Construct Decision Tree for Weather data and classify it using WEKA

Classification is the process for finding a model that describes the data values and concepts for the purpose of Prediction.

Decision Tree:

A decision Tree is a classification scheme to generate a tree consisting of root node, internal nodes and external nodes. Root nodes representing the attributes. Internal nodes are also the attributes. External nodes are the classes and each branch represents the values of the attributes. Decision Tree also contains set of rules for a given data set; there are two subsets in Decision Tree. One is a Training data set and second one is a Testing data set. Training data set is previously classified data. Testing data set is newly generated data.

1.Draw Decision Tree using weather Dataset.

Procedure for Decision Trees:

1) Open Start → Programs → Weka-3-4 → Weka-3-4

2) Open **explorer**.

3) Click on **open file** and select **weather.arff**

4) Select **Classifier option** on the top of the Menu bar.

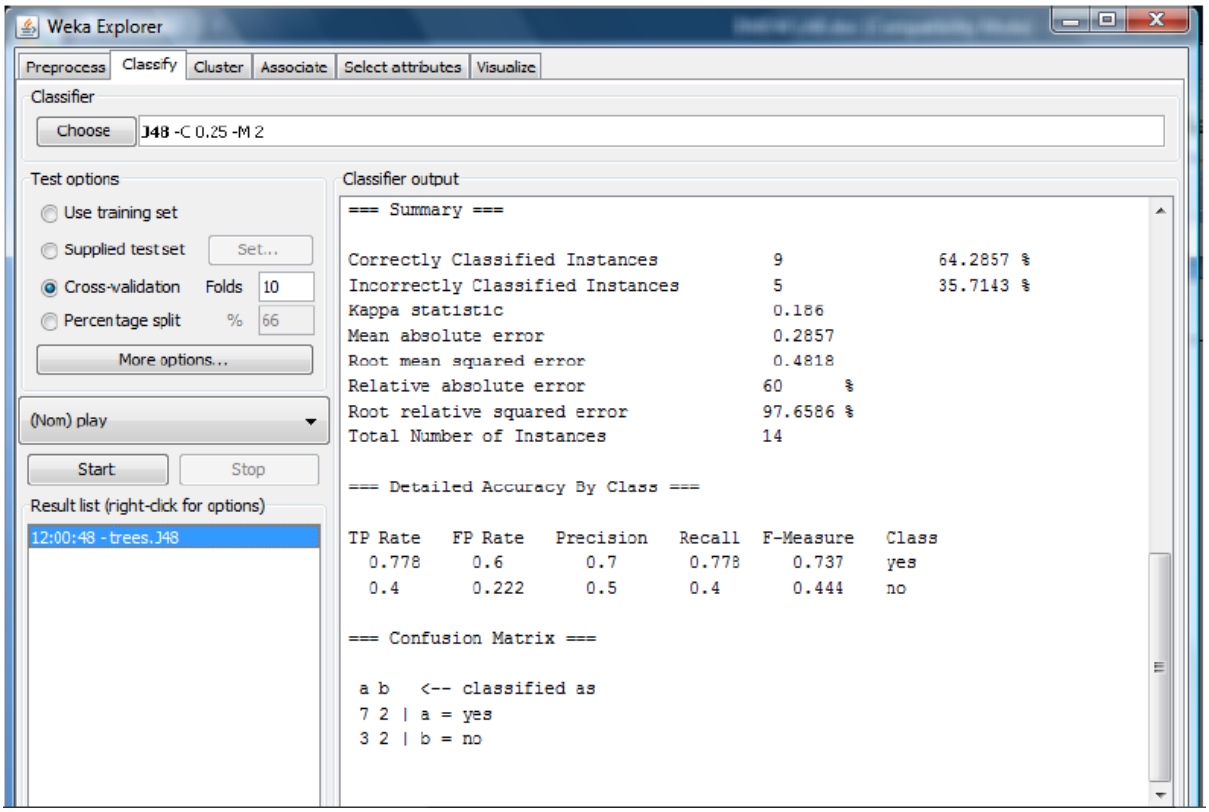
5) Select **Choose button** and click on **Tree option**.

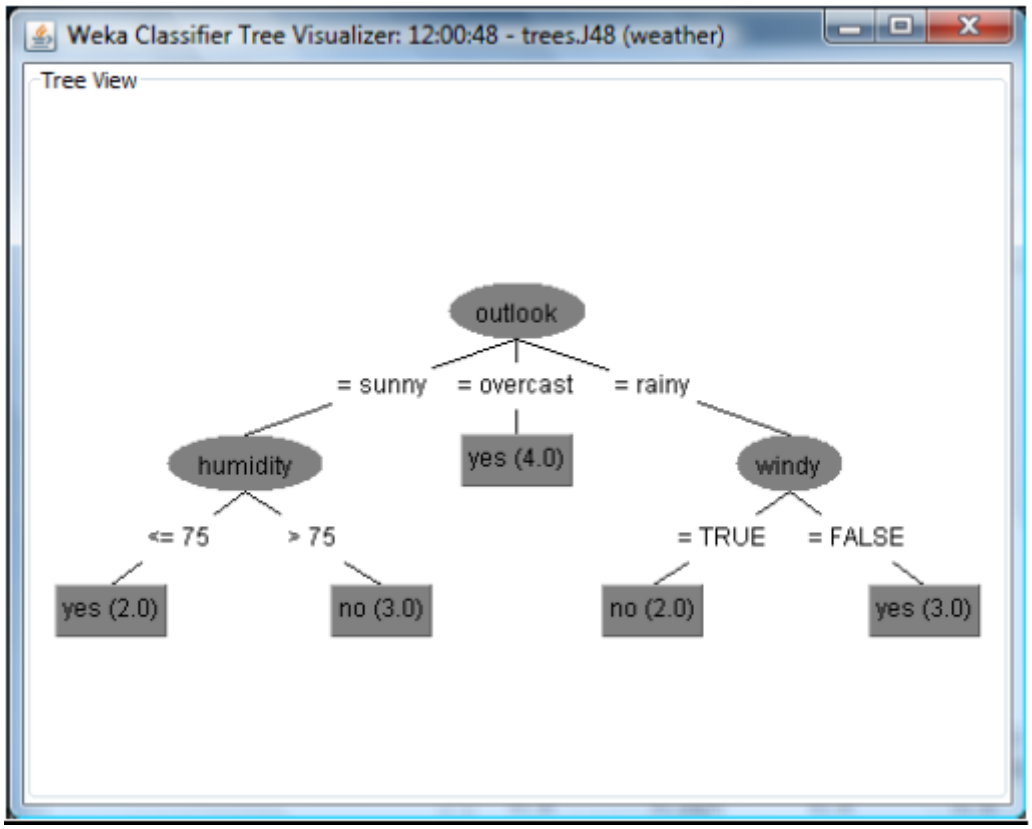
6) Click on **J48.**

7) Click on **Start button** and output will be displayed on the **right side** of the window.

8) Select the **result list** and **right click** on result list and select **Visualize Tree option**.

9) Then **Decision Tree** will be displayed on **new window**.


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
J48 -C 0.25 -M 2
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) play
Start
Stop
Result list (right-click for options)
12:00:48 - trees.J48
Classifier output
=== Summary ===
Correctly Classified Instances 9 64.2857 %
Incorrectly Classified Instances 5 35.7143 %
Kappa statistic 0.186
Mean absolute error 0.2857
Root mean squared error 0.4818
Relative absolute error 60 %
Root relative squared error 97.6586 %
Total Number of Instances 14
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure Class
0.778 0.6 0.7 0.778 0.737 yes
0.4 0.222 0.5 0.4 0.444 no
=== Confusion Matrix ===
a b <-- classified as
7 2 | a = yes
3 2 | b = no


Weka Classifier Tree Visualizer: 12:00:48 - trees.J48 (weather)
Tree View
outlook
= sunny
= overcast
= rainy
humidity
yes (4.0)
windy
<= 75
> 75
= TRUE
= FALSE
yes (2.0)
no (3.0)
no (2.0)
yes (3.0)

1. Draw Decision Tree using following dataset Dataset.

RID	age	income	student	credit_rating	Class: buys_computer
1	youth	high	no	fair	no
2	youth	high	no	excellent	no
3	middle_aged	high	no	fair	yes
4	senior	medium	no	fair	yes
5	senior	low	yes	fair	yes
6	senior	low	yes	excellent	no
7	middle_aged	low	yes	excellent	yes
8	youth	medium	no	fair	no
9	youth	low	yes	fair	yes
10	senior	medium	yes	fair	yes
11	youth	medium	yes	excellent	yes
12	middle_aged	medium	no	excellent	yes
13	middle_aged	high	yes	fair	yes
14	senior	medium	no	excellent	no

3. Construct the Decision Tree using the following customer dataset

@relation customer
@attribute name {x,y,z,u,v,l,w,q,r,n}
@attribute age {youth,middle,senior}
@attribute income {high,medium,low}
@attribute class {A,B}
@data
x,youth,high,A
y,youth,low,B
z,middle,high,A
u,middle,low,B
v,senior,high,A
l,senior,low,B
w,youth,high,A
q,youth,low,B
r,middle,high,A
n,senior,high,A

Other Dataset : Location

@relation location
@attribute age {21,24,25}
@attribute location {hyd,blr,kdp}
@data
21,hyd
21,hyd
24,blr

24,blr
24,blr
24,blr
21,hyd
25,kdp
25,kdp
25,kdp

^^^^^^^^^^^^^^^^^^^^^^