ASSIGNMENT NO. 7

Title: 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 (outof 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.

Software/Libraries Used:

- Python

- pandas

- scikit-learn

Theory/Methodology:

Decision Trees are a popular supervised learning technique used for classification and regression tasks. In this practical, we'll utilize a Decision Tree classifier to predict whether a student will get admission based on their GRE

scores, TOEFL scores, University rating, Statement of Purpose strength, Letter of Recommendation strength, Undergraduate GPA, and Research Experience.

Advantages:

- Decision Trees are easy to understand and interpret.
- They can handle both numerical and categorical data.
- Decision Trees implicitly perform feature selection.

Limitations/Examples:

- Decision Trees are prone to overfitting, especially with complex datasets.
- They may not perform well with imbalanced datasets.

Working/Algorithm:::

1. Data Loading: Download the admission data and load it into a DataFrame.
2. Data Preprocessing: Perform label encoding if necessary and handle missing values.
3. Feature Selection: Select relevant features for the classifier.
4. Data Preparation: Split the data into training and testing sets.
5. Model Training: Initialize and train a Decision Tree classifier using the training data.
6. Model Evaluation: Evaluate the performance of the classifier using metrics such as accuracy, precision, recall, or F1-score.

Conclusion:

By applying a Decision Tree classifier to the admission data, we can predict whether a student will get admission based on their academic scores. Decision Trees offer a simple and interpretable solution for classification tasks. However, it's essential to evaluate the model's performance and consider potential limitations such as overfitting. Additionally, feature selection and data preprocessing play crucial roles in improving the accuracy and generalization of the model.