

GRADUATE ADMISSION PREDICTION USING MACHINE LEARNING TECHNIQUES

BHANU PRASAD GORANTLA¹, K BHAVYA GOUD², P JASMI³, M MANASA⁴

ASSISTANTPROFESSOR¹, UG SCHOLAR^{2,3&4}

DEPARTMENT OF CSE, MALLA REDDY ENGINEERING COLLEGE FOR WOMEN,MAISAMMAGUDA, DHULAPALLY KOMPALLY, MEDCHAL RD, M, SECUNDERABAD, TELANGANA 500100

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

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

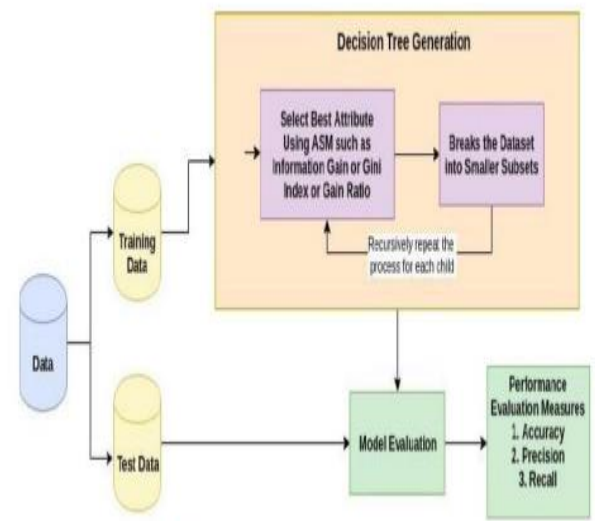


Figure 1: Existing System Architecture

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

LITERATURE SURVEY

Acharya et 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. NavoneelChakrabarty et 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.

IN “WILL I GET IN? MODELING THE GRADUATE ADMISSION PROCESS FOR AMERICAN UNIVERSITIES

AUTHORS: NUNO CARNEIRO, G. FIGUEIRA, M. COSTA”

We study the graduate admission process in American universities from students' perspectives. Our goal is to build a decision support model that provides candidates with pertinent information as well as the ability to assess their choices during the application process. This model is driven by extensive machine learning based analysis of large amounts of historic data available on the web. Our analysis considers factors such as standardized test scores and GPA as well as world knowledge such as university reputation. The learning problem is modeled as a binary classification problem with latent variables that account for hidden information, such as multiple graduate programs within the same institution. An additional contribution of this paper is the

collection of a new dataset of more than 25,000 students, with 6 applications per student on average and, hence, amounting to more than 150,000 applications spanning across more than 3000 source institutions. The dataset covers hundreds of target universities over several years, and allows us to develop models that provide insight into student application behavior and university decision patterns. Our experimental study reveals some key factors in the decision process of programs that provide applicants the ability to make an informed decision during application, with high confidence of being accepted.

In “An empirical comparison of supervised learning algorithms

AUTHORS: RJ Bolton, DJ Hand”

A number of supervised learning methods have been introduced in the last decade. Unfortunately, the last comprehensive empirical evaluation of supervised learning was the Statlog Project in the early 90's. We present a large-scale empirical comparison between ten supervised learning methods: SVMs, neural nets, logistic regression, naive bayes, memory-based learning, random forests, decision trees, bagged trees, boosted

trees, and boosted stumps. We also examine the effect that calibrating the models via Platt Scaling and Isotonic Regression has on their performance. An important aspect of our study is the use of a variety of performance criteria to evaluate the learning methods..

In “A Comparison of Regression Models for Prediction of Graduate Admissions.

AUTHORS:

Zhang,Xinwei;abHan,Yaocia,WeiXu,Wa

ngQilia” In this paper, the gradient boosting regressor-based machine learning model is developed in which parameters necessary for the admission purpose like GRE Score, TOEFL Score, University Rating, Statement of Purpose and Letter of Recommendation Strength, Undergraduate GPA and Research Experience are taken into consideration. The trained model, when tested on unknown test instances, fetches appreciable statistical results for the chance (probability) value prediction of admission and hence gives an impartial notion of the selection criteria. Previously, there have been several experimental attempts involving analytical techniques for developing a statistically successful and reliable graduate admission prediction system. A few of them include: - Acharya et al. 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. - Gupta et al. [2] constructed a decision support system powered by machine learning for prediction of graduate admissions in USA via classification by taking certain parameters like standardized test scores, GPA and Institute Reputation into consideration.

MODULES DESCRIPTION:

STUDENT:

The student can register the first. While registering he required a valid student email and mobile for further communications. Once the student registers, then the admin can activate the student. Once the admin activates the student then the student can login into our system. After login students will write the exams. based on students exams score we can predict chance of admission in particular university .

ADMIN:

Admin can login with his credentials. Once he logs in he can activate the students. The activated student only login in our applications. The admin will perform machine learning models to predict the chance of the admin in graduate schools.

Admin implement support vector machine learning and logistic regression and random forest models,k-nearest neighbour,artificial neural network

MACHINE LEARNING:

Machine learning refers to the computer's acquisition of a kind of ability to make predictive judgments and make the best decisions by analyzing and learning a large number of existing data. The representation algorithms include deep learning, artificial neural networks, decision trees, enhancement algorithms and so on. The key way for computers to acquire artificial intelligence is machine learning. Nowadays, machine learning plays an important role in various fields of artificial intelligence. Whether in aspects of internet search, biometric identification, auto driving, Mars robot, or in American presidential election, military decision assistants and so on, basically, as long as there is a need for data analysis, machine learning can be used to play a role.

CONCLUSION

Among our classification algorithms, Artificial Neural Network again performs best based on Accuracy and FScore (0.95 and 0.91). Even though it does not guarantee to converge to the global minima, it

certainly does find a good optimum based on some random initialization of weights [3]. The work has yielded good results and are in accordance with the approach and performance has been satisfactory.

REFERENCES

- [1] Raghunathan, Karthik. "Demystifying the American graduate admissions process." Web. 2010.
- [2] Gupta, Narender, Aman Sawhney, and Dan Roth. "Will I get in? modeling the graduate admission process for American universities." 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW). IEEE, 2016.
- [3] Caruana, Rich, and Alexandru Niculescu-Mizil. "An empirical comparison of supervised learning algorithms." Proceedings of the 23rd international conference on Machine learning. ACM, 2006.

- [4] Mohan Acharya, Asfia Armaan, Aneeta Antony : A Comparison of Regression Models for Prediction of Graduate Admissions, International Conference on Computational Intelligence in Data Science 2019
- [5] Géron, Aurélien. Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems. " O'Reilly Media, Inc.", 2017.
- [6] Dietterich, Thomas G. "Ensemble methods in machine learning." International workshop on multiple classifier systems. Springer, Berlin, Heidelberg, 2000.
- [7] Schölkopf, Bernhard, Alexander J. Smola, and Francis Bach. Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT press, 2002.
- [8] Russell, Stuart J., and Peter Norvig. Artificial intelligence: a modern approach. Malaysia; Pearson Education Limited, 2016.