LITERATURE SURVEY

ON

APPLIED DATA SCIENCE - UNIVERSITY ADMIT ELIGIBILITY PREDICTOR

S.Sneha

G.Avanthika

S.Vijayalakshmi

S.Prashanth

Tagore Engineering College

ABSTRACT

Many students who aspire to pursue a Master's degree program from a suitably good university turn towards famous coaching institutes and let them take care of everything, like preparing for exams, building an SOP and LOR, and training for visa interviews and searching for the right universities as well. A few of you may prefer to do all these things on your own. In such situations, searching for the right university is a very daunting task. We search for universities that fit our profile on those so-called "university hunt" websites with all the data about universities around the world. These websites have a section known as "University Predictor," which is most of the time a paid section you need to fill your information to make use of that section. I present how to build your own University Admit Predictor, which gives your chances of getting admitted to the desired university. You can also use this model before giving exams to know beforehand what the required score is to gain admission to your dream university. Accordingly, you can set your targets for studies.

By the end of this tutorial, you will be able to build and train a linear regression model to predict the chance of admission to a particular university.

Book/journal	Author's name	Inference
(Bibodi et al. (n.d.))	George J Kahaly	The literature review of the work that has previously done on predicting the chances of students enrolment in universities. There have been several project and studies performed on topics related to students admission into universities. It is used multiple machine learning models to create a system that would help the students to shortlist the universities suitable for them also a second model was created to help the colleges to decide on enrolment of the student. Nave Bayes algorithm was used to predict the likelihood of success of an application, and multiple classification algorithms like Decision Tree, Random Forest, Nave Bayes and SVM were compared and evaluated based on their accuracy to select the best candidates for the college. Limitation of this research as that it did only relied on the GRE, TOEFL and Undergraduate Score of the student and missed on taking into consideration other important factors like SOP and LOR documents quality, past work experience, technical papers of the students etc

(Thi et al. (2007))	Bernadette Biondi	Bayesian Networks were used 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 in the university and had similar profile 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 in 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))	R Paul Robertson	A model that can provide the list of universities/colleges where the which 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 a similar research that predicted the chance of a student getting admission in 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.

(Mishra and Sahoo (2016))	Caraccio	A research from a university point of view to predict the likelihood of a student enrolling in the university after the have enquired about of courses in the university. They used K-Means algorithm for clustering the students based on different factors like feedback, family income, family occupation, parents qualification, motivation etc. to predict if the student will enroll at the university or not. Depending upon the similarity of the attributes among the students they were grouped into clusters and decisions were made. The objective of the model was to increase the enrolment of the students in the university.
(Eberle et al. (n.d.))	Rotondi	Machine learning and predictive modelling to develop a model that to evaluate the admission policies and standards in the Tennessee Tech University. A well know version of the C4.5 algorithm, J48 was used to create the model. Like the models mentioned above they used the different factors of the student profile to evaluate the chances of their admission in the university. The model worked well in predicting the true positive scenarios where the student was had good profile to secure the admission, but it failed in efficiently identifying the true negatives because of which student who does not satisfy the defined criteria.

(Jamison (2017))	Monzani	In research conducted the yield of college admission was predicted using machine learning techniques. Yield rate can be defined as the rate at which the students who have been granted admission by the university actually enrol for the course. Multiple machine learning algorithms like Random Forest, Logistic Regression and SVM were used to create the model; the models were compared based on their performance and accuracy, Random Forest outperformed the other models with 86% accuracy and was thus used to create the system. The factors that proved to be significant in predicting successful application were also highlighted.

		communication and several sensors. Real-time data access can be done by using remote monitoring and Internet of Things (IoT) technology. Data collected at the apart site can be displayed in a visual format on a server PC with the help of Spark streaming analysis through Spark ML-lib, Deep learning neural network models, Belief Rule Based (BRB) system and is also compared with standard values.