# Classification Analysis Practical 7

## 1) Problem Statement

## Assignment on Classification technique

## Every year many students give the GRE exam to get admission in foreign Universities. The data set contains GRE Scores (out of 340), TOEFL Scores (out of 120), University Rating (out of 5), Statement of Purpose strength (out of 5), Letter of Recommendation strength (out of 5), Undergraduate GPA (out of 10), Research Experience (0=no, 1=yes), Admitted (0=no, 1=yes). Admitted is the target variable.

## Data Set: https://www.kaggle.com/mohansacharya/graduate-admissions

## The counselor of the firm is supposed to check whether the student will get an admission or not based on his/her GRE score and Academic Score. So to help the counselor to take appropriate decisions, build a machine learning model classifier using a Decision tree to predict whether a student will get admission or not.

## a) Apply Data pre-processing (Label Encoding, Data Transformation....) techniques if necessary.

## b) Perform data-preparation (Train-Test Split)

## c) Apply Machine Learning Algorithm

## d) Evaluate Model.

## 2) Libraries Used

Python:  
Pandas: For data manipulation and analysis.  
Matplotlib: For creating visualizations.  
Scikit-learn: For implementing machine learning algorithms.

## 3) Theory

Decision Trees are a non-parametric supervised learning method used for classification and regression. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features.

A red line with blue dots

Description automatically generated

A graph with a bar and a number of bars

Description automatically generated with medium confidence

## 4) Methods

The method involves preprocessing the data, splitting it into training and testing sets, applying a Decision Tree Classifier, and then evaluating the model using accuracy, precision, recall, and F1-score.

## 5) Advantages

Advantages of using Decision Trees include:

* • Simple to understand and interpret.
* • Requires little data preparation.
* • Able to handle both numerical and categorical data.

## 6) Disadvantages

Disadvantages of Decision Trees include:

* • Decision-tree learners can create over-complex trees that do not generalize well from the training data.
* • Decision trees can be unstable because small variations in the data might result in a completely different tree being generated.
* • Decision tree learners are often biased toward splits that have more levels.

## 7) Working

The working involves loading the data, preprocessing it, splitting into training and testing sets, fitting a Decision Tree model, making predictions, and evaluating the results.

## 8) Conclusion

The decision tree model provided insights into factors influencing admissions but did not perform effectively with the given data, highlighting the importance of model selection and data sufficiency.