The following 10 Steps were followed to complete this task:

Step 1: Implemented and tested the decision tree learning algorithm

Step 2: Downloaded the two datasets available here.

Given dataset description: Each dataset is divided into three sets: the training set, the validation set and the test set. Data sets are in CSV format. The first line in the file gives the attribute names. Each line after that is a training (or test) example that contains a list of attribute values separated by a comma. The last attribute is the class-variable. Assume that all attributes take values from the domain [0,1].

Step 3: Implemented the following two heuristics for selecting the next attribute:

- Information gain heuristic
- Variance impurity heuristic

Step 4: Implemented the post pruning algorithm as given below (Algorithm 1):

Algorithm 1: Post Pruning

```
Input: An integer L and an integer K
Output: A post-pruned Decision Tree
begin
    Build a decision tree using all the training data. Call it D;
    Let D_Best = D;
    for i = 1 to L do
        Copy the tree D into a new tree D';
        M = a random number between 1 and K;
        for j = 1 to M do
            Let N denote the number of non-leaf nodes in the decision
            tree D'. Order the nodes in D' from 1 to N;
            P = a random number between 1 and N;
            Replace the subtree rooted at P in D' by a leaf node.
            Assign the majority class of the subset of the data at P to
            the leaf node.;
            /* For instance, if the subset of the data at P
               contains 10 examples with class = 0 and 15
               examples with class = 1, replace P by class = 1
               */
        end
        Evaluate the accuracy of D' on the validation set;
        /* accuracy = percentage of correctly classified
           examples                                          */
        if D' is more accurate than D_Best then
            D_Best = D';
        end
    end
    return D_Best;
end
```

Step 5: Implemented a function in Python to print the decision tree to standard output.

Step 6: The written python code can be run with the following command (if you are using command line option):

python dtree.py <L> <K> <training-set> <validation-set> <test-set> <to-print>

Step 7: The above command will print the the accuracies on the test set for decision trees constructed using the two heuristics as well as the accuracies for their post-pruned versions for the given values of L and K. If to-print equals yes, it should print the decision tree in the format described in the problem statement.

Step 8: A "README.txt" file is included which provides all the instructions for compiling the code and A "RESULTS_REPORT.pdf" file showing the outputs for both the given data sets.

Step 9: In the "RESULTS_REPORT.pdf" file, report of the accuracy on the test set for decision trees constructed using the two heuristics is mentioned.

Step 10: In "RESULTS_REPORT.pdf" file, report of the accuracies for the post-pruned decision trees constructed using the two heuristics is given with 10 suitable values for L and K (not 10 values for each, just 10 combinations).