REPORT

QUESTION-1(Regression Trees)

Algorithm for building the tree:

```
build_tree(rows, attributes)
    if attributes is empty:
        return Node with value average(rows)

    root.attr, root.partitionVal=best_attribute(rows, attributes)

    if the best attr is numeric:

        partition rows, according to  root.partitionVal

    else:
        partition rows by value of  root.attr
        remove (root.attr) from attributes

    if left_rows is  not empty
        root.left = build_tree(left_rows, attributes.copy())
    else:
        root.left=Node(average(rows))

    if right_rows is not empty:
        root.right = build_tree(right_rows, attributes.copy())
    else:
        root.right=Node(average(rows))
    return root
```

Some examples including all relevant informations:

EXAMPLE-1

Best achievable test error = 4690.3935
Depth corresponding to this error = 4

Variation in error with depth while pruning:-

DEPTH	ERROR
4	4690.3935
2	3257.9247

The depths for which the model overfits are: 4

The pruned tree looks like:-

Level 0:- Size by Inch<10.5?
Level 1:- Leaf Node,Result=612.5 Extra Mushroom?
Level 2:- Leaf Node,Result=814.286 Leaf Node,Result=716.667

Representation of the pruned tree in If-Else form:-

```
if Size by Inch<10.5
  Price = 612.5
else
  if Extra Mushroom == True
    Price = 814.2857142857143
  else
    Price = 716.6666666666666
```

EXAMPLE-2
Best achievable test error = 7083.3333
Depth corresponding to this error = 7

Variation in error with depth while pruning:-

DEPTH	ERROR
7	7083.3333
6	6712.963
5	6712.963
4	6712.963

The depths for which the model overfits are: 7 6 5

The pruned tree looks like:-
Level 0:- Size by Inch<13.5?
Level 1:- Extra Spicy? Extra Spicy?
Level 2:- Extra Mushroom? Size by Inch<8.5? Leaf Node,Result=900.0 Leaf Node,Result=750.0
Level 3:- Leaf Node,Result=900.0 Leaf Node,Result=716.667 Leaf Node,Result=600.0 Extra Cheeze?
Level 4:- Leaf Node,Result=700.0 Leaf Node,Result=650.0

Representation of the pruned tree in If-Else form:-

```
if Size by Inch<13.5
  if Extra Spicy == True
    if Extra Mushroom == True
      Price = 900.0
```

```
      else
        Price = 716.6666666666666
    else
      if Size by Inch<8.5
        Price = 600.0
      else
        if Extra Cheeze == True
          Price = 700.0
        else
          Price = 650.0
  else
    if Extra Spicy == True
      Price = 900.0
    else
      Price = 750.0
```

EXAMPLE-3
Best achievable test error = 5578.7037
Depth corresponding to this error = 4

Variation in error with depth while pruning:-

DEPTH	ERROR
4	5578.7037
3	1695.3704

The depths for which the model overfits are: 4

The pruned tree looks like:-
Level 0:- Size by Inch<10.5?
Level 1:- Leaf Node,Result=610.0 Extra Mushroom?
Level 2:- Extra Cheeze? Leaf Node,Result=725.0
Level 3:- Leaf Node,Result=812.5 Leaf Node,Result=816.667

Representation of the pruned tree in If-Else form:-

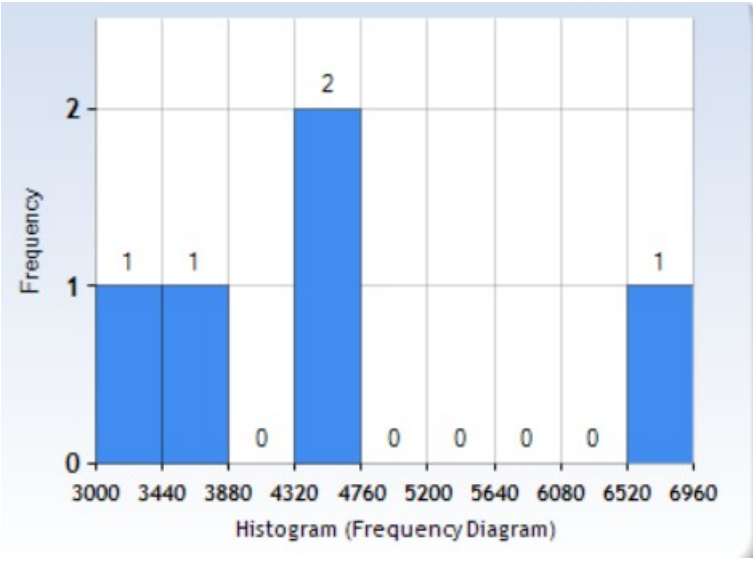
```
  if Size by Inch<10.5
    Price = 610.0
  else
    if Extra Mushroom == True
      if Extra Cheeze == True
        Price = 812.5
      else
```

Price = 816.6666666666666
else
Price = 725.0

Data conatinign tree depths vs error in both overfitting and pruned tree

OVERFITTING TREE DEPTH	OVERFITTING TREE ERROR
4	4690.3935
4	5578.7037
5	5416.6667
6	6712.963
7	7083.3333

PRUNED TREE DEPTH	PRUNED TREE ERROR
2	3257.9247
2	3807.8704
3	4375.0
3	4375.0
4	6712.963



Frequency of Error in pruned tree Error across some tests

QUESTION-2 (Naive Bayes)