**IBM PROJECT 2022**

**PROJECT REPORT DOCUMENTATION**

**TEAM ID: PNT2022TMID32434**

**UNIVERSITY ADMIT ELIGIBILITY PREDICTOR**

**INTRODUCTION**

**Project Overview**

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.

**Purpose**

● Students often feel difficulty in shortlisting the universities to apply for which they tend to wonder if their profile matches the requirements of a certain university.

● Moreover, the cost of applying to a university is extremely high making it critical that students shortlist universities based on their profile.

● A university admission prediction system is quite useful for students to determine their chances of acceptance to a specific university.

● This system reduces dependence on educational consultancies, who charge loads of money to analyze a candidate’s profile and determine the universities he/she should apply to.

**LITERATURE SURVEY**

**Existing Problem**

There exists no tools or technologies for study-abroad aspirants to get their profiles evaluated for free. Hence, this Predictor aims to consider various factors involved in an application and use ML models to predict the chances of the

individual.

**References**

1. Graduate Admission Prediction Using Machine Learning

Aljasmi, et. all. talk about the student admission problem which is very important in educational institutions. This paper addresses machine learning models to predict the chance of a student to be admitted to a master's program. They propose a system that will assist students to know in advance if they

have a chance to get accepted. The machine learning models used are multiple linear regression, k-nearest neighbor, random forest, and Multi layer Perceptron. Experiments show that the Multi layer Perceptron model surpasses other models.

2. HRSPCA: Hybrid recommender system for predicting college admission

Ragab et.all., present a new college admission system using hybrid recommender based on data mining techniques and knowledge discovery rules, for tackling college admissions prediction problems. This is due to the huge numbers of students required to attend university colleges every year. The proposed HRSPCA system consists of two cascaded hybrid recommenders working together with the help of college predictor, for achieving high performance.

3. University Admissions Predictor Using Logistic Regression

Fathiya and Sadath perform a novel study on a predictor for university admissions that allows students to assess their chances of being admitted to an institution. Real student data is gathered in order to construct this. The information is kept in the form of a training set that may be used by the logistic regression classifier that was designed to predict admissions.

4. A Machine Learning Approach for Graduate Admission Prediction

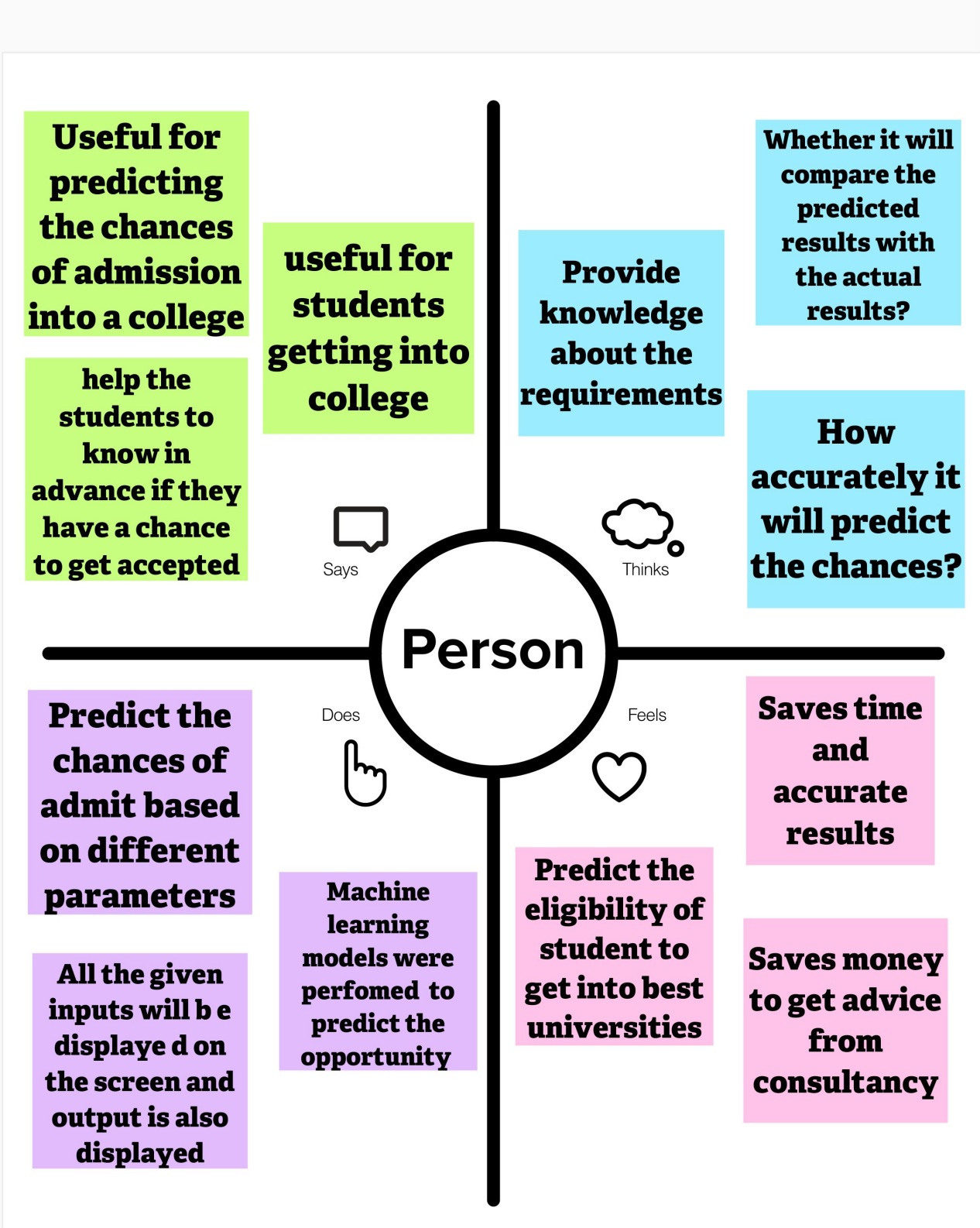
AlGhamdi et.all., evaluate three learning strategies of regression to predict the university rate given the students' profile; namely, linear regression, decision tree, and logistic regression model. This paper evaluates these models to select the best model in terms of the highest accuracy rate and the least error. It was determined that the Logistic Regression model shows the most accurate prediction and hence this model was employed to predict the future applicant's university chance of admission.

**Problem Statement Definition**

Concerns about getting into college are common among students. This project's goal is to assist students in narrowing down institutions based on their profiles. The anticipated results offer them a good indication of their prospects of admission to a particular university. This analysis ought to provide better insight for students who are or will be preparing for exams. **IDEATION & PROPOSED SOLUTION**

**Empathy Map Canvas:**

An empathy map is a simple, easy-to-digest visual that captures knowledge about a user’s behaviours and attitudes. It is a useful tool to helps teams better understand their users. Creating an effective solution requires understanding the true problem and the person who is experiencing it. The exercise of creating the map helps participants consider things from the user’s perspective along with his or her goals and challenges.

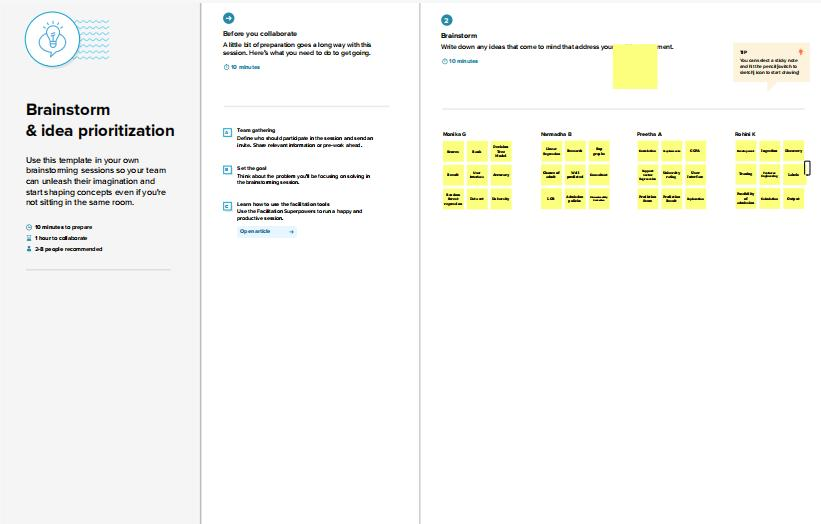


**Brainstorm & Idea Prioritization Template:**

Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich amount of creative solutions. Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you're not sitting in the same room.

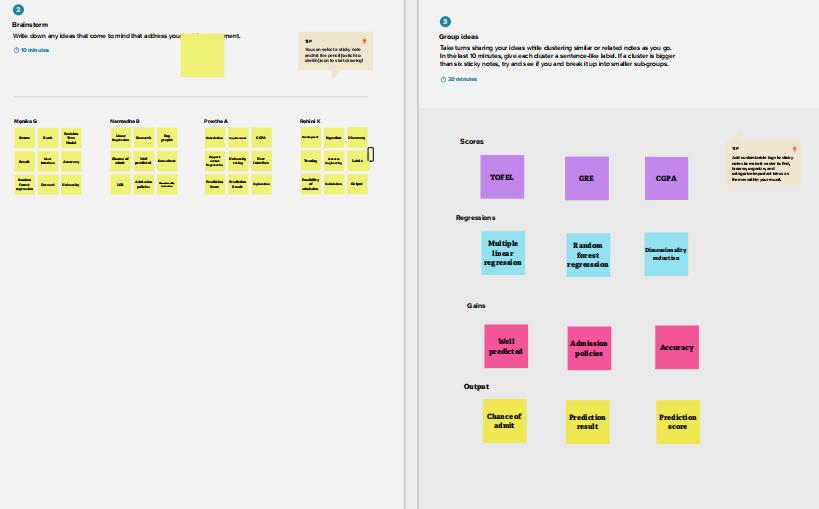
**Step-1:**

**Team Gathering, Collaboration and Select the Problem Statement**



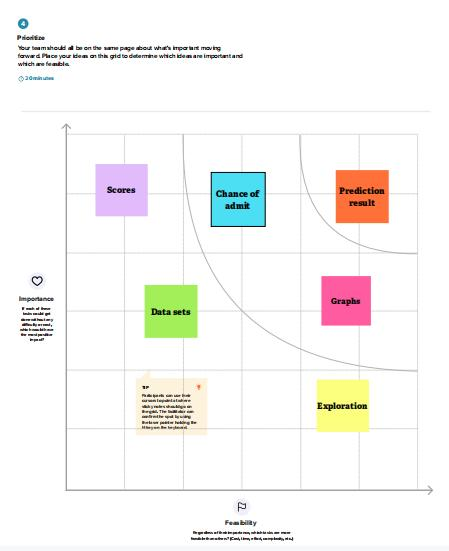
**Step-2:**

**Brainstorm, Idea Listing and Grouping**



**Step-3:**

**Idea Prioritization**



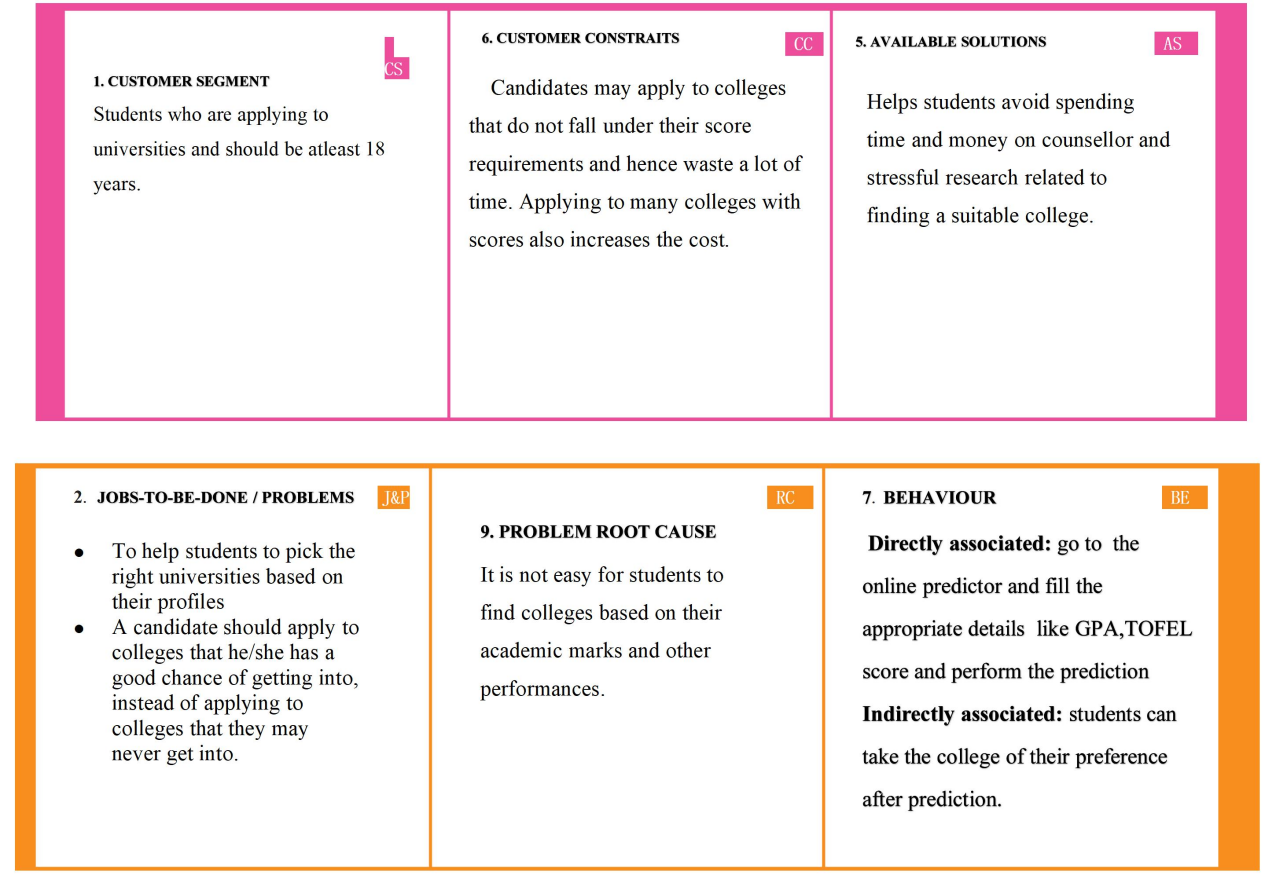
**Proposed Solution**

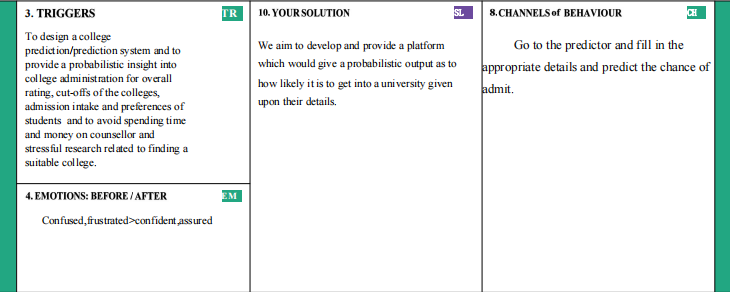
The below table represents the Proposed Solution.

|  |  |  |
| --- | --- | --- |
| **S.NO** | **PARAMETER** | **DESCRIPTION** |
| 1. | Problem Statement  (Problem to be solved) | The Students who are seeking admission into universities needs a way to predict the possibility of admission and accurately predict the chance of admit. |
| 2. | Idea / Solution description | Creating a machine learning model predicting the eligibility of students getting admission in best university based on their Test attributes likeGRE,TOEFL,LOR,CGPA etc. according to their scores the possibilities of chance of admit is calculated. |
| 3. | Novelty / Uniqueness | We will achieve this aim by using the Linear Regression model. Based on the data that we have, we will split out data into training and testing set. The Training set will have features and labels on which our model would be trained. Once our model is trained, we will use the trained model and run it on the test set and predict the output. Then we will compare the predicted results with the actual results that we have to see how our model performed. |
| 4. | Social Impact / Customer  Satisfaction | Since it predicts the chances of admit earlier, it helps the students to save money to get advice from consultancy and accurately predict the results. |
| 5. | Social Impact / Customer  Satisfaction | Student admission problem is very important in educational institutions. This project addresses machine learning models to predict the chance of a student to be admitted. This will assist students to know in advance if they have a chance to get accepted. |
| 6. | Scalability of the Solution | This model can be expanded to include more  prediction for more accurate detection . Training the model with even more attributes will increase the efficiency further. |

**Problem Solution fit**

The below images represent the Problem Solution Fit of the project





**REQUIREMENT ANALYSIS**

**Functional requirement**

Following are the functional requirements of the proposed solution.

|  |  |  |
| --- | --- | --- |
| **FR No.** | **Functional Requirement (Epic)** | **Sub Requirement (Story / Sub-Task)** |
| FR-1 | User Registration | In order to prevent unauthorized access to the system, users must be able to log into their accounts using the system by providing their email and password. |
| FR-2 | User Confirmation | Confirmation via Email  Confirmation via OTP |
| FR-3 | Data Management | This application enables the user to CRUD (Change, Read, Update, and Delete) data. |
| FR-4 | Web Service Management  Process | Web Service Management is the process of registering a web client to provide SSO (Single sign-on) or member data transmission |
| FR-5 | Data retention | The proposed application system handles historical data archiving, retrieval, and retention |
| FR-6 | User Deliverables | Submission of relevant documents - Required Entrance Exam Marksheet, Curriculum vitae(CV), Personal Information, Letter of Recommendation |
| FR-7 | User Profile | Applicant’s dashboards - Personal information, wishlist, skills and course, percentage |

**Non-functional Requirements:**

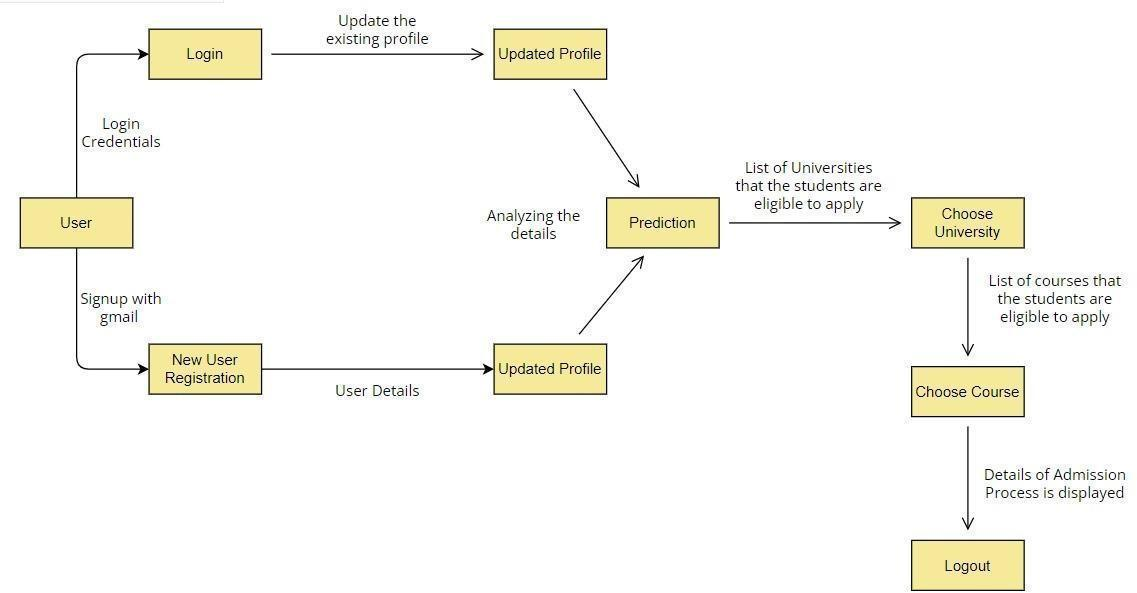
Following are the non-functional requirements of the proposed solution.

|  |  |  |
| --- | --- | --- |
| **FR No.** | **Non-Functional Requirement** | **Description** |
| NFR-1 | **Usability** | A logical interface is required to make the system easy to use and to speed up typical processes. The mistake rate of users providing their information on the checkout page must not exceed 10%. |
| NFR-2 | **Security**  **Reliability** | Authorization access scenarios and definitions, as well as student record handover processes between universities. Utilize certain cryptographic techniques. When the application is validating the user or licence, communication must be limited. |
| NFR-3 | **Reliability** | Data corruption is avoided by employing backup methods and strategies. At the moment of input, all data stored for user variables will be committed to the database |
| NFR-4 | **Performance** | The availability results of the requested college should be supplied to the student in little more than two seconds, and data retrieval should be trustworthy because each student will be granted a maximum of 10 minutes, accessing the database should be done at a reasonable speed. |
| NFR-5 | **Availability** | The system should be available at all times, allowing the user easy access. If the hardware or database fails, a substitute page will be displayed, and the database should be obtained from the data folder. |
| NFR-6 | **Scalability** | Determines the highest workloads under which the system will still run satisfactorily. Deals with the measurement of the system's reaction time under varied load circumstances. |

**PROJECT DESIGN**

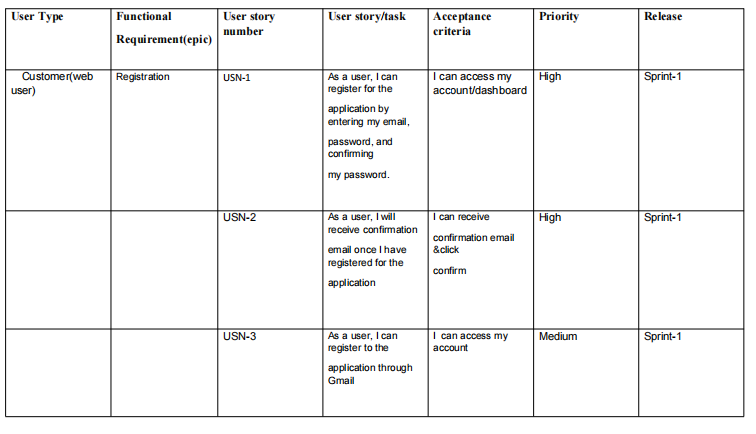
**Data Flow Diagrams**

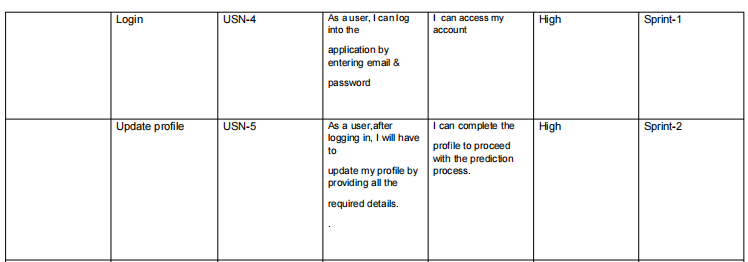
A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

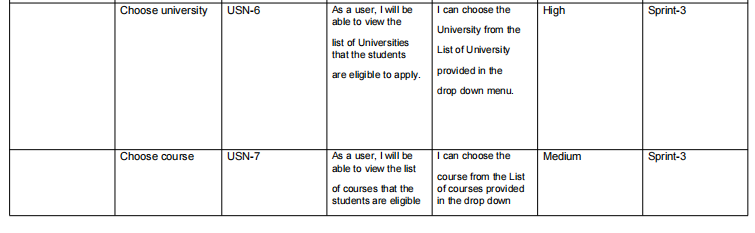


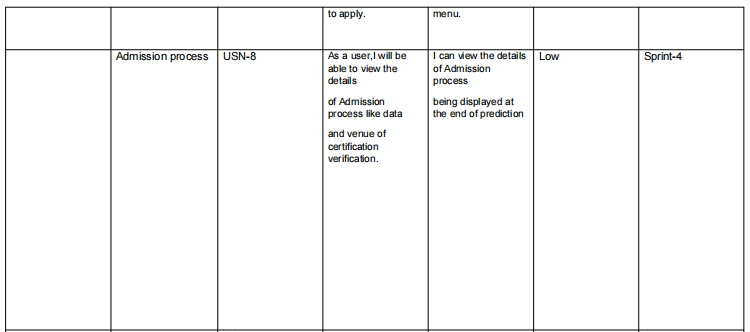
**User Stories**

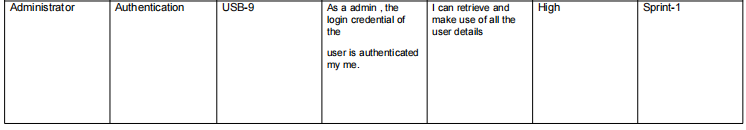
Use the below template to list all the user stories for the product.

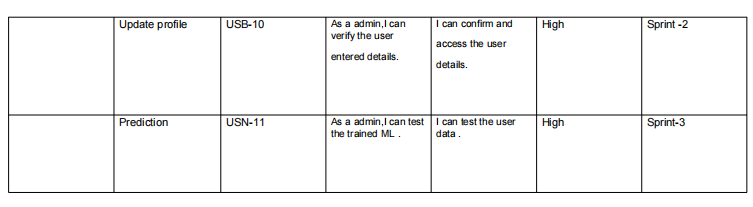






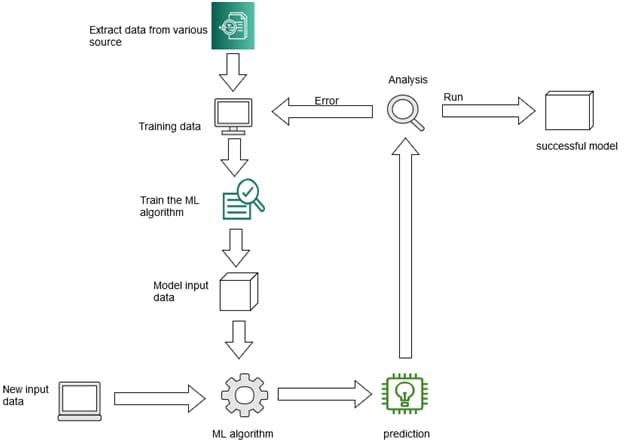






**Solution & Technical Architecture**

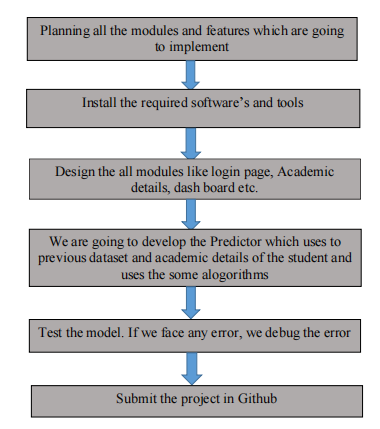
The below figure represents the Solution & Technical Architecture of the project



**PROJECT PLANNING & SCHEDULING**

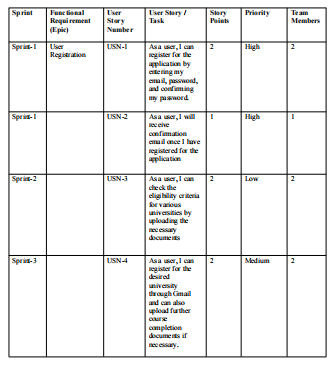
1. Sprint Planning & Estimation

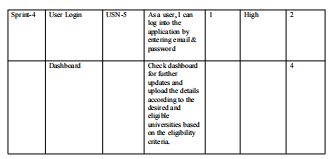
The below figure represents the Sprint Planning & Estimation details.



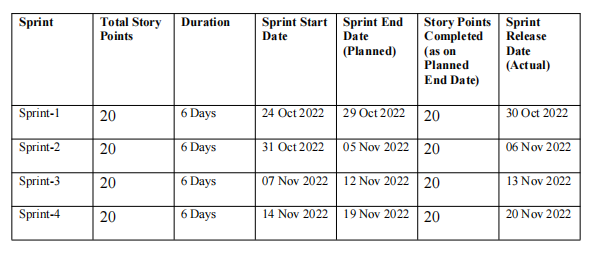
**Sprint Delivery Schedule**

The below figure represents the Sprint Delivery Schedule details.

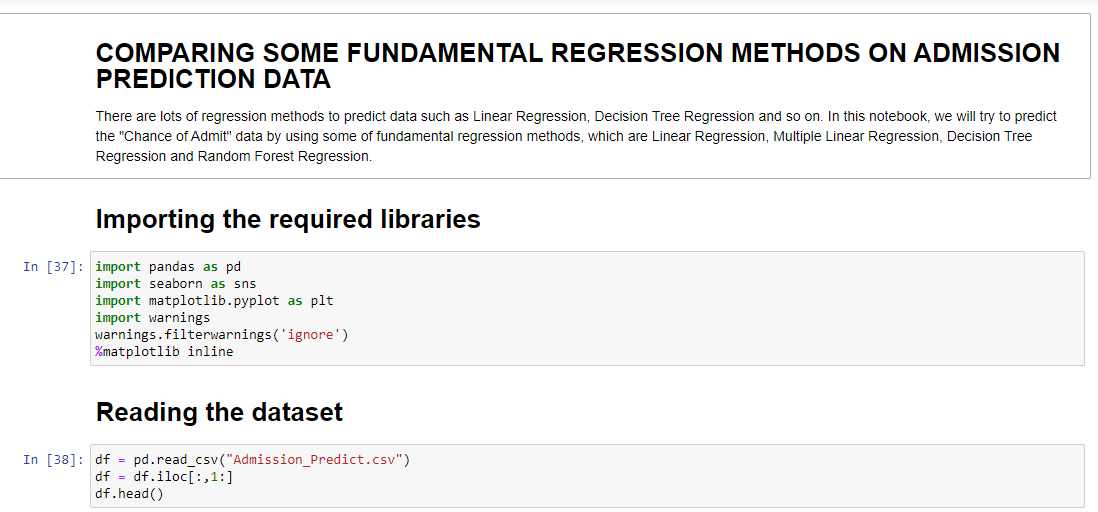


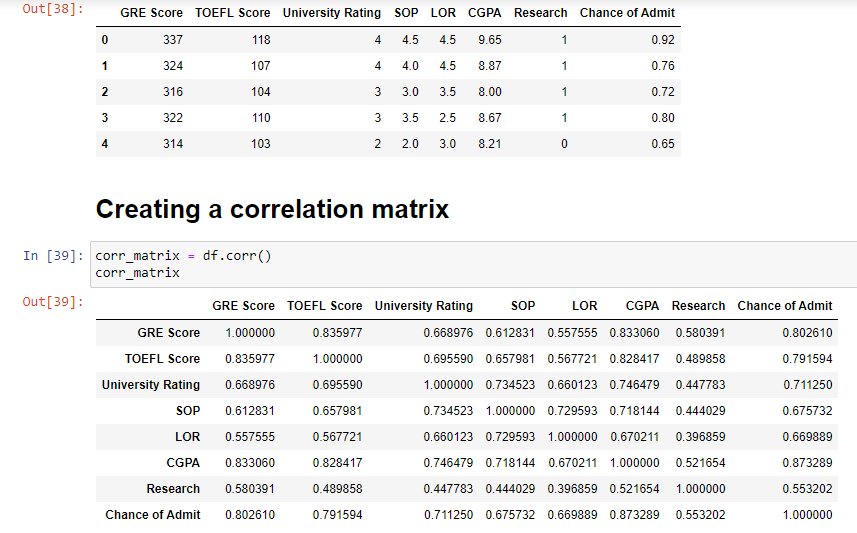


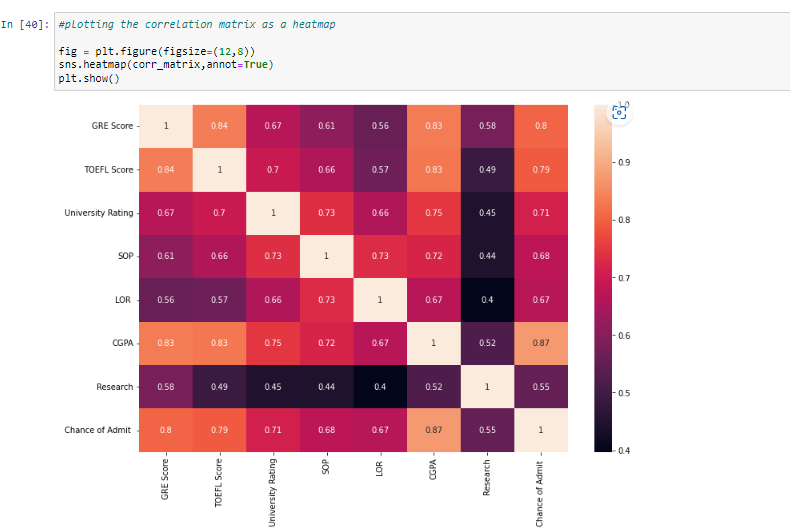
**Project Tracker, Velocity & Burndown Chart:**

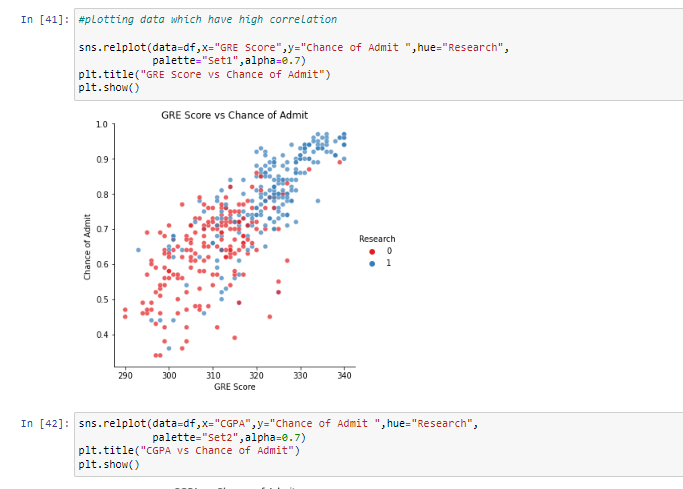


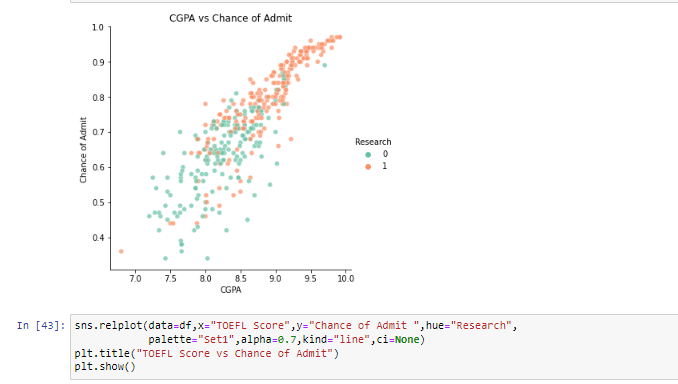
**MODEL BUILDING**

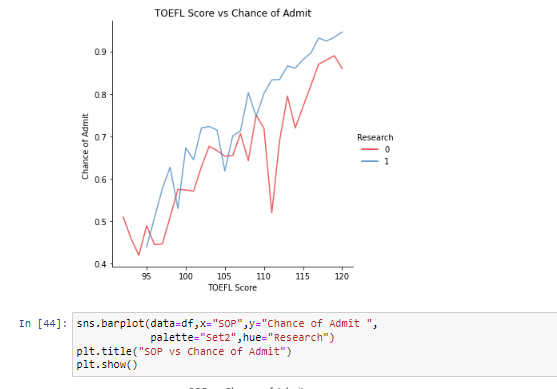


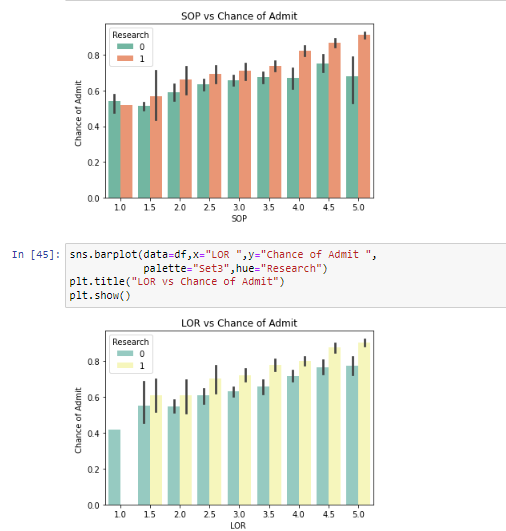


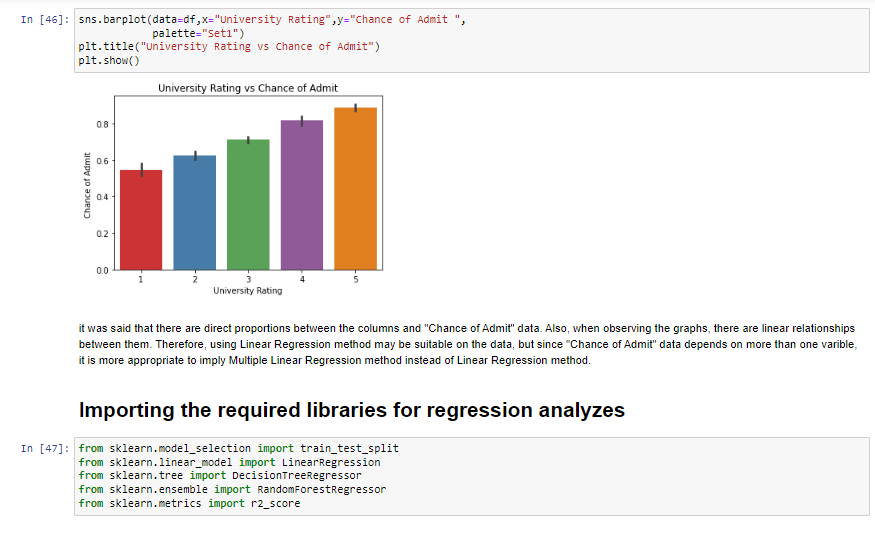


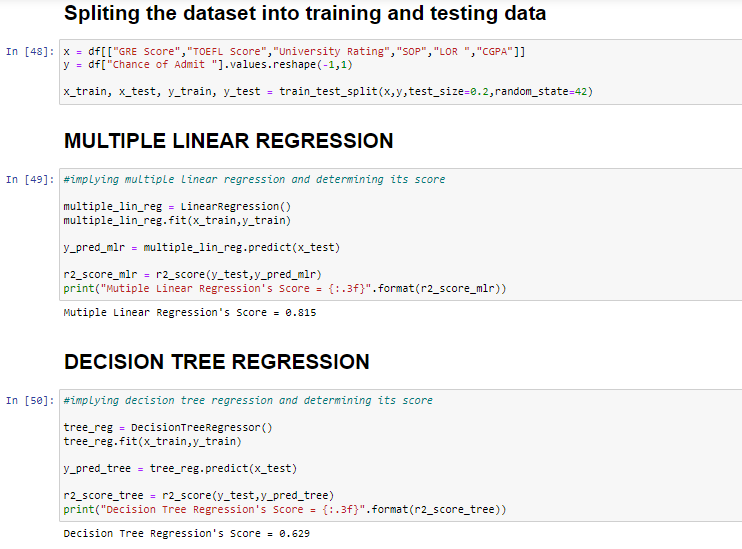


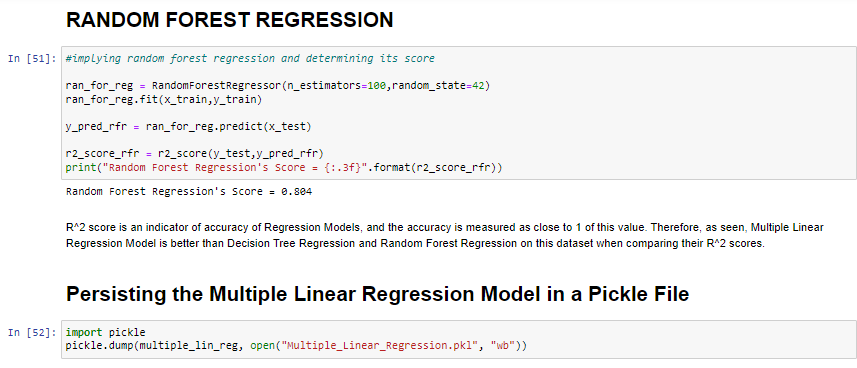


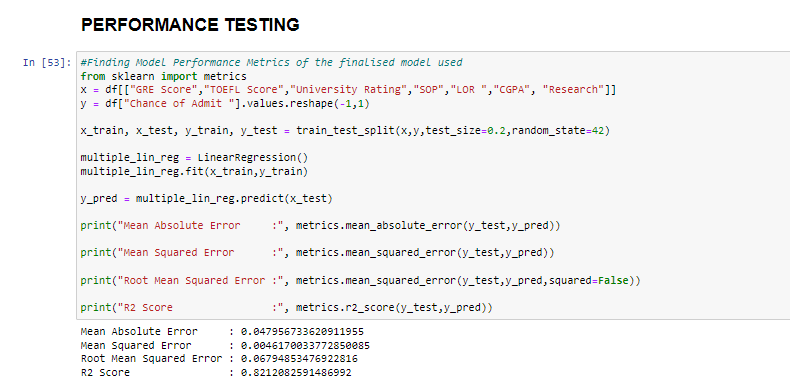












**ADVANTAGES & DISADVANTAGES**

**Advantages:**

● The system helps the applicant to save loads of money in the university application process by predicting the percentage chances of admit.

● The easy to use and elegant UI will attract the applicants to use the web application.

● Graphical visualizations and univariate analysis of the data will help the candidate to understand and analyse what factors play a crucial role in increasing his chances of admittance.

● The final deliverable is hosted as a web app, leveraging SaaS cloud principles that enables the end users to access the application from any convenient device.

**Disadvantages:**

● Since the end product is hosted in cloud as a SaaS, a stable internet connection is required.

● The current solution is hosted in a free tier cloud hosting account, which can handle a limited load. The web app may not handle the requests if the number of users increase at an exponential rate.

**CONCLUSION**

A model was developed to determine the admission of a student to the interested universities. The following parameters were taken into consideration:

● GRE Score

● TOEFL Score

● University Ranking

● SOP

● LOR

● CGPA

From the validations, we can find out that the above parameters greatly contributed in determining the “Chance of Admit” into an university. Different models - Multiple Linear Regression, Decision Tree Regression, Random Forest Regression were taken into consideration. Out of the 3 models, Multiple Linear Regression outperformed other models with a R2 score of 0.819. Hence Multiple Linear Regression was adopted in predicting the results. This model could likely be improved by gathering additional data of students from different universities which has similar selection criteria to choose the candidates for Master’s program.

**FUTURE SCOPE**

Future Scope of the project can be as follows:

● A future update could have chat space where candidates, faculties, current students of the university and alumni can interact and candidates can get their doubts resolved instantly.

● Get in touch with grad-schools’ and professors and determine other important factors that play a key role in increasing the chances of admission.

● To deal with huge volumes of data in the future (Both - applicants and university details), cloud based storages (IBM cloud, AWS, GCP, AZURE) and NoSQL databases (MongoDB, Redis, etc.) could be used instead of the traditional RDBMS storage.

● Alternatively, distributed big-data processing techniques could be explored if the no. of users using the website increases exponentially during the course of time.