Literature Survey

University Admit Eligibility Predictor

Researchers are working towards the modernization of the education system using education data mining technology. Numerous programs and studies have been carried out on topics relating to university admission using many machine learning models which help the students in the admission process to their desired universities. Previous research done in this area used the Naive Bayes algorithm which will evaluate the success probability of student application into a respective university but the main drawback is they didn't consider all the factors which will contribute to the student admission process like TOEFL/IELTS, SOP, LOR, and undergraduate score. Bayesian Networks algorithms have been used to create a decision support network for evaluating the application submitted by foreign students of the university. This model was developed to forecast the progress of prospective students by comparing the score of students currently studying at university. The model thus predicted whether the aspiring student should be admitted to university based on various scores of students. Since the comparisons are made only with students who got admission into the universities but not with students who got their admission rejected so this method will not be that much accurate.

There is a survey paper that depicts the most relevant studies using educational data mining. The researchers concentrate on the field of educational data mining as recent studies show that it is used for analyzing students' performance. Several past studies focused on predicting admission to colleges or universities. A brief literature review of those studies is presented as follows.

Binu et al. proposed a cloud-based data analysis and prediction system for predicting university admission. There were two modules in the proposed framework, i.e. A Hadoop MapReduce data storage module and an Artificial Neural Network to predict the chances. The data collected had attributes such as status, rank, board, quota, etc. The system did not use academic qualifications in the forecasting process. The neural network had two input nodes, one hidden layer with two nodes, and one output layer with two nodes.

Ghai developed an American Graduate Admission Prediction model that allows students to choose an apt university by predicting whether or not they will be admitted to the university. Gupta et al. developed a machine learning decision support system for the prediction of graduate admissions in the USA by taking account of certain parameters in standardized tests.

Acharya et al. proposed a comparative approach by developing four machine learning regression models: linear regression, support vector machine, decision tree, and random forest for predictive analytics of graduate admission chances. Then compute error functions for the developed models and compare their performances to select the best performing model out of these developed models the linear regression is the best performing model with an R2 score of 0.72. Janani Pet al. proposed a developed project that uses machine learning techniques specifically a decision tree algorithm based on the test attributes like GRE, TOEFL, CGPA, research papers, etc. According to their scores, the possibility of a chance of admission is calculated. The developed model has 93% accuracy.

NavoneelChakrabartyet al. proposed a comparison of different regression models. The developed models are gradient boosting regress and linear regression models. Gradient boosting regresses or has to score of 0.84. That surpasses the performance of the linear regression model. They computed different other performance error metrics like mean absolute error, mean square error, and root mean square error. ChithraApoorva et al. proposed different machine learning algorithms for predicting the chances of admission. The models are K-Nearest Neighbor and Linear Regression, Ridge Regression, and Random Forest. These are trained by features that have a high impact on the probability of admission. Out of the generated models, the linear regression model has 79% accuracy.

Bayesian Networks were used by (Thi et al. (2007)) to create a decision support system for evaluating the application submitted by international students in the university. This model was designed to predict the performance of the aspiring students by comparing them with the performance of students currently studying at the university and who had similar profiles during their application. In this way based on the current students' profile, the model predicted whether the aspiring student should be granted admission to the university. Since the comparisons were made only with the students who were already admitted to the university and the data of the students who were denied admission were not included in the research this model proved to be less efficient due to the problem of class imbalance.

(Abdul Fatah S; M (2012)) developed a model that can provide a list of universities/colleges were the best suitable for a student based on their academic records and college admission criteria. The model was developed by applying data mining techniques and knowledge discovery rules to the already existing in-house admission prediction system of the university. (Mane (2016)) conducted similar research that predicted the chance of a student getting admission to college based on their Senior Secondary School, Higher Secondary School, and Common Entrance Examination scores using the pattern growth approach to association rule mining. The performance of both the models was good the only drawback was the problem statement was single university-centric.