RESEARCH ARTICLE OPEN ACCESS

University Admission Trend Analysis and Insight Generation System

M. Vasuki[1], Sanjay K[2]
[1] Professor, Department of Master Computer Applications, Sri Manakula Vinayagar Engineering College, Pondicherry-605 107.
[2] PG Student, Department of Master Computer Applications, Sri Manakula Vinayagar Engineering College, Pondicherry-605 107

Abstract:
The University Admission Trend Analysis and Insight Generation System (UAT) is a machine learning-based application designed to help students evaluate their chances of getting admission to universities. The system analyzes important academic factors such as GRE score, TOEFL score, CGPA, Statement of Purpose (SOP), Letter of Recommendation (LOR), university ranking, and research experience to predict admission possibilities. It uses historical admission data and machine learning algorithms to generate accurate predictions and useful insights. The application includes an interactive dashboard that displays admission trends and visualizations, helping students understand the factors that influence admission decisions and identify areas for improvement. Developed using Python, Streamlit, Pandas, NumPy, Scikit-learn, and Plotly, the system provides a user-friendly platform for data analysis and decision-making. By offering predictive insights and trend analysis, UAT assists students in selecting suitable universities and planning their academic goals more effectively.

Keywords — **Machine Learning, Admission Prediction, Data Analytics, Streamlit, Predictive Analytics, Data Visualisation.**

INTRODUCTION
Education is an important piece to building future job prospects as well as advancing one's career. Each year many students are applying to universities in various countries. Due to the increased competitiveness as well as differing criteria for admission, many students are unsure of their likelihood of getting accepted to a university.
Traditional methods of analyzing admission have been performed primarily through manual methods involving historical comparisons. These methods may not yield sufficiently accurate or relevant data. The introduction of machine learning and other advanced data analytic technologies has allowed predictive models to machine analyze numerous previous admission trends in a timeframe enabling students to receive data-driven insight into their likelihood of being admitted to a university.
The "University Admission Trend Analysis and Insight Generation System" utilizes machine learning techniques that give students the ability to predict their probability of admission, based upon their respective performance in multiple academic indicators/areas such as GRE score, TOEFL score, CGPA, strength of SOP, strength of letter(s) of recommendation (LOR), university ranking/ratings, and research experience.
Furthermore, the system provides interactive data visualizations to help students understand the relationship between academic performance and their probability of admission. In addition to producing predictive analytics insight, the system provides visual insight so that students can make informed decisions regarding their future academic objectives.

LITERATURE SURVEY
Machine learning and data mining have made immense improvements in predicting student's acceptance into universities, as well as subsequent

performance in an academic sense, over the past several years. Several research teams have built different systems to analyze student's academic data in creating predictive models that will help the admissions personnel to make better-educated decisions regarding the admission of students into a higher institution.

Kumar and Singh (2021) developed a prediction system for predicting students who would receive admission to their school, which utilized the following predictive models: K-means clustering, linear regression, Naïve Bayes classification and multiple regression analysis. Their findings continue to demonstrate the variables that have the highest level of impact on how students are predicted to gain admission and include GRE score, TOEFL score and CGPA. This research provides evidence that predictive analytics may also assist students during the admissions process.

Patel et al. (2022) developed a framework for predictive analysis of the university admissions process using both algorithms of Random Forest and Support Vector Machines (SVM). This group analyzed multiple historic datasets of university admission decisions to develop an based upon their analysis how well different predictive models performed when predicting student admissions. Their analysis concluded that combining predictive models using ensemble modelling produced better performance than using any individual predictive model.

Sharma and Verma (2023) continue to investigate the role that educational data visualization has in analyzing the performance of students. They stress that there is/are number of benefits that visualization provides when analyzing the performance of students.

Research on machine learning algorithms has been shown to provide insights into a person's probability of acceptance to graduate programs by analyzing their GPA and involvement in extracurricular activities. Similar research has successfully demonstrated that the creation of an AI powered recommendation system using historical admissions data will help graduate schools identify students with the best chance of being accepted into graduate level programs. The data collected by Lee and Kim (2021) provides insight into the rapid evolution of intelligent decision systems within an educational context.

Research regarding educational data mining will continues through 2024. Ongoing findings will demonstrate the increasing value of cloud-based analytical solutions and real-time dashboards for educational institutions dealing with large amounts of educational data. Cloud-based solutions and real-time dashboards provide educational institutions with enhanced capacity, availability, and efficiency and enable all stakeholders to improve the academic performance decision-making processes by providing accurate and timely information.

The existing research provides empirical insight into how universities can use machine learning and analytics to predict university admissions, improving prediction accuracy, automating analysis, and adding meaning to predictions through interactive visualisation.

Machine Learning as a Smart Predictor of Admission

There are many examples of the use of machine learning in educational analytics. One example is creating a system called the "Trend Analysis and Insight Generation System for University Admissions," which will use machine learning to determine the probability of a student being admitted based on both their academic and personal profiles and help clarify to the student how they can improve their chances of being admitted.

This project will use both historical admission data and predictive model creation through supervised machine learning methodologies (and all the above variables as predictors): GRE Scores, TOEFL Scores, University Ratings, Strength of your Personal Statement (SOP), Strength of your Letters of Recommendation (LOR), Cumulative GPA (CGPA); and Research Experience.

Numerous academic variables are identified as independent variables (inputs) and one dependent variable (output) as the model's output. The model's processing layer performs all prediction calculations, preprocessing of data, and presents its results (admission probability) via an output layer. Additionally, the system displays analytical insights and visualizations associated with the admission probability.

Furthermore, the model includes interactive visual analytic tools that will allow users to study admission trends (important for strategic planning) and also the relationships among the various independent variables of the model (academic variables). Examples of these interactive visual analytic tools include correlation heat maps, scatterplots, bar graphs, and trend graphs. Such visualizations provide user understanding of the academic variables and how they relate to admissions and will therefore assist in the data-driven planning efforts of educational institutions.

As the trend towards adopting data-driven methodologies continues within the educational community, machine learning-based university admission models are capable of providing institutions of higher education with scalable and intelligent solutions for academic analytic purposes. In addition, such models have demonstrated an ability to reduce manual analysis time, produce more accurate predictions, and support prospective students in making the most informed decision possible regarding attendance to a postsecondary institution.

3.1 Developing the University Admission Trend Model Architecture

Data Collection Module

The data collection module gathers past admissions data that includes both the students' past academic achievements and the probability that they will gain admission to a given program In each dataset are the following: GRE and TOEFL test scores, grade-point average(GPA), statement of purpose(SOP), letter of recommendation(LOFW), ranking of the school attended, research experience.

Data Preprocessing Module

The preprocessing module prepares the dataset for machine learning methods. It: handles missing values, cleans data, selects features, normalizes features, splits data into training and test sets.

The use of the StandardScaler allows model feature values to be standardized so that all models will be comparable.

Machine Learning Module

The machine learning module creates prediction models through the use of historical admissions data based on the academic factors that usually determine a student's probability of gaining admission. It provides a means to measure the accuracy of the models.

Prediction Module

The prediction module takes as input the student's academic information and returns a prediction of the student's probability of gaining admission via the previously collected data.

Visualization Module

The visualization module generates analytics to provide visual displays of trends in the data and relations within the academic factors via graphs/charts. The types of visual data available within these systems are: correlation matrix, scatterplots, bar graphs, admissions trend graphs, performance analysis graphs.

Insight Generation Module

The insights module listens to the user's input and provides them with insights to improve their likelihood of gaining admission.

User Interface Module

Streamlit is used as a tool for creating user interaction through an interactive dashboard. The dashboard has four main features: Academic details entry capability, Predictive analysis capability, Visual report analysis, and Trend analysis of admissions to various schools/universities. The modularity of the application allows for interoperability and scalability for the purpose of educational analytics and prediction.

4. Transforming Educational Analytics through Machine Learning

The education industry is undergoing dramatic changes through the application of machine-learning technologies in the form of sophisticated prediction systems, personalized analytics, and automated decision-support tools. Educational professionals, Schools, and universities, as well as students leveraging these systems to act on currently available predictive analytics to measure student performance and improve their planning.

University admission is one of the largest areas where machine learning is having a major impact. Predictive models use historical academic records to create a profile of an individual student's scientific and systematic estimates on the likelihood that they will be admitted to an

institution will help inform students regarding how well they measure up to the application criteria of colleges and universities.

Systems that provide educational analytics also help identify the correlation between academic performance and undergraduate success through analyzing academic factors such as the CGPA, GRE scores, TOEFL scores, and research experience. By allowing students to understand how academic factors influence admission outcomes and quality of life post-graduation to define their academic strengths and weaknesses, this type of analysis will assist students in preparing for the next phase of their educational journey.

Educational analytics systems utilize interactive visualization tools that allow students and other users to visualize complex datasets by generating analytics through graphical representation of their data. Tools such as Streamlit and Plotly, which are utilized to create dashboards, enable users to better understand current trends in admission and how those trends may fluctuate over time.

By supporting comparisons of academic performance with predictive models rather than manually comparing academic records, predicting what will be successful in the future is also enhanced because education analytics systems utilize machine learning systems that streamline the workload of performing an educational analysis. Education analytics' predictive models generate insights and recommendations much faster than does performing manual analysis; therefore, machine-learning-generated analytics provide greater efficiency, accuracy, and scalability than does education analytics, which do not.

5. Ethical Considerations, Challenges, and Future Enhancements

Despite the advantages offered by machine learning (ML) educational systems, there is an obligation to address the difficulties faced and their ethical implications for use in a responsible manner.

Challenges of Admission Prediction Systems:

Quality of Data

Due to the importance of the size and quality of the training data set, the predictions will be inaccurate if the data is missing or inconsistent.

Bias in Prediction

A biased data set will cause falsely predicted results. If the historical data has bias or imbalance, the model will inaccurately prescribe an order of recommendations.

Limited Input Factors

There are many additional factors that influence admission in the real world (i.e., extracurricular activities, interviews of potential students, personal achievements, and other non-academic information).These do not necessarily all exist within the data set.

Limitations of the Model

Linear regression assumes that there are linear relationships between variables, which may not adequately characterize the intricacies of the admissions process.

Ethics Considerations

Fairness

The prediction systems must be accessible and administer fair and unbiased results for any and every user, without regard to their background.

Transparency

It should be clear to all users how their predictions are derived and what elements contribute to being accepted into a particular institution.

Privacy and Security

Academic data for admission to institutions must be stored securely and safeguarded from unauthorized access.

Responsible Use of AI

The products of a machine-learning system should provide support for decision making, not completely take the place of human evaluation.

Future Improvements

The following improvements could tremendously expand the potential of this system:

- Linking real-time databases created by universities, to the system
- Machine learning will enable admission prediction models
- An option for recommending universities based on personal criteria
- Developing mobile apps
- Implementing cloud solutions
- Creating analytical dashboards
- Ability to compare multiple universities based on user-defined criteria
- Including modules to help with career and academic planning

Future education analytic systems will use machine learning, visualization technology and automated

recommendations as tools to assist students with making better informed choices concerning their education and career..

Conclusion

When machine learning and analytics are integrated into the process of analyzing student-related academic data (or analytics), they allow for enhanced analysis of the ability of an applicant to obtain admission into a university or school. For example, the "University Admission Trend Analysis and Insight Generation System" shows detailed predictive analytics and visualization capabilities that provide students with the opportunity to see where they stand academically and make adjustments in order to improve their admission chances.

Machine Learning uses algorithms to process large amounts of information, and the analytical techniques used with the algorithms produce an accurate prediction or insight. In addition, the information generated through machine learning will provide a framework for developing an automated system for conducting admission analysis and analyzing academic trends through the visualizations produced by the usage of machine learning algorithms for educational purposes.

The evolution of the technology of machine learning will lead to the development of more accurate, scalable, and personalized systems of educational analytics. Improvements in predictive modeling and artificial intelligence will result in the continued enhancement of the current utilization of admission prediction systems, thus, improving educational planning.

Reference

• Kumar, R. & Singh, P. (2021). Machine Learning Based Admission Prediction System in Universities using Regressional Models; International Journal of Educational Data Mining.
• Patel, S., Sharma, A. & Verma, R. (2022). Student Admission Prediction using Random Forest & Support Vector Algorithms; International Journal of Advanced Research in Computer Science.
• Ahmed, K., Rahman, T. & Ali, M. (2022). Predictive Analytics for Higher Education through Machine Learning Models; Journal of Educational Technologies and Analytics.
• Lee, J. & Kim, H. (2021). AI-Based Recommendation System for Predicting Graduate Admissions; International Conference on Artificial Intelligence in Education.
• Sharma, D. & Verma, S. (2023). Data Visualization from Education & Analysis of Student Performance through Interactive Dashboards; Journal of Data Science Applications.
• Géron, A. (2019). Practical Machine Learning with Scikit-learn, Keras and TensorFlow; O'Reilly Media Press.
• Pedregosa, F. et al. (2011). Scikit-learn: Machine learning in Python; Journal of Machine Learning Research.
• McKinney, W. (2022). Data Analysis Using Python; O'Reilly Media.
• Streamlit Documentation - Streamlit - Creating Interactive WEB applications using the PYTHON programming language; O'Reilly Media.
• Plotly Documentation - Plotly - library for data visualization and interactive graphs; O'Reilly Media..

Aknowledgement

The University Admission Trend Analysis and Insight Generation System project would not have been possible without the help of people and organizations.

We want to say thank you much to our professors and project supervisors for helping us all the time giving us useful ideas and showing us the technical parts of the project. The help and support they gave us made the University Admission Trend Analysis and Insight Generation System better and more effective.

We are also grateful to the staff and administration at Sri Manakula Vinayagar Engineering College for giving us the things we needed like a place to work and the right tools, to finish the University Admission Trend Analysis and Insight Generation System project.

We also want to thank all the people who have worked on or studied machine learning and analytics and data science. Their research and writings and the code they shared helped us understand the ways we could predict things, which we used to make the University Admission Trend Analysis and Insight Generation System.

We want to say a thank you to the people who made the tools and libraries we used for the

University Admission Trend Analysis and Insight Generation System project, like Python and Streamlit and Scikit-learn and Pandas and NumPy and Plotly. They made it possible for us to make the University Admission Trend Analysis and Insight Generation System work.

