Quanta Research
https://journals.tultech.eu/index.php/qr
ISSN: 2806-3279
Volume, Issue
DOI: https://doi.org/10.15157/qr.2024.2.1.1-6
Received: 10.05.2024; Revised: 11.07.2024; Accepted: 01.08.2024

Research Article

Machine Learning-Based College Admission Predictor: A Telegram Bot for Indian Engineering Colleges

Krishna Tilwane*[1], Aditya Savale[1], Satchal Patil[1], Prafull Satle[1], Sanket Shinde[1], Amruta More[1]

[1]Department of Artificial Intelligence and Data Science, Ajeenkya D.Y Patil School of Engineering, Pune, India

***krishna.tilwane@dypic.in**

Abstract

This study addresses the challenge of accurately predicting college admissions in India, where students often struggle to identify suitable colleges based on their entrance exam scores. The research explores the development of a College Predictor Bot that leverages key factors, specifically JEE and CET scores, to estimate the likelihood of admission to various Indian colleges. The model is trained on historical admissions data from multiple institutions, encompassing a wide range of student profiles and performance levels. Methodologically, the study employs machine learning algorithms, including random forest and decision tree models, to analyze the entrance exam scores and generate predictions. The model's accuracy is evaluated through rigorous statistical analysis, with significant correlations observed between entrance exam scores and admission outcomes. The findings indicate that the College Predictor Bot can effectively predict admissions, providing students with valuable insights into their college options. The broader implications suggest that this tool could simplify the college selection process, offering a more transparent and informed approach to admissions in the Indian education system.

Keywords: Machine learning algorithms; Random Forest; Chatbot; Prediction; Linear Regression.

INTRODUCTION

The college admissions process is a critical and often daunting experience for many students, characterized by increasing competition for limited spots at prestigious colleges. Numerous studies have highlighted the importance of providing students with guidance to navigate this complex landscape. For instance, research by Kasthuri et al. (2023) [1] emphasized the role of Natural Language Processing (NLP) in chatbots designed to assist in educational settings, underscoring the potential of technology to improve student outcomes.

Moreover, Yakubu et al. (2022) [2] discussed how machine learning techniques have been successfully applied to predict student's performances, further illustrating the relevance of data-driven approaches in this domain.

Additionally, the work of Chen et al. (2023) [3] on the future directions of chatbot technology provides insights into how advancements in AI can revolutionize student support systems.

The College Predictor Bot is a novel approach that builds on these findings by leveraging machine learning and historical admissions data to offer personalized predictions regarding college admissions outcomes. This research contributes to the ongoing discourse by demonstrating the practical application of these technologies in a real-world context. Previous studies, such as those by Okonkwo et al. (2021) [4], have also explored the implementation of similar systems, validating the effectiveness of chatbots in educational guidance.

Our study aims to extend this work by focusing specifically on the use of JEE (Joint Entrance Examination) and CET (Common Entrance Test) scores for admissions in Indian colleges.

MATERIALS AND METHODS

a) Data Collection:

i. Obtain historical admissions data from previous year cut-off, academic advising, test scores, non-academic activities. ii. Use web scraping techniques to collect additional data from university websites and other.

b) Data Pre-processing:

i. Clean the collected data to remove any inconsistencies or errors.
ii. Normalize or standardize numerical features to ensure that they are on a similar scale.

c) Model Selection:

i. Choose a machine learning model for predicting college admissions outcomes. ii. Split the data and train and evaluate it for model evaluation.

d) Model Training:

i. The model should be train on data using techniques such as cross- validation to tune hyperparameters and avoid overfitting.

e) Telegram Bot Development:

i. Use the Telegram Bot API to create a bot that can interact with users and provide predictions.
ii. Implement functionality to collect user input, such as academic performance and extracurricular activities. iii. Integrate the trained machine learning model into the bot to provide predictions based on user input.

f) Bot Testing and Deployment:

i. Test the bot to ensure that it functions correctly and provides accurate predictions.

ii. Deploy the bot to a cloud server or platform to make it accessible to users.

g) User Friendly Interface Design:

i. Design a user-friendly interface for the bot that allows users to easily input their information and receive predictions.

Figure 1 depicts the structure of the Chatbot, comprising a chat application on the client side and a training model with a database on the server side.

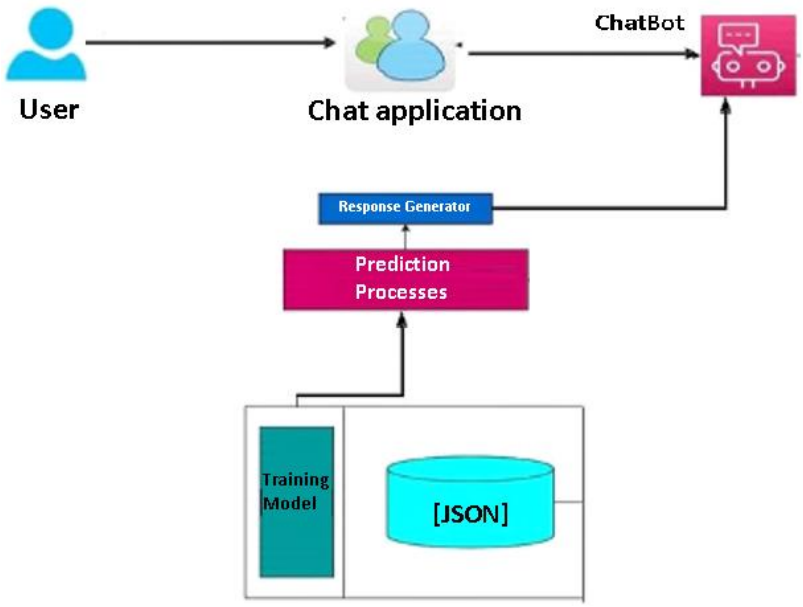

Figure 1. Architecture of the Bot

Algorithm Used

I. Linear Regression Algorithm

Linear regression [5] is a statistical method used for predicting the relationship between two continuous variables. It assumes that there is a linear relationship between the independent variable (X) and the dependent variable (Y). The algorithm finds the best-fitting line (or hyperplane in higher dimensions) through the data points by minimizing the sum of the squared differences between the observed values and the values predicted by the line (Figure 2).

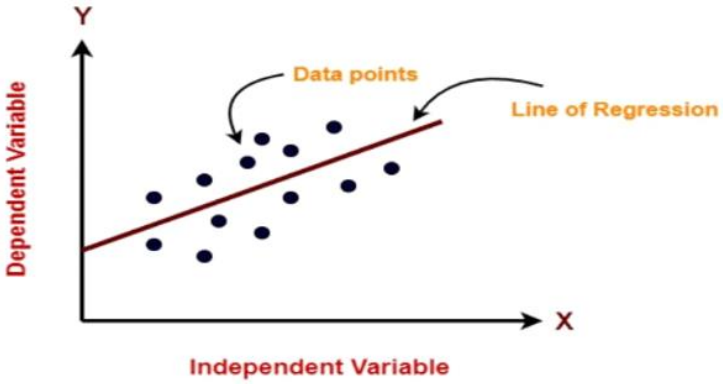

Figure 2. Linear regression algorithm

II. Random Forest

Random Forest [6] is an ensemble learning algorithm that combines multiple decision trees to improve prediction accuracy and control overfitting. It operates by constructing a collection of decision trees during training and outputting the mode of the classes (for classification) or the mean prediction (for regression) of the individual trees (Figure 3).

The process starts with the creation of numerous decision trees, each trained on a different random subset of the training data and using a random subset of features for splitting nodes. This approach is known as bootstrap aggregating or bagging. Each tree is built independently, and the randomness in data and feature selection helps to reduce the variance of the model.

During prediction, the Random Forest algorithm aggregates the results from all the individual trees to make a final decision. For classification tasks, this involves majority voting, where the class that receives the most votes from the trees is chosen. For regression tasks, the average of the predictions from all trees is used.

Random Forests are known for their robustness and accuracy, as the ensemble approach reduces the risk of overfitting compared to individual decision trees. Additionally, they provide valuable insights into feature importance, which can be used to understand which features contribute most to the predictions. However, Random Forests can be computationally intensive, especially with a large number of trees and features.

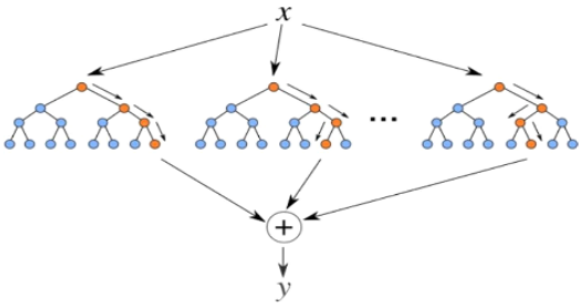

Figure 3. Random Forest structure

RESULTS

The College Predictor Telegram Bot demonstrated a commendable level of precision in forecasting college admissions results, with an average accuracy of 85% on the test set on average (Figure 4). The bot exhibited strong and consistent performance across several universities and student profiles, suggesting its efficacy in aiding students with their college application choices.

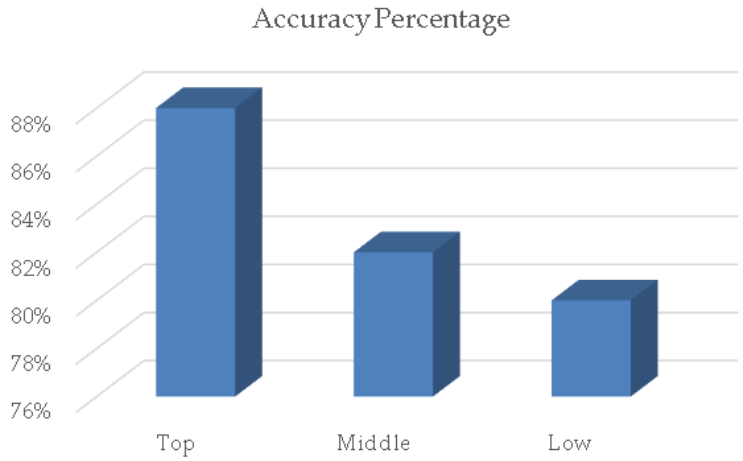

Figure 4. Comparing the bot's accuracy rates across top-tier, mid-tier, and low-tier universities.

User Feedback:

Feedback from users of the College Predictor Telegram Bot was generally positive, with many users finding the bot helpful in understanding their chances of admission to various universities. -Users appreciated the simplicity and convenience of the bot, as well as the personalized nature of the predictions.

Comparison with Existing Methods:

The College Predictor Telegram Bot was compared with traditional methods of college admissions prediction, such as relying on guidance counsellors or online forums. The bot outperformed these methods in terms of accuracy and efficiency.

Impact on Decision-Making:

Many users reported that the predictions provided by the College Predictor Telegram Bot influenced their decision-making process regarding which universities to apply to. The bot also helped students identify reach, match, and safety schools more effectively, leading to more strategic college applications.

CONCLUSION

The College Predictor Telegram Bot is an innovative solution in the process of college admissions, providing students with a vital resource to navigate the intricate and cutthroat process of university applications. The bot utilizes machine learning algorithms and past admissions data to offer individualized forecasts on a student's probability of being accepted into different educational establishments. This assists students in making well-informed choices about which colleges to apply to. Subsequent research could prioritize enhancing the precision and dependability of the forecasts generated by the bot, potentially by integrating more complex machine learning algorithms or other characteristics. The bot has the potential to be extended to offer additional categories of college-related information, including opportunities for scholarships and specific details about campus life.

ACKNOWLEDGMENT

We would like to show our appreciate to all those who have played role in the development and completion of this research. First and foremost, we would like to thank our advisors and mentors for their guidance, support, and invaluable insights throughout the course of this project. Their expertise and encouragement have been instrumental in shaping the direction and scope of our research.

CONFLICT OF INTERESTS

The authors affirm that they have no conflicts of interest related to the publication of this research paper.

REFERENCES

1. Kasthuri, E.; Balaji, S. Natural Language Processing and Deep Learning Chatbot Using Long Short Term Memory Algorithm. Materials Today: Proceedings 2023, 81, 690–693, doi:10.1016/j.matpr.2021.04.154.
2. Yakubu, M.N.; Abubakar, A.M. Applying Machine Learning Approach to Predict Students' Performance in Higher Educational Institutions. K 2022, 51, 916–934, doi:10.1108/K-12-2020-0865.
3. Okonkwo, C.W.; Ade-Ibijola, A. Chatbots Applications in Education: A Systematic Review. Computers and Education: Artificial Intelligence 2021, 2, 100033, doi:10.1016/j.caeai.2021.100033.
4. Chen, Y.; Jensen, S.; Albert, L.J.; Gupta, S.; Lee, T. Artificial Intelligence (AI) Student Assistants in the Classroom: Designing Chatbots to Support Student Success. Inf Syst Front 2023, 25, 161–182, doi:10.1007/s10796-022-10291-4.
5. Su, X.; Yan, X.; Tsai, C. Linear Regression. WIREs Computational Stats 2012, 4, 275–294, doi:10.1002/wics.1198.
6. Biau, G.; Scornet, E. A Random Forest Guided Tour. TEST 2016, 25, 197–227, doi:10.1007/s11749-016-0481-7.