

Prediction Probability of Getting an Admission into University Using ML

¹Prof.Namrata Ganesh Daware, ²Ms. Neha Jaiswar, ³Ms. Pallavi Nikumbhe, ⁴Ms. Meghana Adhikari, ⁵Mr. Akram Ansari

¹Asst.Prof., ^{2,3,4,5}UG Student, ^{1,2,3,4,5}Computer Engg. Dept. Shivajirao S. Jondhle College of Engineering & Technology, Asangaon, Maharashtra, India.

¹namratashinde9881@gmail.com, ²nehajaiswal1612@gmail.com, ³pallavinikumbhe7063@gmail.com, ⁴meghadadhikari7@gmail.com, ⁵ansariakram218bwd@gmail.com

Abstract-In present conditions, student having difficulties in finding a satisfied institution to pursue higher studies based on their profile. There are some advisory administrations and online applications that recommend universities but they charge consultancy fees in huge amount and online apps are not accurate. So, the main aim of this research is to develop an accurate model that can predict the percentage of chances of getting an admission into the university accurately. This model also provides the analysis of scores versus chances of prediction based on historical data so that students can understand properly whether their profile is suitable for the particular institution or not. Linear Regression and random forest algorithms are uses in this model but cat boost algorithm give the highest accuracy. [1]

Keywords- Machine Learning, random forest, Logistic Regression, cat boost.

I. INTRODUCTION

The quality of one's education determines their destiny, hence it plays a significant part in one's life. After graduating, individuals frequently have a number of questions in their minds about going back to school and selecting the finest university. Most students favor institutions of higher learning that are well-known worldwide. Therefore, a greater proportion of Indian students prefer to pursue higher education in the United States. Even though India is home to some well-regarded universities, graduate students sometimes struggle to gain admission to these institutions. [8] Students spend time and money seeking advice because they are unsure of which university is the best. In addition to consulting firms and counselors, there are some blogs and websites that inform students about their possibilities of admission and offer encouragement, but these sources are not particularly reliable, so you shouldn't rely entirely on them. [1] This paper establishes a machine learning model, which takes into account boundaries such as GRE Score, TOEFL Score, the University Ranking, the Proposal Statement and the Recommendation Letter Power, the Undergraduate GPA and the Study Experience. After getting all the inputs, it predicts the chance of admission. On obscure test occasions, the prepared model has substantial factual findings for the (like) estimate of the probability of confirmation and, accordingly, offers an unprejudiced impression of measurement. [2]

This model also provides the analysis of scores versus chances of prediction based on historical data so that students can understand properly whether their profile is suitable for the particular institution or not.

II. AIMS AND OBJECTIVES

a) Aim

The goal of this research is to construct a model. This model also offers an analysis scores versus probability prediction based on historical so that students can decide whether or not their profile is appropriate.

b) Objective

The major goal of this initiative is to assist students in reducing the amount of time and money they must spend at educational consulting companies. Additionally, it will assist students in reducing the number of applications they submit by suggesting universities where they stand the best chance of being admitted, saving them more money on application fees.

III. LITERATURE SURVEY

1) College Admission Predictor AUTHORS: A. Satya Raghava 3, Nagineni Dharani 2, Annam Mallikharjuna Roa

Students can enter their grades and personal information into the College Admission Predictor System, a web-based application system. This aids in predicting their college admissions. The allocation is within the control of the

administrator. The system records the information once the administrator adds the designated seats to a file. The admission allocation process moves more quickly, made simpler and takes less time overall. In which students can register with their personal and academic information to forecast college acceptance, and administrators can assign seats to students. Administrators can add batch and college information. Utilizing this Application made entrance seat allocation simpler and allowed for system implementation. The entire admission allocation process took less time overall, and the allocation process moved more quickly. It aids students in picking the appropriate college. [5]

2) An automated admissions system based on quality for the educational sector

AUTHORS: Brojo Kishore Mishra, Santosh Satapathy, Soumya Sahoo, Sushruta Mishra

Many educational institutions have started to gather speed in the last 20 years, and many of them are already self-supporting. In order to improve the caliber of student admissions in any higher educational institution, this study suggests using machine learning techniques in the educational arena. This research focuses on identifying those admissions inquiries that are most likely to result in admissions. The analysis's findings will help academic planners concentrate their attention on the group of applicants who are most likely to enroll in the university after initial inquiry. [7]

3) Use of Artificial Intelligence in Education and Student Performance Evaluation

AUTHORS: N. M. Saravana Kumar, D.O.

AI allows for the simulation of human intelligence processes. Computer systems enable the learning, reasoning, and self-correction capabilities. In addition to Artificial Intelligence, other technologies are skillfully coupled to produce outstanding applications. To develop a customized teaching-learning environment, we integrate the evolving role of Artificial Intelligence and its approaches into new educational paradigms. For a thorough learning and assessment process, features like recognition, pattern matching, decision making, reasoning, problem solving and so on are used besides knowledge-based systems and supervised machine learning for a complete learning and appraisal process. [4]

IV.EXISTING SYSTEM

Following graduation, people frequently have a number of questions about continuing their education and selecting the ideal university. Most students favor institutions of higher learning that are well-known worldwide. Therefore, a greater proportion of Indian students prefer to pursue higher education in the United States. Even though India is home to some well-regarded universities, graduate students sometimes struggle to gain admission to these institutions. Additionally, because there aren't many job prospects available, finding a position is also a challenge.

V.COMPARATIVE STUDY

Sr. No	Author	Project title	Publication	Technology	Purpose
1.	Annam Mallikharjuna rao,Nagineeni Dharani	College Admission Predictor	IEEE 2018	JNCET	This helps to predict their admission in college
2.	Susshruta Mishra,Soumya Sahoo	A quality based automate admission system for education domain	IEEE 2016	SCOPES	To enhance the quality of student admissions in higher educational institute
3.	Dr.N.M. Saravana kumar	Implementation of AI in imparting education and evaluating student performance	IEEE 2019	AI	Simulation of human intelligence process in made possible with the help of AI

VI. PROBLEM STATEMENT

Educational institutions have always been crucial to the growth and development of every individual in society. Designing a university prediction/prediction system and offering a probabilistic perspective on college administration for overall ranking, cut-offs of the schools, admission intake, and student preferences are the problems that are thus being addressed. Additionally, it saves students from having to spend money and time on tedious university research and time with counselors. Educational institutions have always been crucial to the growth and

development of every individual in society. Finding the ideal college and course for continuing their education has always been a difficult task for students. Sometimes students are certain of the stream they want to enter, but finding universities that match their academic standing and other achievements is difficult for them. The main goal is to create and offer a location that would produce a probabilistic result regarding how likely this paper create and offer a system that would calculate a person's probability of admission to a university based on their information.

VII. PROPOSED SYSTEM

The dataset is constantly divided into a train and test set of 80% and 20% for the purpose of developing the model. Test set has 100 profiles, while the train set has 400. This is the structure of the modeling dataset. Preprocessing is an essential stage of the process. The objective is to purge the data and get it ready for a prediction algorithm. The data acquired from Occidental College needs just minor adjustments to be suitable for the suggested machine learning techniques. [2]

	GRE Score	TOEFL Score	University Rating	SOP	LOR	CGPA	Research	Chance of Admit
0	337	118	4	4.5	4.5	9.65	1	0.92
1	324	107	4	4.0	4.5	8.87	1	0.76
2	316	104	3	3.0	3.5	8.00	1	0.72
3	322	110	3	3.5	2.5	8.67	1	0.80
4	314	103	2	2.0	3.0	8.21	0	0.65

Fig 1: Training Dataset

One of the most frequent issues in data cleansing is figuring out how to handle missing data. It's crucial to detect missing entries, locate them, and apply a treatment based on the variable form that enables us to use the data in the model because the function in question may be a good predictor of the algorithm's outcome. A training set and a testing set were created from the pre-processed data and randomly divided into the two classes. This paper chose 7976 entries from the total as our training collection, or 80%. [8]

VIII. ALGORITHM

Step 1: Start

Step 2: Collect the dataset in step two.

```
path = settings.MEDIA_ROOT + "/" + "Admission_Predict.csv"
df = pd.read_csv(path)
df = df.rename(columns={'Serial No.': 'no', 'GRE Score': 'gre', 'TOEFL Score': 'toefl', 'University Rating': 'rating', 'SOP': 'sop', 'LOR ': 'lor', 'CGPA': 'gpa', 'Research': 'research', 'Chance of Admit ': 'chance'})
df.drop(['no'], axis=1, inplace=True)
```

Step 3: Use a machine learning classifier to train the model.

```
calc_logistic_regression():
    print("*" * 25, "Logistic Regressions")
    # Fitting logistic regression model
    from sklearn.linear_model import LogisticRegression
    lr = LogisticRegression()
    lr.fit(x_train, cy_train)
    #print(f'{x_test}====={cy_test}')
```

Step 4: Run the model and test it.

```
# Predict probabilities on test set
y_prob = lr.predict_proba(X_test)[:,-1]
# Set admission threshold
threshold = 0.5
# Convert probabilities to binary predictions
y_pred = [1 if prob >= threshold else 0 for prob in y_prob]
```

Step 5: Display Stored Results

```
# Evaluate model performance
accuracy = accuracy_score(y_test, y_pred)
print('Accuracy:', accuracy)
```

Step 6: Finish

IX. MATHEMATICAL MODEL

Logistic Regression:

The logistic regression model uses a logistic function to estimate the probability of an event occurring. In this case, the event is getting admission to a university. The logistic function takes in a linear combination of the input features and maps it to a probability between 0 and 1.

Random Forest:

The random forest algorithm creates multiple decision trees and combines their predictions to make a final prediction. Each decision tree is built on a random subset of the input features and a random subset of the training data.

Cat Boost:

Cat Boost is a gradient boosting algorithm that uses categorical features. It works by iteratively adding decision trees to the model and adjusting the weights of misclassified samples.

X.SYSTEM ARCHITECTURE

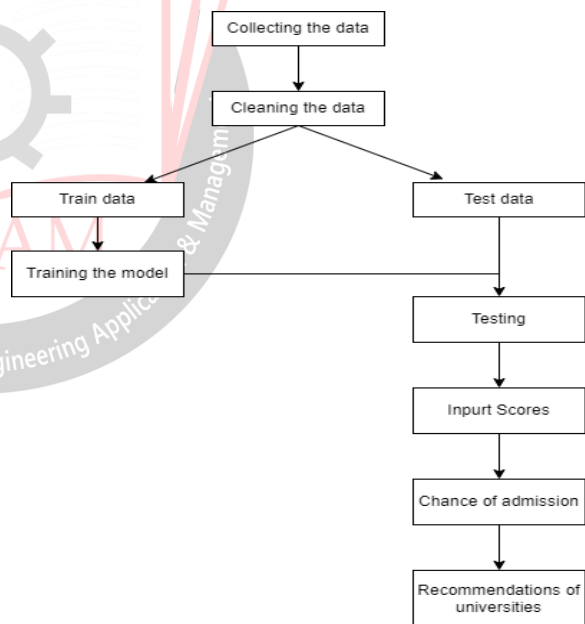


Fig 2: System Architecture

Dataset also be viewed as the collection of the data objects, which are also often called as records, samples, or entities. The Data objects are mainly described by the numbers at the different features that mostly captures. The most basic characteristics object. Mass of a physical object or the time at which an event occurred, etc. Features can be variables, characteristics, fields, attributes, or dimensions. The data preprocessing in this forecast uses techniques like removal of noise in the data, the expulsion of missing information,

modifying default values if relevant and grouping of attributes for prediction at various levels.

XI.DESIGN DETAILS

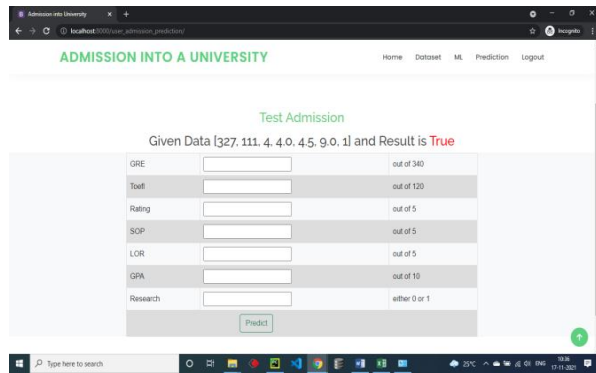


Fig 3: Testing Model

CNN Model	Train set		
	Accuracy on Test 1 (%)	Accuracy on Test 2 (%)	Accuracy on Test 3 (%)
Logistics Regression	96.00%	97.00%	95.00%
Random Forest	97.00%	96.00%	98.00%
Catboost	97.00%	95.00%	96.00%

Fig 4: Machine learning algorithm test result

XII.CONCLUSION

Thus we have tried to implement the paper “prediction probability of getting an admission into university using machine learning” A. Sivasanagri, V Shivani, Y. Bindh D. Deepa, R. Vignesh, IEEE 2021, and according to the implementation the conclusion is to predict the probability of percentage of getting an admission seat into university. Logistic Regression, Random Forest, Cat boost these algorithms are used in this research. The system has performed prediction of the students’ scores and data & compares it with historical data of the institution. So that students can come to know whether they are eligible or not in a particular institute.

REFERENCES

[1]A.Sivasanagri,V.Shivani,Y.BindhuD.Deepa, R.Vignesh “prediction probability of getting an admission into university using machine learning” IEEE 2021.

[2] B. Uday Kiran, B. Simon Paul, T. Pavan Kumar,Mr.M. Sathya Narayana,”Admission Prediction for MS in Foreign Universities using Machine Learning”2022

[3] Acharya MS Arman A, Antony AS (2019) build a model named “A comparison of Regression models for prediction of graduate admissions”. For the students accurately.

[4] Kumar, NM Saravana. "Implementation of artificial intelligence in imparting education and evaluating

student performance." Journal of Artificial Intelligence 1, no. 01(2019): 1-9.

[5] Rao, Annam M, et al. Modeling a “ College Admission Predictor.” Journal of Network Communications and Emerging Technologies (JNCET), Volume 8, Issue 4, April (2018).

[6]Prof. Vishal R.Shinde “College Information Retrieval” in IJREAM, ISSN: 2494-9150 Special Issue-03, April 2017.

[7] Mishra. S and Sahoo. S. (2016). A Quality Based Automated Admission Process System for Educational Domain, pp. 221–223, International conference on Signal Processing, Communication, Power and Embedded System (SCOPE5)-2016.

[8] Austin Waters and Risto Miikkulainen, proceedings of the Twenty-Fifth Innovative Applications of Artificial Intelligence Conference,2013.