Machine Learning: Project 1: Decision Trees (Individual Submission)

Problem Statement:

We want to decide on whether to buy a car based on its attributes. Our decision can be one of the following four: unacceptable, acceptable, good, and very good. We want to use previous customer feedbacks to come to this decision. Write a program to learn a decision tree for this purpose. Decision tree learning should use information gain as the criterion for choosing the attribute for splitting. Tree pruning should not be performed. The tree should be tested on few test samples. The tree structure should be printed as output.

Data Set Description:
Training Data Filename: project1.data (csv file)
Test Data Filename: project1_test.data (csv file)
Training Data Description:
Class Values: unacc, acc, good, vgood (last column of data file)

Attributes and their values: (first six columns of data file in the following order)
price: vhigh, high, med, low.
maint: vhigh, high, med, low.
doors: 2, 3, 4, 5, 6.
persons: 2, 4, 6.
lug_boot: small, med, big.
safety: low, med, high.

The constructed tree structure should be printed as output. For output, you can choose how to draw the tree so long as it is clear what the tree is. Do not use any graphics package. You should use formatted text. You might find it easier to use indentation to show levels of the tree as it grows from the left.

For example:

```
safety = low
| maint = vhigh: unacc
| maint = high: unacc
| maint = med: acc
| maint = low: good
safety = med: vgood
safety = high
| lug_boot = small: unacc
| lug_boot = med: acc
| lug_boot = big: vgood
```

Submission Guidelines:

The program can be written c/C++/java/Python programming language from scratch. No machine learning/data science/statistics package/library should be used. Packages to read data may be used. Your program should be standalone and should not use any special purpose library.

You should name your file as <rollnumber_p1.extension> (e.g., 20CS10001_p1.c). You should submit the program file and not the output/input file. The submitted single program file should have the following header comments: # Roll # Name # Assignment number # Specific compilation/execution flags (if required). You should also submit a one-page readme about your submission.

Please submit the program in MS Teams by February 13, 2022 midnight (hard deadline). Copying from friends/web will lead to strict penalties.