International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/CSEIT228534

Prediction of Graduate Admission Using Machine Learning

Saurin Patel*, Harsh Waghela, Pratham Gupta, Nirbhay Rajgor
Information Technology, Mumbai University, Mumbai, Maharashtra, India

Article Info

Publication Issue :
Volume 8, Issue 5
September-October-2022

Page Number : 184-189

Article History
Accepted: 01 Oct 2022
Published: 11 Oct 2022

ABSTRACT

The goal of this study is to create a model that may assist students in selecting the best institutions based on their scores and their profiles. We can evaluate candidates across a broad range of disciplines, such as Master of Science (international), Master of Technology (India), and Masters in Business Administration (India and international). We intend to build a machine learning model in order to produce outcomes that may benefit the students in choosing the right University. The dataset includes facts on the university and student profiles, together with a field that indicates whether or not the admission was successful. Key performance indicators have been used to compare the predictions made using a variety of algorithms, including Ensemble Machine Learning (KPI).The dependent variable, or the likelihood of admission to a university, is then evaluated using the model that is performing the best. The chances of admit variable, which has a range of 0 to 1, represents the anticipated likelihood of being accepted to a university. Additionally, we want to build a portal that sorts universities according to their acceptance range before listing them.

Index Terms - Key Performance Indicators, M.Tech, MBA, Machine Learning, Dependent Variable.

I. INTRODUCTION

It would be challenging for someone seeking postgraduate studies to choose which college they may apply to based on their GPA, Quantitative, Verbal, TOEFL, and AWA Scores.Instead of applying to schools where they have a chance of getting in, people might apply to many schools that prefer applicants with higher test scores. Their future would suffer as a result of this. Instead of applying to colleges they may never get into, candidates should focus on those where they have a strong chance of being accepted. There aren't many effective methods for quickly learning which colleges one can get into.

A person can use the Education Based Prediction System to help them choose which colleges to apply to based on their test results. The parameters that make up the processing data set are as follows: Name of the university, GRE Quantitative and Verbal Scores, TOEFL, and AWA Scores. Many universities and graduate programme use the GRE Test (Graduate

Copyright: © the author(s), publisher and licensee Technoscience Academy. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited

Record Examinations) as a standard assessment tool for graduate admissions throughout the world.

While submitting an application to a college, other factors are also taken into account, such as Letters of Recommendation (a letter that is prepared on behalf of a candidate by someone who can attest to their abilities in the classroom or in their line of work.), Statement of Purpose (a vital component of a graduate school

Most of the time, it is quite difficult for someone who has completed their undergraduate degree and wishes to pursue a postgraduate degree in a field of their choice to determine which colleges they should apply to with their GRE and TOEFL scores along with their GPA at the time of graduation. Many applicants might submit applications to colleges that don't fit their score requirements, wasting their time. The price goes up when you submit scores for numerous college applications. An education predictor system has been developed because there are not many effective ways to address this problem.

One can enter their scores in the proposed system by the corresponding fields offered. The system then evaluates the data entered and generates a list of universities a person may apply to based on their test results. This is fairly quick the student will end up saving their money and time. In order to accomplish this, we have suggested a novel technique using algorithms for machine learning. To increase our model's accuracy, we have considered several machine learning algorithms, not just one. These include Neural Networks, Linear Regression, and other algorithms Random Forest and Decision Tree. The Algorithms section of this paper will go into more detail on these algorithms. The algorithm having the best KPI will be considered to develop the application system after these algorithms have been compared. We also anticipate grouping universities based on a profile and categorizing them as less likely, highly likely, etc. to accept students.

II. PROBLEM STATEMENT

Educational institutions have always been crucial to the growth and development of every individual in society. There are numerous college prediction websites and apps available today, but using them can be time-consuming to some extent due to the lack of clear information about colleges and the effort required to find the best college for you.

Designing a college prediction/prediction system and offering a probabilistic insight into college administration for overall rating, cut-offs of the colleges, admission intake, and student preferences is the problem statement currently being addressed. Additionally, it saves students from having to spend money and time on stressful college research and counselling sessions.

Finding the ideal college and course for continuing their education has always been a difficult process for students. Sometimes they are certain of the stream they want to enter, but finding colleges that match their academic standing and other achievements is difficult for them. We want to create and offer a system that would calculate a person's probability of admission to a university based on their information.

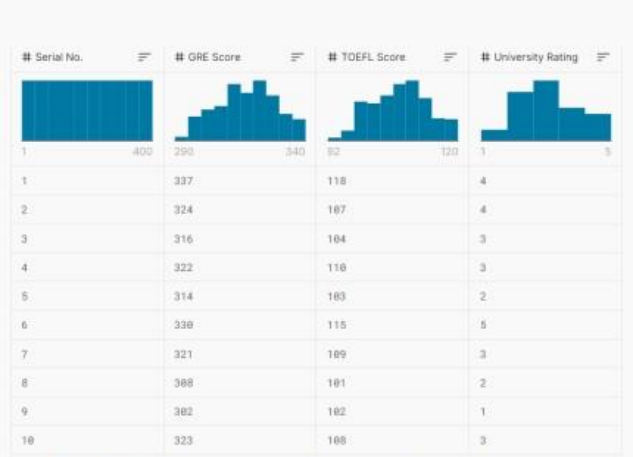
III. Dataset

The data set includes a variety of variables that can influence choosing the best university. Data from 100 different students is included.

The data set is divided into nine different parameters that are significant for masters applications.

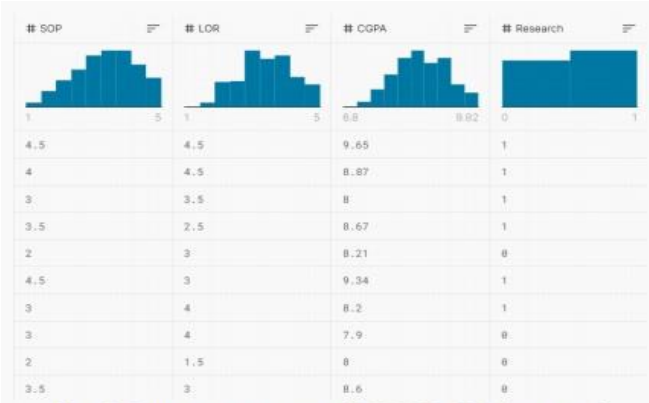
The following criteria must be met: gre, toefl, university ranking, statement of purpose, recommendation letter, undergraduate GPA, research

paper, and chance of admission.



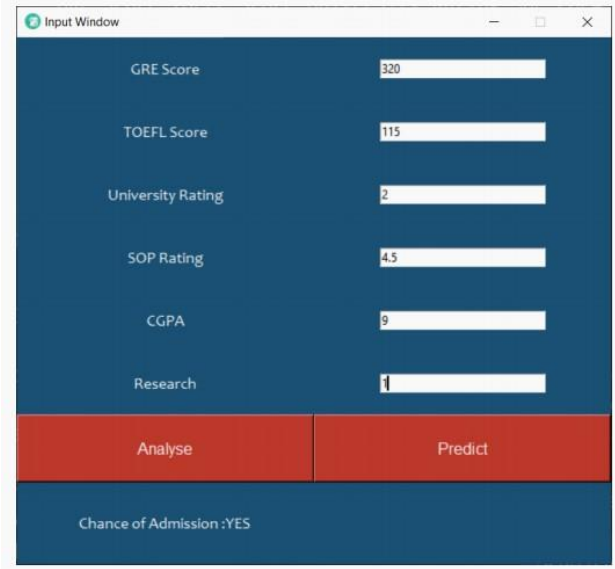
# Serial No.	# GRE Score	# TOEFL Score	# University Rating
1	337	118	4
2	324	107	4
3	316	104	3
4	322	110	3
5	314	103	2
6	330	115	5
7	321	109	3
8	308	101	2
9	302	102	1
10	323	108	3

Fig -1: Data set containing *Gre score, Toelf score and university ranking*



# 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

Fig -2: Data set containing *SOP, LOR, CGPA, research*



IV. Algorithms Used

A. Linear Regression

Regression models use observed data to draw a line that depicts the relationship between various variables. In logistic and non-linear regression models, curved lines are used instead of straight lines, and vice versa. A technique based on supervised learning, the linear regression model responds to just one feature. Every prediction value in a regression model is based on independent variables. It is used to figure out how two quantitative variables relate to one another. Different regression models have different ways of relating independent and dependent variables, as well as different numbers and types of variables that they take into account. It is assumed that these variables have a linear relationship. From this point forward, we attempt to define a linear function that accurately predicts the response value (y) as a function of the feature or independent variable (x). College admission requirements, student preferences, and cutoff dates. Additionally, it saves students from having to spend money and time on stressful college research and counselling sessions.

B. Artificial Neural Network

Neural networks are a collection of algorithms that attempt to identify underlying relationships in a data set by mimicking how the human brain works. A system of neurons is what is referred to as a neural network.

Artificial neural network systems were influenced by biological neural network architecture. Without having been given any specific rules for the task, they learned how to complete it by studying examples and various data sets. The system's ability to generate identification characters from previously processed data without programming or a foundational understanding of these data sets serves as its primary source of inspiration. The various artificial neurons that make up artificial neural networks are referred to as units. Layers of units that together form artificial neural networks are arranged in a particular order. Depending on the complexity of the system, each layer should only have a few millions or dozen units.

The entire Artificial Neural Networks in a system are made up of these units, which are arranged in a

number of layers. Depending on the system complexity, a layer can have a few dozen units or millions of units. Artificial neural networks frequently have hidden layers in addition to input, output, and other layers. The input layer receives information that the neural network needs to analyse or learn from the outside world. Then, after passing through one or more hidden layers, this data is transformed into useful information for the output layer. Last but not least, the output layer delivers an output in the form of an artificial neural network's response to input data.

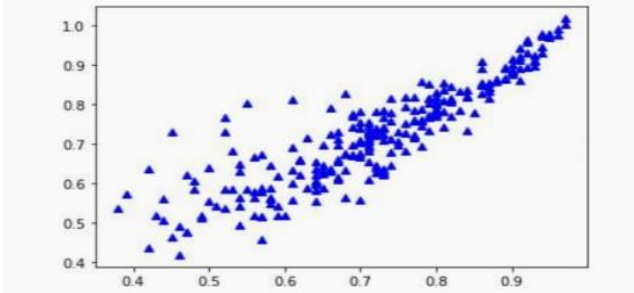


Fig -5 (a): Linear Regression

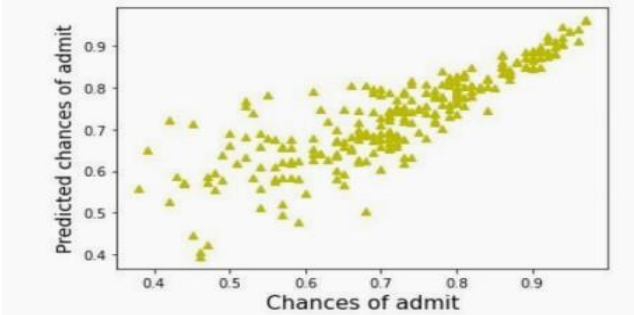


Fig -5 (b): Neural Networks

C. Decision Trees

Learning and prediction are the two steps in the classification process. In the first step of its learning process, the model is created using the provided training data. The model is then used in the prediction step to predict the response for the supplied data. It is among the most well-liked classification algorithms and simple to learn and comprehend. Regression and classification issues are also solved using the decision tree algorithm. For prediction, decision trees use a class label; for a record, it starts at the tree's base. then contrast the root's values with the record attribute. Following the comparison, it moves on to the next node by following the branch that corresponds to the value. Decision trees come in two types: continuous variable and categorical variable. Continuous variables have a continuous target variable, whereas categorical variables have a categorical target variable. Decision nodes, chance nodes, and end nodes are the three different types of nodes in a decision tree. Each leaf node is given a class label by a decision tree. Even the root and internal nodes, which are non-terminal nodes, have attribute test conditions to distinguish records with various characteristics.

D. Random Forests

An algorithm for machine learning called the random forest is frequently employed in classification and regression issues. Decision trees are constructed using a variety of samples and then average out their votes and divide them into two groups. Random forest is capable of handling a data set with continuous variables in the event of regression and categorical variables in case of classification. Consequently, it produces positive results for classification issues. In industry jargon, how the algorithm for the forest works such good is:

A large number of relatively uncorrelated models (trees) operating as a committee will outperform any of the individual constituent models.

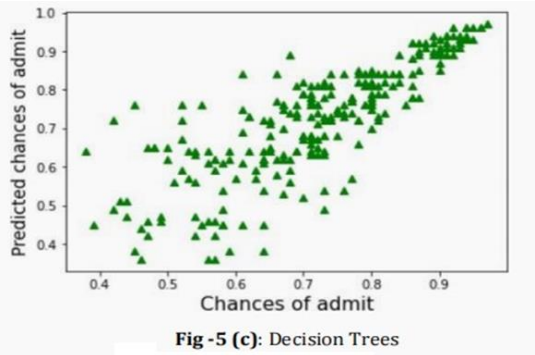


Fig -5 (c): Decision Trees

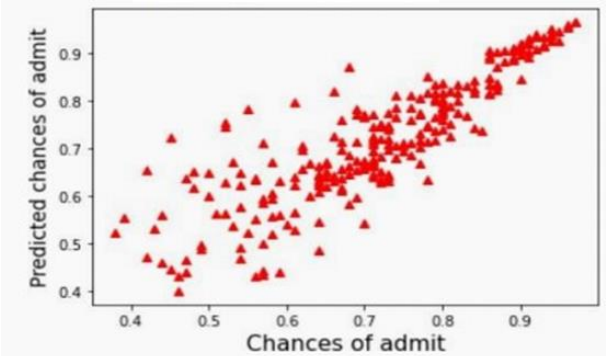
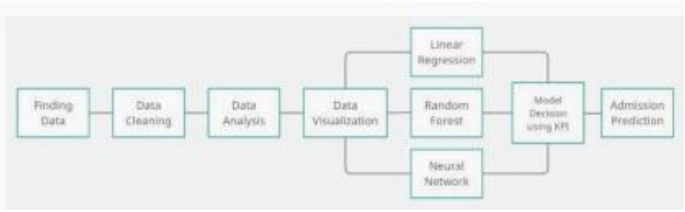


Fig -5 (d): Random Forests

V. Methodology And Implementation

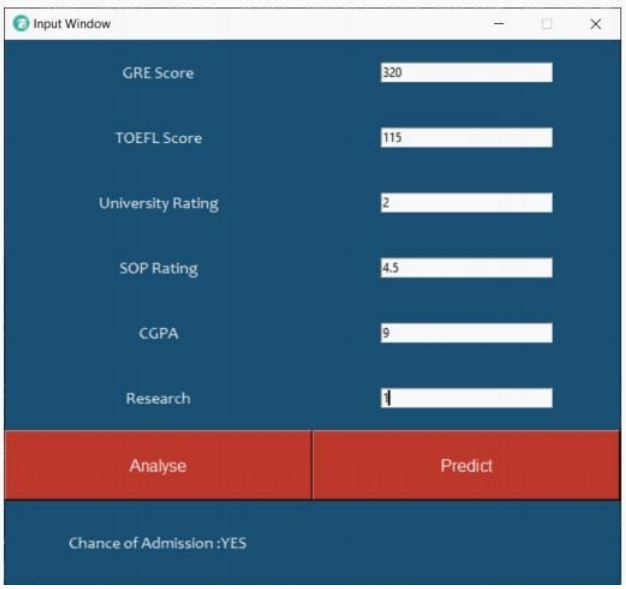
Selecting the appropriate data set is the first basic step in creating a model for our use case. We selected a data set for our predictions that includes all the crucial characteristics that would influence the likelihood of admission. Data cleaning comes next, where we deal with fields that have missing values. We employ a variety of tools and libraries to visualize the data and carry out analysis once it is ready for analysis. The correlation matrix and bar graphs are two examples of this.



We separate the data into training and testing data once it is prepared for processing. The machine learning algorithms namely 1) Linear regression 2) random forest 3) neural network will be used for this. We evaluate these models using key performance indicators after they have been constructed over the data set. These indicators assist us in selecting the best model to determine an applicant's chances of admission.

VI. RESULTS AND CONCLUSION

Millions of students apply to universities each year to start their academic careers. Most of them lack the necessary tools, background knowledge, and caution, which leads to a host of issues like applying to the wrong university or college, which further wastes their time, money, and effort. We have made an effort to assist students who are having trouble selecting the best university for them with the aid of our project. Applying to colleges where one has a good chance of admission is crucial, as opposed to applying to schools where one may never be accepted. Students will only apply to universities where they have a high chance of admission, which will help to reduce costs.



For such individuals, our prepared models function with a satisfactory level of accuracy and could be very

helpful. For individuals in our age bracket who wish to pursue further education at the institution of their choice, this project has a bright future.

VII. REFERENCES

[1]. Borah M.D., Application of knowledge based decision technique to predict student enrolment decision , Recent Trends in Information Systems (Re TIS), 21-23 Dec. 2011,180- 184.

[2]. Ragab, A.H.M. , Hybrid recommender system for predicting college admission, Intelligent Systems Design and Applications (ISDA), 29 Nov. 2012, 107-113. [3] Rensong Dong,The module of prediction of College Entrance Examination aspiration, Fuzzy Systems and Knowledge Discovery (FSKD), 31 May 2012 ,1559-1562.

[3]. E. Gottlieb, "Using integer programming to guide college admissions decisions: a preliminary report", Journal of Computing Sciences in Colleges, Volume 17, Issue 2, Pages: 271 – 279, 2001.

[4]. J.S. Moore, "An expert system approach to graduate school admission decisions and academic performance prediction", ScienceDirect ,Volume 26, Issue 5, October 1998, Pages 659–670

[5]. Reddy, M. Y. S., & Govindarajulu, P. (2018). College Recommender system using student'preferences/voting: A system development with empirical study. INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND NETWORK SECURITY, 18(1), 87-98.

[6]. Deokate monali, Gholave Dhanashri, Jarad Dipali, Khomane Tejaswini (2018). College Recommendation System for Admission. International Research Journal of Engineering and Technology, 9(3), 187-175.

Cite this article as :

Saurin Patel, Harsh Waghela, Pratham Gupta, Nirbhay Rajgor, "Prediction of Graduate Admission Using Machine Leering", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 5, pp. 184-189, September-October 2022. Available at doi : https://doi.org/10.32628/CSEIT228534

Journal URL : https://ijsrcseit.com/CSEIT228534