Date: 14/09/2022

Assignment No. 04

Title: Extracting the rules from the Decision tree

NAME: Saurabh Nagre

PRN: 2019BTECS00080

**Objective/Aim:** To perform following tasks:

1. Design the rule based classifier : Extract the rules from decision tree build in assignment no. 3.
2. Tabulate the results and evaluate the performance of rules generated using following metrics :

a. Coverage

b. Accuracy

c. Toughness (size)

1. Use the following categorical data sets from UCI machine learning repository : a. Balance Scale data set

b. Car evaluation data set

c. Breast-cancer data set

**Introduction:**

This assignment is about induction of Decision Trees from a given dataset. We divide the dataset into two parts, training dataset (80%) and testing dataset(20%). We put attributes of a given training dataset at different levels of a tree following a measure for attribute selection. Once the tree is built, we test its performance using the testing dataset. For evaluating the performance we make use of various metrics provided in the PS. Lastly we built a suitable menu driven GUI platform for the user to interact with this functionality. Based on the Decision tree produced in the assignment 3, we are generating the rules.

**Theory/Algorithm:**

Classification: classification refers to a predictive modeling problem where a class label is predicted for a given example of input data.

Decision Tree: A decision tree is a flowchart-like structure in which each internal node represents a "test" on an attribute (e.g. whether a coin flip comes up heads or tails), each branch represents the outcome of the test, and each leaf node represents a class label (decision taken after computing all attributes). The paths from root to leaf represent classification rules.

Decision trees are represented by internal nodes and external nodes

Internal Nodes: Rectangular External Nodes: Oval

Decision tree classifiers are a popular method of classification—it is easy to understand how decision trees work and they are known for their accuracy. Decision trees can

become large and difficult to interpret.

Rules are in the form of IF-THEN conditions.

To extract rules from a decision tree, one rule is created for each path from the root

to a leaf node.

Each splitting criterion along a given path is logically ANDed to form the

rule antecedent (“IF” part). The leaf node holds the class prediction, forming the rule

consequent (“THEN” part).

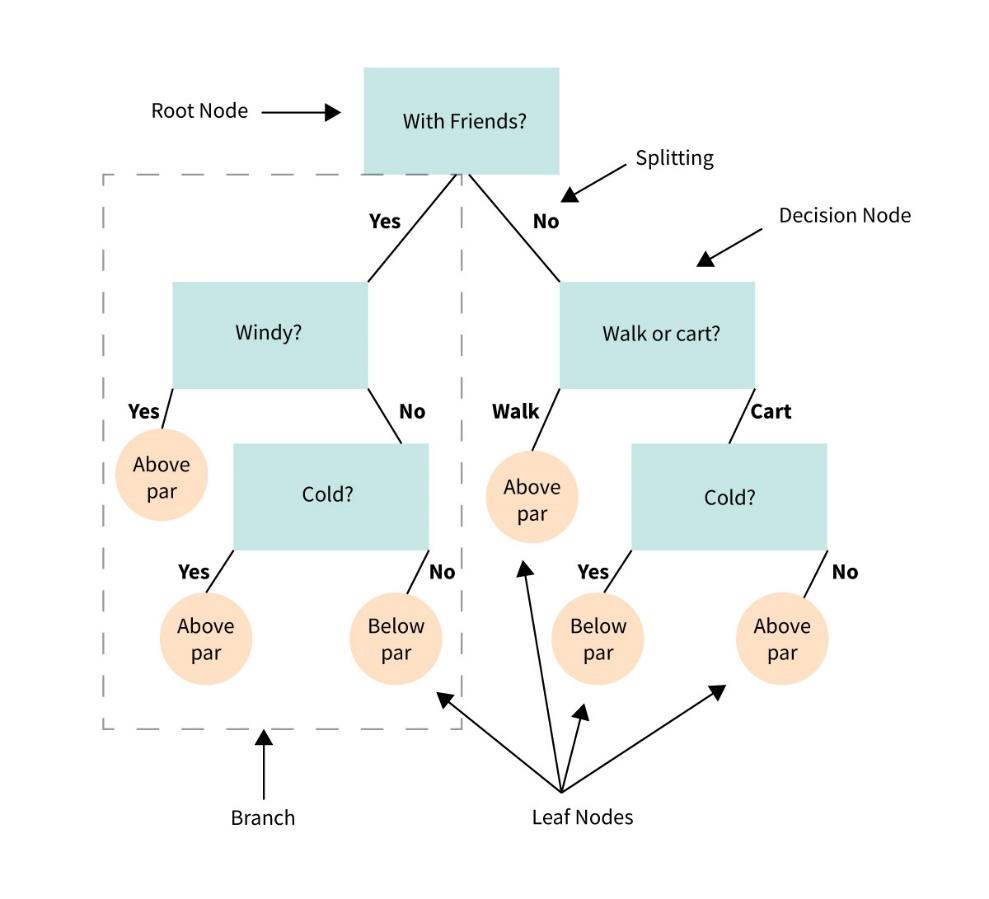
R1: IF age = youth AND student = no THEN buys computer = no

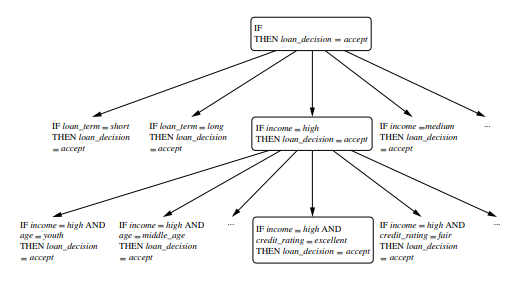
R2: IF age = youth AND student = yes THEN buys computer = yes

R3: IF age = middle aged THEN buys computer = yes

R4: IF age = senior AND credit rating = excellent THEN buys computer = yes

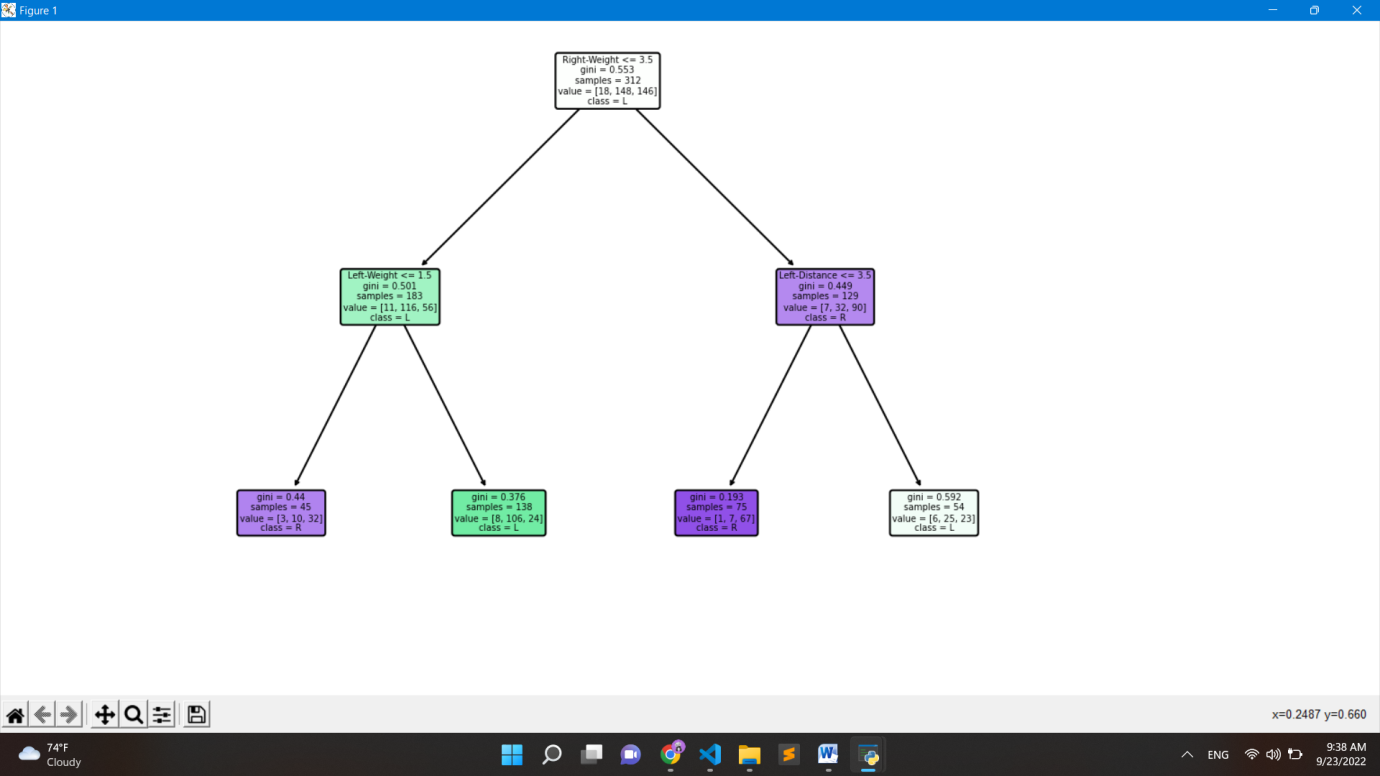
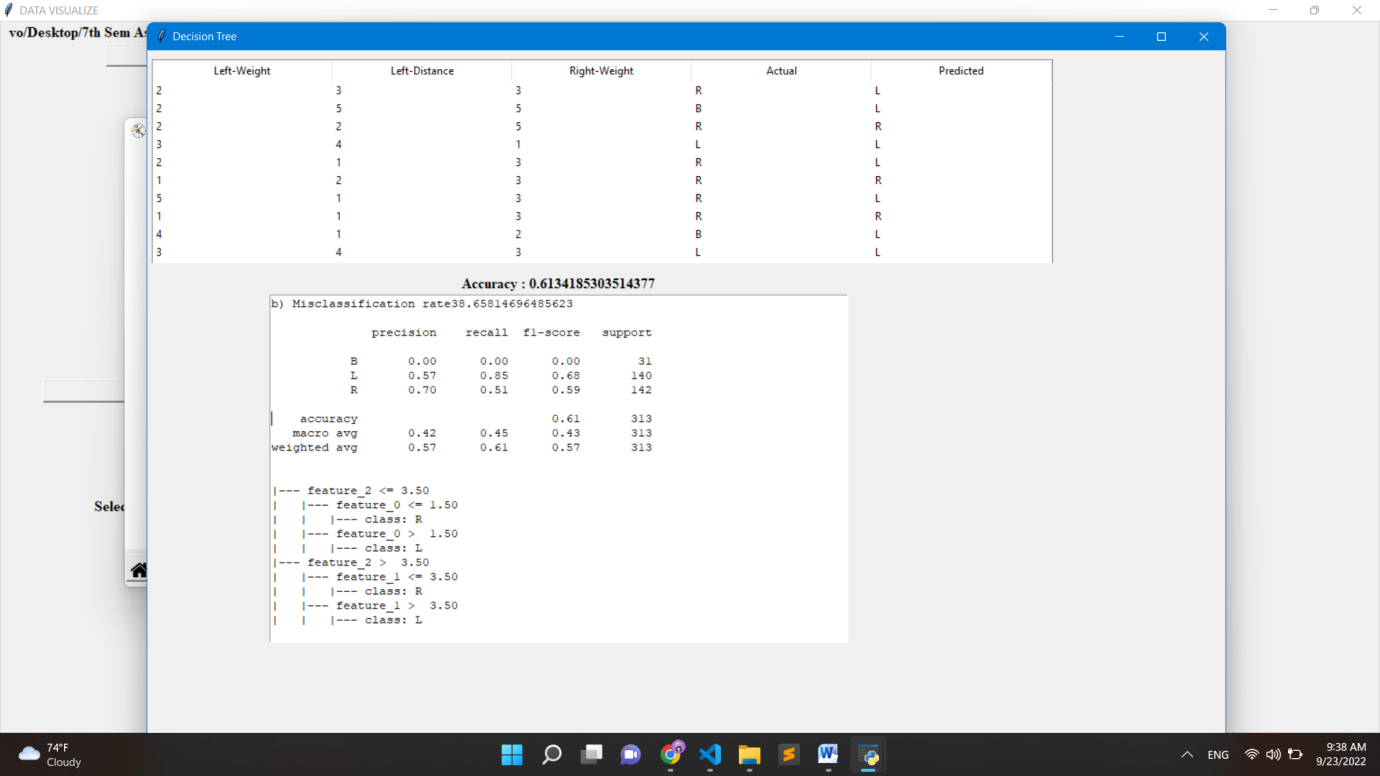
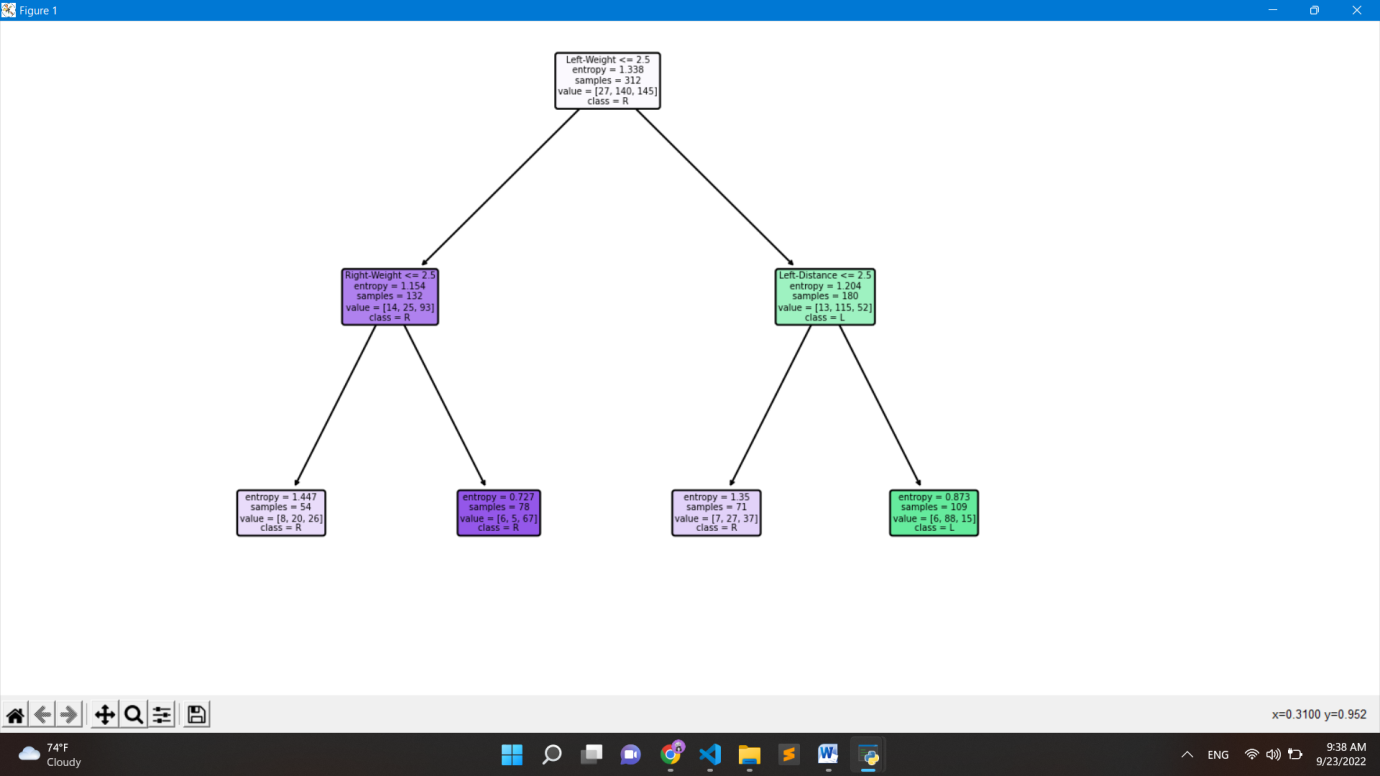
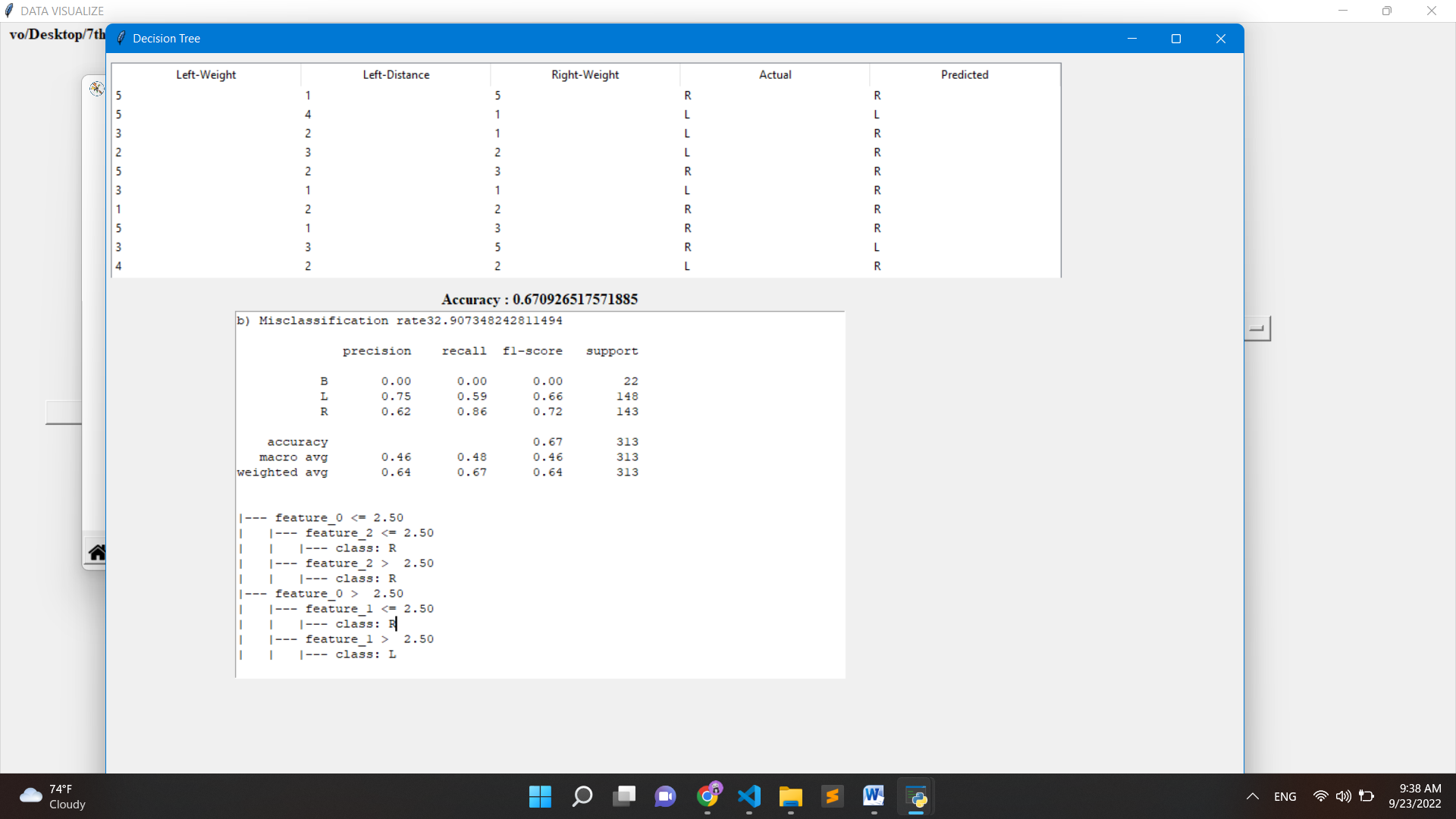
R5: IF age = senior AND credit rating = fair THEN buys computer = no





**Actual Experimentation/ simulation/ result/ Observation:**

The GUI:



**Conclusion:** In this assignment we learnt:

1. How to build a dynamic gui using tkinter
2. For to parse .csv files
3. Building Decision Tree using various attribute selection methods.
4. Rule Extraction from decision tree.

**References:**

1. <https://www.vedantu.com/commerce/gaining-ratio>
2. <https://towardsdatascience.com/>
3. <https://online.datasciencedojo.com/blogs/a-comprehensive-tutorial-on-classification-using-decision-trees>
4. <https://en.wikipedia.org/wiki/Decision_tree>