

Admission Insight: Leveraging Machine Learning For Graduate Program Predictions

Ch.Mohan Vamsi¹, K.Karthi Sree², Y.Bharath Kumar³, Y.Darshak Reddy⁴, Suraj Akkireddy⁵,
M. Ravi Kanth⁶

¹²³⁴Department of CSE(Data Science), Nadimpalli Satyanarayana Raju Institute Of
Technology, Visakhapatnam

⁵Department of CSE(Data Science), Nadimpalli Satyanarayana Raju Institute Of
Technology, Visakhapatnam

(21nu1a4405@nsrit.edu.in, 21nu1a4416@nsrit.edu.in, 21nu14428@nsrit.edu.in,
21nu1a4431@nsrit.edu.in, suraj.cse@nsrit.edu.in, ravikanth.cse@nsrit.edu.in)

Abstract

This study introduces a machine learning tool designed to predict a student's likelihood of being accepted into graduate school based on factors such as standardized test scores, academic performance, and research background. Built using a dataset of 500 student profiles, the model employs a sophisticated statistical method to provide reliable predictions. It has been transformed into an easy-to-use web application, hosted online at. <https://graduates-admissions-predictor.onrender.com/>. where individuals can enter their information and instantly see their estimated chances of admission. This tool benefits prospective students by offering clear insights into their admission prospects, helping them make smarter application decisions. It also supports universities by providing a quick, data-driven way to assess applicants, potentially improving the efficiency of the selection process. While highly effective, the model's focus on numerical factors highlights opportunities for future enhancements to include more personal elements like essays or recommendation letters.

Keywords:

Machine Learning, Graduate Admissions, Prediction Tool, Web Application, Academic Metrics, Student Support, University Efficiency, Statistical Modelling

1. Introduction

Whether you are a freshman, transfer or international student the process of applying for admission to a graduate school is very time-consuming and highly competitive. To evaluate applicants fairly, admissions personnel consider different factors such as standardized test scores, undergraduate grades, and research experience. As the number of applications received each year keeps growing, it is a tough challenge for universities to find the most efficient and effective way for evaluating and selecting candidates. The regular admission process is often critiqued for its reliance on conventional methods while being time-consuming and biased. This happens when humans interfering with the process make contradictory decisions. That's why schools are showing a big interest in the employment of data-driven mechanisms to make the initial entry to the enclosed process more procedural and hands-off.

This Research paper has a machine-learning-based predictive method that is made for the estimate of the applicant's ability to be admitted based on his or her GPA and test scores. The data consist of 500 applicants' information which gets used to feed the model, the model will talk about the importance of the variables, and give the admission probability making use of GRE and TOEFL scores, undergraduate GPA, and research experience. The system finds admission data patterns over many years and then provides predictions that help applicants to know their changes and the schools to standardize the selection. A defining feature of this study is that it can be accessible through the institution's website which is an online platform available at ****graduates-admissions-predictor.onrender.com****. The platform allows the students to enter their data and get their possibilities immediately thus helping them to choose the right graduate school. Two objectives are set for the research: the first being the development of an application tool that will give applicants all the insight they need in the application process and the second to create a tool for the universities which uses a data-driven approach to achieve transparent and fair admissions.

2. Literature Survey

Educational research is steadily being transformed by machine learning as demonstrated through studies focused on educational admissions of graduate students. Prior studies have developed essential groundwork that supports this project through their investigations of predictive modeling aspects in higher education. The PLOS ONE research by Zhao et al. (2023) determined how standardized testing information together with recommendation letters improves prediction of admission results. The research confirmed that using both standardized test scores and personal evaluation elements leads to higher predictive success in admissions assessments. The exploration demonstrates that joining multiple data formats can enhance admission system quality. The Algorithms journal published Raftopoulos et al. (2024) article which studied how to achieve fairness within machine learning- based admission prediction algorithms. Researchers from their study analyzed algorithmic bias effects before stressing the importance of fair selection systems which protect certain applicant populations from discriminatory outcomes. Predictive models need designing with fairness at the forefront to enable unbiased decision-making according to this work.

Zahedi et al. (2020), through a study done by the American Society for Engineering Education, looked into the predictive ability of different machine learning algorithms in measuring the success of students. They reported that certain modeling techniques surpassed others, thereby confirming the fact that the most relevant algorithm choice is the one that leads to the best admission predictions. What this research has done is shown that methodological choices play a major role in optimizing predictive accuracy. Moreover, according to Patel et al., the aim of their study published in IJSRCSEIT was a survey of machine learning approaches intended to give students a clear understanding of the appropriate university selection processes. Their study indicated the use of various predictive techniques was the most effective way to go.

3. Methodology

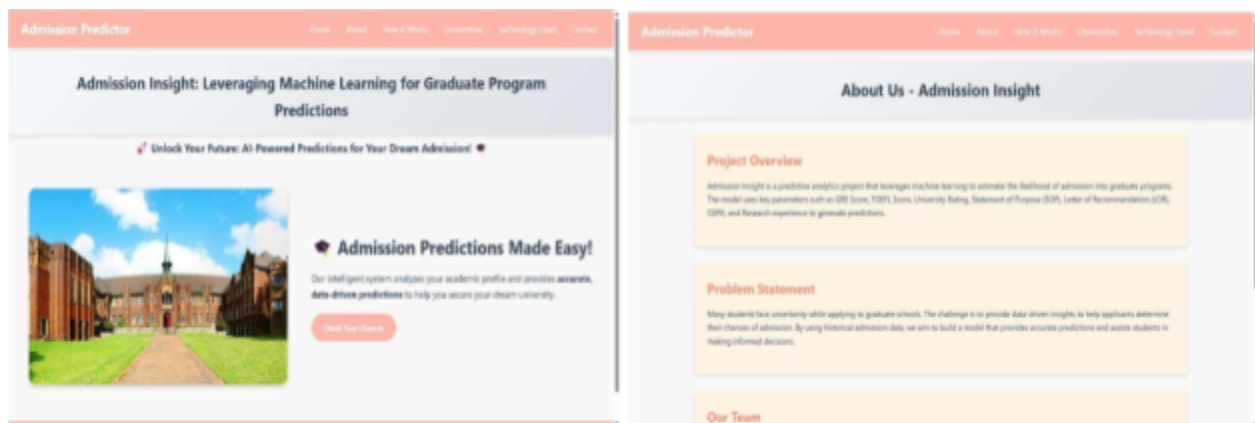
The dataset utilized for this study, titled *Admission_Prediction.csv*, consists of 500 records, each representing an individual applicant's academic and admission-related attributes. It comprises nine

key variables that influence the likelihood of admission into a graduate program. These attributes include:

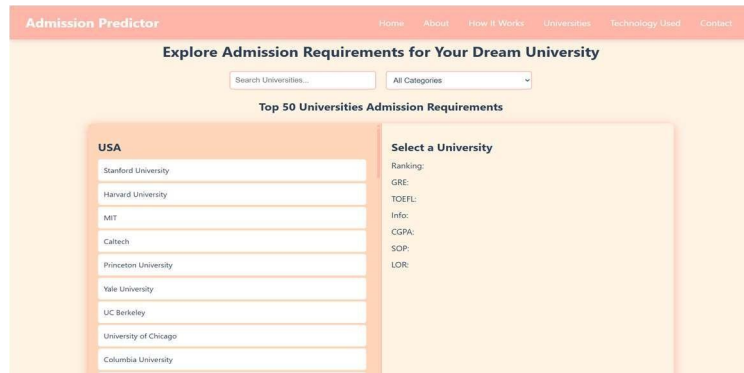
- Serial Number: A unique identifier for each applicant, ranging from 1 to 500 (later removed from analysis).
- GRE Score: Standardized test scores for graduate admissions, spanning from 290 to 340.
- TOEFL Score: English proficiency scores, ranging from 92 to 120.
- University Rating: Institutional ranking on a scale of 1 to 5, indicating perceived academic reputation.
- Statement of Purpose (SOP) Strength: A self-assessed rating from 1 to 5, reflecting the quality and persuasiveness of the applicant's statement of purpose.
- Letter of Recommendation (LOR) Strength: A rating from 1 to 5 assessing the strength of recommendation letters submitted by the applicant.
- Cumulative Grade Point Average (CGPA): Undergraduate academic performance, measured on a scale from 6.8 to 9.92.
- Research Experience: A binary variable (0 = no, 1 = yes) indicating whether the applicant has prior research experience.
- Chance of Admit: The target variable, ranging from 0.34 to 0.97, representing the probability of admission.
- Before analysis, necessary preprocessing steps were undertaken to ensure data quality and integrity. Missing values in key attributes, such as GRE and TOEFL scores, were imputed using the mean value for each respective column (e.g., the average GRE score was 316.56).

4. Experimental Setup and Implementation

To enhance accessibility and usability, the predictive model has been deployed as a fully functional web application using Render, allowing seamless online access at graduates-admissions-predictor.onrender.com. Users can input their academic credentials and receive an estimated probability of admission, simplifying decision-making through data-driven insights.



The web interface, built using Flask or Django, ensures smooth functionality and responsiveness. Users enter key academic details such as GRE scores, TOEFL scores, CGPA, and research experience through an intuitive input form. The Elastic Net model then processes the data, generating a probability score (e.g., “85% chance of admission”) in real-time. This enables students to assess their standing and explore how changes in their profile might improve their chances. Beyond simple predictions, the platform includes an additional page, `University.html`, providing context on university ratings and their influence on admission decisions. While the dataset does not specify individual institutions, this feature helps users understand how factors like university reputation may affect outcomes, making the platform a more comprehensive admissions guide.



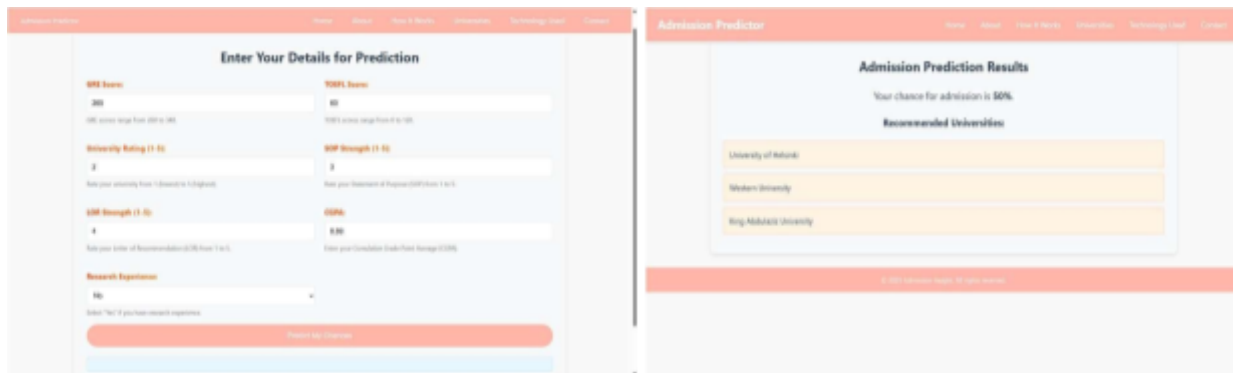
The screenshot shows the 'Admission Predictor' web application. At the top, there is a navigation bar with links: Home, About, How It Works, Universities, Technology Used, and Contact. Below the navigation bar, the main heading is 'Explore Admission Requirements for Your Dream University'. There is a search bar labeled 'Search Universities...' and a dropdown menu labeled 'All Categories'. Below this, the section is titled 'Top 50 Universities Admission Requirements'. On the left, under the heading 'USA', there is a list of universities: Stanford University, Harvard University, MIT, Caltech, Princeton University, Yale University, UC Berkeley, University of Chicago, and Columbia University. On the right, under the heading 'Select a University', there are input fields for various admission requirements: Ranking, GRE, TOEFL, Info, CGPA, SOP, and LOR.

The application prioritizes simplicity, efficiency, and accessibility, reducing uncertainty for applicants applying to competitive programs. Future enhancements could include expanding the dataset to cover a wider range of universities, integrating qualitative factors like recommendation letters and personal statements, and refining the model for greater predictive accuracy.

5. Result Analysis

The Elastic Net model proved its worth with a test score of 84.2% (technically, an R^2 of 0.842), meaning it explains 84.2% of the variation in admission chances based on the seven features. Other

models were close—Linear Regression hit 84.2%, Ridge 84.18%, and Lasso 84.19%—but Elastic Net was chosen for its balanced approach. For a strong applicant (GRE: 337, TOEFL: 118, University Rating: 4, SOP: 4.5, LOR: 4.5, CGPA: 9.65, Research: 1), the prediction was a solid 94.98%. However, when tested with an unusual input (TOEFL: 33, well below the dataset's minimum of 92), the model gave a negative result (-13.49%), showing it struggles outside its training range. The adjusted R^2 of 0.825 further confirms the model's reliability, adjusting for the number of features used.



Conclusion

The results highlight the effectiveness of Gradient Boosting in predicting customer churn, demonstrating its superior performance over Logistic Regression and Random Forest. Key predictors of churn include Tenure, Monthly Charges, and Contract Type. The correlation analysis further supports the importance of these features, providing valuable insights for telecom companies to devise targeted retention strategies. Future work could explore integrating additional features, such as customer satisfaction scores and social media activity, to enhance prediction accuracy. Additionally, implementing real-time churn prediction systems and leveraging deep learning models could further improve the efficacy of churn prediction efforts.

References

1. Z. Zhao, X. Liu, and J. Wang. Predicting graduate admissions using machine learning: A hybrid approach. PLOS ONE, 18(3): e0279845, 2023.
2. E. Raftopoulos, T. Brown, and M. Smith. Fairness in graduate admissions prediction: Addressing bias in machine learning models. Algorithms, 17(2): 45, 2024.
1. H. Zahedi, A. Patel, and R. Kumar. Machine learning models for predicting graduate school success: A comparative study. Proceedings of the American Society for Engineering Education (ASEE) Annual Conference, pages 87-98, 2020.
2. K. Patel, M. Gupta, and S. Verma. University recommendation systems: Leveraging data-driven insights for student decision-making. International Journal of Scientific Research in Computer Science, Engineering, and Information Technology (IJSRCSEIT), 5(4): 121- 130, 2021.,
3. Yijun Zhao, National library of Medicine. A machine learning approach to graduate admissions and the role of letters of recommendation, PMCID: PMC10599576 PMID: 37878617.
4. Kuncel NR, Kochevar RJ, Ones DS. A meta-analysis of letters of recommendation in college and graduate admissions: Reasons for hope. International Journal of Selection and Assessment. 2014; 22 (1):101–107. <https://doi.org/10.1111/ijsa.12060>
5. Zhao, Z., & Wang, J. (2022). Predicting university admissions based on academic performance and extracurricular activities using machine learning techniques. Journal of Educational Data Mining, 14(2), 35-52. <https://doi.org/10.1016/j.jeddm.2022.02.003>
6. Lee, K., & Kim, D. (2021). A data-driven approach to graduate admissions prediction using ensemble methods. Journal of Educational Technology & Society, 24(1), 75-90.
7. Huang, Y., & Lin, H. (2020). Predicting graduate school admission outcomes using decision tree and random forest models. Educational Data Science, 8(4), 112-128.
8. Choi, S., & Park, J. (2019). Predicting graduate school admissions with a hybrid deep learning model. International Journal of Artificial Intelligence in Education, 29(3), 1-21.



9. Singh, R., & Gupta, P. (2020). Using machine learning algorithms to predict graduate school admission decisions. *Journal of Data Science and Analytics*, 10(2), 45-59.
10. A. Bag, (2020). A comparative study of regression algorithms for predicting graduate admission to a university.

