LITERATURE SURVEY

This section includes the literature review of previous research on the assessment of student enrolment opportunities in universities. Numerous programs and studies have been carried out on topics relating to university admission used many machine learning models which helps the students in the admission process to their desired universities. Previous research done in this area used Naïve 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 in the student admission process like TOEFL/IELTS, SOP, LOR and under graduate score(Chithra Apoorva D.A march 2020).

Problem Understanding:
Initially first we have to spend some time on what are the problems or concerns students having during their pre admission period and we should set the solutions to those problems as objectives of this research.
Data Understanding: Data should be collected from multiple sources like yocket and also consider all the factors including which will play a tiny role in student admission process.
Data Preparation: Data should be cleaned that is removing the noise in the data and filling the missing values or extreme values and finalising the attributes/factors which will have crucial importance in student admission process.
Building Models: several ML models have to be developed using various machine learning algorithms for admission to a particular university and the user interface has to be developed to access those models.
Evaluation: Developed models are evaluated according to their accuracy scores. Once the model is finalised that model will be merged with node red for final deployment.(Malepati chanduNath, march2020)

Data collection

The way toward get together information relies upon sort of undertaking, for a ML project, real time information is utilized. The information index can be gathered from different sources like a document, data set, sensor and different sources and some free informational collections from web can be utilized. Kaggle and UCI

Machine learning repository are the storehouses that are utilized the most for information assortment for Machine learning models. Kaggle is quite possibly the most visited sites that is utilized for gathering informational collections.

Pre-Processing

Information pre-processing is a cycle of cleaning the raw information i. e the information is gathered in reality and is changed over to a perfect dataset. There are certain steps executed to change

Over the data into a little clean data collection and make it practical for examination, this piece of the

Interaction is called as information pre-processing (Krithika cs October 2021).

The greater part of this present reality information is chaotic, as :

- Missing Data
- Noisy Data
- Inconsistent Data
- Conversion of Data
- Ignoring the missing Qualities
- Filling the missing Qualities
- Detection of Exceptions
- Feature Extraction

Data virtualization

Data Virtualization is the representation of data in a graph, chart, or other visual format.

Data virtualization by using the histogram and analyzing the data by the graph.

The top three featues that affect the chance of Admit are:

- CGPA
- GRE Score
- TOEFL Score