

UNIVERSITY ADMIT ELIGIBILITY PREDICTOR

IBM NALAIYA THIRAN

Project Report submitted by

1.D.Shrinidhi(412519106029)

2.Gouthami.S(412519106032)

3.Gragalakshmi.L(412519106035)

4.Janani Priyadharshini.S(412519106053)

BACHELOR OF ENGINEERING

in

ELECTRONICS AND COMMUNICATION ENGINEERING

Project Report

Date	18 November 2022
Team ID	PNT2022TMID03986
Project Name	University Admit Eligibility Predictor

1.INTRODUCTION

Project Overview

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

The students or parents require a way to analyze and list the universities available for their cut-off mark so that the student or parent will be directly benefited by preparing themselves in par with the university's requirement. This can be implemented using Machine learning, Data analyzing and etc.

2. LITERATURE SURVEY

Existing Solution and Problem

[1] **Authors:** S. Sridhar, S. Mootha and S. Kolagati, ,.

Title: "A University Admission Prediction System using Stacked Ensemble Learning," 2020

Abstract: For an aspiring graduate student, shortlisting the universities to apply to is a difficult problem. Since an application is extremely dynamic, students often tend to wonder if their profile matches the requirement 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. The system could make use of data related to previous applicants to various universities and their admit or reject status. Earlier models of such prediction systems suffer from several drawbacks such as not considering important parameters like GRE (Graduate Record Exam) scores or research experience. Further, the accuracy reported by earlier models is also not sufficiently high. In this paper, a stacked ensemble model that predicts the chances of admit of a student to a particular university has been proposed. The proposed model takes into consideration various factors related to the student including their research experience, industry experience etc. Further, the system proposed has been evaluated against various other machine learning algorithms including other deep learning methods. It is observed that the proposed model easily outperforms all other models and provides a very high accuracy.

Observation: An effective method has been proposed to predict the chances of a student being admitted to a

specific university. In addition, we have compared the performance of various machine learning algorithms to the proposed method in predicting admits. It is observed that the proposed method provides the best performance with an accuracy of 91%.

[2] **Authors:** S. Fong, Y. -W. Si and R. P. Biuk-Aghai,

Title: "Applying a hybrid model of neural network and decision tree classifier for predicting university admission," .Abstract: Predicting university admission is a complex decision making process that is more than merely relying on test scores. It is known by researchers that students' backgrounds and other factors correlate to the performance of their tertiary education. This paper proposes a hybrid model of neural network and decision tree classifier that predicts the likelihood of which university a student may enter, by analysing his academic merits, background and the university admission criteria from that of historical records. Our prototype system was tested with live data from sources of Macau secondary school students. In addition to the high prediction accuracy rate, flexibility is an advantage as the system can predict suitable universities that match the students' profiles and the suitable channels through which the students are advised to enter. Our model can be generalized with other attributes and perform faster when compared to using a neural network alone.

Observation : A hybrid model upon which the Recommender System of Admission to University (RSAU) prototype is built. It analyses secondary school pupils' data from numerous sources to forecast their prospects of admission to institutions. It assists secondary school administrators, instructors, and senior secondary students in making recommendations to universities.

References

[1] S. Sridhar, S. Mootha and S. Kolagati, "A University Admission Prediction System using Stacked Ensemble Learning," 2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA), 2020, pp. 162-167, doi: 10.1109/ACCTHPA49271.2020.9213205.

[2] S. Fong, Y. -W. Si and R. P. Biuk-Aghai, "Applying a hybrid model of neural network and decision tree classifier for predicting university admission," 2009 7th International Conference on Information, Communications and Signal Processing (ICICS), 2009, pp. 1-5, doi: 10.1109/ICICS.2009.5397665

[3] Ahammad, Khalil & Chakraborty, Partha & Akter, Evana & Fomey, Umme & Rahman, Saifur. (2021). A Comparative Study of Different Machine Learning Techniques to Predict the Result of an Individual Student Using Previous Performances. International Journal of Computer Science and Information Security,. 19. 5-10. 10.5281/zenodo.4533374.

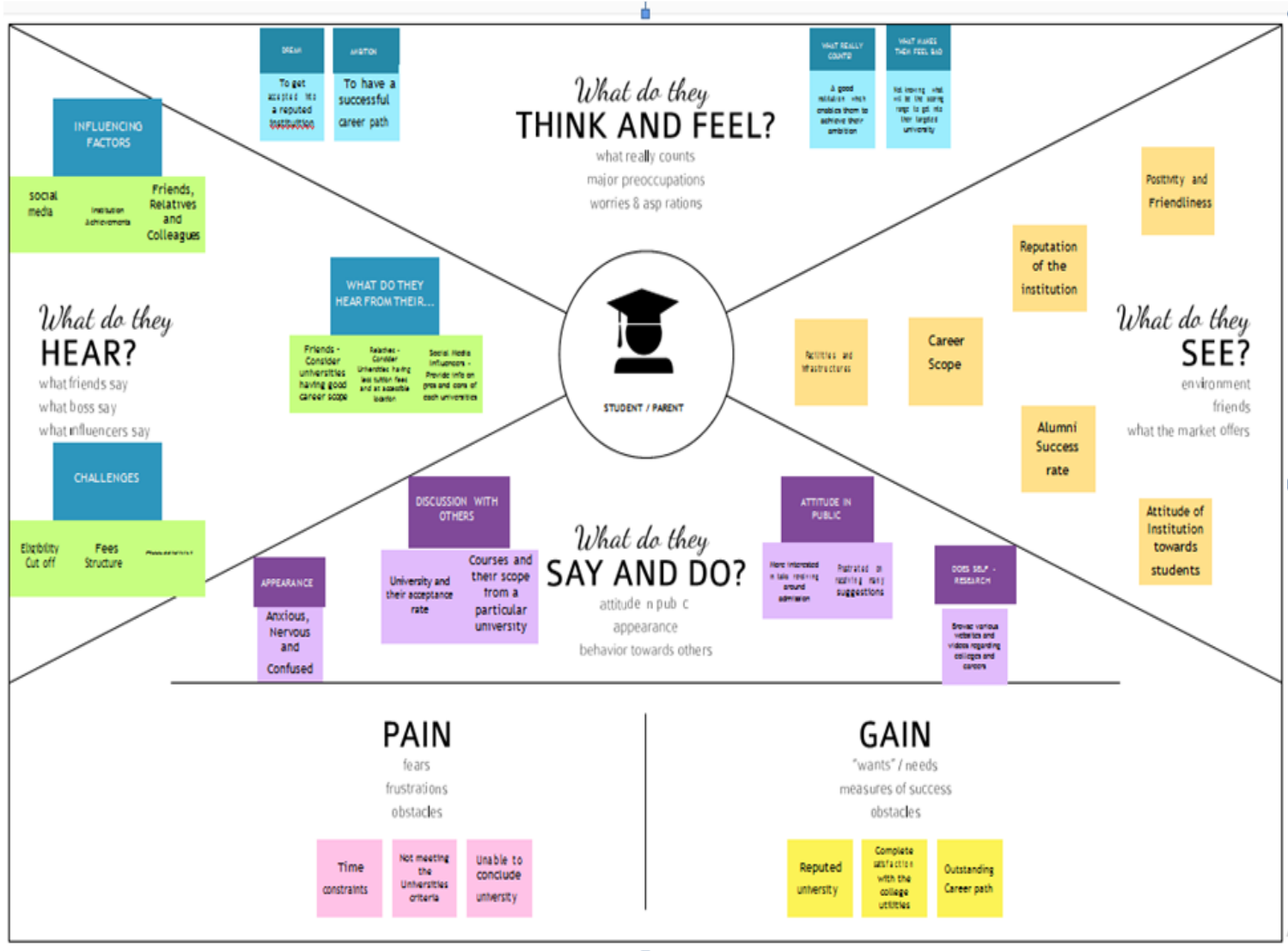
Problem Statement Definition

Students are often worried about their chances of admission to University. The aim of this project is to help students in short listing 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 .The students or parents require a way to analyze and list the universities available for their cut-off mark so that the student or parent will be directly benefited by

preparing themselves in par with the university’s requirement. This can be implemented using Machine learning, Data analyzing and etc.

IDEATION & PROPOSED SOLUTION

Empathy Map Canvas



Ideation & Brainstorming

Brainstorming

Gragalakshmi

UG / PG	If UG Ask State board/ Central	Enter cut- off
If PG, mention UG Details	Cost	Get Details of International Exam Scores
Fetch university details	Scholarships and funding	Facilities and infrastructure

Gouthami

Chatbot	ML model to be built	Get datasets
Enhance model based on feedback	List various universities based on input	Calculate overall cost
	Give outputs based on initial parameters	

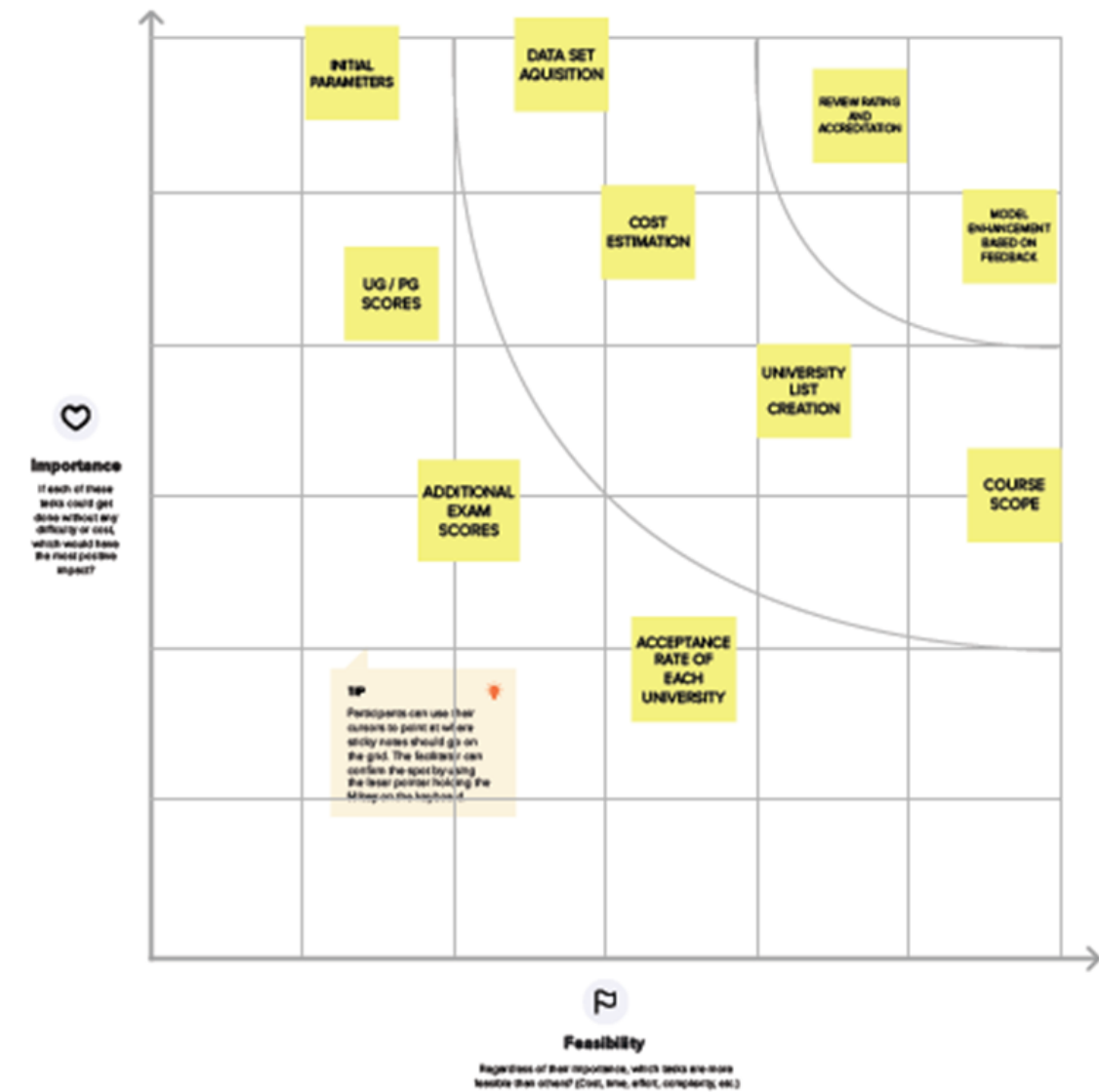
Shrinidhi

Targeted college Vs Received college	College reviews and ratings	Pictures and Videos of College
Facilities and infrastructure Information	College Ranking and Accreditation	Estimated Fees and Cost of Living
Available seats	% of acceptance	

Janani

User friendly webpage	Chatbot assistance	Initial Parameters for evaluation
Cut-off / scores	Get help from Chatbot if needed	Get details of university
	Leave feedback	

Ideation Prioritization



a. Proposed Solution

S.No	Parameter	Description
1.	Problem Statement (Problem to be solved)	The students and parents require a way to analyze and list the universities available for their cut-off mark so that the student or parent will be directly benefited by preparing themselves in par with the university's requirement.
2.	Idea / Solution description	A website is created where the users can login and give their details as input. Based on the data collected, universities above the ranking, universities on their range and below the range are displayed. Apart from that, data such as scholarships, accommodations, and cost of living are also displayed.
3.	Novelty / Uniqueness	University and User Login. Universities can upload what are their requirement from students. Mock tests scores are uploaded and based on which the predicted university list is given. A checkbox to add a university to preference and a dropbox to select the why they prefer.
4.	Social Impact/ Customer Satisfaction	Better performance of students are expected based on the results due to mock tests. Anxiety and Stress levels are reduced amidst users.
5.	Business Model (Revenue Model)	Subscription based model. Revenue from marketing.
6.	Scalability of the Solution	The solution proposed can be used by parents as well as students to know their cadre. Apart from this, universities can also make use of this website to know how many students are preferring the universities based on what criteria. Hence the scalability of the solution is high and widespread.

a.Problem Solution fit

Problem-Solution fit canvas 2.0		University Admission Predictor	
Define CS, fit into CC	1. CUSTOMER SEGMENT(S) CS <ul style="list-style-type: none"> Student Community Universities Parents 	6. CUSTOMER CONSTRAINTS CC <p>Students:</p> <ul style="list-style-type: none"> Cut-off mark criteria unknown University Requirements unknown <p>Parents:</p> <ul style="list-style-type: none"> Cost of living, tuition fee budget Details about the locality <p>University:</p> <ul style="list-style-type: none"> Why students do & do not prefer What are the factors that hold back students from applying What are required to enable students apply to universities 	5. AVAILABLE SOLUTIONS AS <p>Websites:</p> <ul style="list-style-type: none"> Unclassified data No proper tier/ ranking is given Inaccurate/ Misleading information on websites Limited number of searches <p>Counselling Services</p> <ul style="list-style-type: none"> Expensive Might not get into preferred universities
	2. JOBS-TO-BE-DONE / PROBLEMS J&P <ul style="list-style-type: none"> Data acquisition Accuracy Maintenance Security and Privacy handling Verification of acquired details Query handling 	9. PROBLEM ROOT CAUSE RC <p>The students or parents require a way to analyse and list the universities available for their cut-off mark so that the student or parent will be directly benefitted by preparing themselves in par with the university's requirement.</p>	7. BEHAVIOUR BE <ul style="list-style-type: none"> Look out for websites with accurate and elaborate details Compare various details and cross verify Give several data as input and improve their search results Ask queries and look out for immediate response Privacy and Security of the website is ensured
Identify strong TR & EM	3. TRIGGERS TR <ul style="list-style-type: none"> Acceptance rate Feedback and reviews Preferred suggestions 	10. YOUR SOLUTION SL <ul style="list-style-type: none"> A website is created where the users can login and give their details as input. Based on the data collected, universities above the ranking, universities on their range and below the range are displayed. Apart from that, data such as scholarships, accommodations, and cost of living are also displayed. University and User Login. Universities can upload what are their requirement from students. Mock tests scores are uploaded and based on which the predicted university list is given. A checkbox to add a university to preference and a drop box to select the why they prefer. 	8. CHANNELS of BEHAVIOUR CH <p>8.1 ONLINE</p> <ul style="list-style-type: none"> Review and Feedback Choice of preference
	4. EMOTIONS: BEFORE / AFTER EM <p>Before</p> <ul style="list-style-type: none"> Anxious & Nervous Fear of failure <p>After</p> <ul style="list-style-type: none"> Confident Clear about career prospect 		<p>8.2 OFFLINE</p> <ul style="list-style-type: none"> Peer discussions Real-time user experience

Explore AS, differentiate

Focus on J&P, tap into BE, understand RC

Extract online & offline CH of BE

REQUIREMENT ANALYSIS

Functional Requirements

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail Registration through LinkedIn
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	User Details	GRE or/and TOEFL Score Sheet Curriculum Vitae (CV)
FR-4	User Requirements	Upload all the necessary documents in the appropriate location in the website . Based on the uploads, the system would see through all the necessary information . The list of all eligible universities for the candidate would be displayed based on the given information with criteria.

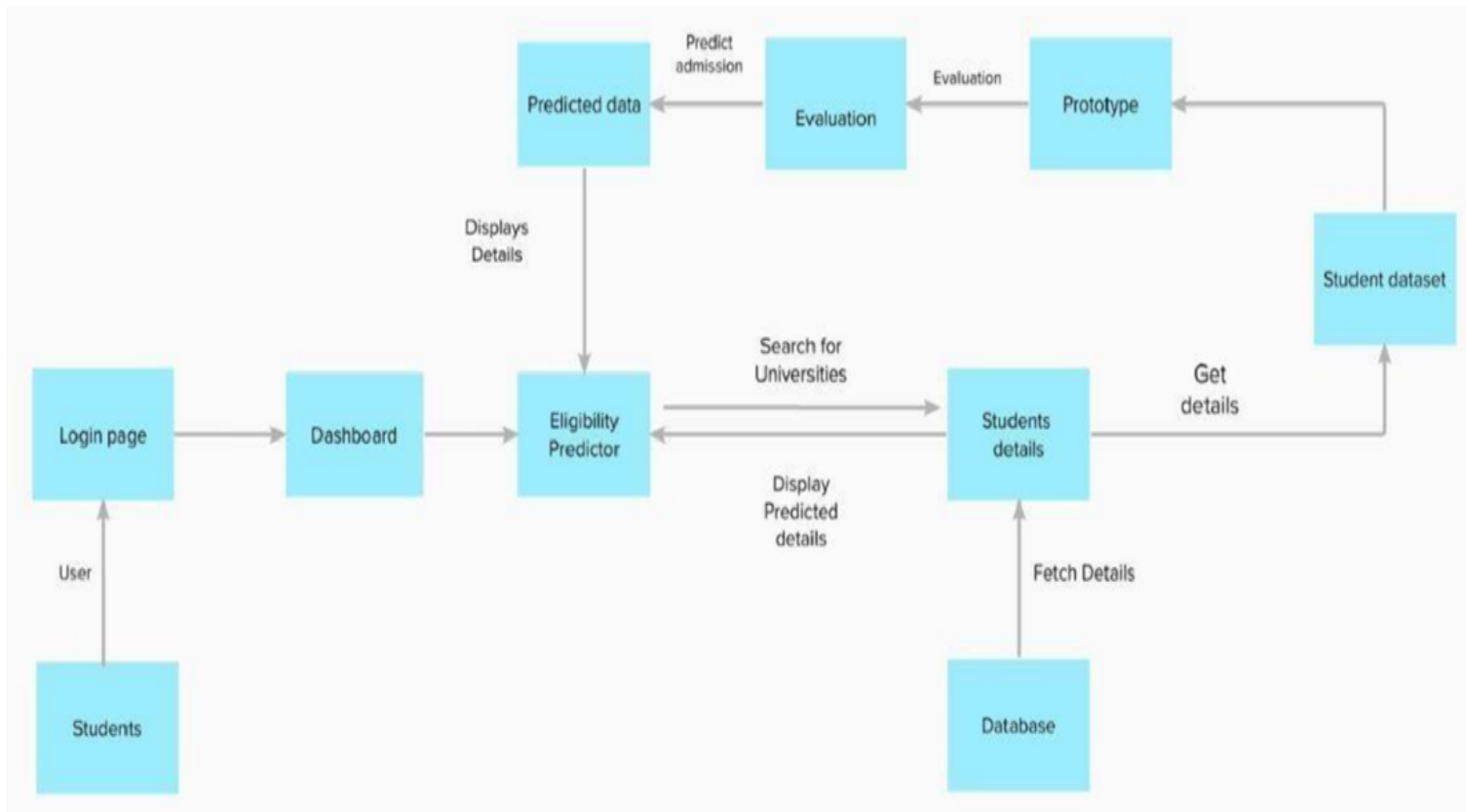
Non-Functional

FR No.	Non-Functional Requirement	Description
NFR-1	Usability	User friendly Easy to use without any prior knowledge More convenient to access and highly efficient in prediction with consuming less time.
NFR-2	Security	Since each user has their own credentials for login, it is more authenticated to access any individual's website. Database is stored and backed up for future use in a more authenticated way.
NFR-3	Reliability	Highly reliable in prediction based on the data provided.

NFR-4	Performance	<p>The website can efficiently handle the traffic by service the request as soon as possible.</p> <p>Viewing this webpage using a 56-kbps modem connection would not exceed 30 seconds (quantitatively, the mean time).</p>
NFR-5	Availability	Fast and efficient
NFR-6	Scalability	<p>Since an academic portal is crucial to the courses that use it, it is crucial that a sizable number of users be able to access the system at the same time.</p> <p>The admission season is probably when the system will be under the most strain.</p> <p>It must therefore be able to manage numerous concurrent users.</p>

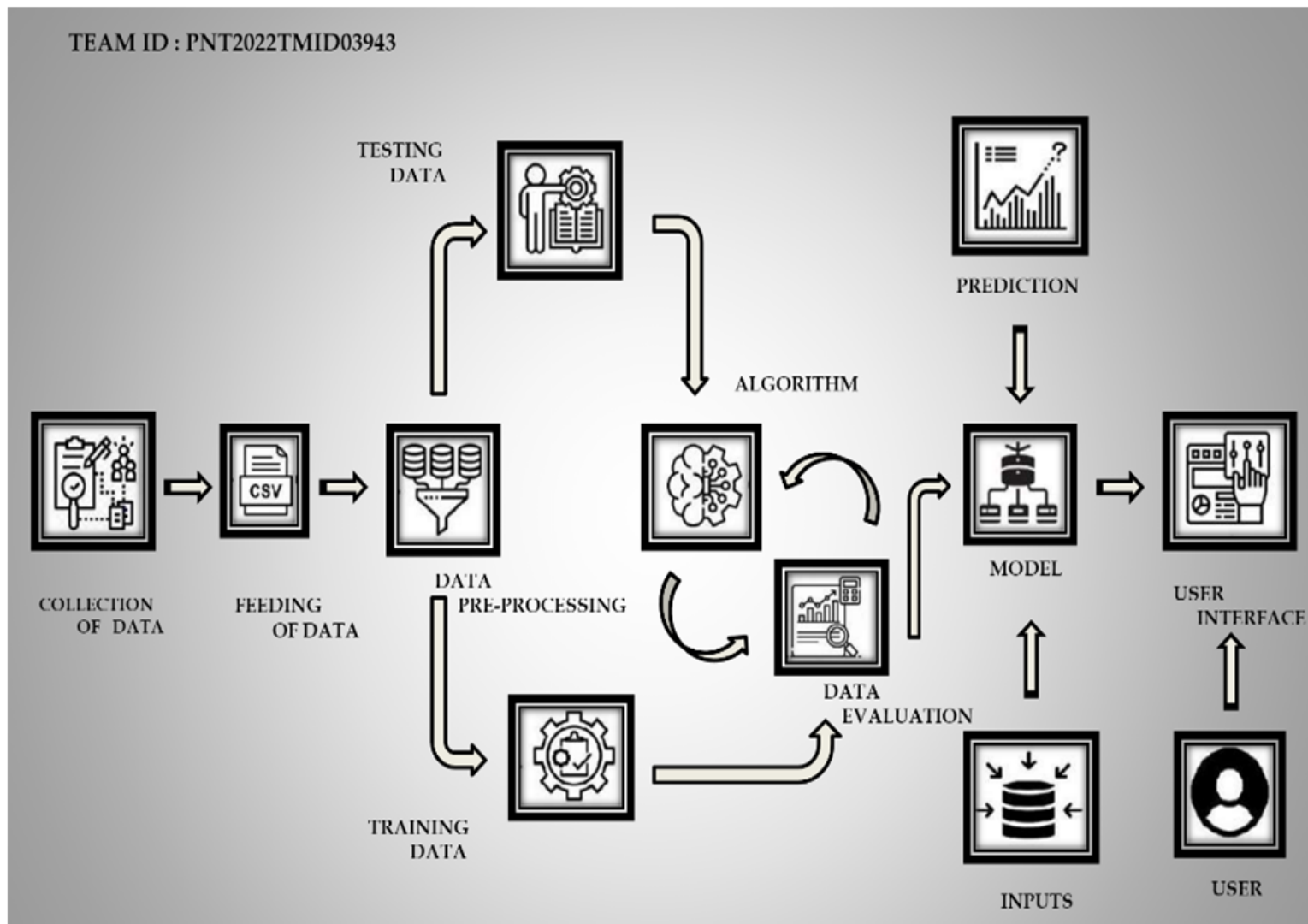
PROJECT DESIGN

a. Data Flow Diagrams



Solution & Technical Architecture

1. Solution Architecture



Technical Architecture

S NO	COMPONENTS	DESCRIPTION	TECHNOLOGY
1	User Interface	How User interacts with the application (For eg : WEB UI, Mobile App, etc)	HTML , CSS , JavaScript, etc
2	Application Logic - 1	Logic for processing application	Python
3	Application Logic - 2	Logic for processing application	IBM Watson STT Service
4	Application Logic - 3	Logic for processing application	IBM Watson Assistant
5	Database	Data type, Configuration ,etc	MySQL , etc
6	Cloud Database	Database Service on Cloud	IBM DB2
7	File Storage	File storage requirements	IBM Block Storage / Other Storage Service
8	Infrastructure (Server / Cloud)	Application Deployment	Local, Cloud Foundry, etc

PROJECT PLANNING & SCHEDULING

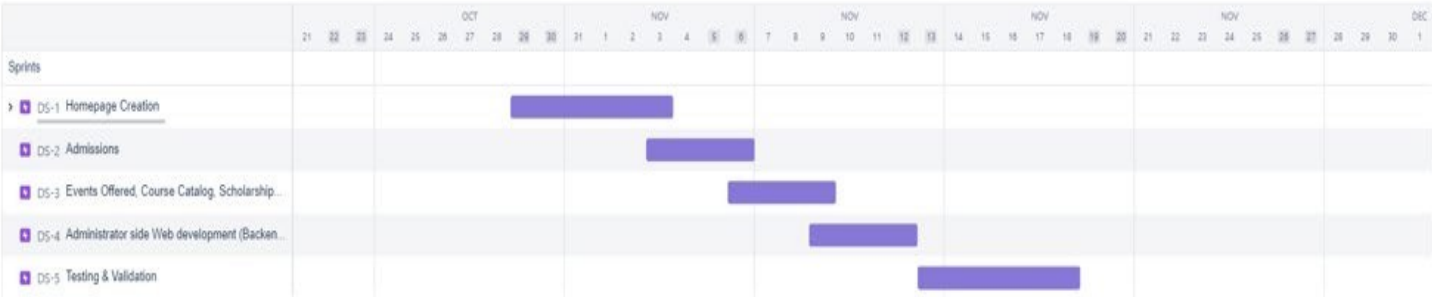
a. Sprint Planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	2	High	3
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	1	High	3
Sprint-2		USN-3	As a user, I have to upload the data which are necessary for prediction.	2	Low	2
Sprint-3		USN-4	As a user, I can able to choose the eligible university based on the uploaded data.	2	Medium	2
Sprint-4	Login	USN-5	As a user, I can log into the application by entering email & password	1	High	4
	Dashboard		Check the dashboard and get to know their eligible criteria based on the uploaded datasets.			4

Sprint Delivery Schedule

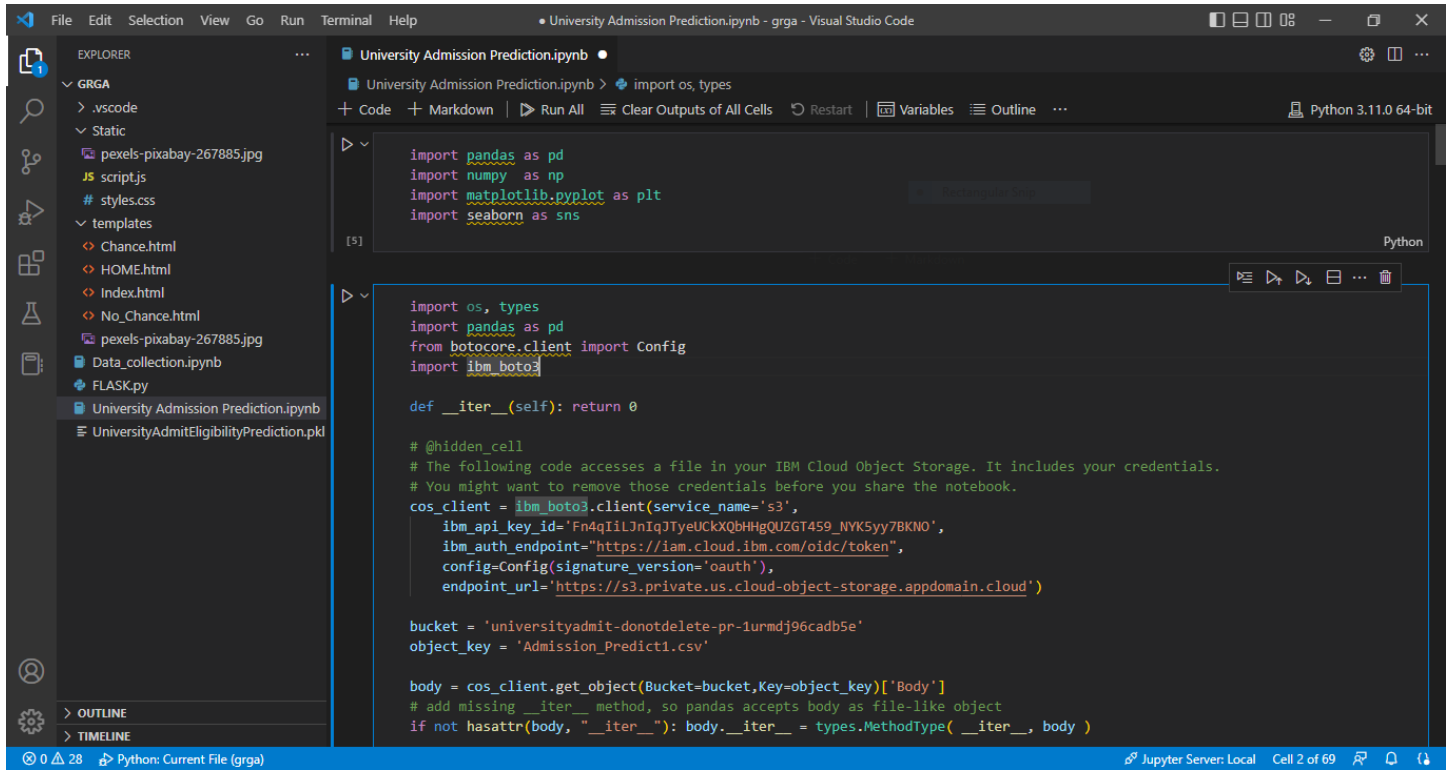
Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	15	07 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	15	15 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	20 Nov 2022

Reports from JIRA



A.CODING & SOLUTIONING:

Modelling:



The screenshot shows a Jupyter Notebook titled "University Admission Prediction.ipynb" in Visual Studio Code. The notebook is running on a Jupyter Server (Local). The code in the notebook is as follows:

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

import os, types
from boto3.client import Config
import boto3

def __iter__(self): return 0

# @hidden_cell
# The following code accesses a file in your IBM Cloud Object Storage. It includes your credentials.
# You might want to remove those credentials before you share the notebook.
cos_client = boto3.client(service_name='s3',
    ibm_api_key_id='Fn4qIilJnIqJTyeUckXQbHHgQUZGT459_NVK5yy7BKN0',
    ibm_auth_endpoint="https://iam.cloud.ibm.com/oidc/token",
    config=Config(signature_version='oauth'),
    endpoint_url='https://s3.private.us.cloud-object-storage.appdomain.cloud')

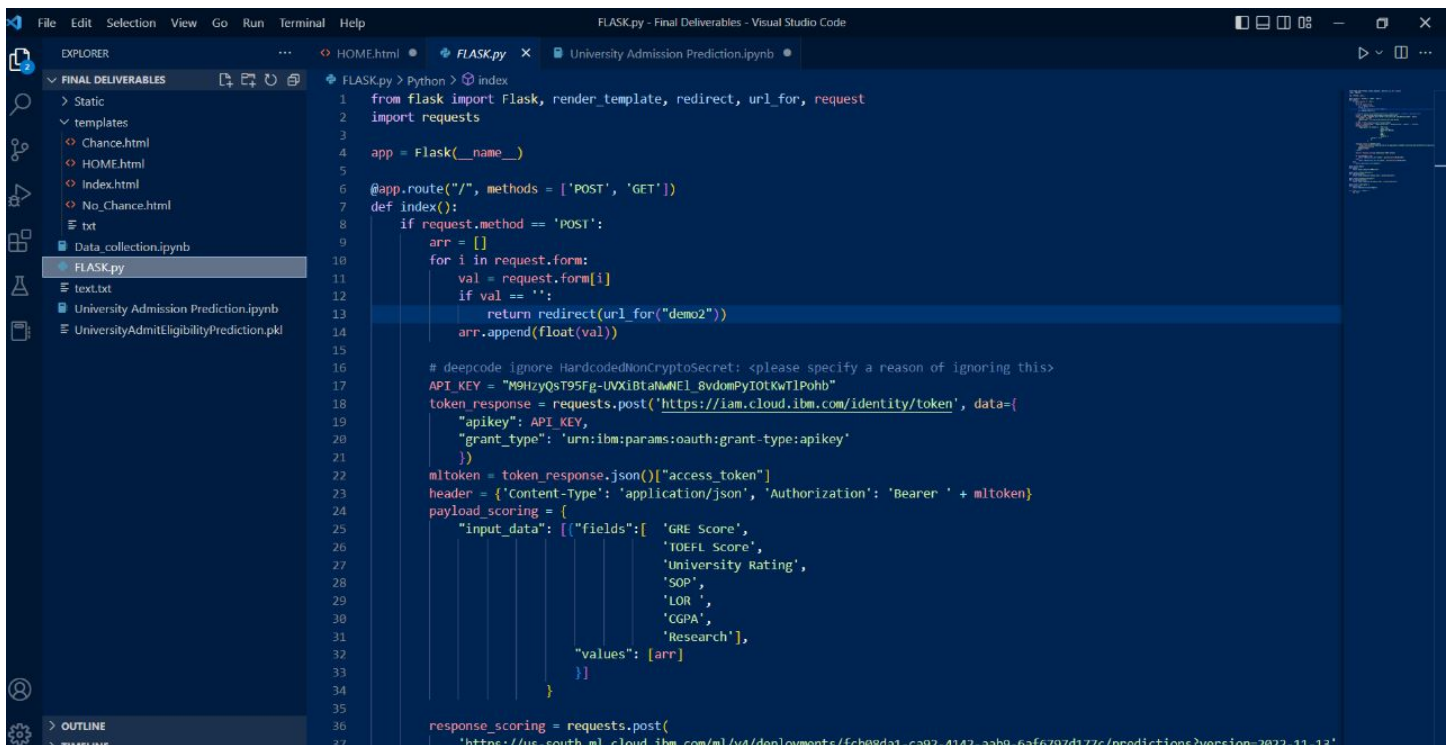
bucket = 'universityadmit-donotdelete-pr-1urmdj96cadb5e'
object_key = 'Admission_Predict1.csv'

body = cos_client.get_object(Bucket=bucket,Key=object_key)['Body']
# add missing __iter__ method, so pandas accepts body as file-like object
if not hasattr(body, "__iter__"): body.__iter__ = types.MethodType( __iter__, body )
```

Index.html:

```
templates > <> Index.html > html
1  <!DOCTYPE html>
2  <html lang="en">
3
4  <head>
5      <meta charset="UTF-8">
6      <meta http-equiv="X-UA-Compatible" content="IE=edge">
7      <meta name="viewport" content="width=device-width, initial-scale=1.0">
8      <meta name="viewport" content="width=device-width, initial-scale=1, maximum-scale=1, user-scalable=no">
9      <link rel="stylesheet" type="text/css" rel="noopener" target="_blank" href="../static/css/styles.css">
10     <link href="https://cdn.jsdelivr.net/npm/bootstrap@5.2.2/dist/css/bootstrap.min.css" rel="stylesheet" integrity="sha384-Zenh87qX5"
11     <title>University Admit Eligibility Predictor</title>
12 </head>
13
14 <body>
15     {% block body %}
16
17     <h1>index page
18     </h1>
19
20     {% endblock %}
21     <script src="https://cdn.jsdelivr.net/npm/bootstrap@5.2.2/dist/js/bootstrap.bundle.min.js" integrity="sha384-OERcA2EqJJCMA+/3y+gx"
22 </body>
23
24 </html>
```

Flask:



```
FLASK.py - Final Deliverables - Visual Studio Code
EXPLORER
  FINAL DELIVERABLES
    Static
    templates
      Chance.html
      HOME.html
      Index.html
      No_Chance.html
    txt
    Data_collection.ipynb
    FLASK.py
    text.txt
    University Admission Prediction.ipynb
    UniversityAdmitEligibilityPrediction.pkl

FLASK.py > Python > index
1  from flask import Flask, render_template, redirect, url_for, request
2  import requests
3
4  app = Flask(__name__)
5
6  @app.route("/", methods = ['POST', 'GET'])
7  def index():
8      if request.method == 'POST':
9          arr = []
10         for i in request.form:
11             val = request.form[i]
12             if val == '':
13                 return redirect(url_for("demo2"))
14             arr.append(float(val))
15
16         # deepcode ignore HardcodedNonCryptoSecret: <please specify a reason of ignoring this>
17         API_KEY = "M9HzysT95Fg-UvXIBtaWNEl_8vdomPyIoTkwTlPohb"
18         token_response = requests.post('https://iam.cloud.ibm.com/identity/token', data={
19             "apikey": API_KEY,
20             "grant_type": 'urn:ibm:params:oauth:grant-type:apikey'
21         })
22         mltoken = token_response.json()["access_token"]
23         header = {'Content-Type': 'application/json', 'Authorization': 'Bearer ' + mltoken}
24         payload_scoring = {
25             "input_data": [{"fields": [
26                 'GRE Score',
27                 'TOEFL Score',
28                 'University Rating',
29                 'SOP',
30                 'LOR ',
31                 'CGPA',
32                 'Research'
33             ],
34                 "values": [arr]
35             }
36         ]
37         response_scoring = requests.post(
38             'https://us-south.ml.cloud.ibm.com/ml/v4/deployments/fcb88da1-c892-4142-aab9-6af6797d177c/predictions?version=2022-11-13',
```


SOLUTIONING:

127.0.0.1:5000/home

IEEE XploreIBMGitHubKaggleMathWorks AccountTamil eBooksSearch | RSNA Case...IBM CloudProject Templates ~...Payment Status | Pa...UniBest

UNIVERSITY ADMIT ELIGIBILITY PREDICTOR

This university eligibility predictor is a tool created for individuals who have completed their undergrad studies and hope to be accepted into prestigious colleges. It gives them the opportunity to apply to several universities, based on specific standards related to their admissions procedure. It might prove to be very beneficial because it can give students some clarification regarding the universities to which they can submit an application for admission.

The Team ID : PNT2022TMID03943

Enter the details

GRE Score:

250 to 340

TOFEL Score:

50 to 120

University Rating:

1 to 5

SOP:

1 to 5

LOR:

1 to 5

CGPA:

5 to 10

Research:

☐ Yes

☒ No

Predict

Eligible Criteria:

127.0.0.1:5000/chance/79.58940853455628

IEEE XploreIBMGitHubKaggleMathWorks AccountTamil eBooksSearch | RSNA Case...IBM CloudProject Templates ~...Payment Status | Pa...UniBest

Congratulations!

Yay!The model has predicted that you have a chance

Eligibility Score:79.58940853455628%

Go Back

Not Eligibile Criteria:

127.0.0.1:5000/nochance/43.79285108902343

IEEE Xplore

IBM

GitHub

Kaggle

MathWorks Account

Tamil eBooks

Search | RSNA Case...

IBM Cloud

Project Templates -...

Payment Status | Pa...

UniBest

Sorry

Oops!The model has predicted that you don't have a chance

Eligibility Score:**43.79285108902343%**

Go Back

Defect Analysis

This report shows the number of resolved or closed bugs at each severity level, and how they were resolved

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	10	4	2	3	20
Duplicate	1	0	3	0	4
External	2	3	0	1	6
Fixed	11	2	4	20	37
Not Reproduced	0	0	1	0	1
Skipped	0	0	1	1	2
Won't Fix	0	5	2	1	8
Totals	24	14	13	26	77

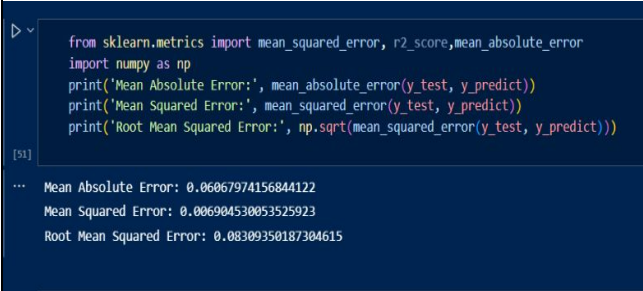
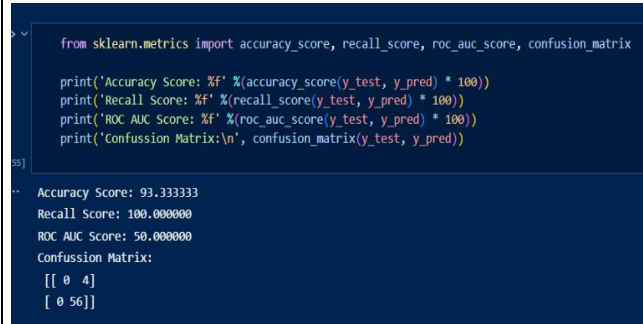
Test Case Analysis

This report shows the number of test cases that have passed, failed,and untested

Section	Total Cases	Not Tested	Fail	Pass
Print Engine	7	0	0	7
Client Application	42	0	0	43
Security	2	0	0	2
Outsource Shipping	3	0	0	3
Exception Reporting	9	0	0	9
Final Report Output	4	0	0	4
Version Control	2	0	0	2

1.RESULTS:

Performance metrics

S.No.	Parameter	Values	Screenshot
1.	Metrics	<p>Regression Model:</p> <p>MAE -</p> <p>0.06867974156844122</p> <p>,MSE</p> <p>-0.006984530053559</p> <p>23</p> <p>RMSE -</p> <p>0.08309350187384615</p> <p>Classification Model:</p> <p>Confusion Matrix -</p> <p>[[0 4]</p> <p>0 56]] , Accuracy</p> <p>Score-93.333333 &</p>	 <pre>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)))</pre> <p>[51]</p> <p>... Mean Absolute Error: 0.06867974156844122 Mean Squared Error: 0.00698453005355923 Root Mean Squared Error: 0.08309350187384615</p>  <pre>from sklearn.metrics import accuracy_score, recall_score, roc_auc_score, confusion_matrix print('Accuracy Score: %f' %(accuracy_score(y_test, y_pred) * 100)) print('Recall Score: %f' %(recall_score(y_test, y_pred) * 100)) print('ROC AUC Score: %f' %(roc_auc_score(y_test, y_pred) * 100)) print('Confusion Matrix:\n', confusion_matrix(y_test, y_pred))</pre> <p>[55]</p> <p>... Accuracy Score: 93.333333 Recall Score: 100.000000 ROC AUC Score: 50.000000 Confusion Matrix: [[0 4] [0 56]]</p>

2	Tune the Model	Hyper parameter Tuning - Validation Method -	<pre> from sklearn import datasets from sklearn.tree import DecisionTreeClassifier from sklearn.model_selection import StratifiedKFold, cross_val_score X, y = datasets.load_iris(return_X_y=True) clf = DecisionTreeClassifier(random_state=42) sk_folds = StratifiedKFold(n_splits = 5) scores = cross_val_score(clf, X, y, cv = sk_folds) print("Cross Validation Scores: ", scores) print("Average CV Score: ", scores.mean()) print("Number of CV Scores used in Average: ", len(scores)) </pre> <p> Cross Validation Scores: [0.96666667 0.96666667 0.9 0.93333333 1. Average CV Score: 0.9533333333333334 Number of CV Scores used in Average: 5 </p>
---	----------------------	--	---

ADVANTAGES AND DISADVANTAGES

Advantages

- It helps student for making decisionfor choosing a right college.
- Here the chance of occurrence of error is less when compared with the existingsystem.
- It is fast, efficientand reliable.
- Avoids data redundancy and inconsistency.
- Very user-friendly.
- Easy accessibility of data.

Disadvantages:

- Required active internet connection.
- System will provide inaccurate results if data entered incorrectly.

● CONCLUSION

Thus it is concluded that our model provides a robust prediction score with Multiple Linear Regression Algorithm & users are able to predict the scores what they expected. User experience of the websiteis a simple one to use & puts the user at ease to utilize the feature.

● FUTURE SCOPE

We intend to enhance the user experience by adding the details of every universitiesso that the user can get to know the culture of the place, alumni reviews, rating of the universities etc.

- **APPENDIX**

Github :<https://github.com/IBM-EPBL/IBM-Project-23614-1659888568>

Project Demo: https://www.youtube.com/watch?v=tkz4m2_RRUw