# Graduate Admission Prediction Using Machine Learning Techniques

**K. JeevanRatnakar, G. Koteswara Rao, B. DurgaPrasanth Kumar, G.prithvi, D.Venkata SaiEswar**

Associate Professor, Department of CSE, KKR & KSR Institute of Technology and Sciences, A.P, India.
B.Tech Student, Department of CSE, KKR & KSR Institute of Technology and Sciences, A.P, India.
B.Tech Student, Department of CSE, KKR & KSR Institute of Technology and Sciences, A.P, India.
B.Tech Student, Department of CSE, KKR & KSR Institute of Technology and Sciences, A.P, India.
B.Tech Student, Department of CSE, KKR & KSR Institute of Technology and Sciences, A.P, India.

**ABSTRACT:**     In India every year lacks of students getting the graduation degree and willing to join post-graduation in other countries. Newly graduate students usually are not knowledgeable of the requirements and the procedures of the postgraduate admission and might spent a considerable amount of money to get advice from consultancy organizations to help them identify their admission chances. Human consultant and calculations might be bias and inaccurate.

This paper helps on predicting the eligibility of Indian students getting admission in best university based on their Test attributes like GRE,TOEFL,LOR,CGPA etc. according to their scores the possibilities of chance of admit is calculated.

**KEYWORDS:** Multiple Linear Regression, Random forest Regression, Multiple Linear Regression, Dimensionality reduction

## I. INTRODUCTION

The world markets are developing rapidly and continuously looking for the best knowledge and experience among people. Young workers who want to stand out in their jobs are always looking for higher degrees that can help them in improving their skills and knowledge. As a result, the number of students applying for graduate studies has increased in the last decade. This fact has motivated us to study the grades of students and the possibility of admission for master's programs that can help universities in predicting the possibility of accepting master's students submitting each year and provide the needed resources.

The dataset [5] presented in this paper is related to educational domain. Admission is a dataset with 500 rows that contains 7 different independent variables which are:
• Graduate Record Exam1 (GRE) score. The score will be out of 340 points.

• Test of English as a Foreigner Language2 (TOEFL) score, which will be out of 120 points.

• University Rating (Uni.Rating) that indicates the Bachelor University ranking among the other universities. The score will be out of 5

• Statement of purpose (SOP) which is a document written to show the candidate's life, ambitious and the motivations for the chosen degree/ university. The score will be out of 5 points.
• Letter of Recommendation Strength (LOR) which verifies the candidate professional experience, builds credibility, boosts confidence and ensures your competency. The score is out of 5 points
• Undergraduate GPA (CGPA) out of 10
• Research Experience that can support the application, such as publishing research papers in conferences, working as research assistant with university professor (either 0 or 1). One dependent variable can be predicted which is chance of admission that is according to the input given will be ranging from 0 to 1.

We are developing four Regression Models which are multiple Linear Regression, Random forest Regression, Multiple Linear Regression using Dimensionality reduction and Random forest Regression using Dimensionality reduction to finding the accuracy of those models. Out of those we use high accuracy models.

In the Existing System, Many machine algorithms are used to the prediction of Graduate Admission. The existing system compares the four machine learning algorithms on the basis of accuracy. The algorithms are Linear Regression, Support Vector Regression, Random forest Regression, Decision Tree Regression. In this system Linear Regression performs the best on the dataset with low MSE and high R2 score. Figure 2, shows sample data set, the dataset contains 500 rows and 7 independent variables of data. In the figure 1, showing the architecture of existing system
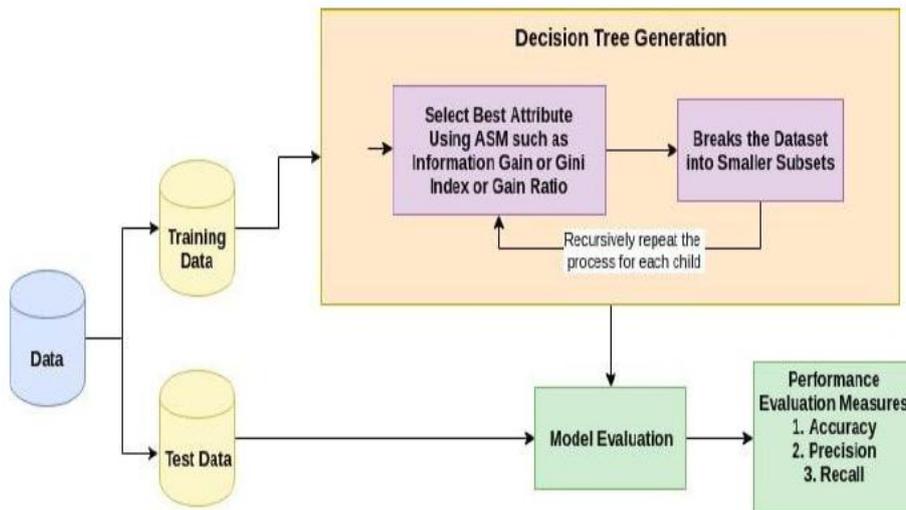


**Figure 1**: Existing System Architecture

| Serial No. | GRE Score | TOEFL Score | University Rating | SOP | LOR | CGPA | Research | Chance of Admit |
|---|---|---|---|---|---|---|---|---|
| 1 | 337 | 118 | 4 | 4.5 | 4.5 | 9.65 | 1 | 1 |
| 2 | 324 | 107 | 4 | 4 | 4.5 | 8.87 | 1 | 1 |
| 3 | 316 | 104 | 3 | 3 | 3.5 | 8 | 1 | 1 |
| 4 | 322 | 110 | 3 | 3.5 | 2.5 | 8.67 | 1 | 0 |
| 5 | 314 | 103 | 2 | 2 | 3 | 8.21 | 0 | 0 |
| 6 | 330 | 115 | 5 | 4.5 | 3 | 9.34 | 1 | 1 |
| 7 | 321 | 109 | 3 | 3 | 4 | 8.2 | 1 | 1 |
| 8 | 308 | 101 | 2 | 3 | 4 | 7.9 | 0 | 1 |
| 9 | 302 | 102 | 1 | 2 | 1.5 | 8 | 0 | 1 |
| 10 | 323 | 108 | 3 | 3.5 | 3 | 8.6 | 0 | 0 |
| 11 | 325 | 106 | 3 | 3.5 | 4 | 8.4 | 1 | 1 |
| 12 | 327 | 111 | 4 | 4 | 4.5 | 9 | 1 | 1 |
| 13 | 328 | 112 | 4 | 4 | 4.5 | 9.1 | 1 | 1 |

**Figure 2**: Sample data set

## II. LITERATURE SURVEY

Acharyaet al. [1] proposed a comparative approach by developing four machine learning regression models: linear regression, support vector machine, decision tree and random forest for predictive analytics of graduate admission chances. Then compute error functions for the developed models and compare their performances to select the best performing model out of these developed models the linear regression is the best performing model with R2 score of

0.72. Janani Pet al. [2] proposed a developed project uses machine learning technique specifically a decision tree algorithm based on the test attributes like GRE, TOEFL, CGPA, research papers etc. According to their scores the possibilities of chance of admit is calculated. The developed model has 93% accuracy.

NavoneelChakrabartyet al. [3] proposed a comparison of different regression models. The developed models are gradient boosting regress or and linear regression model. Gradient boosting regress or have to score of 0.84. That surpassing the performance of linear regression model. They computed different other performance error metrics like mean absolute error, mean square error, and root mean square error. ChithraApoorva et al. [4] proposed different machine learning algorithms for predicting the chances of admission. The models are K- Nearest Neighbor and Linear Regression, Ridge Regression, Random Forest. These are trained by features have a high impact on the probability of admission. Out of the generated models the linear regression model have 79% accuracy.

### III. PROPOSED SYSTEM

The proposed system consist of four regression models. Out of those we use Linear Regression using Dimensionality Reduction which is also a high accurate model.

A user interface is provided through which an actor can interact with the system. The algorithm with improved accuracy will act as a backend for the user interface. Whenever any actor (Student/Consultancy) provides the data to the user interface it will show the result of Chance of Admission which is ranging 0 to 1.
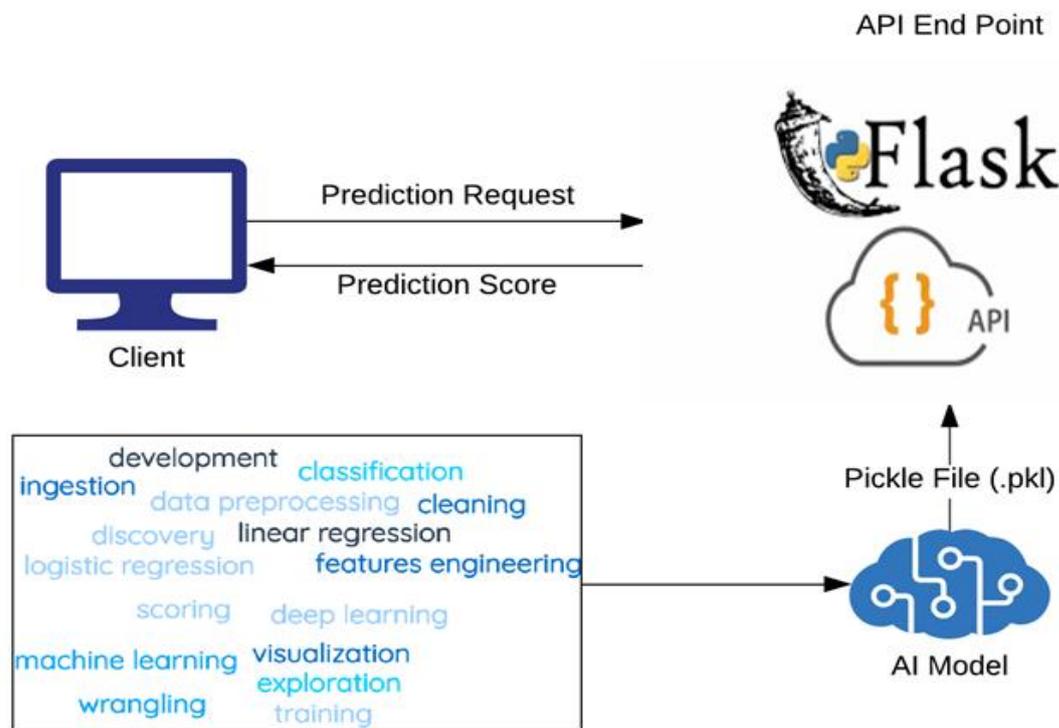


**Figure 3**: Architecture of Propose System

User Manual There are several steps for using this:
1. Initially the user has to open our website and provide all the requested values.
2. The user has to give his GRE score within the range of 270 to 340.
3. The user has to give his TOEFL score within the range of 100 to 120.
4. The user has to give his LOR number within the range of 1 to 5
5. The user has to give his CGPA score within the range of 6.5 to 10
6. In the same way he/she has to select values within given ranges or given options.
7. All the given inputs are displays in screen and prediction is also displayed output is ranging from 0 to 1.

**Benefits of the Proposed System**

- High R2 score
- Low Root Mean Square Error (RMSE)
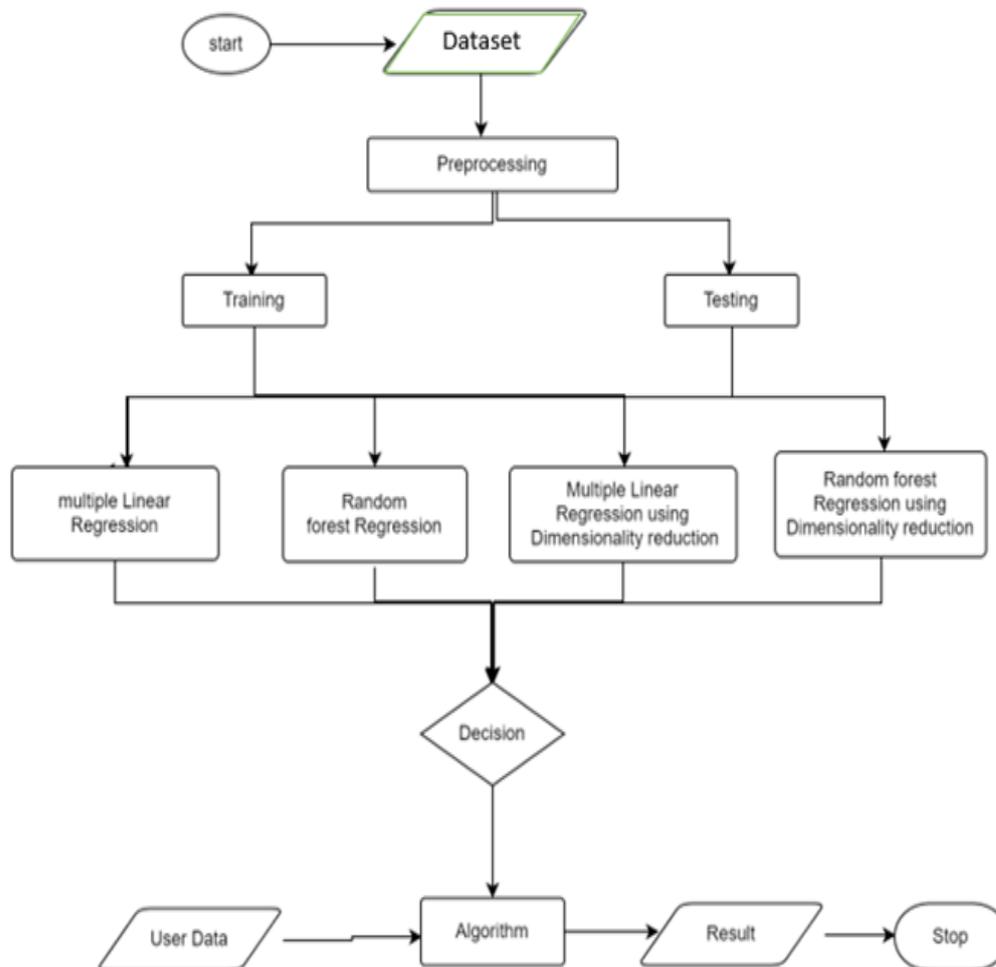


**Figure 4**: Flow chart

### IV. CONCLUSION

Student admission problem is very important in educational institutions. In this project addresses machine learning models to predict the chance of a student to be admitted. This will assist students to know in advance if they have a chance to get accepted. Machine learning models were performed to predict the opportunity of a student to get admitted to a master's program. The machine learning models included are multiple linear regression, random forest, Multiple Linear Regression with Backward Elimination and random forest regression with backward elimination. Experiments show that the Linear Regression model surpasses other models.

Our aim would be to predict the "Chance of Admit" based on the different parameters that are provided in the dataset. We will achieve this aim by using the Linear Regression model. Based on the data that we have, we will split

out data into training and testing sets. The Training set will have features and labels on which our model would be trained. The label here is the "Chance of Admit". If you think from a no-technical standpoint then label is basically the output that we want and features are the parameters that drive us towards the output. Once our model is trained, we will use the trained model and run it on the test set and predict the output. Then we will compare the predicted results with the actual results that we have to see how our model performed. This whole process of training the model using features and known labels and later testing it to predict the output is called Supervised Learning.

## REFERENCES

1. Acharya MS, Armaan A, Antony AS (2019) a comparison of regression models for prediction of graduate admissions. In: 2019 IEEE International conference on computational intelligence in data science (ICCIDS). IEEE

2. Janani P, HemaPriya V, MonishaPriya S, Prediction of MS Graduate Admissions using Decision Tree Algorithm ,International Journal of Science and Research (IJSR) ISSN: 2319-7064 ResearchGate Impact Factor (2018): 0.28 | SJIF (2018): 7.426.

3. Gupta N, Sawhney A, Roth D (2016) Will I get in? Modeling the graduate admission process for American universities. In: 2016 IEEE 16th international conference on data mining workshops (ICDMW). IEEE.

4. NavoneelChakrabarty, Siddhartha Chowdhury, Srinibas Rana ons, A Statistical Approach to Graduate Admissions' Chance Prediction,in book: Innovations in Computer Science and Engineering (pp.333-340), march 2020.

5.https://www.researchgate.net/publication/339653123_A_Statistical_Approach_to_Graduate_Admissions'_Chance_Prediction.

6. ChithraApoorva D A, MalepatiChanduNath, Peta Rohith, BinduShree.S, Swaroop.S modelling the Prediction for University Admission using Machine Learning. International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-6 March 2020.