

### **AIM:**

Construct a Decision Tree for any data set of your choice and classify it.

### **DESCRIPTION:**

#### **Classification & Prediction:**

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.

### **CONSTRUCTION:**

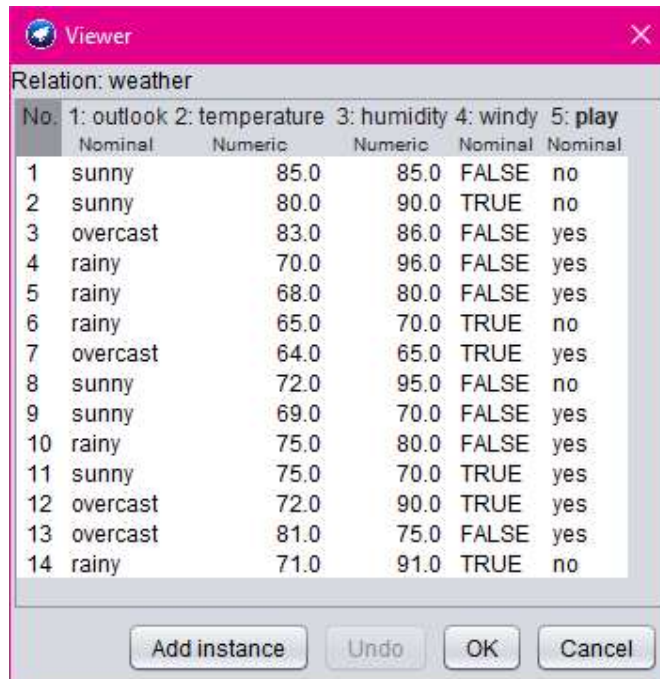
A tree can be “*learned*” by splitting the source set into subsets based on an attribute value test. This process is repeated on each derived subset in a recursive manner called *recursive partitioning*. The recursion is completed when the subset at a node all has the same value of the target variable, or when splitting no longer adds value to the predictions. The construction of decision tree classifier does not require any domain knowledge or parameter setting, and therefore is appropriate for exploratory knowledge discovery. Decision trees can handle high dimensional data. In general decision tree classifier has good accuracy. Decision tree induction is a typical inductive approach to learn knowledge on classification.

### **PROCEDURE (WEKA):**

1. Open Start -> Programs -> Weka-3-4 -> Weka-3-4
2. Open **Explorer**.

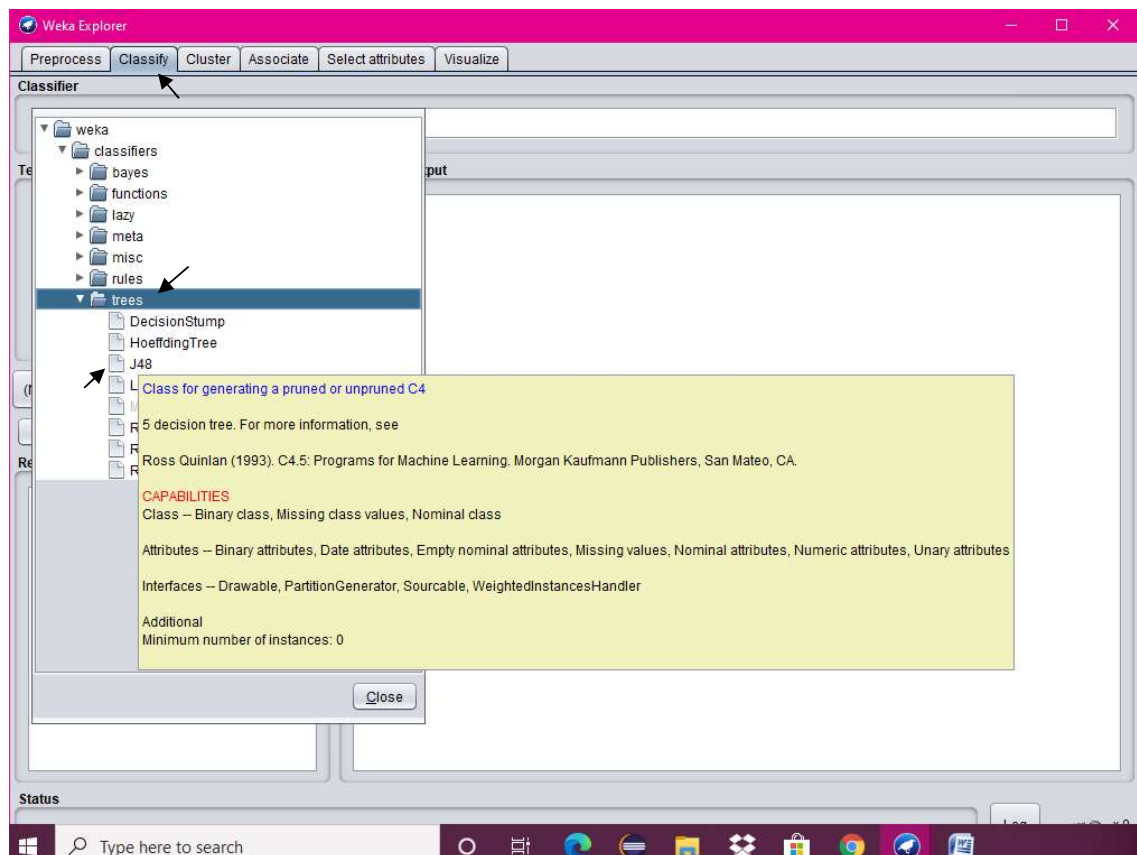


3. Click on **open file** and select **weather.arff**



No.	1: outlook	2: temperature	3: humidity	4: windy	5: play
	Nominal	Numeric	Numeric	Nominal	Nominal
1	sunny	85.0	85.0	FALSE	no
2	sunny	80.0	90.0	TRUE	no
3	overcast	83.0	86.0	FALSE	yes
4	rainy	70.0	96.0	FALSE	yes
5	rainy	68.0	80.0	FALSE	yes
6	rainy	65.0	70.0	TRUE	no
7	overcast	64.0	65.0	TRUE	yes
8	sunny	72.0	95.0	FALSE	no
9	sunny	69.0	70.0	FALSE	yes
10	rainy	75.0	80.0	FALSE	yes
11	sunny	75.0	70.0	TRUE	yes
12	overcast	72.0	90.0	TRUE	yes
13	overcast	81.0	75.0	FALSE	yes
14	rainy	71.0	91.0	TRUE	no

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.

The screenshot shows the Weka Explorer Classifier window. The 'Test options' section on the left has 'Cross-validation' selected with 'Folds' set to 10. The 'Start' button is highlighted with a mouse cursor. The 'Classifier output' section on the right displays the following information:

Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===

=== Summary ===

Metric	Value	Percentage
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	97.6596	%
Total Number of Instances	14	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
Weighted Avg.	0.778	0.600	0.700	0.778	0.737	0.189	0.789	0.847	yes
	0.400	0.222	0.500	0.400	0.444	0.189	0.789	0.738	no

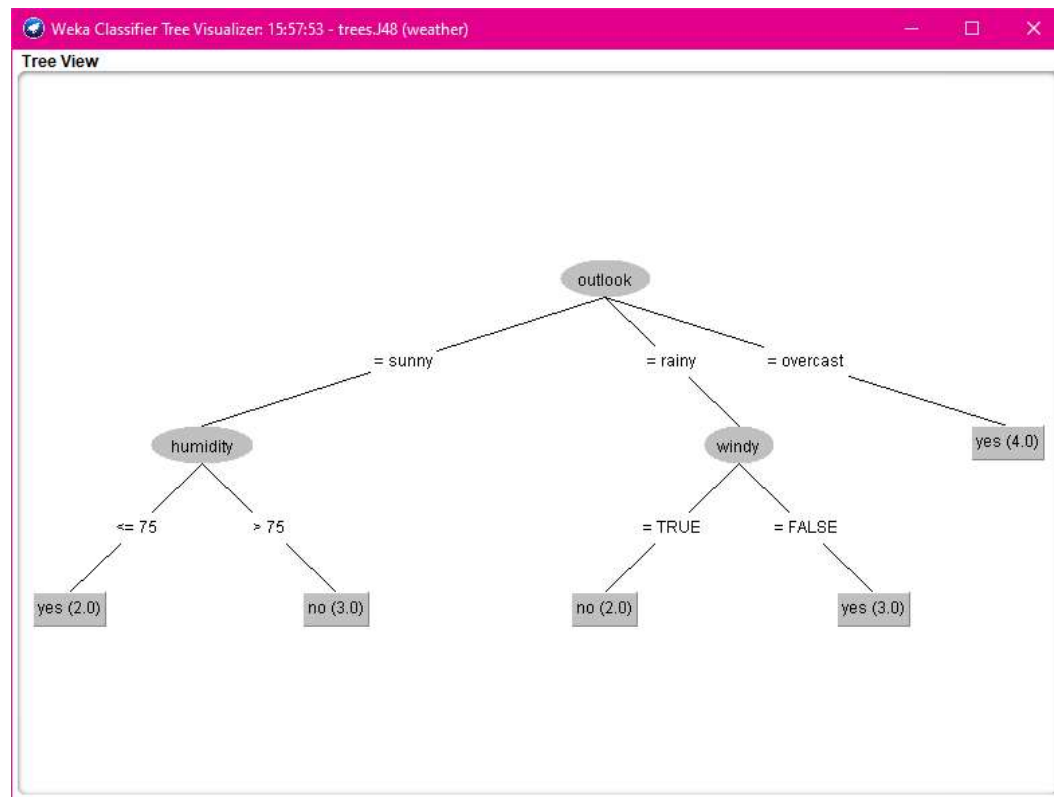
=== Confusion Matrix ===

a b <-- classified as

	a = yes	b = no
7 2		
3 2		

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**.



Submitted By-  
Kavita Maurya  
2016UCO1579