INFORMATION TECHNOLOGICAL UNIVERSITY



Department of Computer Science

Academic year

(Spring 2019)

**Optimize**

**institution’s student admission process**

By

Priya Phapale

Student Id : 94662

INTRODUCTION

Now a day’s higher education is touted to lead to a number of benefits, including financial security and a prosperous career. Almost all student has plan to complete their graduation as well as post-graduation courses. Every student tries to get selected in top universities so that they will have better job opportunities and a bright future. This can be a turning point of life for each student.

Each year millions of students go through admission process which is tedious task to do. University also has a big responsibility to decide which student should get admit and which student should reject. For taking this decision universities also have their own standards and policies which they need to follow.

We all know this decision is hard and time taking as there are many attributes to be considered like good GPA, GRE, TOEFL scores, impressive SOPs, Letter of recommendations, extracurricular, Outstanding Achievements, Projects and Research etc. As there are millions of profiles, we cannot even imagine doing it manually. To make this process easier I have created such a model which will consider all attributes and produce result.

STRATEGIC PLAN

My goal is to optimize institution’s student admission process which can be addressed using the decision tree techniques. To achieve this I need to follow below steps-

1. Data collection and visualize the data- We are using data set which has four attributes- Admit, GRE, GPA, Rank. We need to visualize the data in table format
2. Apply some statistics on data like sum of each column, mean value etc.
3. Visualize the data for GRE, GPA attributes in graphical form using seaborn library.
4. Train our model using sigmoid function and find accuracy of it.
5. Then we need to normalize GRE and GPA columns value as they have large values.
6. Split into features and targets and find prediction accuracy.
7. Building decision tree model using sklearn module
8. Prediction of a new data on any model



DECISION TREE MODEL

A decision tree is useful and popular methodology. It is a basically a flowchart-like tree structure where an node represents feature or attribute, the branch represents a decision rule, and each leaf node represents the outcome. This leaf node structure helps us in decision making. Root node is a topmost node in a decision tree. It learns to partition on the basis of the attribute value. It partitions the tree in recursively manner call recursive partitioning. Decision tree model is popular as it is inspired by human level thinking and widely used for making predictions.

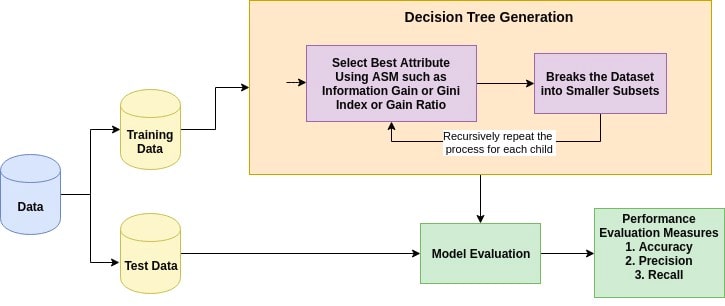
The decision tree is a distribution-free or non-parametric method, which does not depend upon probability distribution assumptions. Decision trees can handle high dimensional data with good accuracy.

The basic idea behind any decision tree algorithm is as follows:

1. Select the best attribute using Attribute Selection Measures (ASM) to split the records.
2. Make that attribute a decision node and breaks the dataset into smaller subsets.
3. Starts tree building by repeating this process recursively for each child until one of the conditions will match:

* All the tuples belong to the same attribute value.
* There are no more remaining attributes.
* There are no more instances

Why decision tree model?



* Decision trees are easy to interpret and visualize.
* It can easily capture non-linear patterns.
* It requires fewer data preprocessing from the user
* It can be used for feature engineering such as predicting missing values, suitable for variable selection.
* The decision tree has no assumptions about distribution because of the non-parametric nature of the algorithm.

DECISION TREE FOR STUDENT ADMISSION PREDICTION

GRE Score

Low

Score < 300

Expected score >= 300

GPA

Expected

GPA >= 3.0

Low

GPA < 3.0

Rank

Low Rank

1 or 2

Expected Rank 3 or 4