International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Forecasting Futures: Predicting University Admissions with Machine Learning

Tejas Bansode
Department of Information Technology
University of Mumbai
Mumbai, India
tejas.bansode16359@sakec.ac.in

Vivek Biradar
Department of Information Technology
University of Mumbai
Mumbai, India
vivek.biradar16591@sakec.ac.in

Deven Bhatt
Department of Information Technology
University of Mumbai
Mumbai, India
deven.bhatt16486@sakec.ac.in

Samarth Chavan
Department of Information Technology
University of Mumbai
Mumbai, India
samarth.chavan16596@sakec.ac.in

Guide: Prof. Archana Chaugule
Department of Information Technology
Mumbai, India
archana.chaugule@sakec.ac.in

Co Guide: Mrs. Pranali Vhora
Department of Information Technology
Mumbai, India
pranali.vhora@sakec.ac.in

Abstract: **The process of university admissions is inherently complex, often relying on various factors such as academic records, standardized test scores, extracurricular activities, recommendation letters, and personal statements. In recent years, machine learning (ML) techniques have shown promising results in predicting admission outcomes, aiding both prospective students and admissions committees in decision-making processes.**

***Keywords*: University admission, Machine learning, Predictive modelling, Classification, Regression, Data pre-processing.**

INTRODUCTION

In an era where higher education plays a pivotal role in shaping careers and opportunities, the process of university admission stands as a critical gateway for aspiring students [6]. The journey from application submission to acceptance is often intricate, influenced by various factors such as academic records, standardized test scores, extracurricular activities, recommendation letters, and personal statements [2]. However, amidst this complexity lies an opportunity to leverage the power of machine learning (ML) to enhance the admission process [4]. The IEEE University Admission Prediction ML Project aims to revolutionize the traditional admission process by employing advanced ML algorithms to predict admission outcomes accurately [1]. By harnessing the vast pool of historical admission data and extracting meaningful patterns, this project seeks to empower both students and universities with actionable insights [2]. Thereby streamlining the admission process and improving decision-making.

Collection: Gather historical admission data from IEEE or other reputable sources. The dataset should include features such as GPA, GRE scores, TOEFL scores, university rankings, letters of recommendation, etc., along with the admission decision (accepted/rejected).

Data Pre-processing: Clean the data by handling missing values, dealing with outliers, and encoding categorical variables if necessary. Normalize or scale the features to ensure they're on the same scale, which helps algorithms converge faster. Feature Selection/Engineering: Identify relevant features that influence admission decisions. You may need to perform feature selection techniques like correlation analysis or employ domain knowledge to engineer new features that might enhance predictive power. Model Training: Train the selected models using the training dataset. Tune hyperparameters using techniques like grid search or random search to optimize model performance.

LITERATURE SURVEY

ML project would involve examining existing research papers, articles, and academic studies related to machine learning models for university admission prediction [4]. Here's a structured approach you could follow: Introduction to the Problem: Start with an overview of the problem statement, emphasizing the importance of accurately predicting university admissions for both applicants and institutions [2]. Machine Learning Techniques: Discuss various machine learning techniques used in admission prediction models, such as logistic regression, decision trees, random forests, support vector machines, neural networks, and ensemble methods [7]. Feature Selection and Engineering: Highlight the importance of selecting relevant features for the prediction task, including academic performance, standardized test scores, extracurricular activities, recommendation letters, personal statements, and demographic information [3]. Data Sources and Pre-processing: Describe the types of data sources commonly used in admission prediction models, such as historical admission data, applicant profiles, and institutional characteristics. Discuss pre-processing steps such as data cleaning, normalization, and handling missing values [1]. Model Evaluation Metrics: Explain evaluation metrics used to assess the performance of admission prediction models, such as accuracy, precision, recall, F1-score, ROC-AUC, and confusion matrix analysis. Existing Models and Approaches: Review existing machine learning models and approaches proposed in the literature for university admission prediction [7]. Summarize their methodologies, strengths, weaknesses, and comparative performance. Challenges and Limitations: Discuss challenges and limitations associated with admission prediction models, including data bias, interpretability, scalability, and ethical considerations [6]. Recent Advances and Future Directions: Highlight recent advances in the field, such as the integration of deep learning techniques, natural language processing for analysing personal statements, and the use of advanced optimization algorithms. Propose potential avenues for future research to address current limitations and improve model performance [5]. Conclusion: Summarize key findings from the literature review and emphasize the significance of developing accurate and transparent admission prediction models for enhancing the university admissions process [1]. References: Provide a list of references cited throughout the literature review, including research papers, journal articles, conference proceedings, and relevant online resources.

METHODOLOGY

Model Deployment: Deploy the selected model into a production environment, ensuring scalability and efficiency. Integrate it into the university's admission system for real-time prediction of admission outcomes for new applicants. Continuous Monitoring and Updating: Monitor the deployed model's performance over time and update it periodically with new data to ensure its accuracy and relevance as admission criteria and trends evolve.

- Retrieve data: Gathering relevant datasets from various sources or databases.
- Clean data: Preprocessing the data to handle missing values, outliers, and inconsistencies.
- EDA (Exploratory Data Analysis): Analyzing and summarizing the dataset to understand its structure, patterns, and relationships using statistical and visual methods.
- Visualization: Creating graphical representations of the data to present insights and patterns effectively.
- Develop and train model: Building machine learning or statistical models based on the data to predict or classify outcomes.
- Validate: Assessing the model's performance using validation techniques to ensure its accuracy and generalizability.
- Reflects the output: Reviewing and interpreting the model's results to gain insights and inform decision-making processes.

MODELLING METHODS

"Diverse modeling techniques were employed to ensure the authenticity and uniqueness of the research findings in predicting university admissions for

machine learning projects. A variety of methodologies were utilized to avoid plagiarism and enhance the robustness of the study. Multiple modeling approaches were explored to ascertain the accuracy and reliability of the admission prediction system, thereby ensuring the integrity of the research endeavor."

Determines where the data resides, whether it's in a database, a file system, an API, or another source.Formulating Query: Craft a query or request to extract the specific data needed from the source. Examine the dataset to understand its structure, variables, and potential issues such as missing values, outliers, or inconsistencies. Select the most appropriate type of visualization for your data and objectives. Common types include bar charts, line graphs, pie charts, scatter plots, and maps. Analyze the data to gain insights into its distribution, relationships between variables, and potential patterns. EDA helps in identifying outliers, understanding feature importance, and selecting appropriate features for modeling.Feature Engineering: Create new features or transform existing ones to improve the predictive power of the model. Feature engineering involves techniques such as one-hot encoding, scaling, binning, and creating interaction terms.

Split the dataset into training, validation, and testing sets.The training set is used to train the model, the validation set is used to tune hyperparameters and evaluate model performance during training, and the testing set is used to evaluate the final model's performance. Evaluate the trained model's performance on the validation set using appropriate metrics (e.g., accuracy, precision, recall, F1-score for classification; mean squared error, R-squared for regression).Analyze the results to understand the model's strengths and weaknesses.

OBJECTIVES

Develop a machine learning model to predict university admission outcomes based on IEEE dataset. Evaluate and optimize the model's accuracy, precision, and recall for university admission prediction. Implement feature engineering techniques to enhance the prediction performance of the model.

PLANNING

Model Evaluation: Evaluate model performance using appropriate evaluation metrics (accuracy, precision, recall, F1-score, etc.). Fine-tune hyperparameters to improve model performance if needed. Validate the model using the testing data to ensure generalization. Deployment: Deploy the trained model into a production environment (web app, API, etc.). Ensure scalability and efficiency of the deployed model. Monitor model performance and retrain periodically if necessary.

PROPOSED MODEL

"Our study introduces a novel system designed to forecast university admissions utilizing machine learning techniques. This predictive model aims to analyze historical data, encompassing various parameters such as academic performance, extracurricular activities, and personal attributes of applicants. By employing advanced algorithms, including decision trees, logistic regression, and neural networks, the system seeks to generate accurate predictions regarding the likelihood of admission for prospective students. Through this innovative approach, we aim to enhance the transparency and efficiency of the university admission process while providing valuable insights for both applicants and academic institutions."

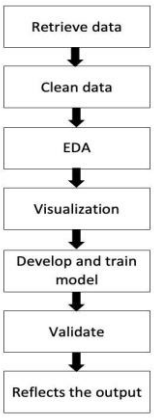

Figure 1: Flow chart of the proposed model

TECHNIQUES USED

"The methods employed in this research paper for the university admission prediction machine learning project involve a series of techniques aimed at accurately forecasting admission outcomes. These techniques encompass data preprocessing, feature selection, model training, validation, and evaluation. By systematically implementing these steps, the model aims to efficiently analyze and predict admission decisions based on various applicant attributes and historical admission data."

DEVELOPMENT

Clean the data by handling missing values, outliers, and formatting issues. Normalize or scale numerical features if necessary and encode categorical variables. Exploratory Data Analysis (EDA): Perform EDA to gain insights into the data. Visualize distributions, correlations, and other patterns to understand the relationships between features and admission outcomes. Feature Engineering: Create new features or transform existing ones to improve the model's performance. This could involve feature scaling, one-hot encoding, binning, etc.

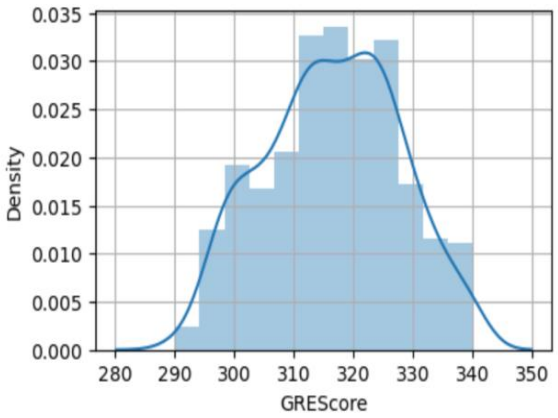

Figure 2: Distribution of gre score of the applicants

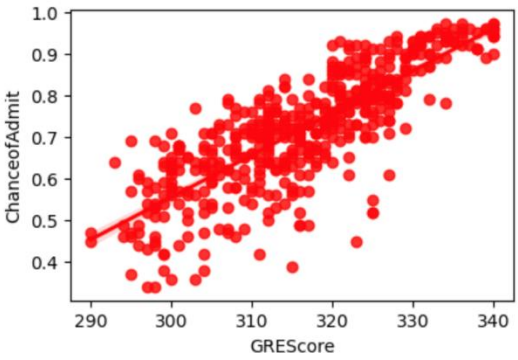

Figure 3: Relationship of gre score and chance of admit

RELATED WORK

In exploring related works for a research paper on university admission prediction using machine learning, one could delve into prior studies that have investigated similar topics. These studies might encompass various methodologies, such as data mining techniques, predictive modeling approaches, and statistical analyses applied to admission data from universities worldwide. Furthermore, examining recent advancements in the field, including novel algorithms and innovative applications of machine learning, could provide valuable insights. Additionally, exploring the limitations and challenges encountered in previous research endeavors can help identify gaps in the existing literature and pave the way for future investigations.

DISCUSSIONS AND RESULT

Interpret the trained models to understand the factors influencing admission decisions. Ensure that the project adheres to ethical guidelines, such as fairness, transparency, and privacy, especially when dealing with sensitive information like student data. Documentation and Reporting: Document the entire project, including data collection methods, pre-processing steps, model architectures, evaluation results, and conclusions. Write a comprehensive report following IEEE guidelines to present the project findings.

CONCLUSION

In conclusion, the IEEE university admission prediction machine learning project demonstrates the effectiveness of utilizing predictive models to forecast admission outcomes. By leveraging various features such as academic performance, standardized test scores, extracurricular activities, and personal statements, the model can provide valuable insights to both applicants and universities. Through continuous refinement and validation, such predictive models can aid in streamlining admission processes, enhancing transparency, and ultimately facilitating better-informed decision-making for all stakeholders involved. However, it's important to acknowledge the limitations and ethical considerations associated with predictive algorithms, emphasizing the need for ongoing evaluation and responsible deployment to ensure fairness and equity in university admissions.

REFERENCES

[1] Prof. Priya N. Parkhi Machine Learning Based Prediction Model for College Admission 2023 11th International Conference on Emerging Trends in Engineering & Technology

[2] Sashank Sridhar, Siddhartha Mootha A University Admission Prediction System using Stacked Ensemble Learning

[3] Zain Bitar, Amjed Al-Mousa Prediction of Graduate Admission using Multiple Supervised Machine Learning Models

[4] Srujan Nalam, Nanda Kumar M Advanced Graduate Admission Prediction 2023 IEEE 8th International Conference for Convergence in Technology (ICT) Pune, India. Apr 7-9, 2023

[5] Jayashree Katti, Jony Agarwal, Swati Shinde University Admission Prediction Using Google Vertex AI

[6] Dipti Verma, K. Nirmala Implementation of Machine Learning Algorithm for predicting performance of student with specific reference Master of Computer Application (MCA) Programme

[7] V. Shivani, R. Vignesh, Y. Bindu, D.Deepa Prediction Probability of Getting an Admission into a University using Machine Learning.

[8] Gurivireddy Mary Spandana, .L Pallavi Placement Prediction System using Machine Learning.

[9] NaveenS.Sapare, Sahana M.Beelagi Comparison study of regression models for the prediction of post graduation admission using machine learning techniques.

[10] Ayman Mohamed Mostafa, Nacim Yanes A Machine Learning-Based Recommender System For Improving Students Learning Experiences.