only relied on the GRE, TOEFL and Undergraduate Score

of the student and missed on taking into consideration other important factors like SOP and LOR documents quality, past work experience, technical papers of the students etc.

References

Abdul Fatah S; M, A. H. (2012). Hybrid Recommender System for Predicting College Admission, pp. 107–113.

Bibodi, J., Vadodaria, A., Rawat, A. and Patel, J. (n.d.). Admission Prediction System Using Machine Learning.

Eberle, W., Simpson, E., Talbert, D., Roberts, L. and Pope, A. (n.d.). Using Machine Learning and Predictive Modeling to Assess Admission Policies and Standards.

Jamison, J. (2017). Applying Machine Learning to Predict Davidson College ' s Admissions Yield, pp. 765–766.

Mane, R. V. (2016). Predicting Student Admission decisions by Association Rule Mining with Pattern Growth Approach, pp. 202–207.

Master Portal (2017). MasterPortal.

URL: *http://www.mastersportal.eu/countries/82/united-states.html*

Mohan S Acharya, Dec 30, 2018, "A Comparison of Regression Models for Prediction of Graduate Admissions, IEEE International Conference on Computational Intelligence in Data Science 2019".
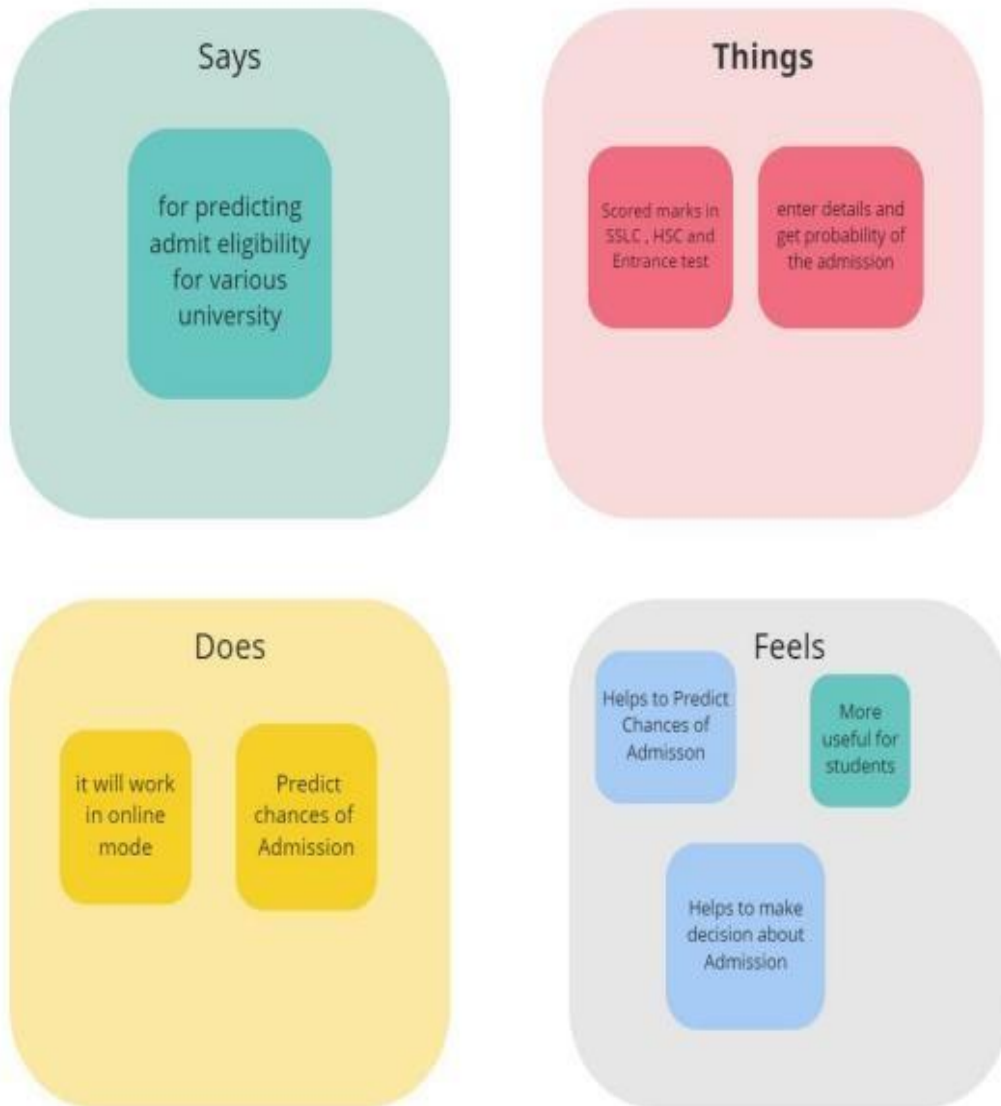
Problem Statement Definition

Educational organizations have always played an important and vital role in society for development and growth of any individual. There are different college prediction apps and websites being maintained contemporarily, but using them is tedious to some extent, due to the lack of articulate information regarding colleges, and the time consumed in searching the best deserving college. The problem statement, hence being tackled, is to design a college prediction/prediction system and to provide a probabilistic insight into college administration for overall rating, cut-offs of the colleges, admission intake and preferences of students. Also, it helps students avoid spending time and money on counsellor and stressful research related to finding a suitable college. It has always been a troublesome process for students in finding the perfect university and course for their further studies. At times they do know which stream they want to get into, but it is not easy for them to find colleges based on their academic marks and other performances. We aim to develop and provide a place which would give a probabilistic output as to how likely it is to get into a university given upon their details.

IDEATION & PROPOSED SOLUTION

Empathy Map Canvas

Says

for predicting admit eligibility for various university

Things

Scored marks in SSLC , HSC and Entrance test

enter details and get probability of the admission

Does

it will work in online mode

Predict chances of Admission

Feels

Helps to Predict Chances of Admisson

More useful for students

Helps to make decision about Admission

Ideation & Brainstorming

We know after 12th board results, the main problem of the students is to find an appropriate college for their future education. It is a tough decision to make for many students as to which college they should apply to. We have built a system that compares the students data with past admission data and suggests college in a sequence of their preference. We have used XGBOOST classifier, Decision tree classifier, etc. as our statistical model to predict the probability of getting admission to college. It was observed that the performance of XGBOOST was achieved highest among all.

Proposed Solution

           Education plays a vital role in today's era. While we talk about career - a person's degree, course, university and the knowledge that he possesses - is the key factor on which the firm hires a fresher. As soon as a student completes his/her Higher Secondary Schooling, the first goal of any student is to get into an appropriate College so that he can get a better education and guidance for his future. For that, students seek help from many sources like online sites or career experts to get the best options for their future. A good career counselor charges a huge amount for providing such solutions. Online sources are also not as reliable as the data from particular sources is not always accurate. Students also perform their analysis before applying to any institutions, but this method is slow and certainly not consistent for getting actual results and possibly includes human error. Since the number of applications in different universities for each year is way too high, there is a need to build up a system that is more accurate or precise to provide proper suggestions to students. Our aim is to use machine learning concepts to predict the probability of a student to get admission into those preferred colleges and suggest a list of collages in a sequence of the probability of getting admission to that specific college. The following are the steps that include the work we have done in sequence of implementation.

Problem Solution Fit

     Educational organizations have always played an important and vital role in society for development and growth of any individual. There are different college prediction apps and websites being maintained contemporarily, but using them is tedious to some extent, due to the lack of articulate information regarding colleges, and the time consumed in searching the best deserving college. The problem statement, hence being tackled, is to design a college prediction/prediction system and to provide a probabilistic insight into college administration for overall rating, cut-offs of the colleges, admission intake and preferences of students. Also, it helps students avoid spending time and money on counsellor and stressful research related to finding a suitable college. Our aim is to use machine learning concepts to predict the probability of a student to get admission into those preferred colleges and suggest a list of collages in a sequence of the probability of getting admission to that specific college. The following are the steps that include the work we have done in sequence of implementation.

REQUIREMENT ANALYSIS

Functional Requirement

In this stage of methodology the researcher identify major requirements pertaining to the prototype to be developed. This phase is very important in prototypes identification of requirements, and the perception of the scope and objective of the problem. In relation to that the knowledge of problem is that the services of QOU admission using online service anywhere any time. In this stage the problem statement, the objective and the scope would be clear enough. Requirements are a formal definition of a system and reflect the needs of users for a system that helps to solve problems. The phase presents the perspective student and staff requirements expected from the Undergraduate Online Admission System for QOU. Data collection is vital in producing a requirement model and prototype that actually fulfilled users' needs. The interview method used to collect data in this research. The researcher conducted meeting with the head of admissions department in the QOU this process was implemented through face to face interview. This interview has been taken as a sample, to know what exactly the how the manual admission process is conducted in the QOU and also to identify the weakness of the current process. Also to know what is the admission office needs from the online web system. Moreover, I discussed with the many applicant of the QOU about their demands and easiness required from the system. The solution for solving the problem in this study is to design and develop a web based system to increase the interaction in admission office and the applicants of QOU, so the student can easily request admission. The

outcome of this stage will bring up a provisional Design. The system analysis and design or model of the system includes Rational Rose diagrams. The Rational Rose diagrams will contain the use case diagrams that can be detailed into sequence diagrams for every use case and class diagram.

Non Functional Requirement

The system shall be completely operational all hours of the day unless system failure or upgradation work is to be performed. Down time after a failure shall not exceed 24 hours. No training is required to use the website. The form, home, about, FAQ and analysis pages load up within 10 seconds. The results from the predictor should not take more than 30 seconds. The system can support any number of users at a time. The mean time to view a web page over a 56Kbps modem connection shall not exceed 5 seconds. The system shall provide password protected access to the website to all users – students and admins both. The system will be able to incorporate more features without major reengineering. The system must interface with the MongoDB database and the web search engine it will be launched on.
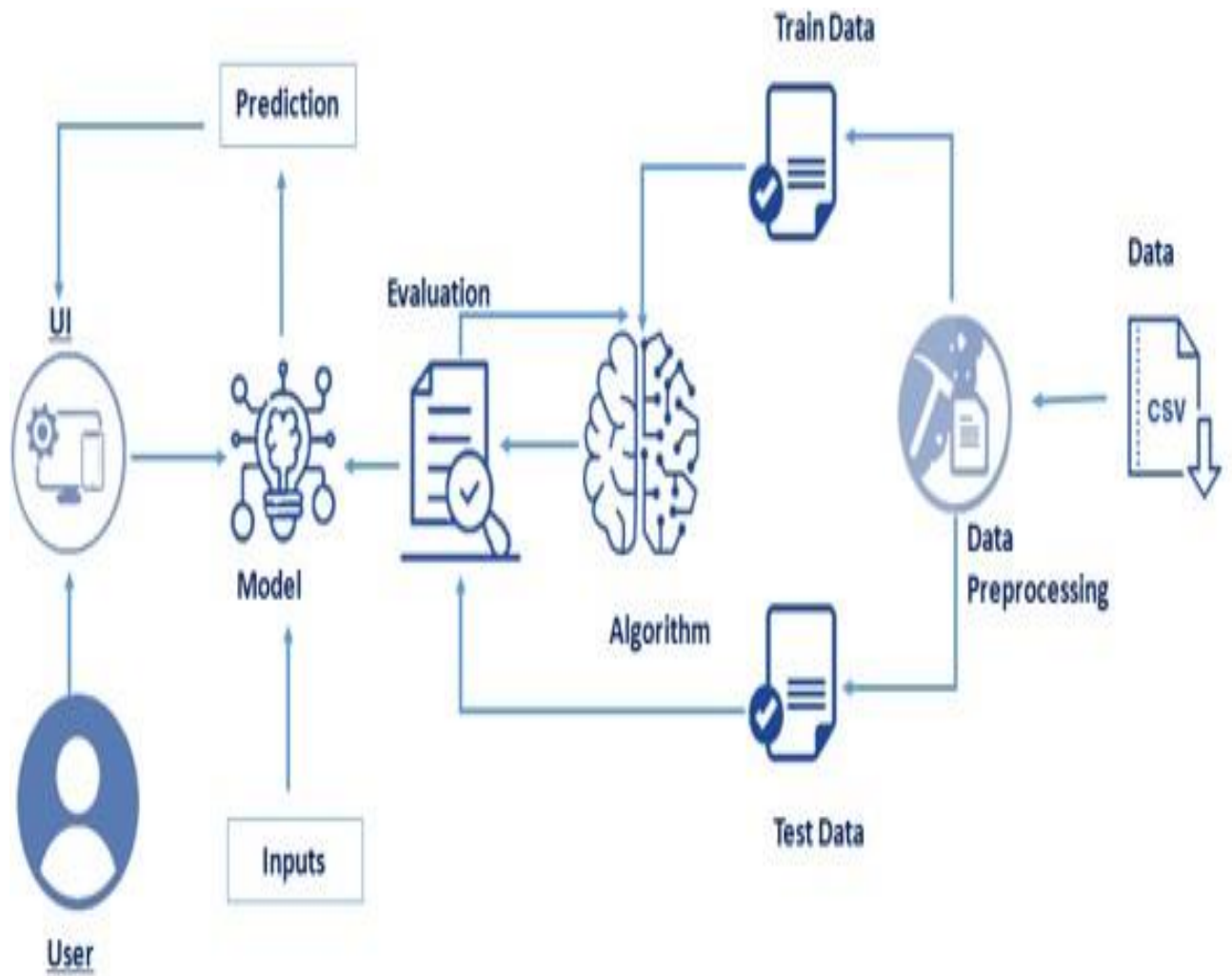
PROJECT DESIGN

Data Flow Diagram
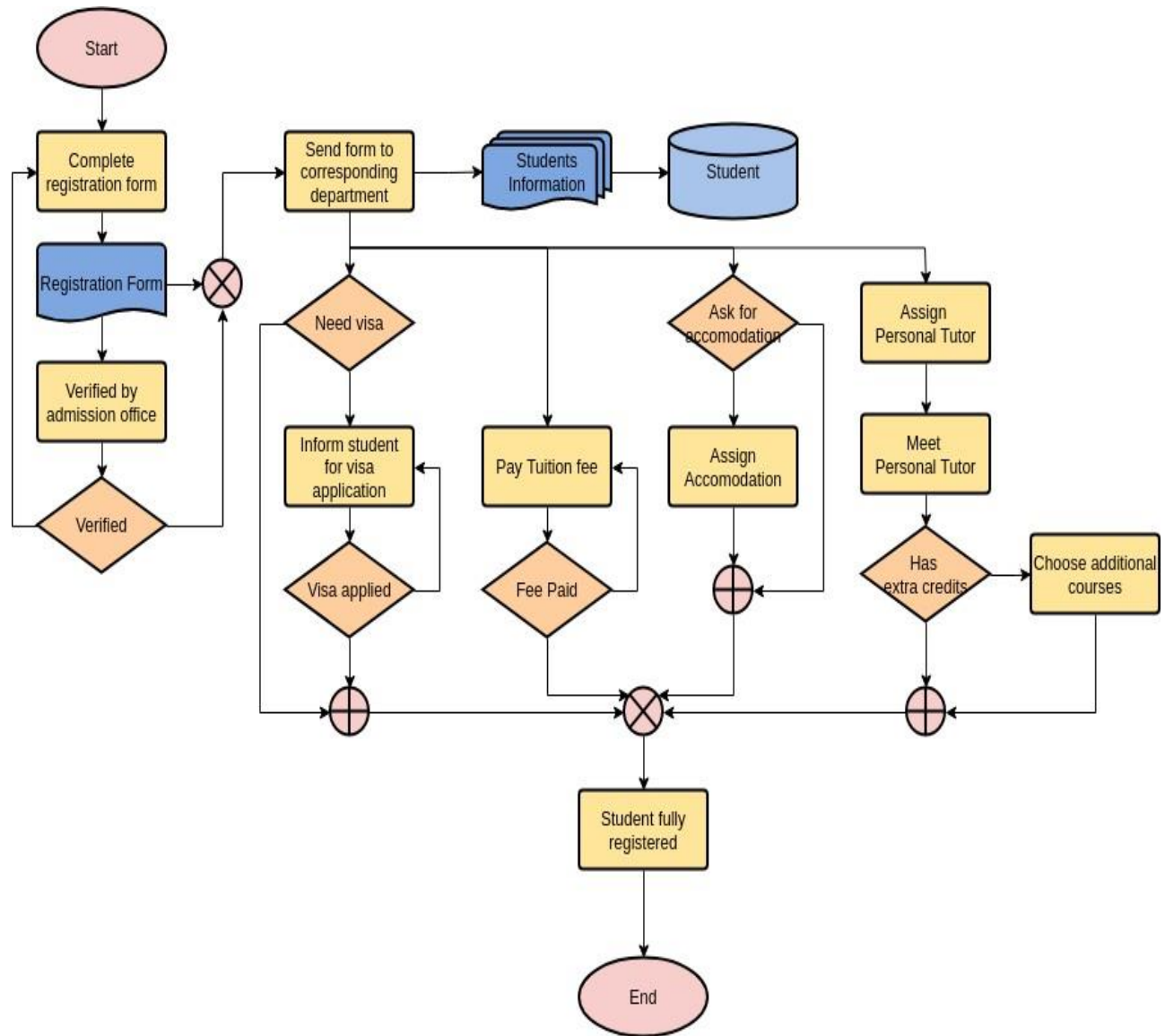
Solution & Technical Architecture

User Stories

The principal objective of the research is to help the students who are aspiring to pursue their education in the USA. The SAP system will help them to evaluate the chances of success in a particular university without being dependent on any education consultancy firm. It will help them in saving a huge amount of time and money spent in the application process. Also, it will help them to limit the number of applications made by the students by suggesting them the best universities where they have high chances of securing admission thereby by saving the amount of money spent by the students by applying in universities where they have less chance to secure admit based on their profile. System is needed so as to answer the queries of students in a compete and concise manner as well as to provide them an as accurate as possible analysis of their chances of admissions to their dream universities. Other factors such as changes in policies by the university or by the country can also affect chances of admissions in a way that is beyond the scope of this project. Admissions also depend on the individual university's policy regarding the intake of foreign students and is not modelled by our system.

PROJECT PLANNING & SCHEDULING

Sprint planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	User Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	2	High	2
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	1	High	1
Sprint-2		USN-3	As a user, I can check the eligibility criteria for various universities by uploading the necessary documents	2	Low	2
Sprint-3		USN-4	As a user, I can register for the desired university through Gmail and can also upload further course completion documents if necessary.	2	Medium	2
Sprint-4	User Login	USN-5	As a user, I can log into the application by entering email & password	1	High	2
	Dashboard		Check dashboard for further updates and upload the details according to the desired and eligible universities based on the eligibility criteria.			4

Sprint Delivery Schedule
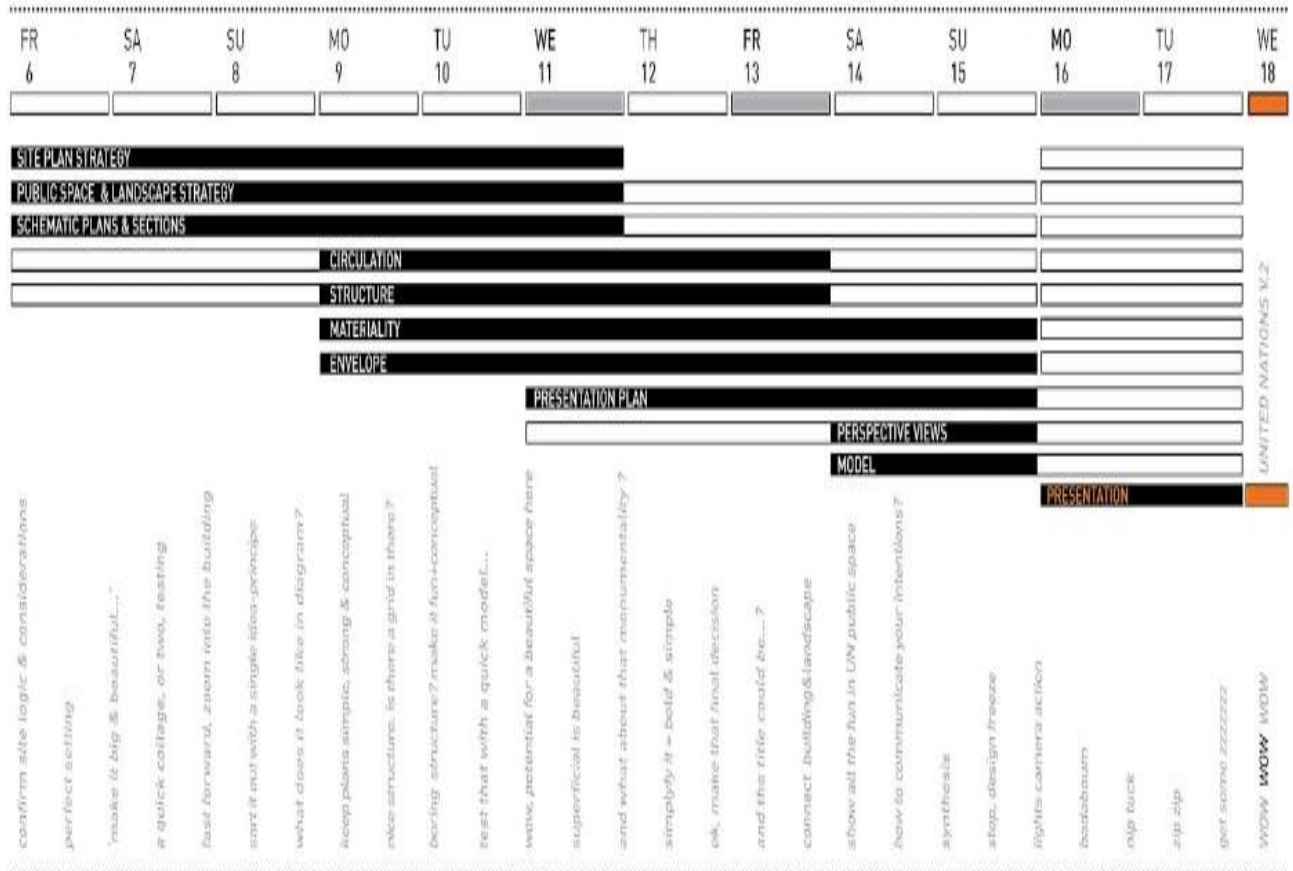
Report for JIRA:
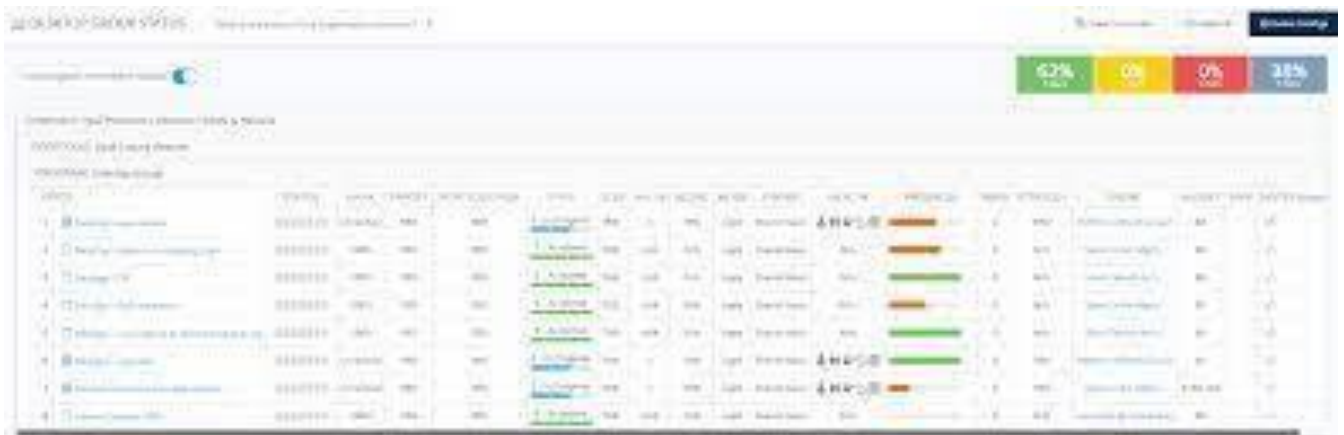
Scrum Project
Software project

Back to project

Reports

All reports

AGILE

SP board
Board

Burndown Chart

Burnup Chart

Sprint Report

Velocity Chart

Cumulative Flow Diagram

Version Report

Epic Report

Control Chart

You're in a company-managed project
Learn more

Agile

Burndown Chart

Track the total work remaining and project the likelihood of achieving the sprint goal. This helps your team manage its progress and respond accordingly.

Burnup Chart

Track the total scope independently from the total work done. This helps your team manage its progress and better understand the effect of scope change.

Sprint Report

Understand the work completed or pushed back to the backlog in each sprint. This helps you determine if your team is overcommitting or if there is excessive scope creep.

Velocity Chart

Track the amount of work completed from sprint to sprint. This helps you determine your team's velocity and estimate the work your team can realistically achieve in future sprints.

Cumulative Flow Diagram

Shows the statuses of issues over time. This helps you identify potential bottlenecks that need to be investigated.

Version Report

Track the projected release date for a version. This helps you monitor whether the version will release on time, so you can take action if work is falling behind.

Epic Report

Understand the progress towards completing an epic over time. This helps you manage your team's progress by tracking the remaining incomplete/unestimated work.

Control Chart

Shows the cycle time for your product, version or sprint. This helps you identify whether data from the current process can be used to determine future performance.

CODING & SOLUTIONING

Feature 1:

We have an algorithm that collects all the relevant data like admits, rejects, test scores, CGPA, courses studied and more provided by previous Yocket users over the years. The algorithm has created a pattern of admits and rejects across courses and universities based on the collected information. When you add information to your profile, it compares and calculates the admit chances based on the pattern it has deciphered. The Machine Learning aspect of the feature takes in your information including your admit and reject record to update its database and results in the future.

You must be a Yocketer. <Get the app> or <Sign up> on the website. Your choice!

We are currently catering only to Master courses in the US or Canada, so you must be applying for Masters in the US or Canada to enjoy this feature.

You must fill in your aptitude scores, UG scores, GRE scores (including not taken), English test scores, experience, technical papers to get your admit predictor scores.

Once completed the above steps, visit the University Program page and check under the Admission section.

Feature 2:

- A high GPA (relative to what admitted students have) and a rigorous curriculum
- Strong test scores (relative to what admitted students have)
- A specific, honest, and well-written personal statement
- A unique extracurricular interest or passion (a "spike," as we like to call it)
- Volunteering experience with measurable impact
- Compelling letters of recommendation written on your behalf
- Work experience, particularly jobs related to your academic or professional interests

#1: Excellent Grades in Challenging Courses

#2: High Test Scores

#3: Sincere, Specific, and Well-Written Essays

#4: A Spike in Your Extracurriculars

#5: Compelling Letters of Recommendation

#6: Volunteering Experience With Measurable Impact

TESTING

Test Cases

The Education Based Prediction System helps a person decide what colleges they can apply to with their scores. The dataset that is used for processing consists of the following parameters: University name, Quants and Verbal Scores(GRE) TOEFL and AWA Scores. The GRE Test (Graduate Record Examinations) is a standardized test used by many universities and graduate schools around the world as part of the graduate admissions process. Other factors are also taken into consideration while applying to colleges, such as Letter of Recommendation (a formal document that validates someone's work, skills or academic performance),Statement of Purpose (a critical piece of a graduate school application that tells admissions committees who you are, what your academic and professional interests are, and how you'll add value to the graduate program you're applying to),Co-curricular activities and Research papers as well(research papers from journals that are not well known or have a high percentage of plagiarism are not taken into consideration for this case). When a person has completed their undergraduate degree and wants to pursue a Postgraduate degree in a field of their choice, more often than not, it is very confusing for the person to figure out what colleges they should apply to with the scores that they have obtained in GRE and TOEFL, along with their GPA at the time of their graduation. Many candidates may apply to colleges that do not fall under their score requirements and hence waste a lot of time. Applying to many colleges with scores also increases the cost. There are not many efficient methods that are

available to help address this issue and hence an Education Predictor System has been developed. In the system proposed, a person can enter their scores in the respective fields provided.

The system then processes the data entered and produces an output of the list of colleges that a person could get into, with their scores. This is relatively quick and helps conserve time and money. In order to achieve this we have proposed a novel method utilising Machine Learning algorithms. To maximize the accuracy of our model, we have taken into consideration not one; but several machine learning algorithms. These algorithms include Neural Networks, Linear Regression, Decision Tree and Random Forest. More about these algorithms will be covered in the Algorithms section of this paper. These Algorithms are then compared and the algorithm which has the best key performance indicators will be used to develop the Prediction System. We also look forward to incorporate clustering of universities based on a profile and then classifying them as less likely, highly likely acceptance etc.

User Acceptance Testing

1. **Data Collection**

Having a university admission relies not only on the exam result of the students but also on many other considerations related to their social, economic, cultural, or geographical factors. First, the authors deeply analyze and recognize the causes that are liable or have an impact on admission. To collect data, it is prepared a set of 27 questionnaires as shown in the following subsection B. Then the authors provide these questionnaires to the students of BSMRSTU's various departments such as Computer Science and Engineering, Electrical and Electronic Engineering, Electronics and Telecommunication Engineering, Applied Chemistry and Chemical Engineering, Mathematics, Statistics, Chemistry, and Environmental Science & Disaster Management. The first four departments' data is considered for those who got admission in the Engineering faculty (also called A Unit) and the rest four departments' data for those who did not get admission in the Engineering faculty. The total collection is 500 students' data. These features are grouped into two main categories without the target factor. These are (1) before engineering faculty admission (obtained marks in admission test is not included) (2) after engineering faculty admission (obtained marks in admission test is included.

2. Data Preprocessing

The authors prepared the collected data in tabular form from the questionnaire feedback of the students who participated during the data collection. They also applied some data cleaning techniques e.g. handling noise, outliers, missing values, and duplicate data to transform the raw data in a useful and efficient format. The authors considered each questionnaire as a distinct variable or feature for the dataset as shown. The authors split these 27 features into two categories before and after the admission test takes place to the BSMRSTU Engineering Faculty (Unit A). They allocate the first 25 variables as input features and the last factor (27thvariable) as the output label to predict the chance of admission in the engineering faculty before the test. We then allocated the first 26 factors as input features and the last factor (27thvariable) as the output label to predict the chance of admission in the engineering faculty after the test happened. The difference between these two categories is of one feature i.e. 'expected scores/marks in A Unit' (26th variable in Table I).For identifying the most relevant input variables (feature selection) to predict undergraduate admission, the authors used the embedded methods which used built-in feature selection methods in machine learning algorithms. All three tree-based machine learning algorithms XGBoost, Light GBM, and GBM used in this investigation have their feature selection method.

3. Description of the Features

A description of the extracted features is given below. Some of the closely related features are discussed together. S.S.C GPA: S.S.C result is an important factor for identifying a student's quality. It reflects the basic science knowledge of a student. To attend any university admission exam a student should have passed it successfully. So, this factor is important for predicting admission test results. The numeric grade is recorded out of 5.00.H.S.C GPA: H.S.C result is the most important factor for identifying a student's quality. It also bears the science knowledge of a student. So this attribute is more important for predicting admission test results. The numeric grade is recorded out of 5.00.

Physics grade: A student's physics grade is the reflection of knowledge on the physics subject. In the BSMRSTU admission test, there are 30 marks on physics for the A unit. So, the grade of physics is more important for getting a chance in the engineering faculty of BSMRSTU. The numeric grade is recorded out of 5.00.

Chemistry grade: Chemistry grade is the reflection of knowledge on the chemistry of a student. In the BSMRSTU admission test of A unit, there are 20 marks on chemistry.
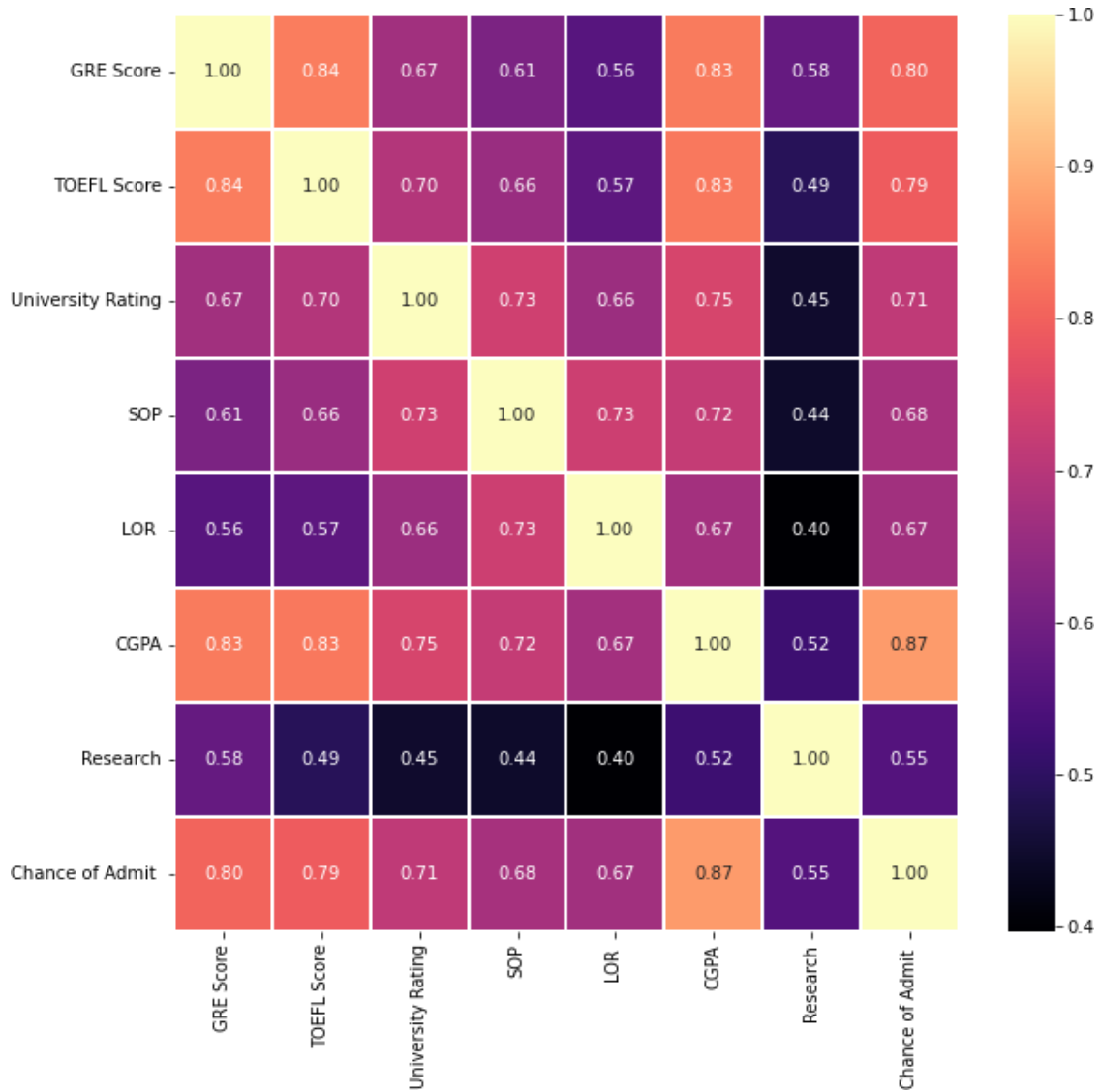
RESULTS

Performance Metrics

The complete system is designed using Python. Data set is classified into 9 different parameters which are considered important during the application for Masters.

Those parameters are: gre scores, toefl scores, university rating, statement of purpose, letter of recommendation, undergraduate gpa, research paper, chance of admit.

# SOP	# LOR	# CGPA	# Research
4.5	4.5	9.65	1
4	4.5	8.87	1
3	3.5	8	1
3.5	2.5	8.67	1
2	3	8.21	0
4.5	3	9.34	1
3	4	8.2	1
3	4	7.9	0
2	1.5	8	0
3.5	3	8.6	0

	GRE Score	TOEFL Score	University Rating	SOP	LOR	CGPA	Research	Chance of Admit
GRE Score	1.00	0.84	0.67	0.61	0.56	0.83	0.58	0.80
TOEFL Score	0.84	1.00	0.70	0.66	0.57	0.83	0.49	0.79
University Rating	0.67	0.70	1.00	0.73	0.66	0.75	0.45	0.71
SOP	0.61	0.66	0.73	1.00	0.73	0.72	0.44	0.68
LOR	0.56	0.57	0.66	0.73	1.00	0.67	0.40	0.67
CGPA	0.83	0.83	0.75	0.72	0.67	1.00	0.52	0.87
Research	0.58	0.49	0.45	0.44	0.40	0.52	1.00	0.55
Chance of Admit	0.80	0.79	0.71	0.68	0.67	0.87	0.55	1.00

Corrleation matrix where The GRE Test (Graduate Record Examinations) is a standardized test used by many universities and graduate schools around the world as part of the graduate admissions process.
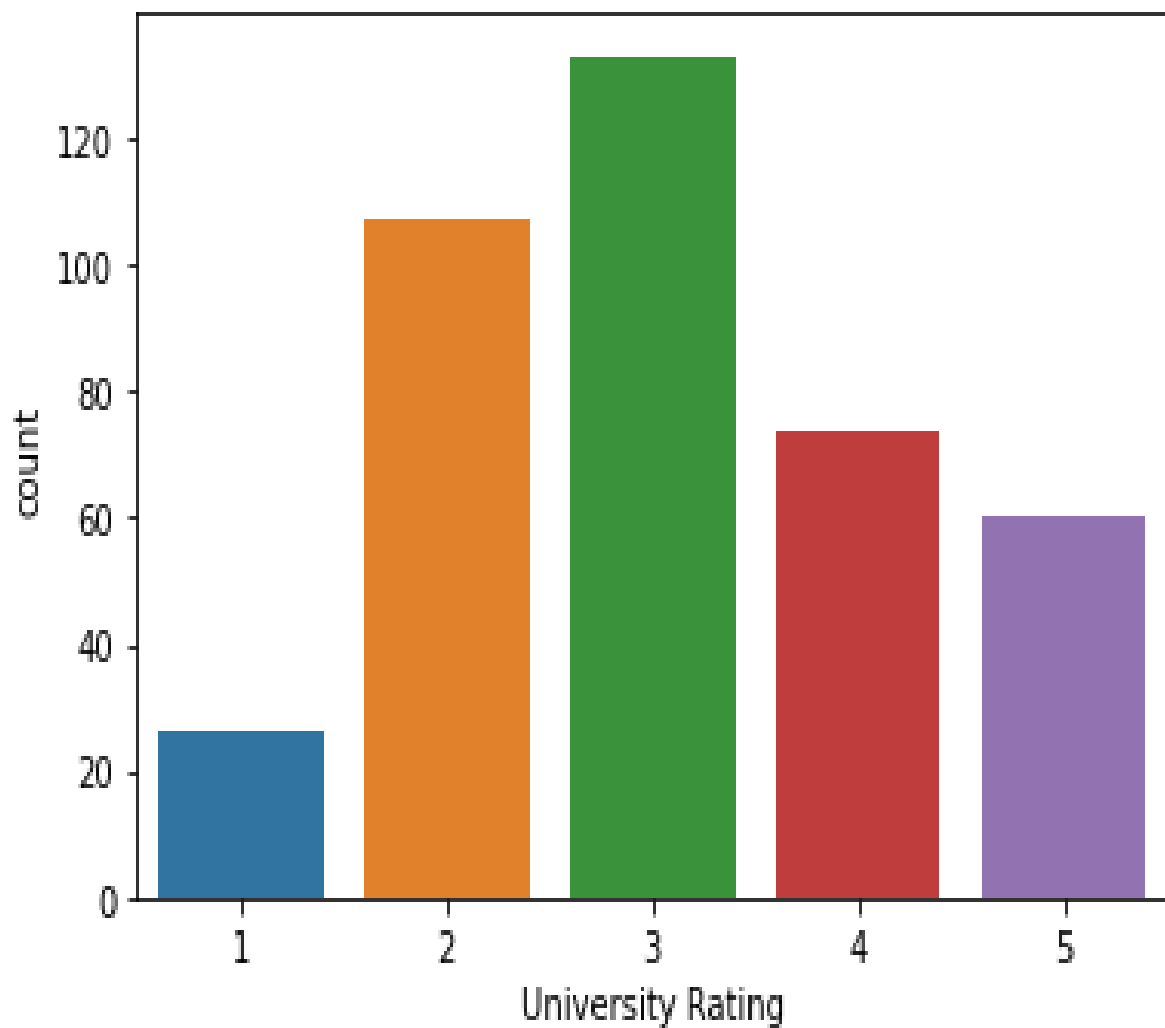
The number of students selected to universities based on parameters like gre scores, toefl scores, university rating, statement of purpose, letter of recommendation, undergraduate gpa, research paper, chance of admit.
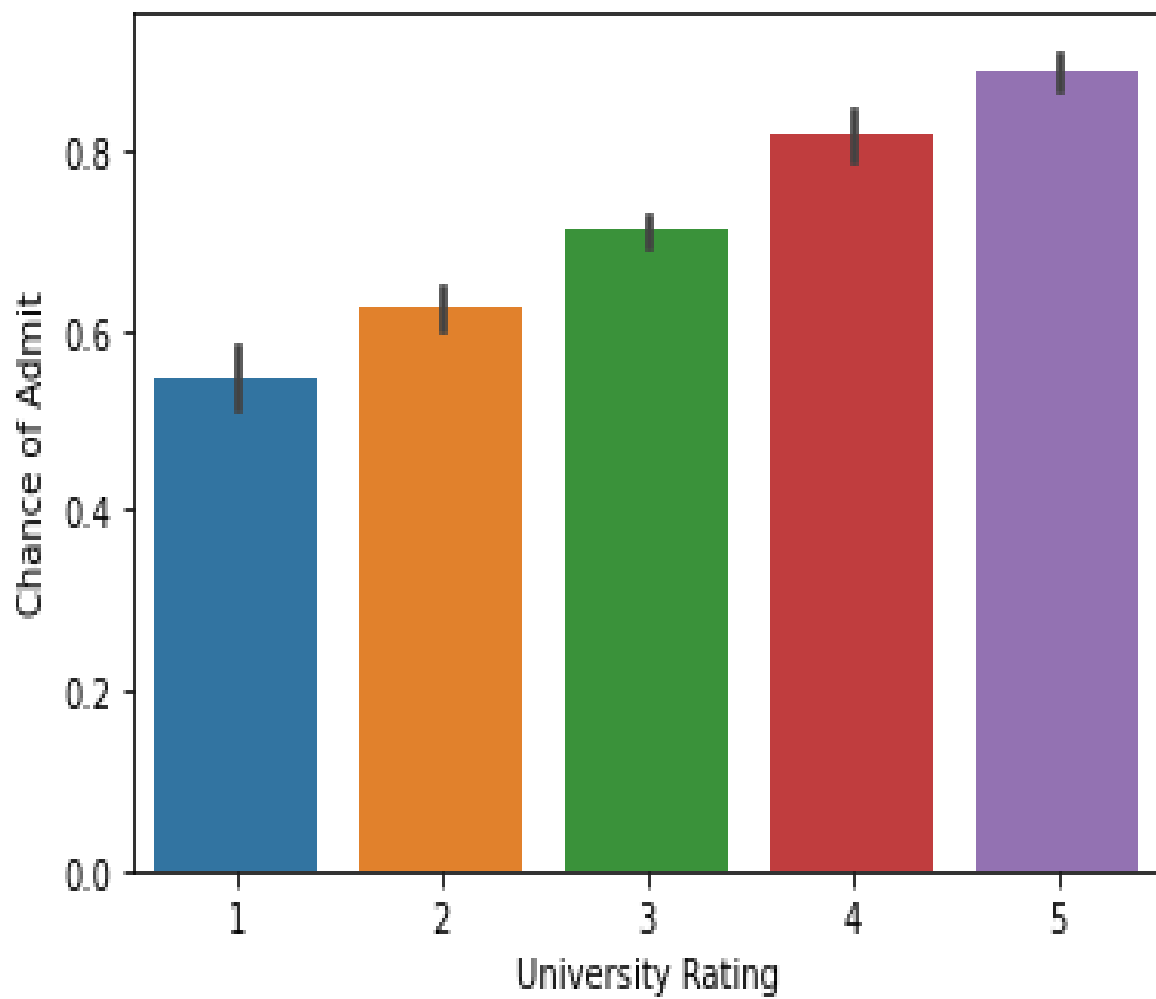
This graph shows that chance to get admission based on marks in top universities based on university rating.

ADVANTAGES & DISADVANTAGES

ADVANTAGES

Reach to geographically scattered student: One of the important objectives of the admission system is communicate with all the students scattered geographically.

Reducing time in activities: Reduce the time taken process the applications of students, admitting a student, conducting the online examination, verify student marks, and send call letters to selected students.

Centralized data handling: Transfer the data smoothly to all the departments involved and handle the data centralized way.

Paperless admission with reduced man power: Reduce the manpower needed to perform all the admission and administration task by reducing the paper works Cost cutting. Reduce the cost involved in the admission process.

Operational efficiency: Improve the operational efficiency by improving the quality of the process.

DISADVANTAGES:

Require much man power i.e. much efforts, much cost and hard to operate and maintain.

Since, all the work is done in papers so it is very hard to locate a particular student record when it is required.
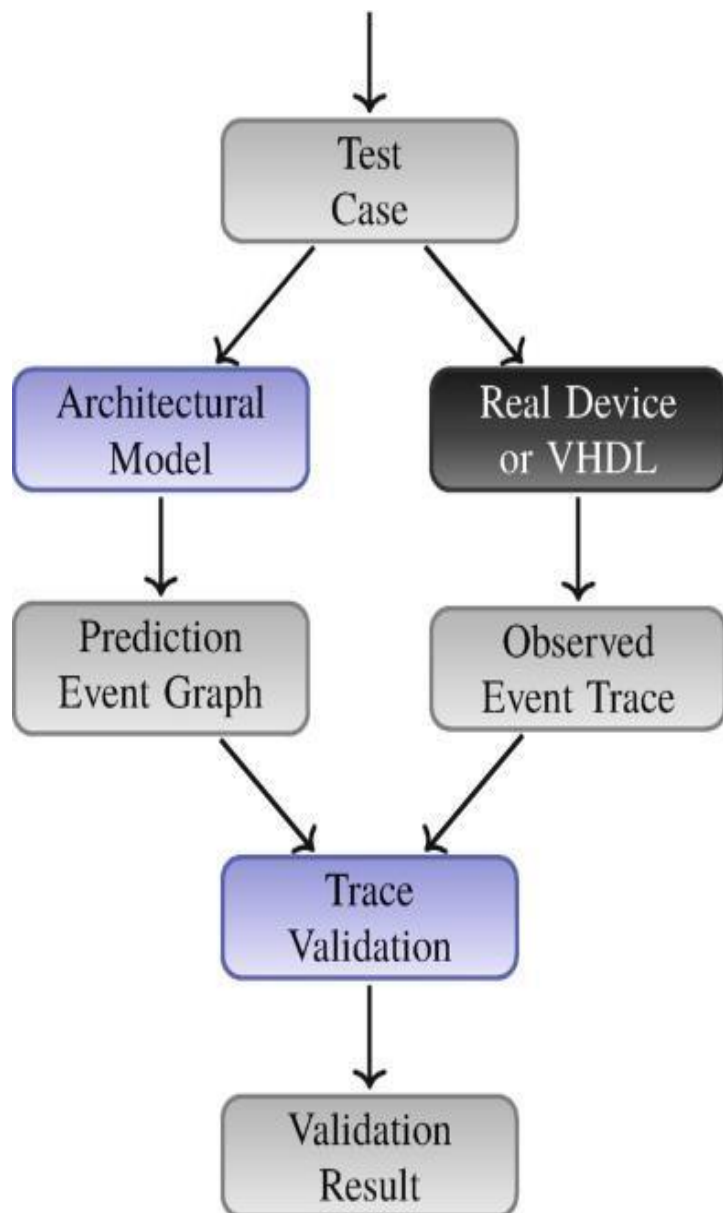
FUTURE SCOPE:

The scope of this project is a web application that allows users to enter their academic data and get predictions of their chances of admissions in the university tier of their choosing.

It also provides them answers to the most common FAQ's that arise when thinking of admissions abroad for Post Graduate studies.

It also provides an analysis based on the data set used that shows how the different parameters affect chances of admissions.

A Database will also be implemented for the system so that students can save their data and review and edit it as they progress with the most recent predictions being saved with their profile.

Issues of web security other than password protection within the website are not part of this project

Test
Case

Architectural
Model

Real Device
or VHDL

Prediction
Event Graph

Observed
Event Trace

Trace
Validation

Validation
Result

Legend:

Hardware

Tool

Data

CONCLUSIONS

The main objective of this research was to develop a prototype of the system that can be used by the students aspiring to pursue their education in the USA. Multiple machine learning algorithms were developed and used for this research. KNN proved to best-fit for development of the system when compared with the Logistic regression model. The model can be used by the students for evaluating their chances of getting shortlisted in a particular university with an average accuracy of 75%. Decision Tree algorithm was used to predict the universities which were best suitable for a student based on their profile. A simple user interface was developed to make the application interactive and easy to use for the users from the non-technical background. Shiny library from R was used to create the user interface. The overall objective of the research was achieved successfully as the system allow the students to save the extra amount of time and money that they would spend on education consultants and application fees for the universities where they have fewer chances of securing admission. Also, it will help the students to make better and faster decision regarding application to the universities.

FUTURE SCOPE

The limitation of the research we have created the models based only on the data of Indian Students studying Masters in Computer Science , we have considered only ten universities with different rankings. In future, more data related to additional universities and courses can be added to the system. Also, the system can be enhanced to a web-based application by making changes to the code. Other classification algorithms can be evaluated to resolve the problem if they perform better than the current algorithm the system can be easily updated to support the new algorithm by changing the server code in the app.

APPENDIX

Source Code

```python
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
```

```python
data=pd.read_csv(r"C:\Users\CMD\Downloads\Admission_Predict.csv")
data
```

```python
data.drop(["Serial No."], axis=1, inplace=True)
```

```python
data.describe()
```

```python
data.info()
data.info()
```

VISUALIZATION

```python
plt.scatter(data['GRE Score'],data['CGPA'])
plt.title('CGPA vs GRE Score')
plt.xlabel('GRE Score')
plt.ylabel('CGPA')
plt.show()
```

```python
plt.scatter(data['CGPA'],data['SOP'])
plt.title('SOP for CGPA')
```

```python
plt.xlabel('CGPA')
plt.ylabel('SOP')
plt.show()




data[data.CGPA >= 8.5].plot(kind='scatter', x='GRE Score', y='TOEFL Score',color="BLUE")

plt.xlabel("GRE Score")
plt.ylabel("TOEFL SCORE")
plt.title("CGPA>=8.5")
plt.grid(True)

plt.show()




data["GRE Score"].plot(kind = 'hist',bins = 200,figsize = (6,6))

plt.title("GRE Scores")
plt.xlabel("GRE Score")
plt.ylabel("Frequency")

plt.show()




p = np.array([data["TOEFL Score"].min(),data["TOEFL Score"].mean(),data["TOEFL Score"].max()])
r = ["Worst","Average","Best"]
plt.bar(p,r)

plt.title("TOEFL Scores")
plt.xlabel("Level")
```

```python
plt.ylabel("TOEFL Score")

plt.show()



g = np.array([data["GRE Score"].min(),data["GRE
Score"].mean(),data["GRE Score"].max()])
h = ["Worst","Average","Best"]
plt.bar(g,h)

plt.title("GRE Scores")
plt.xlabel("Level")
plt.ylabel("GRE Score")

plt.show()



plt.figure(figsize=(10, 10))

sns.heatmap(data.corr(), annot=True, linewidths=0.05, fmt=
'.2f',cmap="magma")

plt.show()



data.Research.value_counts()

sns.countplot(x="University Rating",data=data)


sns.barplot(x="University Rating", y="Chance of Admit ", data=data)
```

TRAINING AND TESTING SPLIT#

X=data**.**drop(['Chance of Admit '],axis=1) *#input data_set*
y=data['Chance of Admit '] *#output labels*
from sklearn.model_selection **import** train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.15)

from sklearn.ensemble **import** GradientBoostingRegressor
rgr = GradientBoostingRegressor()
rgr**.**fit(X_train,y_train)

GradientBoostingRegressor()

rgr**.**score(X_test,y_test)

0.8184271504939563

y_predict=rgr**.**predict(X_test)

from sklearn.metrics **import** mean_squared_error,
r2_score,mean_absolute_error
import numpy **as** np
print('Mean Absolute Error:', mean_absolute_error(y_test, y_predict))
print('Mean Squared Error:', mean_squared_error(y_test, y_predict))
print('Root Mean Squared Error:', np**.**sqrt(mean_squared_error(y_test,
y_predict)))

Mean Absolute Error: 0.04310629034892705
Mean Squared Error: 0.0047012186514370 5
Root Mean Squared Error: 0.06856543335702801

y_train = (y_train>0.5)
y_test = (y_test>0.5)

from sklearn.linear_model._logistic **import** LogisticRegression

```python
lore = LogisticRegression(random_state=0, max_iter=1000)

lr = lore.fit(X_train, y_train)

y_pred = lr.predict(X_test)

from sklearn.metrics import accuracy_score, recall_score, roc_auc_score,
confusion_matrix

print('Accuracy Score:', accuracy_score(y_test, y_pred))
print('Recall Score:', recall_score(y_test, y_pred))
print('ROC AUC Score:', roc_auc_score(y_test, y_pred))
print('Confussion Matrix:\n', confusion_matrix(y_test, y_pred))
```
Accuracy Score: 0.8833333333333333
Recall Score: 0.9807692307692307
ROC AUC Score: 0.6153846153846154
Confussion Matrix:
 [[2 6]
 [1 51]]

```python
import pickle

pickle.dump(lr, open("university.pkl", "wb"))

import pickle

lr = pickle.load(open("university.pkl", "rb")) #logistic regression model

pip install -U ibm-watson-machine-learning


from ibm_watson_machine_learning import APIClient
import json
```

```
uml_credentials = {
    "url": "https://us-south.ml.cloud.ibm.com",
    "apikey": "poJ22ua6BCG9qY33B8fkgnz1bnP1f9DZqUlF9NkBM1bZ"
}

client = APIClient(uml_credentials)
```

Output

Step 1

User login page to fill the required datas

Step 2

Processing the data after required fields are filled

University Admission Eligibility Prediction System

Enter your details and get probability of your admission

Students are often worried about their chances of admission to University. The aim of this project is to help students in shortlisting universities with their profiles. The predicted output gives them a fair idea about their admission chances in a particular university. This analysis should also help students who are currently preparing or will be preparing to get a better idea.

Enter the details

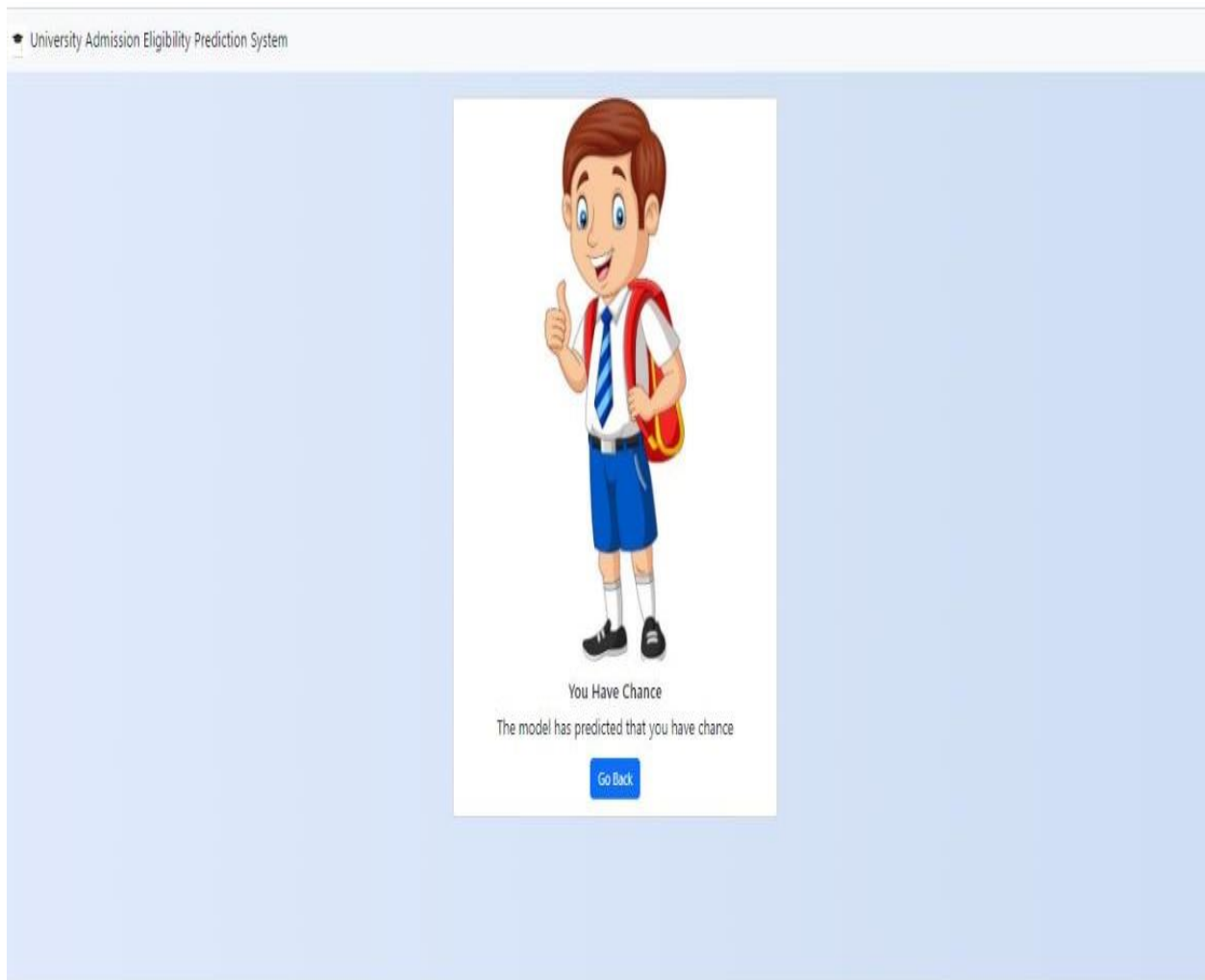|	
GRE Score:	279
TOFEL Score:	68
University Rating:	3.50
SOP:	2.50
LOR:	4.51
CGPA:	8.70
Research:	○ Yes ● No

Predicting...

Step 3

If the person had chance to get admission, this page is shown to student

Step 4

If the marks not meet required CGPA, You are not eligible to applied university

You have a LOW / NO chance

GitHub

https://github.com/IBM-EPBL/IBM-Project-32372-1660209444

Demo Video Link

https://uploadnow.io/f/lHQxkfq