PROJECT REPORT

Date	19 November ,2022
Team ID	PNT2022TMID21492
Project Name	University Admit Eligibility Predictor

UNIVERSITY ADMIT ELIGIBILITY PREDICTOR

1. INTRODUCTION

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 to a particular university. This analysis should also help students who are currently preparing or will be preparing to get a better idea. College admission predictor is a boon to many students. This helps the student not only to help in filling out theapplication forms but also give the students an idea about their future college by calculating their cut off.

1.1 Project Overview

The university prediction mostly depends the initial steps that are being carried out. Firstly, the dataset collection. The data that is collected for the prediction must be accurate and concise in nature. Any discrepancies in the dataset will cost the accuracy of the university prediction directly. The second step is Data preprocessing. The collected data is improper i.e., those data will have outliers, missing values and the number of attributes may also be huge. At times the data can also be unstructured. In order to solve this issue, the data must be cleaned and preprocessed in a proper manner. The next important issues arise with the data consistency, the university prediction data must be consistent.

1.2 Purpose

The university prediction may help the school student's hugely to predict good possibility of good college. This university prediction can help the students hugely to plan ahead and save their valuable time. The reputation of the university can be majorly dependent on these university prediction because the prediction has a direct hand in determining the society trust on the application. Proper data preprocessing helps the model to get a high yield accuracy.

2. LITERATURE SURVEY

2.1 Existing Problem

S.	Author	Title of the	Publicati	Description	Advantage
No	Name	Paper	on		
			Year		
1	Sara Graduate	2020	Student	The machine	
	Aljasmi,	Admission		admission	learning
	Ali Bou	Prediction		problem is	models
				very important	included are
	Nassif,	Using		in educational	multiple
	Ismail	Machine		institutions.	linear
	Shahin,	Learning		This paper	regression,
	Ashraf	3		addresses	k-nearest
				machine	neighbor,
	Elnaga			learning	random
	r.			models to	forest, and
				predict the	Multilayer
				chance of a	Perceptron.
				student to be	Experiments
				admitted to a	show that
				master's	the
				program. This	Multilayer
				will assist	Perceptron
				students to	model
				know in	surpasses
				advance if	other
				they have a	models.
				chance to get	
				accepted.	

2	Sharan	Predicting	2020	The primary	This
	Kumar	Student		purpose is to	model
	Paratala	University		discuss the	would
		Admission		prediction of	likely be
	Rajagopa	Using Logistic		student	greatly
		Regression		admission to	improved
				university	by the
				based on	gathering
				numerous	of
				factors and	additional
				using logistic	data of
				regression.	students
				The admission	from
				decision	different
				depends on	universiti
				criteria within	es which
				the particular	has
				college or	similar
				degree	selection
				program. The	criteria to
				independent	choose the
				variables in	candidates
				this study will	for
				be measured	Master's
				statistically to	program.
				predict	
				graduate	
				school	
				admission.	

3	<u>Jia Qing</u>	Introduction	2021	The objective	Linear
		to Modelling		of this analysis	bestregressi
		Tabular Data:		is to explore	on seems to
				the most	perform the
		Predicting a		important	compared
		student's		factors for a	to the
		chance of		student to get	neural net
		gaining		into graduate	and the
				school and to	random
		admission		select the most	forest
		using ML		accurate	which
				model to	proves that
				predict a	complicat
				student's	ed models
				chances of	doesn't
				gaining	always
				admission into	produce
				Graduate	better
				school	results.
					ensembling
					the models
					produced a
					better result

M Ragab, Abdul Fatah S. Mashat, Ahmed M Khedra	Recommender system for predicting college admission	presents 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.	prediction accuracy rate, flexibility is an advantage, as the system can predict suitable colleges that match the students' profiles and the suitable track channels through which the students are advised to enter. The system is adaptive, since performing trusted needed tasks

2.2 References

[1]Sara Aljasmi,Ali Bou Nassif, Ismail Shahin,Ashraf Elnagar "Graduate Admission Prediction Using Machine Learning" , 2020

[2] Sharan Kumar Paratala Rajagopa, "Predicting Student University Admission

Using Logistic Regression",2020

- [4] Abdul Hamid M Ragab, Abdul Fatah S. Mashat, Ahmed M Khedra_,"HRSPCA:Hybrid Recommendersystem forpredicting college admission", 2012
 - 2.3 Problem Statement Definition

3.IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

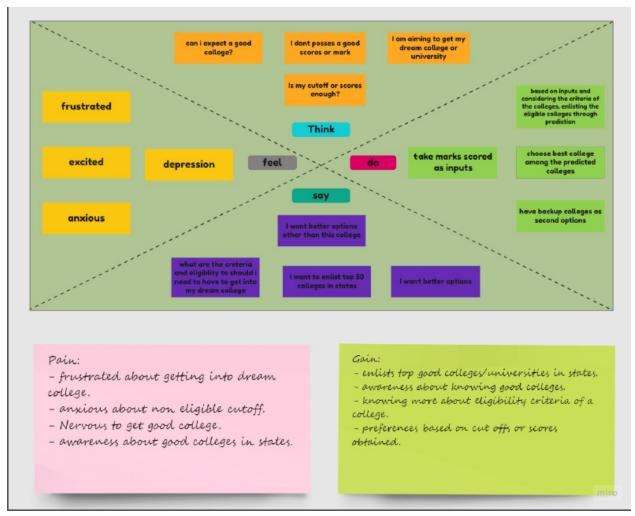


Figure 1

This Figure 1 shows a empathy map that will gives a collaborative visualization for the end user, what they can perform using this project, what challenges have been faced and also what is the

3.2 Ideation & Brainstorming

3.2.1 Brainstorm

Brainstorming phase, the ideas from every group members are gathered.

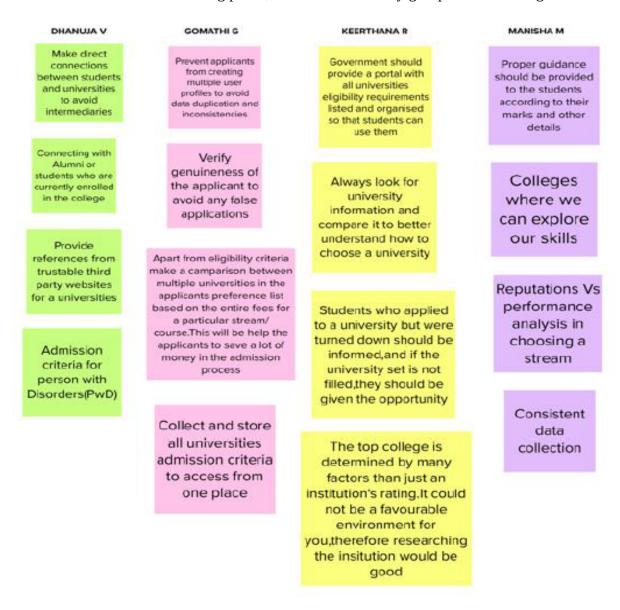


Figure 2 Brainstorm

3.2.2 Group ideas

Grouping the ideas under the suitable topics for better understanding



Figure 3 Group ideas

3.2.3 Prioritize

Prioritizing the ideas or the features and performing the feasibility study on it.

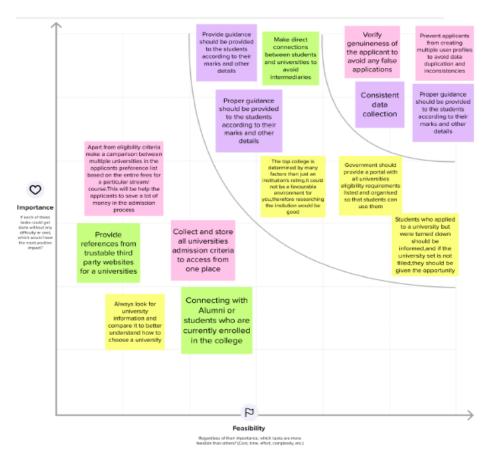


Figure 4 Prioritize

3.3 Proposed Solution

S.No.	Parameter	Description
1	Duahlam Statement (Duahlam to	Ctudents are often averaged about their
1.	Problem Statement (Problem to besolved)	Students are often worried about their
	besorved)	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 chancesin a particular university. This analysis should also help students who are currently preparing or will be preparing to get abetter idea. It also aims to make a direct connection between the students and the universities andavoid any intermediaries.
2.	Idea / Solution description	This project intends to calculate the probability of acceptance in a particular grad-school afterassessing the candidate's profile. The key attributes that will be considered formaking the decisions are: a. GRE & TOEFL Scores b. Undergrad CGPA c. SOP & LOR d. Corporate Work Experience / Research Experience e. Extracurriculars For determining the % of acceptance, we will be using various ML models such as Logistic Regression, Multiple Linear Regression, Decision Tree & Random Forest and assess which model gives the highest accuracy with the help of performance metrics like accuracy-score, precision andrecall.
3.	Novelty / Uniqueness	1. We intend to develop a novel deep learning- based hybrid model that has a better accuracy than the existing traditional ML models 2. The web-app will also provide feedback on the parameters where

		the candidate is lacking so that he can improve on thoseareas.
4.	Social Impact/ Customer Satisfaction	Students often feel difficult in shortlisting theuniversities to apply which they tend to wonderif their profile matches the requirement of a certain university.
		2. Moreover, the cost of applying to a university is extremely high making it critical that students shortlist universities basedon their profile.
		3. A university admission prediction system is quite useful for students to determine their chancesof acceptance to a specific university.
		4. This system reduces dependence oneducational consultancies, who charge loads of money to analyze a candidate's profile and determine the universities he/she should apply to.
5.	Business Model(Revenue Model)	Advertisements of different universities couldbe placed in the web-app to generate revenue through ads.
		2. In the future, a separate premium plan couldbe created where the students can directly interact withthe professors and alumni of the university through video calls.

6. Scalability of the Solution 1. A future update could have chat space wherecandidates, faculties, current students of the university and alumni can interact and candidates can get their doubts resolved instantly. 2. 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.

3.4 Problem Solution fit



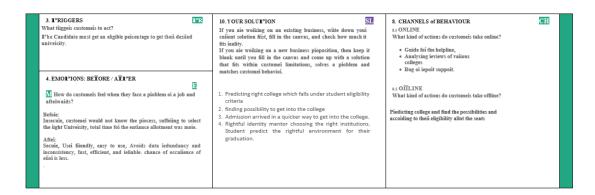


Figure 5 Problem Solution fit

4. REQUIREMENT ANALYSIS

4.1 Functional requirement

Following are the functional requirements of the proposed solution

FR NUMBER	FUNCTIONAL	SUB TASKS
	REQUIREMENT	
FR 1	Upload the basic marksand cutoff scoresfor theparticular college/university.	enter the marks based on subjectwise along with cut off scores for criteria satisfaction.
FR 2	Predict colleges basedon marks and cut off.	Predicts the colleges with the help of trained dataset.
FR 3	Top good university prediction based on marks.	Enlists all the top good universities basedon marks and cut offscores.

4.2 Non-Functional requirement

Following are the non-functional requirements of the proposed solution

FR NUMBER	NON- FUNCTIONAL REQUIREMENT	SUB TASKS
NFR 1	Usability	 No training is required to use thewebsite. The results from the predictor should not take more than 30 seconds.
NFR 2	Security	 The system shall provide password protected access to the website to all users –students & admins both.
NFR 3	Reliability	 University Application process itself being a tedious task students needs lots of determination for completing overall application process. It seems students have to work on lots ofthings when he/she prepares for application process.
NFR 4	Performance	 This system can support any number of users at a time. The mean time to view a webpage overa 56 Kbps modem connection shall not exceed5 seconds.

5.PROJECT DESIGN

5.1 Data Flow Diagrams

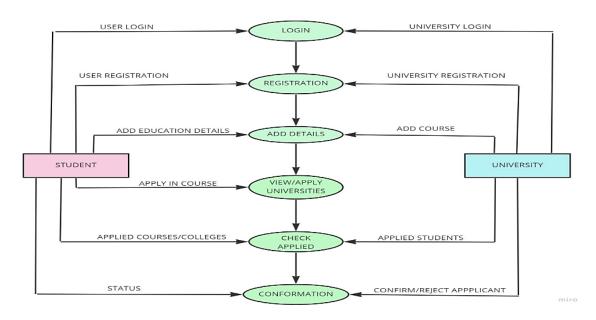


Figure 6 Data flow diagram

5.2 Solution & Technical Architecture

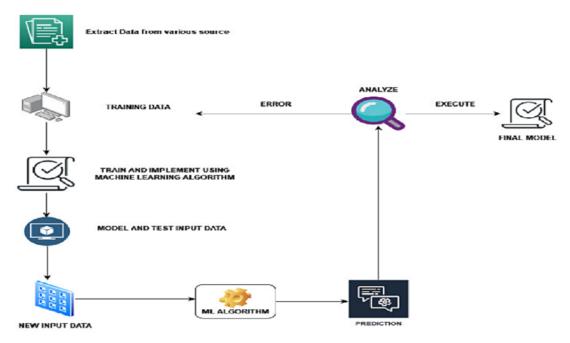


Figure 7 Solution & Technical Architecture

5.3 User Stories

User Type	Requireme nt (Epic)	User Story Numb er	User Story / Task	-		Release
Custo mer (Mobi le user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account /dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation emailonce I haveregistered for the application	I can receive confirmationemail &click confirm	High	Sprint-1
		USN-3	As a user, I can register for the application through Facebook	I can register &access the dashboard with Facebook Login	Low	Sprint-2
		USN-4	As a user,I can register for the application through Gmail	I can loginusing my emailand password	Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password	I can log into the application by entering myemail and password	High	Sprint-1
	Dashboard	USN-6	As a user I can register into the application	I can use my personaldetails to register	Medium	Sprint-1

		USN-7	As a user I can apply details	I can usemy 12th cut-off	High	Spint-1
		USN-8	As a userI can view/apply universities	I can checkeligibili tycollege and apply	High	Sprint-1
		USN-9	As a user I can checkapplied	I can check availability course in eligible college and apply	High	Sprint-1
		USN-10	As a user I can check my conformation	I can clickconfirm/decline	High	Sprint-1
Customer (Webuse r)	Registration	USN-1	As a user, I can register for the application by entering my email,password, and confirming my password.	I can access my account /dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation emailonce I haveregistered for the application	I can receive confirmationemail &click confirm	High	Sprint-1
		USN-3	As a user, I can register for the application through Facebook	I can register & access the dashboard with Facebook Login	Low	Sprint-2
		USN-4	As a user,I can register for the application through Gmail	I can loginusing my emailand password	Medium	Sprint-1

	Login	USN-5	As a user, I can log into the application by entering email & password	I can log into the application by entering myemail and password	High	Sprint-1
	Dashboard	USN-6	As a user I can register into the application	I can use my personaldetails to register	Medium	Sprint-1
		USN-7	As a user I can apply details	I can usemy 12th cut-off	High	Spint-1
		USN-8	As a user I can view/apply universities	I can checkeligibili tycollege and apply	High	Sprint-1
		USN-9	As a user I can checkapplied	I can check availability course in eligible collegeand apply	High	Sprint-1
		USN-10	As a user I can check my conformation	I can clickconfirm/decline	High	Sprint-1
Customer CareExec utive	Mail	USN-1	As a customer care executive ,I can access customer's information and to solve their queriesand issues	I can access customer's information and to solve theirqueries and issues	Medium	Sprint-1
Administra tor	Dashboard	USN-1	As a admin I can checkthe login credential ofuser	I can checkthe login credential of user and allow the user into the app	High	Sprint-1

	USN-2	As a adminI can check the registration details of user	I can check theregistration details of user	Medium	Spint-1
	USN-3	As a adminI can add courses to user	I can add courses to user	Medium	Sprint-1
	USN-4	As a admin I can check total number of userapplied to particular college and courses	I can check totalnumber of user applied to particular college and courses	High	Sprint-1
	USN-5	As a adminI can accept and reject the eligibleand not eligible user respectively	I can publish the result's ofuser	High	Sprint-1

6.PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

Sprint	Functional	User	User Story / Task	Story	Priority	Team
	Requirement (Epic)	Story Number		Points		Members
Sprint-1	Registration	USN-1	As a user, I will beable to register myapplication by entering my email,password, and confirming my password.	2	High	Dhanuja V

Sprint-1		USN-2	As a user, I will be able to receive an email confirmation afterregistration.	1	High	Gomathi G
Sprint-2		USN-3	As a user, I can register for the application through Gmail.	2	Low	Keerthana R
Sprint-1		USN-4	As a user, I can register for the application by entering details byself.	2	Medium	Manisha M
Sprint-1	Login	USN-5	As a user, I can loginto the application by entering email & password	1	High	Gomathi G
	Dashboard					

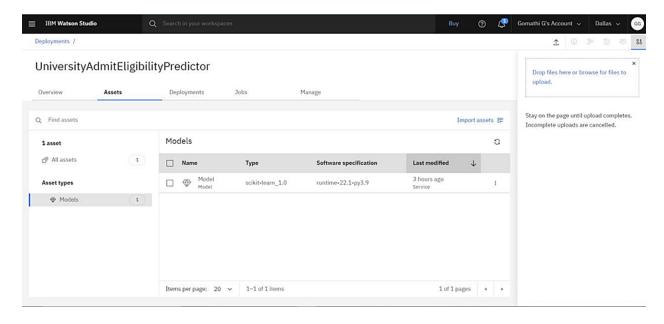
6.2 Sprint Delivery Schedule

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on	Sprint Release Date
			Dute	(L'amieu)	Planned End 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	20	04 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	20	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	19 Nov 2022

7.CODING & SOLUTIONING (Explain the features added in the project along with code)

7.1 Feature 1

The model has been deployed using IBM Watson Machine learning service.



7.2 Database Schema (if Applicable)

```
In [9]:
        data.info()
       RangeIndex: 400 entries, 0 to 399
       Data columns (total 8 columns):
        # Column
                      Non-Null Count Dtype
        0 GRE Score
                            400 non-null
                                            int64
                            400 non-null
        1 TOEFL Score
                                          int64
                                           int64
        2 University Rating 400 non-null
           SOP
                             400 non-null
                                           float64
        4
           LOR
                             400 non-null
                                           float64
           CGPA
                             400 non-null
                                            float64
           Research
                             400 non-null
                                            int64
           Chance of Admit
                             400 non-null
                                            float64
       dtypes: float64(4), int64(4)
       memory usage: 25.1 KB
```

```
In [24]: df['result']=pd.cut(df.Chance_of_Admit_,bins=[0,0.80,1],labels=['No','Yes'])
Out[24]:
             GRE_Score TOEFL_Score University_Rating SOP LOR_ CGPA Research Chance_of_Admit_ result
                  337
                             118
           0
                                              4 4.5
                                                      4.5 9.65
                                                                                   0.92
                                                                                         Yes
                  324
                             107
                                              4 4.0 4.5 8.87
                                                                                  0.76
                                                                                         No
           2
                  316
                                              3 3.0
                              104
                                                           8.00
                                                                                   0.72
                                                      3.5
                                                                                         No
                  322
                                              3 3.5 2.5 8.67
                              110
                                                                                   0.80
                                                                                         No
                              103
                  314
                                              2 2.0
                                                     3.0 8.21
                                                                                   0.65
                                                                                         No
         395
                   324
                              110
                                              3 3.5 3.5 9.04
                                                                                   0.82
         396
                   325
                              107
                                              3 3.0 3.5 9.11
                                                                                   0.84
         397
                   330
                                                      4.5
                   312
         398
        400 rows × 9 columns
```

8. TESTING

8.1 Test Cases

Test Id	Feature Type	Compone nt	Test Scenario	Prerequisit e	Steps To Execute	Test Data	Expected Result	Actual Result	Statu s	Comment s	TC for Automation(Y/ N)	Bu g ID	Execute d By
IndexP a ge_TC_ 0 01	UI	Index Page	Verify if user could able to give input in the input field	HTML_CSS	1. Enter URL and click go 2.Check all the elements in the UI		UI component s and the input fields have to be function properly	Working as Expecte d	Pass	Working Good	N		Manisha M
IndexP. a. ge_TC_ 0 02	Function al	Index Page	Verify if user could be able to authenticate d into the home page	HTML,CSS	1.Check with all the elements in the UI that it is displayed or not 2. Need to click on the predictor tab in the navigation bar.		UI component s has to work Properly and user could be able to navigate to the predictor tab properly	Working as Expecte d	Pass	Working Good	N		Dhanuja V
Predict Page_T C_003	UI	Predict Page	Verify if the UI elements in the Prediction page are function properly or not.	HTML,CSS, FLASK	1.Check with all the elements in the UI that it is displayed or not		UI component s has to work Properly	Working as Expecte d	Pass	Working Good	N		Manisha M
Predict P age_TC _ 004	Function al	Predict Page	Enter the values in the input fields and click on	FLASK	1. Need to enter values for each attribute	187 96 4 4	Navigate to the prediction page and do predict	Working as Expecte d	Pass	Working Good	N		Dhanuja V

			predict		s and click on predict.	9 No Research	the accurate result					
Output P age_TC _ 005	Function al	Chance Page	Verify if it is navigates to chance page only if the appropriate values are entered.	FLASK	Need to enter values for each attribute s and click on predict. If prediction equals one, chance page is displayed.	Prediction n: you have a chance	Redirect to Chance Page	Working as Expecte d	Pass	Working Good	N	Keerthan a R
Output P age_TC _ 006	Function al	No Chance Page	Verify if it is navigates to no chance page only if the appropriate values are entered.	FLASK	Need to enter values for each attribute s and click on predict. If prediction n equals zero, chance page is displayed.	Predictio n: you don't have a chance	Redirect to No chance page	Working as Expecte d	Pass	Working Good	N	Gomathi G

8.2 User Acceptance Testing

8.2.1 Purpose of the document

The purpose of this document is to briefly explain the test coverage and open issues of the University Admit Eligibility Predictor project at the time of the release to User Acceptance Testing

8.2.2 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	12	3	4	3	20
Duplicate	1	0	3	0	4
External	3	3	0	1	7
Fixed	10	2	4	19	36
Not Reproduced	0	0	1	0	1
Skipped	0	0	1	1	2
Won't Fix	0	5	2	1	8
Totals	26	13	15	25	78

8.3 Test Case Analysis

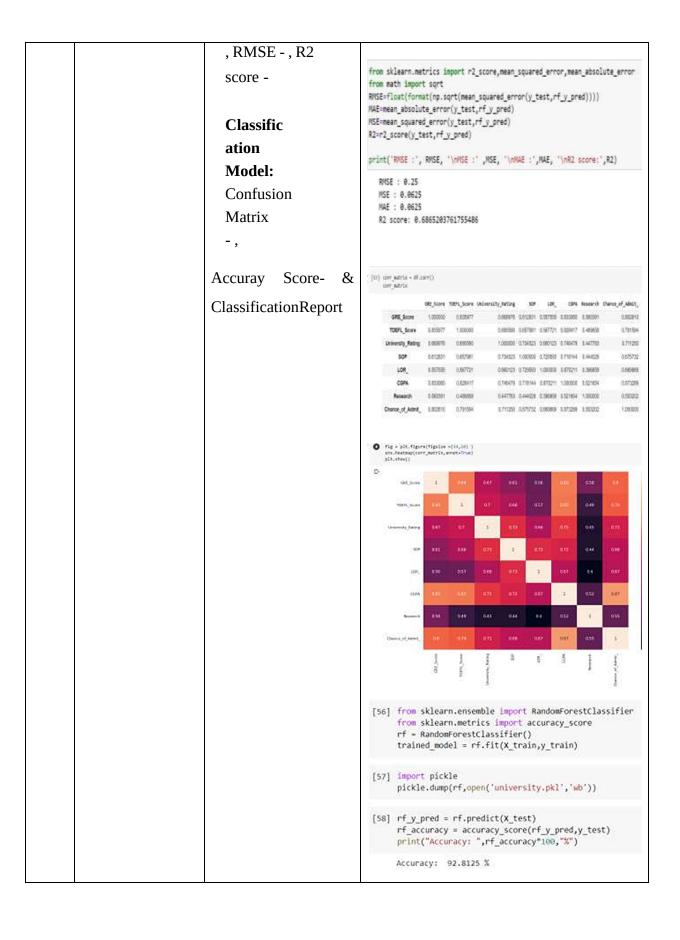
This report shows the number of test cases that have passed, failed, and untested

Section	Total Cases	Not Tested	Fail	Pass
Index page	10	0	0	10
Predict page	4	0	0	4

9. RESULTS

9.1 Performance Metrics

S.N O	PARAMETERS	VALUES	SCREENSHOTS
1	Metrics	Regression Model: MAE - , MSE -	



2	Tune the Model	Hyperparamet er	<pre>from sklearn.linear_model import LogisticRegression from sklearn.model_selection import GridSearchCV</pre>
		Tuning - Validation	<pre>c_space = np.logspace(-5, 8, 15) param_grid = {'C': c_space}</pre>
		Method	<pre>logreg = LogisticRegression()</pre>
			<pre>logreg_cv = GridSearchCV(logreg, param_grid, cv = 5)</pre>
			<pre>logreg_cv.fit(X_train, y_train)</pre>
			<pre>print("Tuned Logistic Regression Parameters: ()".format(logreg print("Best score is {}".format(logreg_cv.best_score_))</pre>
			Tuned Logistic Regression Parameters: {'C': 0.4393970560760 Best score is 0.9

10. ADVANTAGES & DISADVANTAGES

ADVANTAGES:

- It gives an overall accuracy of 94%, which is really high.
- The dataset consists of all possible attributes needed for prediction.
- Confidence booster if results are positive.
- Students can also change their scores to see how they affect the overall prediction results and focus more on that area.

DISADVANTAGES:

- The model is built in such a way that the prediction is positive only if the chance of admit percent is greater than 80%. Even if the result is 79%, the prediction would be NO.
- The complexity of the examinations is not considered. Therefore the results may vary every year for the same set of attribute.
- The dataset used for training the model is of comparatively small size. Therefore, the model cannot be relied on to take accurate real-time decisions.

11. CONCLUSION

The dataset is trained with different ML model. The ML models used to train our dataset are KNN, Logistic Regression, Random Forest, SVM each having accuracy of 86%, 88%, 93%, 89% respectively.Random forest algorithm is finally selected to be used in our model. The Machine Learning model is integrated using flask for our web application. At long last, understudies can have an open-source AI model which will assist the understudies with knowing their opportunity of entrance into a specific college with high exactness

12. FUTURE SCOPE

A real-time project can be developed by gathering data from institutions. The data can be processed and trained using big data frameworks like spark and MLLib can be used to train the data using different machine learning models.

13. APPENDIX

Source

App.py

Model.py

Index.html

Chance.html

```
| File | Sirk | Wew | Window | Sirk | Shouthfund | Mindowshim | Mindow
```

Nochance.html

GitHub link: https://github.com/IBM-EPBL/IBM-Project-34836-1660277900

Project Demo Link: https://youtu.be/SgjHKKobVKc