Predicting the University Eligibility using Data Science

Authors: Ms. Savitha R, M.E., B.E., Shruthi PG, Vino S, Mayakannan L, Sindhuja V

Abstract

The problem statement is to design a college prediction/ prediction system and to provide a probabilistic insight into college administration for overall rating, cut-offs of the colleges, admission intake and preferences of students. It has always been a troublesome process for students in finding the perfect university and course for their further studies. At times they do know which stream they want to get into, but it is not easy for them to find colleges based on their academic marks and other performances. We aim to develop and provide a place which would give a probabilistic output of how likely it is to get into a university given their details.

Introduction

Students are often worried about their chances of admission to University. The aim of this project is to help students in shortlisting universities with their profiles. The predicted output gives them a fair idea about their admission chances to a particular university. This analysis should also help students who are currently preparing or will be preparing to get a better idea.

The education sector has been witnessing a significant increase in the number of students aspiring to pursue higher education in foreign universities. The challenges faced by students in this process are many, starting from the daunting task of selecting the right universities, understanding the admission requirements, submitting applications, and finally, waiting for an acceptance letter. One of the biggest hurdles that students face while applying to foreign universities is the uncertainty around their chances of getting accepted. This leads to a significant amount of stress and anxiety, not to mention the financial burden of applying to multiple universities. Our platform's competitive edge lies in its accuracy and reliability. Unlike other platforms that provide only general information, our algorithm takes into account the specific admission requirements of each university, giving students a more accurate understanding of their chances of getting in. Additionally, our platform will offer personalized advice on how students can improve their chances of getting accepted. This can be achieved using KNN Algorithm.

Literature Survey

K-Nearest Neighbors (KNN) is a widely used machine learning algorithm for classification and regression tasks. It is a non-parametric and instance-based algorithm, meaning that it does not make any assumptions about the underlying distribution of the data and relies on the training data instances to make predictions.

KNN belongs to the family of lazy learning algorithms, where the training data is memorized and used during the prediction phase. It is a simple and intuitive algorithm that is easy to implement and has been widely studied in the literature.

In the KNN algorithm, the class or value of an unseen instance is predicted based on the k-nearest neighbors in the training data. The distance metric used to find the nearest neighbors can vary, but the most commonly used metric is Euclidean distance. The value of k, the number of neighbors to consider, is typically chosen through cross-validation or other model selection techniques.

KNN has been applied to a wide range of applications, including image and speech recognition, recommender systems, and anomaly detection. Despite its simplicity, KNN has shown competitive performance compared to more complex algorithms in many scenarios.

This literature survey will provide an overview of the KNN algorithm, its variations, strengths and weaknesses, as well as its applications in various fields. We will also review some of the recent advancements and challenges in KNN research, including techniques for improving its efficiency and scalability, handling high-dimensional data, and dealing with imbalanced datasets.

Objectives

The objective is to develop a web application that simplifies the college decision-making process for students by leveraging the skills we have learned, including regression and classification techniques for data analysis, data visualization for gaining insights, and applying different algorithms to improve the accuracy of predictions. The application will use evaluation metrics to ensure the quality of results, and will be built using the Flask framework. The goal of the project is to reduce the anxiety associated with choosing the right university by providing an accessible and user-friendly tool that increases the likelihood of selecting a good fit institution.

To achieve the stated objective, we will employ a data-driven approach that involves collecting and analyzing relevant data points such as academic performance, personal interests, financial constraints, location, and university rankings. We will then use regression and classification algorithms to build predictive models that can accurately recommend universities that match the student's preferences and constraints.

To ensure the models' accuracy, we will use evaluation metrics such as mean squared error (MSE), root mean squared error (RMSE), and accuracy scores. By employing these metrics, we can fine-tune the models' parameters and select the best algorithms that perform well on the given data.

We will also incorporate data visualization techniques to provide students with a more intuitive understanding of the recommendations. By using charts and graphs to display the results, students can easily grasp the reasoning behind the recommendations and adjust their preferences accordingly.

Furthermore, we will use the Flask framework to build a web application that makes the process of selecting universities simple and accessible. The application will have an intuitive user interface that allows students to enter their preferences and constraints, view recommendations, and compare universities side by side. The application will also have a backend system that uses the predictive models to generate recommendations and store user data securely.

The end result will be a web application that provides an easy-to-use and accessible tool for students to make informed decisions about their higher education. By reducing the stress and anxiety associated with choosing a university and increasing the likelihood of selecting a good fit institution, we hope to empower students and improve their educational outcomes.

Architecture Diagram
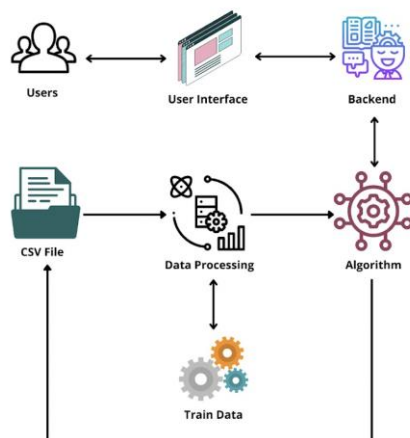
Methodology

1. Ideation and Brainstorming

   The ideation and brainstorming phase is crucial for developing a clear understanding of the problem and defining the scope of the project. In this phase, we will identify the key stakeholders and their needs, analyze the market landscape, and determine the feasibility of the project. We will also brainstorm ideas and concepts for the application, and create a high-level roadmap for the project.

2. Data Collection

   The success of the application depends heavily on the quality of data collected. In this phase, we will gather relevant data points such as academic performance, personal interests, financial constraints, location, and university rankings. We will also identify sources of data such as publicly available datasets, surveys, and web scraping. We will then clean, preprocess, and validate the data to ensure that it is accurate and reliable.

3. Design Process

The design process involves creating a user interface and user experience (UI/UX) that is intuitive and user-friendly. In this phase, we will create wireframes and prototypes of the application, conduct user research and usability testing, and refine the design based on feedback. We will also define the architecture of the application and identify the technologies and frameworks required.

4. Development Process

The development process involves implementing the design and building the application. In this phase, we will create the frontend and backend of the application, integrate the predictive models, and implement features such as user authentication, database management, and data visualization. We will also conduct extensive testing and debugging to ensure that the application is reliable and free of errors.

5. Migration and Deployment

Once the application is built, we will migrate it to a production environment and deploy it to a web server. We will configure the server and set up a database to store user data securely. We will also optimize the application for performance, security, and scalability, and conduct load testing to ensure that it can handle a large number of users.

6. Launch

The launch phase involves promoting the application and making it accessible to the target audience. In this phase, we will develop a marketing strategy and create promotional materials such as social media posts, emails, and advertisements. We will also conduct user testing and gather feedback to further refine the application and improve the user experience. Finally, we will launch the application to the public and continue to monitor and improve it based on user feedback and analytics.

KNN Model:

importing libraries

import numpy as nm

import matplotlib.pyplot as mtp

import pandas as pd

#importing datasets

data_set= pd.read_csv('user_data.csv')

#Extracting Independent and dependent Variable

x= data_set.iloc[:, [2,3]].values

y= data_set.iloc[:, 4].values

Splitting the dataset into training and test set.

from sklearn.model_selection **import** train_test_split

x_train, x_test, y_train, y_test= train_test_split(x, y, test_size= 0.25, random_state=0)

#feature Scaling

from sklearn.preprocessing **import** StandardScaler

st_x= StandardScaler()

x_train= st_x.fit_transform(x_train)

x_test= st_x.transform(x_test)

College Admission Predictor: The KNN algorithm can be used to develop a web-based application that predicts a student's chances of getting admission into a university based on their academic performance, standardized test scores, extracurricular activities, and other relevant factors. The application could provide a user-friendly interface for students to enter their details and receive a prediction of their chances of getting into their preferred university.

Scholarship Eligibility Predictor: Many universities offer scholarships to students based on various criteria such as academic excellence, financial need, and extracurricular achievements. The KNN algorithm can be used to develop an application that predicts a student's eligibility for scholarships based on their academic performance and personal characteristics. This could help students identify scholarships they may be eligible for and increase their chances of receiving financial aid.

Campus Life Recommender: Students often have different preferences and priorities when it comes to selecting a university. The KNN algorithm can be used to develop an application that recommends universities to students based on their interests and preferences. The application could take into account factors such as location, size, campus culture, and available majors and suggest universities that match the user's preferences.

College Rankings Predictor: College rankings can have a significant impact on a student's decision to apply to a particular university. The KNN algorithm can be used to develop an application that predicts a university's ranking based on various factors such as academic performance, research output, and student satisfaction. This could help students identify universities that are likely to be highly ranked in the future and make more informed decisions about where to apply.

Future Scope:

Campus Life Recommender: Students often have different preferences and priorities when it comes to selecting a university. The KNN algorithm can be used to develop an application that recommends universities to students based on their interests and preferences. The application could take into account factors such as location, size, campus culture, and available majors and suggest universities that match the user's preferences.

College Rankings Predictor: College rankings can have a significant impact on a student's decision to apply to a particular university. The KNN algorithm can be used to develop an application that predicts a university's ranking based on various factors such as academic performance, research output, and student

satisfaction. This could help students identify universities that are likely to be highly ranked in the future and make more informed decisions about where to apply.

Conclusion:

Our platform offers a much-needed solution to the uncertainty faced by students aspiring to pursue higher education in foreign universities. With its accuracy and reliability, we believe that our platform will become an essential tool for students, universities, and education agencies worldwide.

References:

1. J. Katti, J. Agarwal, S. Bharata, S. Shinde, S. Mane and V. Biradar, "University Admission Prediction Using Google Vertex AI," 2022 First International Conference on Artificial Intelligence Trends and Pattern Recognition (ICAITPR), 2022, pp. 1-5, doi: 10.1109/ICAITPR51569.2022.9844176.

2. M. Omaer Faruq Goni, A. Matin, T. Hasan, M. Abu Ismail Siddique, O. Jyoti and F. M. Sifnatul Hasnain, "Graduate Admission Chance Prediction Using Deep Neural Network," 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), 2020, pp. 259-262, doi: 10.1109/WIECON- ECE52138.2020.9397988.

3. B. Ujkani, D. Minkovska and L. Stoyanova, "A Machine Learning Approach for Predicting Student Enrollment in the University," 2021 XXX International Scientific Conference Electronics (ET), 2021, pp. 1-4, doi: 10.1109/ ET52713.2021.9579795.