International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 04 | Apr 2023 www.irjet.net p-ISSN: 2395-0072

Graduate Admission Predictor

Sonali Pandey[1], Pratik Patil[2], Rakesh Kumar[3], Madhuri Gedam[4]

[1,2,3,4] *Department of information Technology, Shree LR Tiwari College of Engineering, Maharashtra, India*

Abstract - *When applying to master's programs, prospective graduate students are sometimes faced with the challenge of selecting their preferred universities. While there are many predictors and consultancies that can help a student, they are not always accurate because decisions are based on a small number of previous admissions. Here, we describe a machine learning-based approach in which, given the student profile, we assess various regression techniques, such as Linear Regression, Decision Trees, and Random Forest. Then, we calculate error functions for each model and evaluate performance to determine which model performs the best. The outcomes then show whether the selected university is a risky or ambitious one.*

Key Words: Graduate Admissions, Predictor, Statistical Model, Probability Estimation, Support Vector Regression.

1. INTRODUCTION

The complicated procedure of applying to a graduate school in the USA results in a very hectic undergraduate schedule for Indian students. The parallel approach of keeping a high Cumulative Grade Point Average, with good GRE, TOEFL scores and publishing research papers, getting good Letters of Recommendation and making a good Statement of Purpose certainly makes every Indian student busy who wants to further excel in academia. We understand that all the things cumulatively are very important for graduate school admissions. But what are the most important factors?

Several Machine learning algorithms will be utilized in our predictor to predict the rate of acceptance as a percentage.

Machine Learning algorithms like Logistic Regression, Linear Regression, Decision Trees, and Random Forest. Regression models will be compared according to their coefficient of determination denoted by R2 whilst classification models will be compared according to their accuracy, precision, and recall.

This paper aims to provide a comprehensive understanding of the Graduate Admission Predictor

The objective is to offer insights into the working principles of the system, its advantages, its applications, and its limitations, providing a foundation for future research and development.

2. PROPOSED SYSTEM

A Graduate Admission Predictor system could potentially be developed using machine learning algorithms, based on data about applicants and their outcomes (i.e., whether they were admitted or rejected). This historical data could include various factors that are typically considered in graduate admissions, such as undergraduate GPA, standardized test scores (e.g., GRE), letters of recommendation and other application materials.

To develop such a system, the first step would be to gather and preprocess the historical data, which might involve cleaning the data, dealing with missing values. Then, a machine learning model would need to be trained on this data to predict the likelihood of admission for new applicants.

I. GRE Score (General Record Examinations); this score measures general knowledge in undergrad Math and English. This score ranges from a value of 260 to 340

II. TOEFL Score (Test of English as a Foreign Language); this score measures student's English abilities. This score has a value between 0 and 120.

III. SOP (Statement of Purpose); a letter written by the applicant explaining their purpose of the application. This is given a score between one and five.

IV. LOR (Letter of Recommendation); tests the weight of the recommendation provided by the applicant. This is given a score between one and five.

V. CGPA (Cumulative GPA); based on the academic performance of the applicant in undergraduate studies. This is scored on a range from one to ten.

VI. University Rating; based on the reputation of the applicant's previous university. This is given a score between one and five.

VII. Research Experience; binary value based on whether the applicant has any research familiarity. This value is either one or zero.

VIII. Chance of Admission; the rate of admission into graduate school. This attribute is the targeted value in which will be predicted as the rate from zero to one.

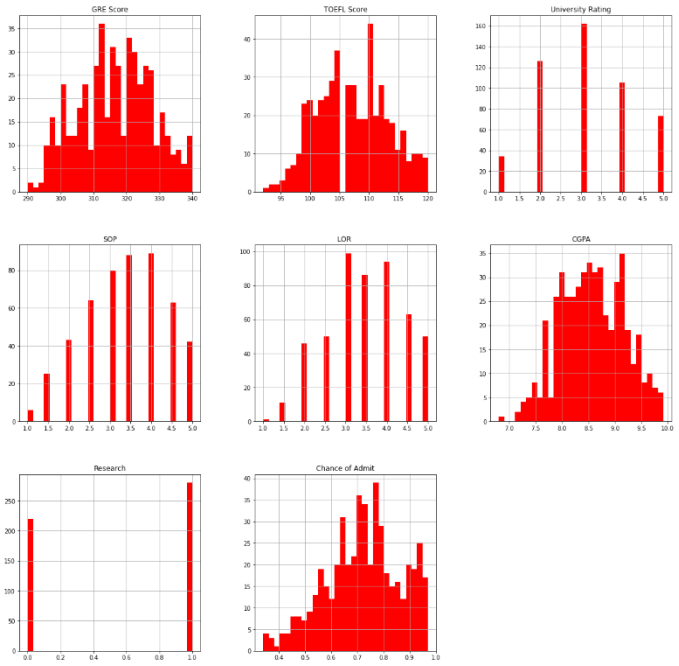


Fig -1: Data Visualization

Fig -2: Heat Map

3. ALGORITHM

Below is a list of algorithms analyzed for the Graduate Admission Process:

3.1 Regression Algorithms

A. Linear Regression

1. In simple linear regression goal is to obtain a relationship model between x and y.

2. We predict the value of one variable y based on the variable x.

3. X is known as independent variable and y is called as dependent variable.

4. It is simple because it examines the relationship between two variables only.

5. It is called as linear regression because when the independent variable increases (or decreases), the dependent variable also increases (or decreases) in a linear fashion.

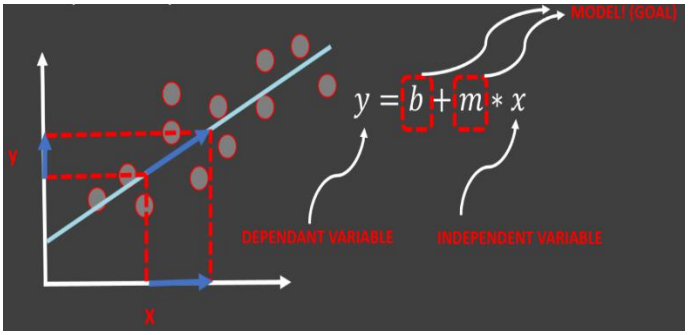


Fig -3: Linear Regression

B. Multiple Linear Regression

1. Multiple Linear Regression examines relationship between more than two variables.

2. It is different from simple linear regression which is a statistical model that examines the linear relationship between two variables only.

3. Each independent variable has its own corresponding coefficient.

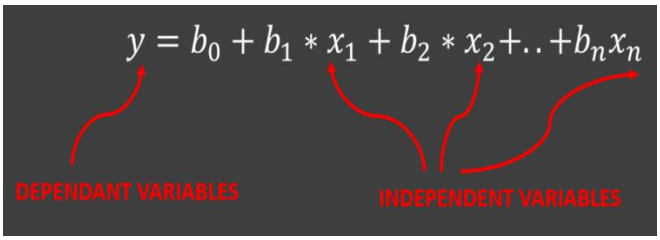


Fig -4: Multiple Linear Regression

3.2 Classification Algorithms

A. Artificial Neural Network

1. Multi-layer perceptron is a class of feedforward artificial neural networks.

2. It usually consists of input layer, hidden layer and a output layer.

3. It uses a supervised learning technique called back-propagation for training.

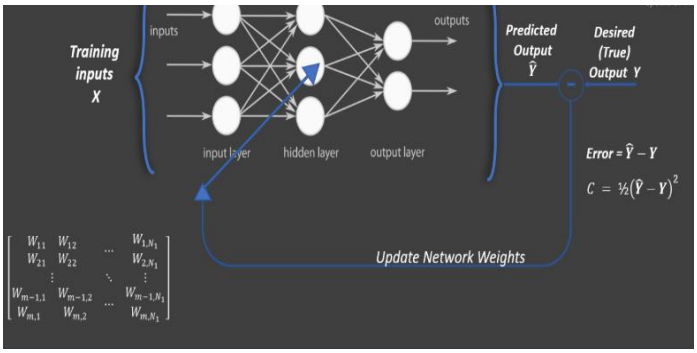


Fig -5: Artificial Neural Network

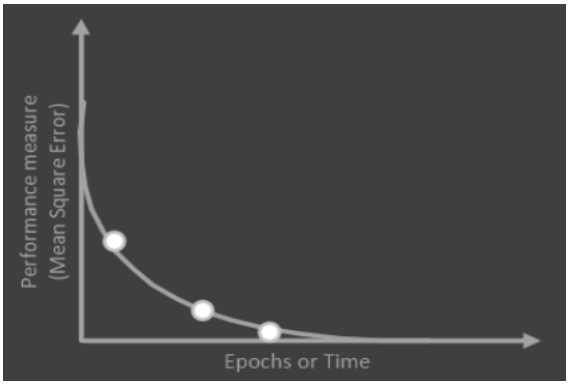


Fig -6: Mean Square Error

3.3 Regression KPIs

3.3.1 Regression Metrics: To assess model performance

After model fitting, we would like to assess the performance of the model by comparing model predictions to actual (True) data.

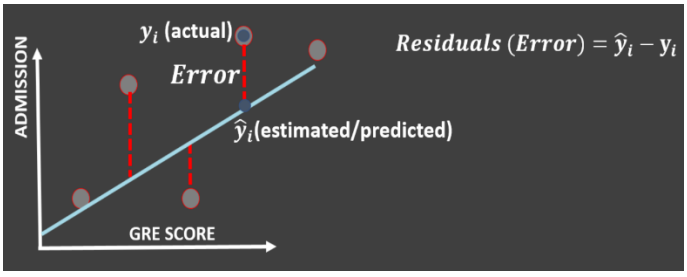


Fig -7: Regression Metrics

A. R SQUARE (R^2)-Coefficient of Determination

The percentage of the variation (of y) that has been explained by the independent variables in the model is shown by the R-square or coefficient of determination.

If R^2=80, this means that 80% increase in the university admission is due to GRE scores (assuming a simple linear regression model).

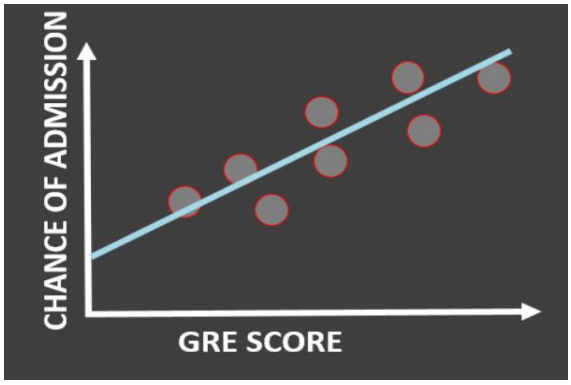


Fig -8: R SQUARE (R^2)-Coefficient of Determination

B. ADJUSTED R SQUARE (R^2)

If R^2=80, this means that 80% increase in the university admission is due to GRE scores.

Let's add another useless independent variable, let's say the heights of the student to the Z-axis.

Now R^2 increases and becomes: R^2=85%

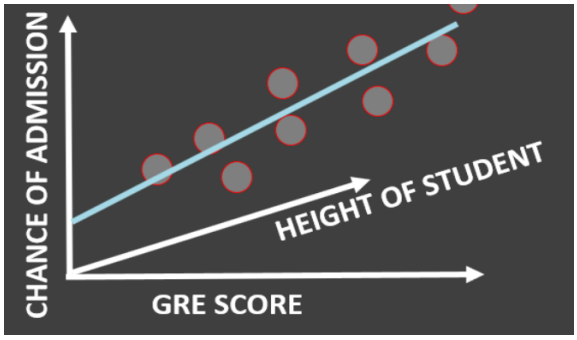


Fig -9: ADJUSTED R SQUARE (R^2)

4. LITERATURE REVIEW

Khan, M. A., Dixit, M., & Dixit, A [1] published their research work on Demystifying and Anticipating Graduate School Admissions using Machine Learning Algorithms. They have offered a cutting-edge strategy for estimating the likelihood of admission to graduate school.

Acharya, M. S., Armaan, A., & Antony, A. S. [2] presented a research paper which gives a comparison of Regression Models for Prediction of Graduate Admissions that compute error functions for the different models and compare their performance to select the best performing model.

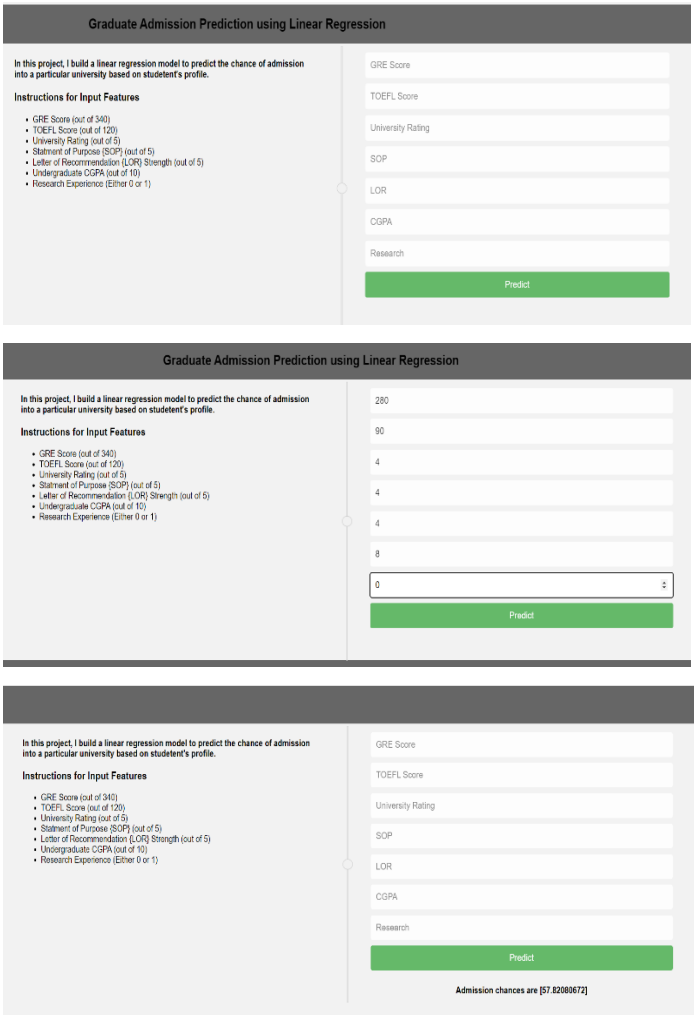
Bitar, Z., & Al-Mousa, A [3] have proposed a comprehensive and insightful overview on " Prediction of Graduate

Admission using Multiple Supervised Machine Learning Models"

Jeganathan, S., Parthasarathy, S., Lakshminarayanan, A. R., Ashok Kumar, P. M., & Khan, M. K. A. [4]. have done a brief study on Predicting the Post Graduate Admissions using Classification Techniques. The authors have applied the classification techniques such as Logistic Regression, KNN Classification, Support Vector Classification, Naive Bayes Classification, Decision Tree Classification and Random Forest Classification on the given academic admission dataset.

apare, N. S., & Beelagi, S. M. [5] have presented a research paper which gives a " Comparison study of Regression Models for the prediction of post-Graduation admissions using Machine Learning Techniques"

5. RESULTS & ANALYSIS



6. CONCLUSIONS

In terms of accuracy, Artificial Neural Network once again outperforms all of our other categorization techniques. With a low MSE and high R2 score, it is evident that Linear Regression outperforms Random Forest on our dataset. Random Forest comes in second place. This is explained by the linear dependencies of the dataset's features. Higher test scores, GPAs, and other metrics typically increase the likelihood of acceptance. The inclusion of a few outliers has impacted the Linear model to some extent.

7. FUTURE SCOPE

We aim to expand our dataset and increase the number of profiles with some variations. The number of outliers (profiles that do not seem impressive but had a high chance of admission) would be significantly increased to reduce the linear dependency of features. We will also use Deep Neural Networks as another plausible model to understand the subjective nature of admission. The predictor presented can be implemented not only in university admission faculties but also at recruiting agencies or human resources departments. By putting this predictor into practice, applicant CV analysis would take less time. In the future, we plan to carry out our study by using different volumes of data & with more attributes, mainly considering neural networks.

8. REFERENCES

[1] Khan, M. A., Dixit, M., & Dixit, A. (2020). Demystifying and Anticipating Graduate School Admissions using Machine Learning Algorithms. 2020 IEEE 9th International Conference on Communication Systems and Network Technologies (CSNT). doi:10.1109/csnt48778.2020.9115788

[2] Acharya, M. S., Armaan, A., & Antony, A. S. (2019). A Comparison of Regression Models for Prediction of Graduate Admissions. 2019 International Conference on Computational Intelligence in Data Science (ICCIDS). doi:10.1109/iccids.2019.8862140

[3] Bitar, Z., & Al-Mousa, A." Prediction of Graduate Admission using Multiple Supervised Machine Learning Models", 2020

[4] Jeganathan, S., Parthasarathy, S., Lakshminarayanan, A. R., Ashok Kumar, P. M., & Khan, M. K. A. (2021). Predicting the Post Graduate Admissions using Classification Techniques. 2021 International Conference on Emerging Smart Computing and Informatics (ESCI). doi:10.1109/esci50559.2021.9396815

[5] apare, N. S., & Beelagi, S. M. " Comparison study of Regression Models for the prediction of post-Graduation admissions using Machine Learning Techniques",2021

[6] Gedam, Madhuri N., and Bandu B. Meshram. "Proposed Secure Content Modeling of Web Software Model." IJETT 6.2 (2019).