CSPC 54 : AI/ML

Assignment 2 - WEKA

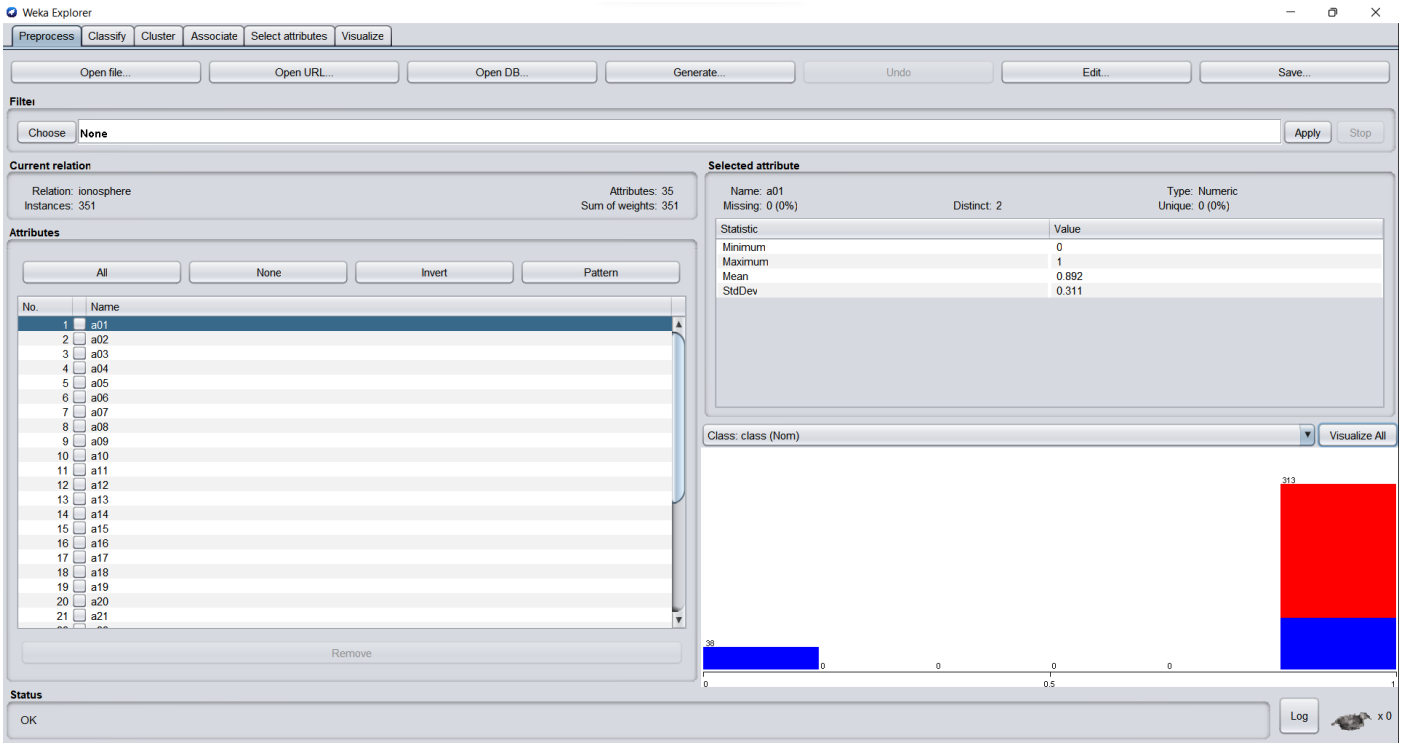
Name : Rajneesh Pandey,

Roll no. 106119100,

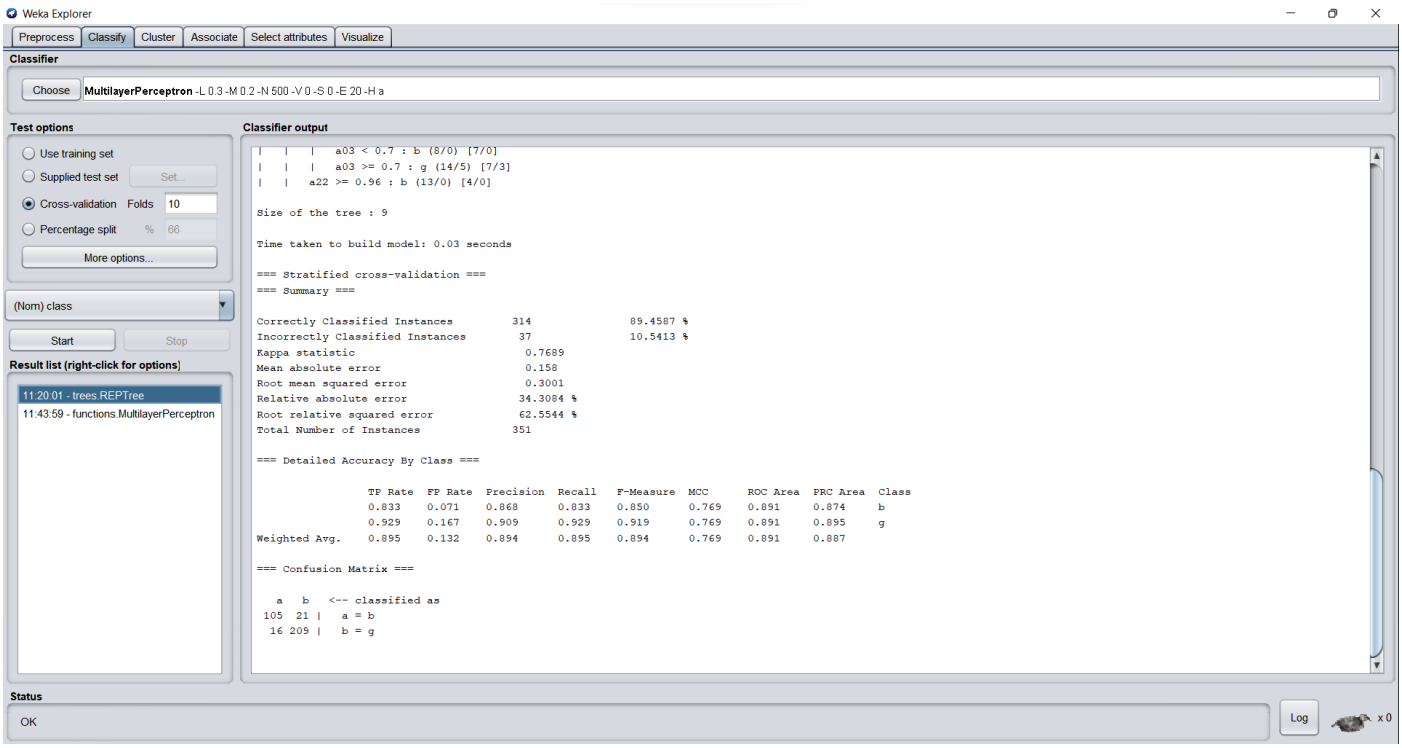
Class : CSE-B

Dataset Chosen : *ionosphere*

About Data Set : Each instance describes the properties of radar returns from the atmosphere and the task is to predict whether there is structure in the ionosphere. There are 34 numerical input variables of generally the same scale.



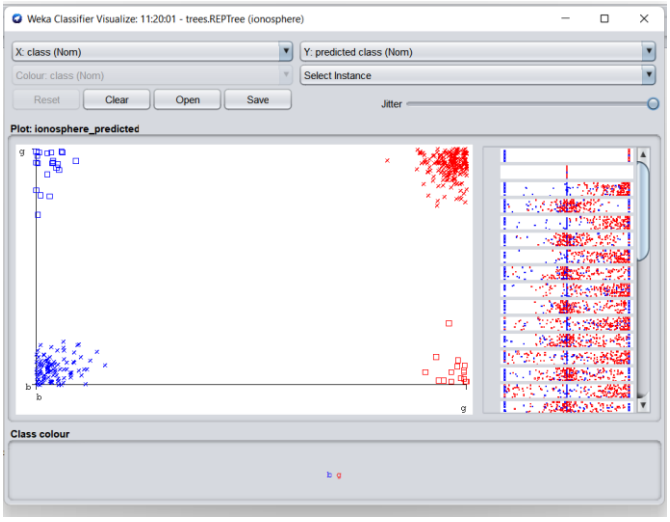
1. Decision Tree :



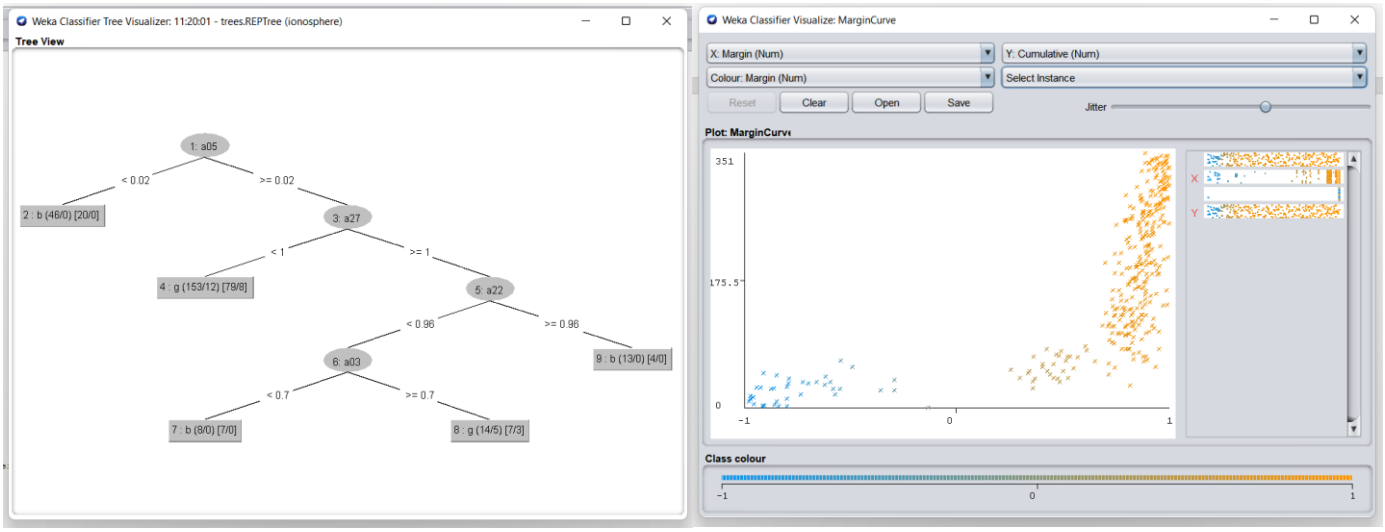

```
=== Run information ===

Scheme:       weka.classifiers.trees.REPTree -M 2 -V 0.001 -N 3 -S 1 -L -1 -I 0.0
Relation:     ionosphere
Instances:    351
Attributes:   35
              a01
              a02
              a03
              a04
              a05
              a06
              a07
              a08
              a09
              a10
              a11
              a12
              a13
              a14
              a15
              a16
              a17
              a18
              a19
              a20
              a21
              a22
              a23
              a24
              a25
              a26
              a27
              a28
```

```
              a29
              a30
              a31
              a32
              a33
              a34
              class
Test mode:    10-fold cross-validation
=== Classifier model (full training set) ===
REPTree
============
a05 < 0.02 : b (46/0) [20/0]
a05 >= 0.02
|   a27 < 1 : g (153/12) [79/8]
|   a27 >= 1
|   |   a22 < 0.96
|   |   |   a03 < 0.7 : b (8/0) [7/0]
|   |   |   a03 >= 0.7 : g (14/5) [7/3]
|   |   a22 >= 0.96 : b (13/0) [4/0]
Size of the tree : 9
Time taken to build model: 0.03 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances         314               89.4587 %
Incorrectly Classified Instances        37               10.5413 %
Kappa statistic                          0.7689
Mean absolute error                      0.158
Root mean squared error                  0.3001
Relative absolute error                 34.3084 %
Root relative squared error             62.5544 %
Total Number of Instances              351

=== Detailed Accuracy By Class ===
                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.833    0.071    0.868      0.833    0.850      0.769    0.891     0.874     b
                 0.929    0.167    0.909      0.929    0.919      0.769    0.891     0.895     g
Weighted Avg.    0.895    0.132    0.894      0.895    0.894      0.769    0.891     0.887

=== Confusion Matrix ===

   a   b   <-- classified as
 105  21 |   a = b
  16 209 |   b = g
```

Visualizing Classifier Errors :

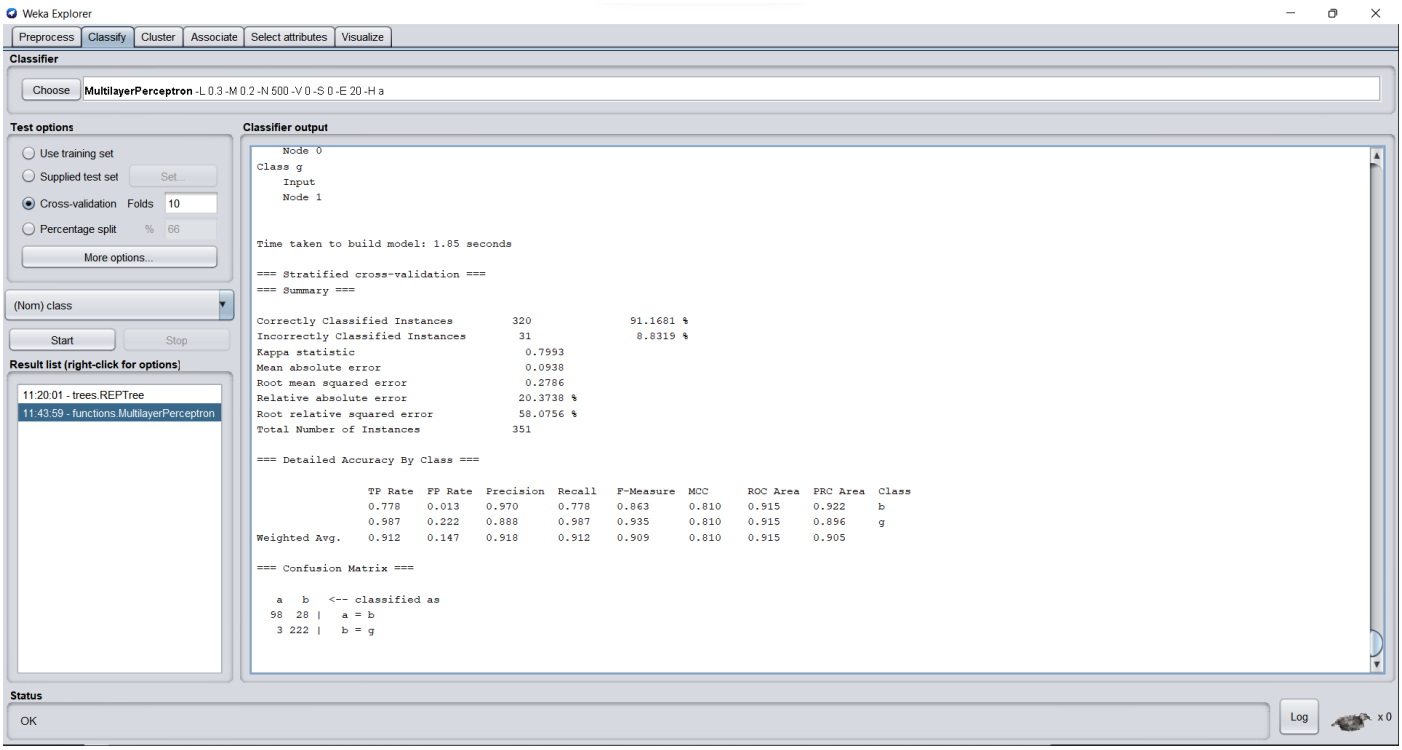


Visualizing Decision Tree: **Margin Curve**



So, the Overall Accuracy by **Decision Tree Algorithm : 89.45 %**

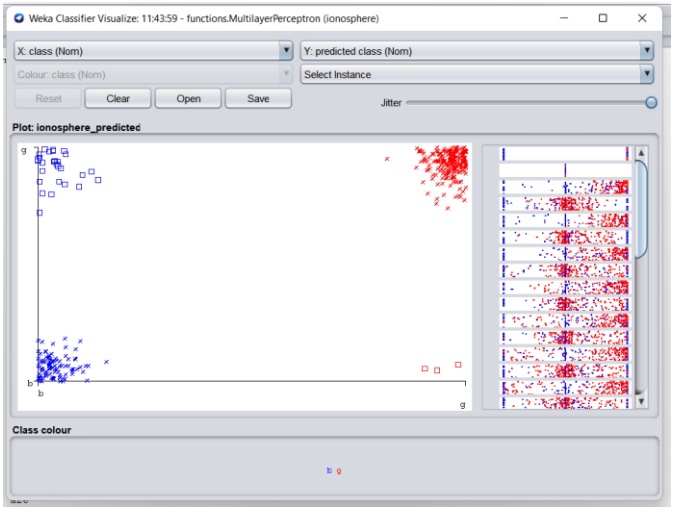
2. ANN (Artificial Neural Networks) :



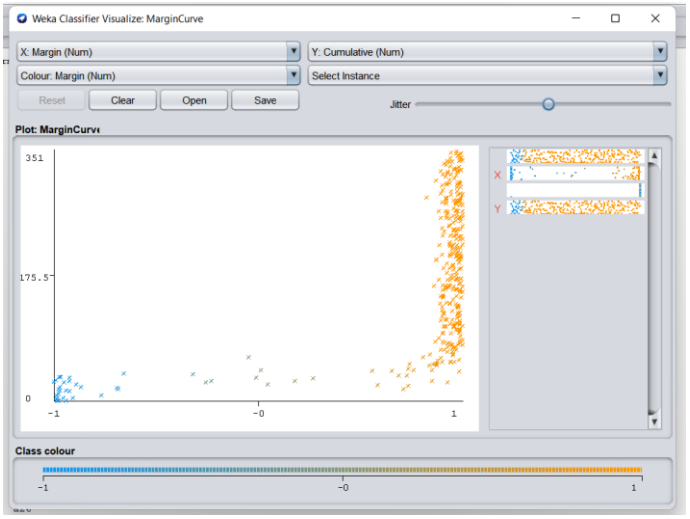
```
=== Run information ===

Scheme:       weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H a
Relation:     ionosphere
Instances:    351
Attributes:   35
              a01
              a02
              a03
              a04
              a05
              a06
              a07
              a08
              a09
              a10
              a11
              a12
              a13
              a14
              a15
              a16
              a17
              a18
              a19
              a20
              a21
              a22
              a23
              a24
              a25
              a26
              a27
              a28
              a29
              a30
              a31
              a32
              a33
              a34
              class
Test mode:    10-fold cross-validation
Class b
    Input
    Node 0
Class g
    Input
    Node 1
Time taken to build model: 1.85 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         320               91.1681 %
Incorrectly Classified Instances        31                8.8319 %
Kappa statistic                          0.7993
Mean absolute error                      0.0938
Root mean squared error                  0.2786
Relative absolute error                 20.3738 %
Root relative squared error             58.0756 %
Total Number of Instances              351

=== Detailed Accuracy By Class ===
```

```
                 TP Rate   FP Rate   Precision   Recall   F-Measure   MCC      ROC Area  PRC Area  Class
                 0.778     0.013     0.970       0.778    0.863       0.810    0.915     0.922     b
                 0.987     0.222     0.888       0.987    0.935       0.810    0.915     0.896     g
Weighted Avg.    0.912     0.147     0.918       0.912    0.909       0.810    0.915     0.905

=== Confusion Matrix ===

   a   b   <-- classified as
  98  28 |   a = b
   3 222 |   b = g
```

Visualizing Classifier Errors :



Margin Curve :



So, the Overall Accuracy by **Decision Tree Algorithm : 91.1681 %** *which is higher than the Decision tree.*