Admission Prediction System

Submitted By:

N Amarnath Rao (AM.EN.U4CSE21267)
G VS Kowshik (AM.EN.U4CSE21271)
R Yaswanth (AM.EN.U4CSE21247)
Sanka Bhargava Sai Krishna (AM.EN.U4CSE21251)

ABSTRACT

Informal Description:

We aim to develop a system that predicts admission outcomes for prospective students when they apply to our educational institution. The system should assess applicants based on their academic history, extracurricular activities, and other relevant factors, providing personalized insights into their likelihood of acceptance.

Formal Description:

In the context of an educational institution's admission process:

Task (T): Predict admission outcomes for applicants based on their academic history, extracurricular activities, and other relevant factors.

Experience (E): Historical data on past admissions, including academic records, extracurricular involvement, and acceptance decisions.

Performance (P): Accuracy of admission predictions, measured by how well the system forecasts acceptance outcomes aligned with the institution's admission criteria.

The goal is to create a machine learning-based Admission Prediction System for an educational institution. The system should learn from historical admission data (Experience E) to predict admission outcomes (Task T). The performance measure (P) would be the accuracy of the predictions, ensuring alignment with the institution's admission criteria. Assumptions include the availability of relevant historical admission data and the assumption that admission criteria will remain consistent within the learning period. Additionally, the system assumes that applicants might not have provided a complete overview of their qualifications.

Assumptions:

- **Applicant Profile Stability**: Assumes that applicants' academic history and extracurricular activities will remain relatively stable during the learning period.
- **Relevant Historical Admission Data**: Assumes the availability of sufficient and relevant historical data to train the prediction system effectively.
- **Incomplete Applicant Information**: Assumes that applicants might not provide a complete overview of their qualifications, and the system can make accurate predictions based on partial data.
- **Applicant Engagement**: Assumes that applicants will appreciate and engage with admission predictions without finding them intrusive or discouraging.

1. INTRODUCTION

Motivation:

The primary motivation for solving the Admission Prediction System for an educational institution is to streamline the admissions process and provide valuable insights to both applicants and the institution. By offering personalized predictions based on academic history and extracurricular activities, we aim to enhance the application experience and increase transparency in the admission process. Additionally, the project serves as an opportunity to explore and implement machine learning techniques in the educational domain, contributing to skill development and understanding.

Benefits of Solution:

The successful implementation of the prediction system will lead to several benefits. Applicants will gain valuable insights into their likelihood of acceptance, enabling them to make informed decisions and potentially improve their profiles. From the institution's perspective, the system can optimize the admissions process, ensuring a more efficient and data-driven selection of candidates. Meaningful benefits include improved applicant satisfaction, increased transparency, and a competitive edge in the educational landscape.

Solution Use:

The solution will be integrated into the educational institution's admission process, operating in real-time to offer predictions based on applicant profiles and relevant factors. The operationalization of the solution involves continuous monitoring and refinement to adapt to changing admission criteria and trends. The ultimate goal is to create a dynamic and self-improving system that becomes an integral part of the admissions process. The solution's lifetime extends beyond the initial development, requiring ongoing maintenance and updates to ensure its effectiveness and relevance. If successful,

the results and insights obtained from the prediction system may be presented in a comprehensive report to stakeholders, showcasing the impact on the admissions process and overall institution performance.

2. Dataset finalization

Graduate Admissions:-

What is the data about?

The data appears to be related to graduate admissions, possibly for master's or doctoral programs. The columns suggest various attributes that might be considered during the admission process, such as GRE score, TOEFL score, university rating, statement of purpose (SOP), letter of recommendation (LOR), CGPA (Cumulative Grade Point Average), and whether the applicant has research experience. The target variable seems to be 'Chance of Admit,' indicating the likelihood of admission.

Number of features and description:

- **Serial No.**: A serial number or identifier for each data point.
- **GRE Score**: The Graduate Record Examination (GRE) score, a standardized test for graduate school admissions.
- **TOEFL Score**: The Test of English as a Foreign Language (TOEFL) score, assessing English proficiency for non-native English speakers.
- **University Rating**: A rating associated with the applicant's undergraduate institution, indicating the perceived quality.
- **SOP**: Statement of Purpose, a document where applicants explain their academic and professional goals and why they are interested in the program.
- **LOR**: Letter of Recommendation, a reference letter from someone who can vouch for the applicant's qualifications.
- **CGPA**: Cumulative Grade Point Average, a measure of a student's academic performance.
- **Research**: A binary variable indicating whether the applicant has research experience (1 for yes, 0 for no).
- **Chance of Admit**: The target variable, representing the probability or likelihood of admission.

Importance of Features:

GRE and TOEFL scores, CGPA, and research experience are likely crucial factors in the admission process, reflecting academic proficiency and research background.

SOP and LOR might provide insights into the applicant's motivations and character, influencing the decision.

University Rating could be a proxy for the reputation of the applicant's undergraduate institution.

Applications of the dataset:

This type of dataset is often used in predictive modeling for graduate admissions. Machine learning models can be trained on historical data to predict the chances of admission for new applicants.

Research in education analytics might use this dataset to understand the impact of different factors on admission probabilities.

It could also be utilized for evaluating the fairness and transparency of the admission process.

Admission through Academics:

What is the data about?

The dataset seems to be related to individuals and their academic details, potentially for admission purposes. The columns include information such as name, age, date of birth (dob), nationality, gender, mother tongue, academic scores, entrance rank, SAT scores, and admission status.

Number of features and description:

- **id**: This is likely an identifier for each individual in the dataset.
- **name**: The name of the individual.
- **age**: Age of the individual.
- **dob**: Date of birth.
- **nationality**: Nationality of the individual.
- **gender**: Gender of the individual.
- **mother_tongue**: The primary language spoken by the individual.
- **10th_score**: Academic score in the 10th grade.
- **12th_score**: Academic score in the 12th grade.
- **main_score**: Main academic score (possibly undergraduate or equivalent).
- **entrance_rank**: Rank obtained in an entrance exam (if applicable).
- **sat_score**: SAT (Scholastic Assessment Test) score.
- **Admission_status**: Indicates whether the individual was admitted or not.

Applications of the dataset:

The dataset appears to be relevant for educational institutions or organizations involved in the admission process.

It could be used for predicting admission outcomes based on various factors like academic scores, entrance ranks, and SAT scores.

Educational research focused on the correlation between academic performance and admission status could leverage this dataset.

It might be used for creating models to understand the factors influencing admission decisions.

Admission Based on Fee and Branch

What is the data about?
This dataset revolves around the admission process in a university, focusing on aspects like gender, admission fee, preferred branch, loan requirements, campus preferences, entrance exam ranks, and the final admission acceptance status. It essentially captures the details of students applying for admission and the factors influencing their admission outcomes.

Number of features and importance

- **admission_id**: A unique identifier for each admission record, aiding in tracking and referencing.

gender: Indicates the gender of the student applying, which might be considered for diverse admission considerations.

- **fee_range**: Encompasses the range of admission fees, providing insights into the financial commitment of the students.
- **fee_slab**: Specifies the specific fee category, offering a more detailed breakdown of admission costs.

loan: Indicates whether the student has applied for a loan, shedding light on their financial situation.

- **req_branch**: Represents the preferred branch of study chosen by the student, a crucial factor in the admission decision-making process.
- **alloted_branch**: Indicates the branch actually allotted to the student, reflecting the outcome of the admission process.
- **campus_req**: Signifies the campus preferred by the student, if multiple campuses exist, providing insights into location preferences.
- **campus_allotted**: Represents the campus where the student is eventually admitted, indicating the final location of study.
- **entrance_rank**: Reflects the rank obtained by the student in the entrance exam, a key determinant in the admission process.
- **admission_acceptance**: Denotes whether the student has accepted the admission offer, shaping the final enrollment status.

Each feature plays a crucial role in understanding the student's profile, preferences, and the factors influencing their admission journey.

Applications:

- **Admission Prediction Models**: Utilized to develop predictive models to forecast admission outcomes based on historical data.
- **Resource Planning**: Assisting in planning for the allocation of resources, including faculty and infrastructure, based on branch preferences.
- **Financial Analysis**: Analyzing the distribution of fee ranges and loan applications to understand the financial dynamics of the student population.
- **Diversity and Inclusion Studies**: Exploring gender-based trends to enhance diversity and inclusion initiatives within the university.

Admission Data of IIT's:

What is the data about?
This dataset revolves around the admission process for the Indian Institutes of Technology (IITs). It encompasses various factors that are considered during the admission procedure.

Number of features and importance?

- Year: The year in which the admission process took place. This is crucial for tracking trends and changes in the admission criteria over time.
- 10th Marks: The marks obtained in the 10th-grade examination. This is an indicator of the candidate's academic performance in the earlier years of schooling.
- 12th Marks: The marks secured in the 12th-grade examination. This holds significant weight as it reflects the candidate's recent academic achievements and preparedness for higher education.
- 12th Division: The division or category in which the candidate falls based on their 12th-grade marks. This might provide additional insights into the academic profile of the applicant.
- AIEEE Rank: The All India Engineering Entrance Examination (AIEEE) rank. This is a standardized test that assesses a candidate's aptitude in physics, chemistry, and mathematics – crucial subjects for engineering.
- College: The specific college within the IIT system to which the admission is sought. This information is pivotal as different IITs may have distinct admission criteria or preferences.

Applications:
- **Predictive Modeling**: Developing models to predict the likelihood of admission based on historical data, helping prospective students understand their chances.
- **Policy Analysis:** Analyzing trends over the years to assess the impact of changes in admission criteria or educational policies.
- **Educational Research**: Studying the correlation between different academic factors (10th marks, 12th marks, AIEEE rank) and success in IIT admissions.

This dataset is a valuable resource for understanding the factors influencing admission into IITs, enabling data-driven insights and informed decision-making in the realm of higher education.

3. MODELS

1. Gradient Boosting Algorithm:

The powerful Gradient Boosting Algorithm is employed to elevate the accuracy and efficiency of an Admission Prediction System in the educational domain. Using decision trees as base learners, the algorithm sequentially corrects prediction errors, constructing a robust model for forecasting admission outcomes. By considering applicant profiles, academic history, extracurricular activities, and other relevant factors, the Gradient Boosting Algorithm offers precise and personalized predictions. This contributes to an informed decision-making process for both applicants and the educational institution, ultimately improving transparency and optimizing the admissions process.

2. Naive Bayes:

The Naive Bayes algorithm, known for its simplicity and efficiency, is leveraged to address the Admission Prediction System for educational institutions. Utilizing probabilistic models, Naïve Bayes classifies applicants based on their profiles and historical admission outcomes. By modeling dependencies between applicant qualifications and admission criteria, the algorithm calculates the likelihood of an applicant being accepted, enabling effective personalized predictions. The Naïve Bayes approach offers a lightweight yet effective solution for improving the transparency and efficiency of the admissions process in educational institutions.

3. Support Vector Machine (SVM):

The robust Support Vector Machine (SVM) algorithm is applied to optimize the Admission Prediction System in an educational environment. By transforming applicant profiles and academic data into a high-dimensional space, SVM constructs hyperplanes that separate and classify applicants based on their likelihood of acceptance. The algorithm excels in capturing intricate patterns and nonlinear relationships, ensuring accurate predictions for admission outcomes. SVM's ability to handle complex datasets makes it a valuable asset for improving the precision and personalization of admission predictions, contributing to enhanced transparency and satisfaction for both applicants and the educational institution.

4. Neural Networks:

Neural networks for admission data mimic human decision-making. They learn patterns from past admission records, analyzing factors like grades and test scores. Input data, resembling a student's profile, undergoes layers of interconnected nodes, adjusting weights to optimize predictions. This mirrors how humans evaluate applicants, identifying key features for acceptance. The model refines its understanding through iterative training, adapting its internal parameters to improve accuracy. By emulating human-like cognition, neural networks discern complex patterns, enhancing their ability to classify applicants effectively based on historical data, making them powerful tools for automating admission decisions.