Data Driven Computing and Networking (DDCN-2019)

Classification using Decision Tree Algorithm

A. Write a python script to perform the following tasks:-

1. Load all the required packages to implement Decision Tree Classifier based algorithm
2. Import packages to load the datasets
3. Load the dataset "balance.csv"
4. Printing Dataset characterstics and five records
5. Slice data set to create feature set X by taking first second to five columns and target set Y as the first column.
6. Spliting feature set and target set both into trainingset and test set
7. Create Decision Tree Classifier using gini index
8. Train the created Decision Tree classifier model on training data set of feature set and target set
9. Make prediction using gini index based Decision Tree Classifier on a random new data [4, 4, 3, 3]
10. Make prediction of target using gini index based Decision Tree Classifier for test data set of feature set data
11. Print Predictions using Gini_index criteria
12. Print accuracy of gini index based Decision Tree Classifier for the test data set of target set
13. Create Decision Tree Classifier using entropy
14. Train Decision Tree classifier model on training data set of feature set and target set
15. Make prediction of target using information gain based Decision Tree Classifier for test data set of feature set data
16. Print Predictions using Entropy Measure
17. Print accuracy of information gain based Decision Tree Classifier for the target set