**EXPERIMENT NO:11**

**Aim:**

To Construct Decision Tree for Customer data 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.

# Creation of Customer Table:

**Procedure:**

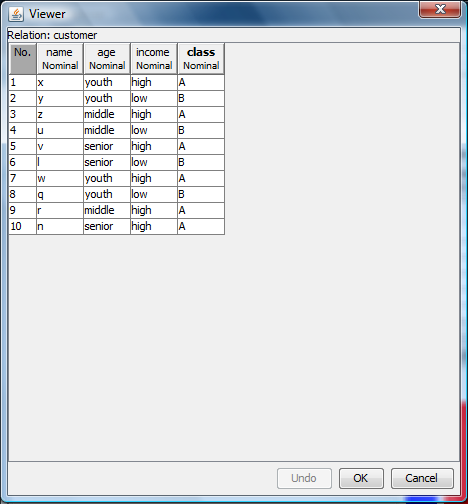
1. Open Start  Programs  Accessories  Notepad
2. Type the following training data set with the help of Notepad for Customer Table. @relation customer

@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

1. After that the file is saved with **.arff** file format.
2. Minimize the arff file and then open Start  Programs  weka-3-4.
3. Click on **weka-3-4**, then Weka dialog box is displayed on the screen.
4. In that dialog box there are four modes, click on **explorer**.
5. Explorer shows many options. In that click on **‘open file’** and select the arff file
6. Click on **edit button** which shows customer table on weka.

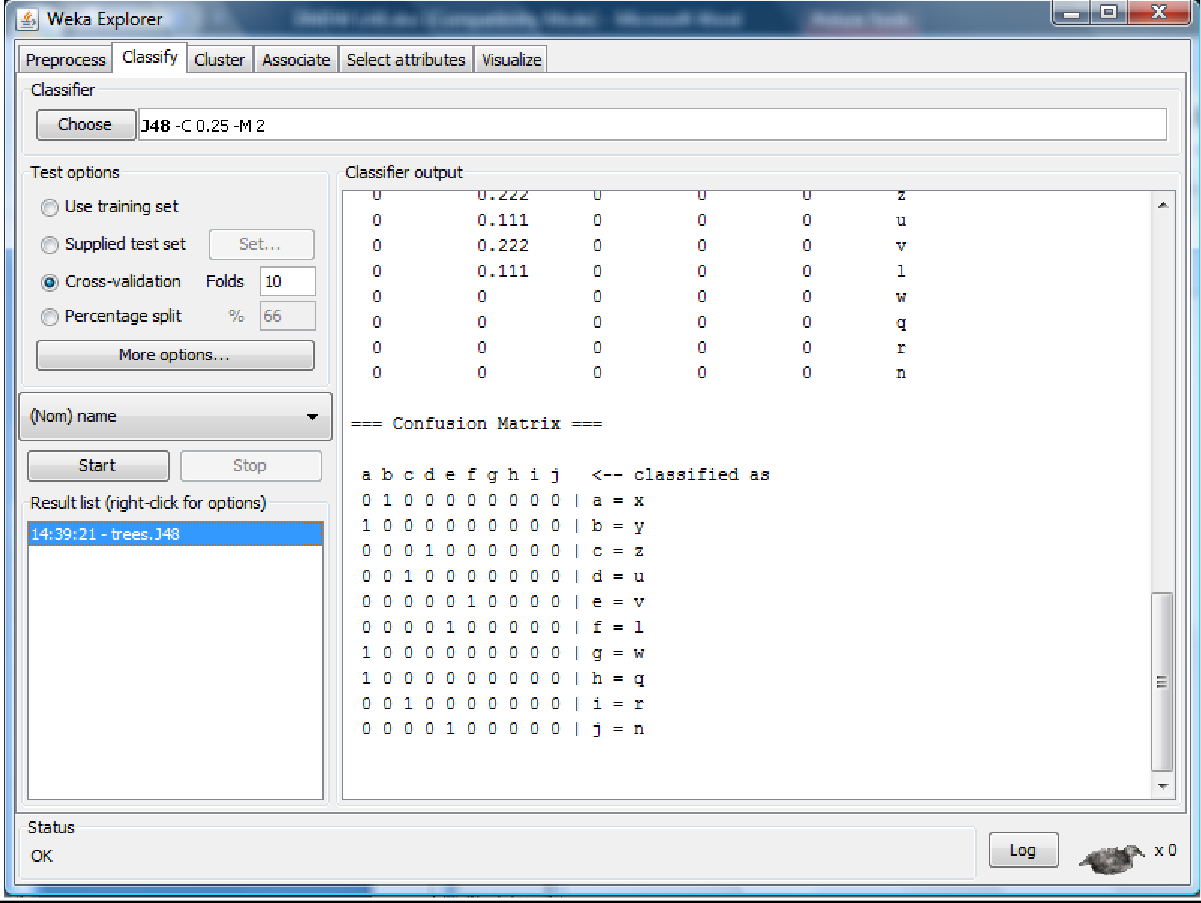
# Training Data Set  Customer Table



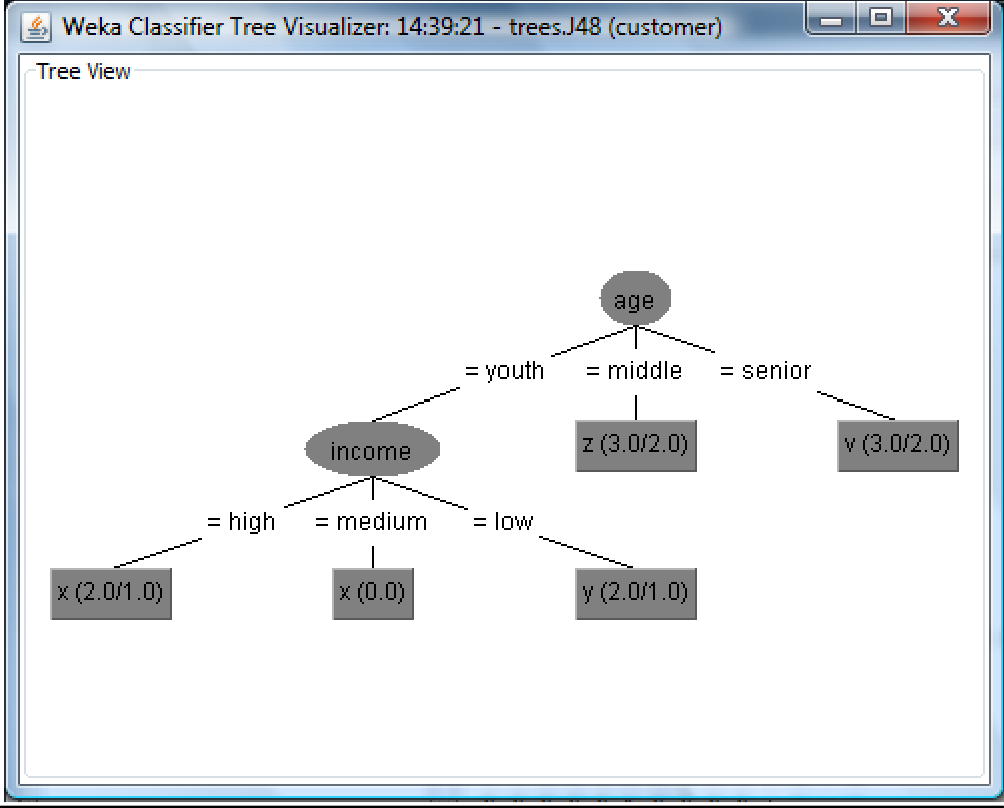
**Procedure for Decision Trees:**

* 1. Open Start  Programs  Weka-3-4  Weka-3-4
  2. Open **explorer**.
  3. Click on **open file** and select **customer.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**.

# Output:



**Decision Tree:**



**Result:** This program has been successfully executed.

**EXPERIMENT NO:12**

**Aim:**

To Construct Decision Tree for Location data 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.

# Creation of Weather Table:

**Procedure:**

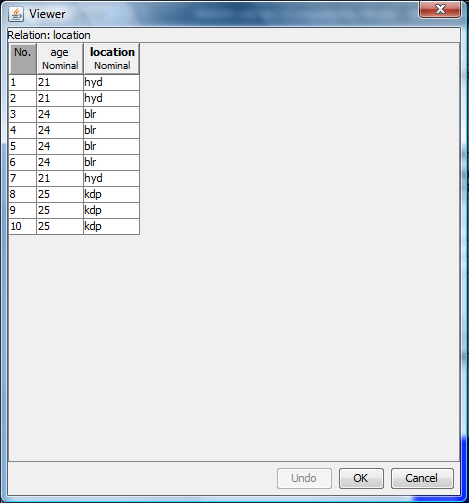
1. Open Start  Programs  Accessories  Notepad
2. Type the following training data set with the help of Notepad for Location Table. @relation location

@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

1. After that the file is saved with **.arff** file format.
2. Minimize the arff file and then open Start  Programs  weka-3-4.
3. Click on **weka-3-4**, then Weka dialog box is displayed on the screen.
4. In that dialog box there are four modes, click on **explorer**.
5. Explorer shows many options. In that click on **‘open file’** and select the arff file
6. Click on **edit button** which shows location table on weka.

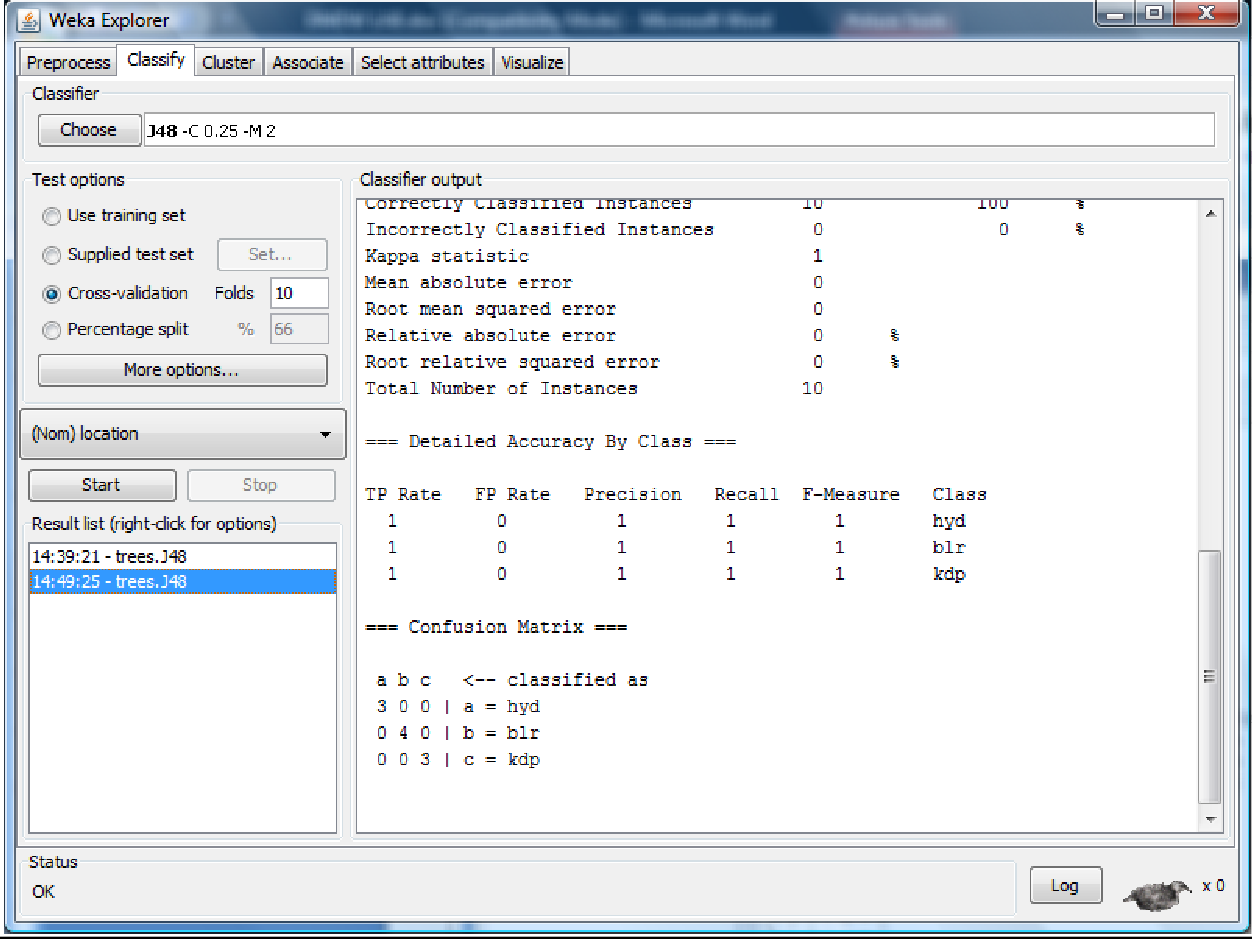
# Training Data Set  Location Table



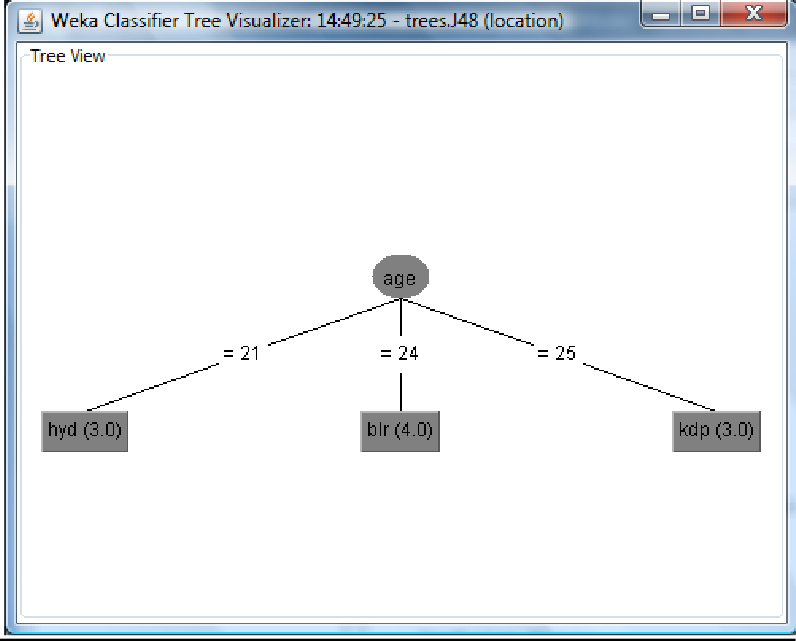
**Procedure for Decision Trees:**

1. Open Start  Programs  Weka-3-4  Weka-3-4
2. Open **explorer**.
3. Click on **open file** and select **location.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**.

# Output:



**Decision Tree:**



# Result:

This program has been successfully executed.

**EXPERIMENT NO:13**

**Aim:**

Write a procedure for Visualization for Weather Table.

# Description:

This program calculates and has comparisons on the data set selection of attributes and methods of manipulations have been chosen. The Visualization can be shown in a 2-D representation of the information.

# Creation of Weather Table:

**Procedure:**

1. Open Start  Programs  Accessories  Notepad
2. Type the following training data set with the help of Notepad for Weather Table. @relation weather

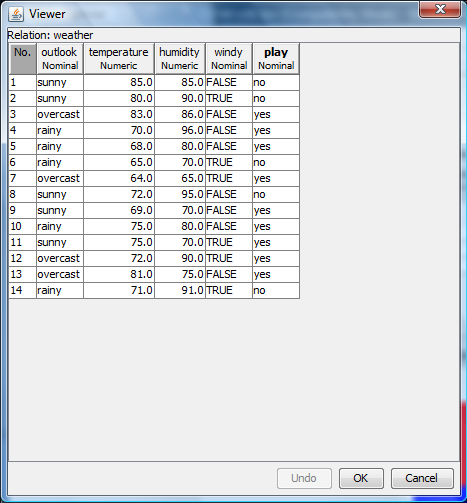
@attribute outlook {sunny, rainy, overcast} @attribute temperature numeric

@attribute humidity numeric @attribute windy {TRUE, FALSE} @attribute play {yes, no}

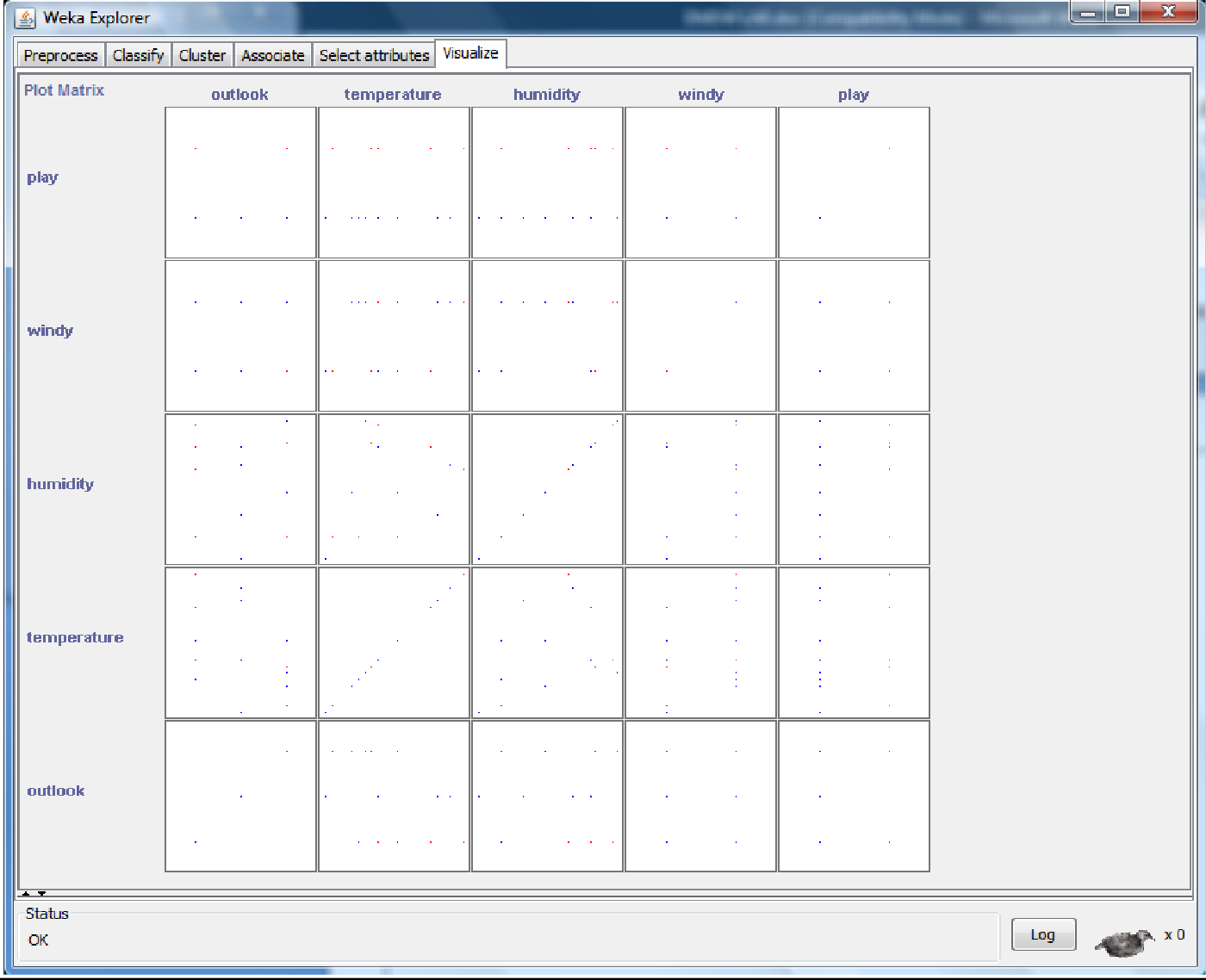
@data sunny,85,85,FALSE,no sunny,80,90,TRUE,no overcast,83,86,FALSE,yes rainy,70,96,FALSE,yes rainy,68,80,FALSE,yes rainy,65,70,TRUE,no overcast,64,65,TRUE,yes sunny,72,95,FALSE,no sunny,69,70,FALSE,yes rainy,75,80,FALSE,yes sunny,75,70,TRUE,yes overcast,72,90,TRUE,yes overcast,81,75,FALSE,yes rainy,71,91,TRUE,no

1. After that the file is saved with **.arff** file format.
2. Minimize the arff file and then open Start  Programs  weka-3-4.
3. Click on **weka-3-4**, then Weka dialog box is displayed on the screen.
4. In that dialog box there are four modes, click on **explorer**.
5. Explorer shows many options. In that click on **‘open file’** and select the arff file
6. Click on **edit button** which shows weather table on weka.

# Training Data Set  Weather Table



* 1. **Plot Matrix:**



# Procedure:

* + 1. Open Start  Programs  Weka-3-4  Weka-3-4
    2. Open the explorer and click on **Preprocess**, then a new window will appear. In that window select

**weather.arff** file then the data will be displayed.

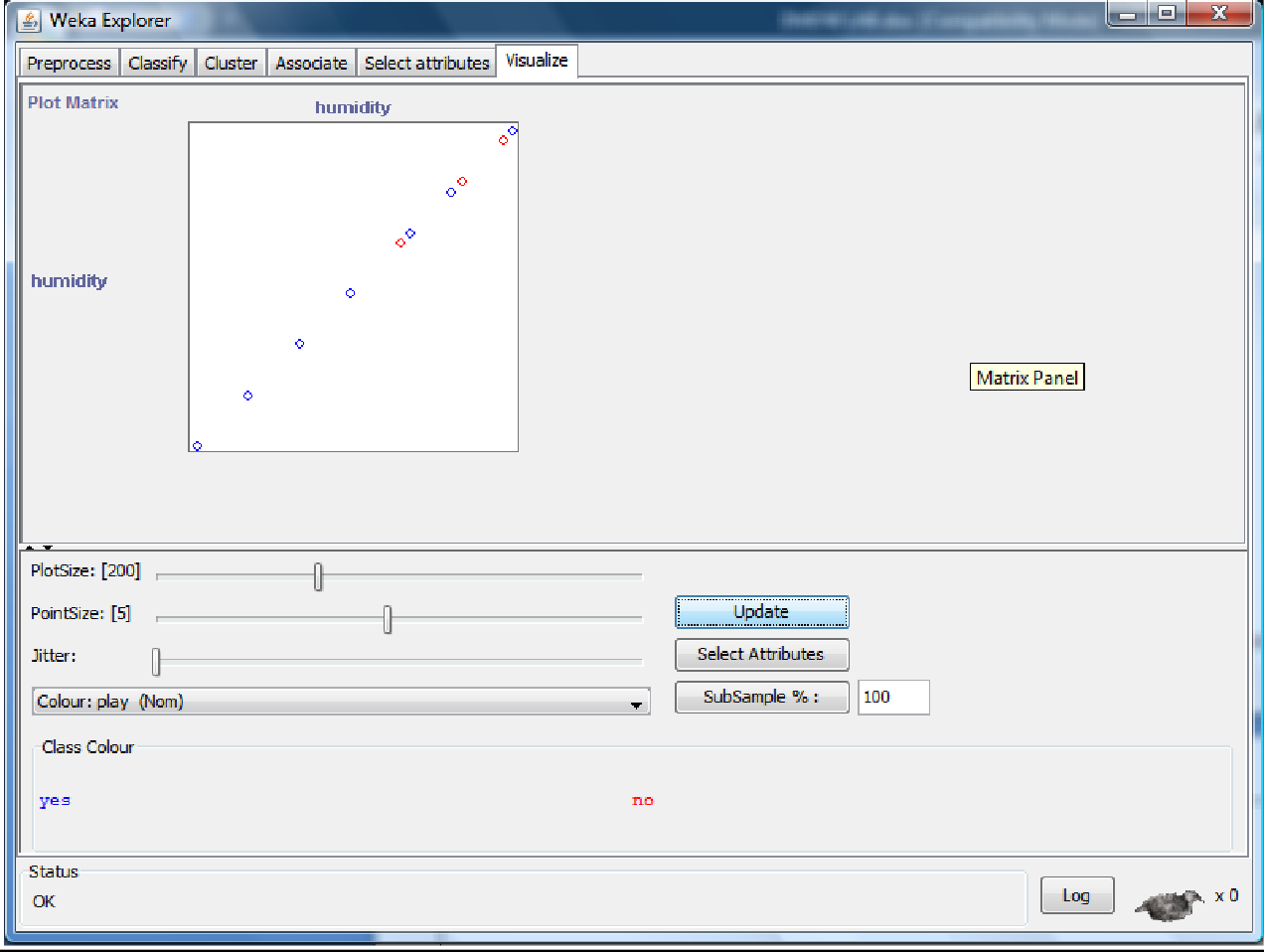
* + 1. After that click on the **Visualize tab** on the top of the Menu bar.
    2. When we select **Visualize tab** then **Plot Matrix** is displayed on the screen.

# Output:

* + 1. After that we select the **Select Attribute button**, then select **Outlook attribute** and clock OK.
    2. Click on the **Update button** to display the output.
    3. After that select the **Select Attribute button** and select **Temperature attribute** and then click OK.
    4. **Increase** the **Plot Size** and **Point Size**.
    5. Click on the **Update button** to display the output.
    6. After that we select the **Select Attribute button**, then select **Humidity attribute** and clock OK.
    7. Click on the **Update button** to display the output.
    8. After that select the **Select Attribute button** and select **Windy attribute** and then click OK.
    9. **Increase** the **Jitter Size**.
    10. Click on the **Update button** to display the output.
    11. After that we select the **Select Attribute button**, then select **Play attribute** and clock OK.
    12. Click on the **Update button** to display the output.

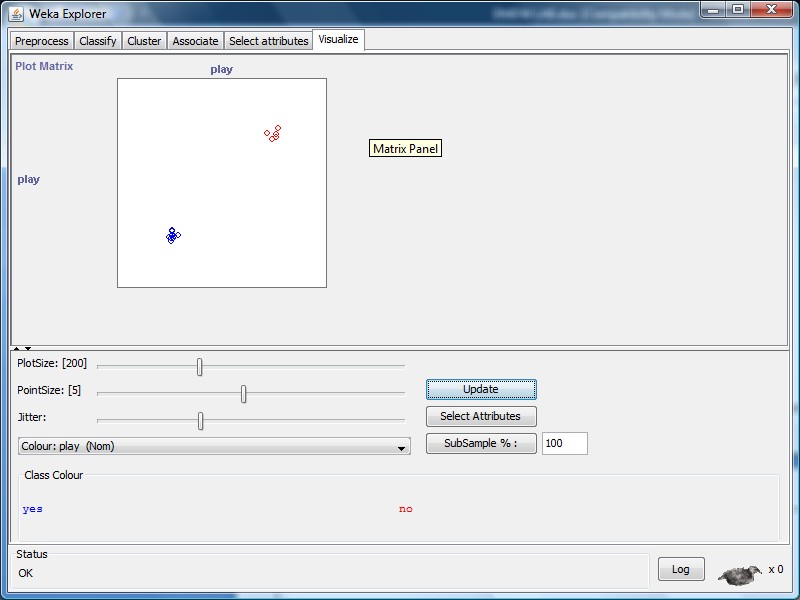
# Output:

**Output:**



# Output:

**Output:**



# Result:

This program has been successfully executed.

**EXPERIMENT NO:14**

**Aim:**

Write a procedure for Visualization of Banking Table.

# Description:

This program calculates and has comparisons on the data set selection of attributes and methods of manipulations have been chosen. The Visualization can be shown in a 2-D representation of the information.

# Creation of Banking Table:

**Procedure:**

* + - 1. Open Start  Programs  Accessories  Notepad
      2. Type the following training data set with the help of Notepad for Banking Table. @relation bank

@attribute cust {male,female} @attribute accno

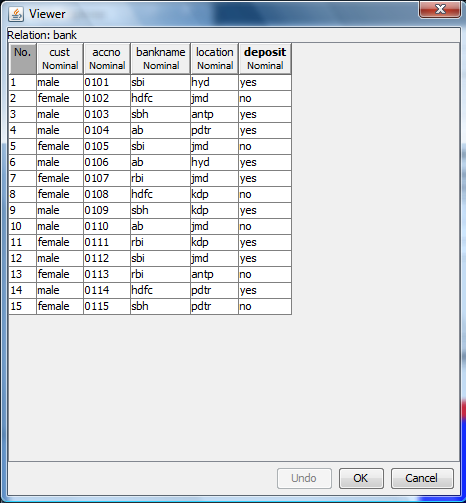
{0101,0102,0103,0104,0105,0106,0107,0108,0109,0110,0111,0112,0113,0114,0115}

@attribute bankname {sbi,hdfc,sbh,ab,rbi} @attribute location {hyd,jmd,antp,pdtr,kdp} @attribute deposit {yes,no}

@data male,0101,sbi,hyd,yes female,0102,hdfc,jmd,no male,0103,sbh,antp,yes male,0104,ab,pdtr,yes female,0105,sbi,jmd,no male,0106,ab,hyd,yes female,0107,rbi,jmd,yes female,0108,hdfc,kdp,no male,0109,sbh,kdp,yes male,0110,ab,jmd,no female,0111,rbi,kdp,yes male,0112,sbi,jmd,yes female,0113,rbi,antp,no male,0114,hdfc,pdtr,yes female,0115,sbh,pdtr,no

* + - 1. After that the file is saved with **.arff** file format.
      2. Minimize the arff file and then open Start  Programs  weka-3-4.
      3. Click on **weka-3-4**, then Weka dialog box is displayed on the screen.
      4. In that dialog box there are four modes, click on **explorer**.
      5. Explorer shows many options. In that click on **‘open file’** and select the arff file
      6. Click on **edit button** which shows Banking table on weka.

# Training Data Set  Banking Table



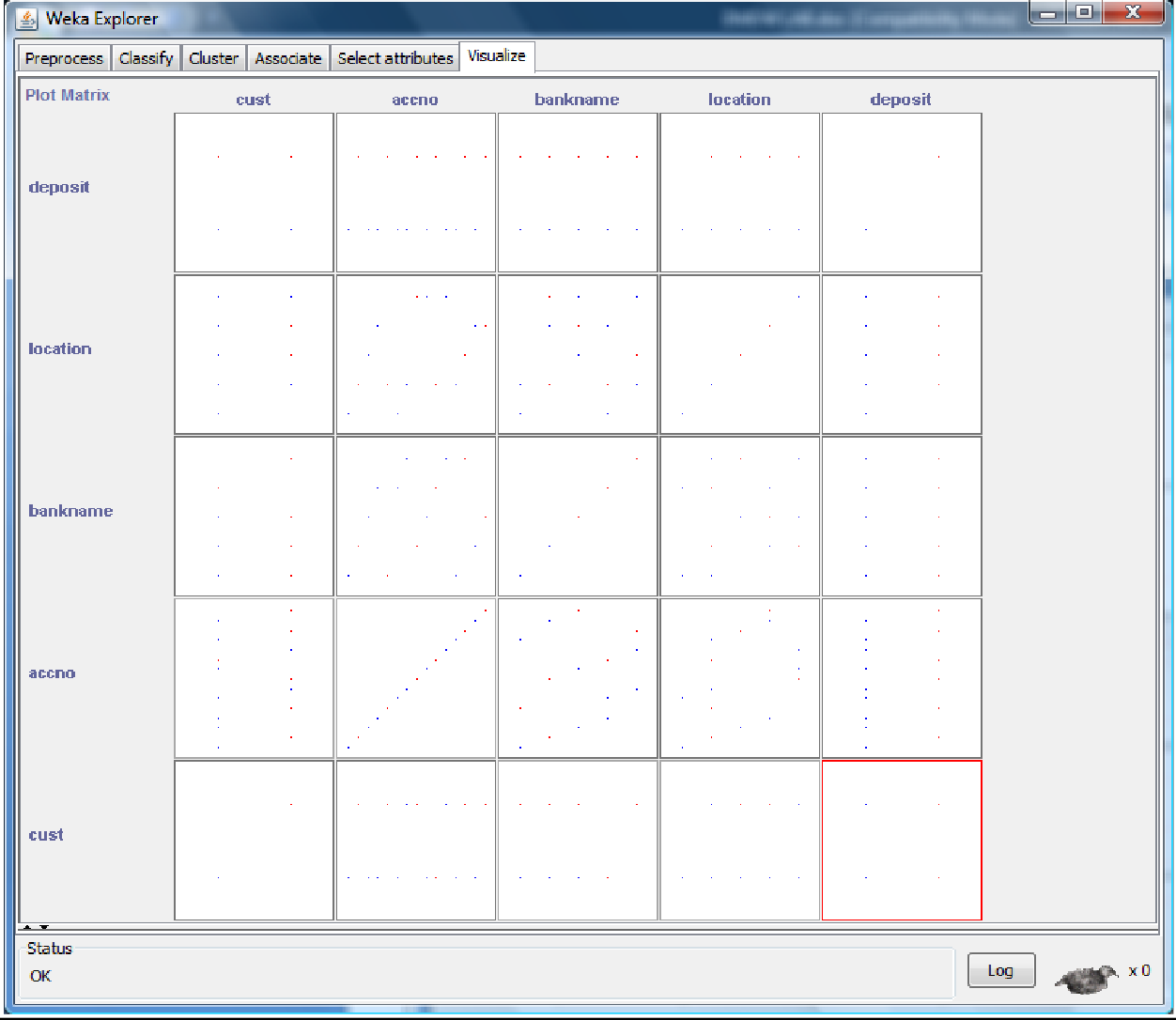
**Procedure:**

1. Open Start  Programs  Weka-3-4  Weka-3-4
2. Open the explorer and click on **Preprocess**, then a new window will appear. In that window select **bank.arff**

file then the data will be displayed.

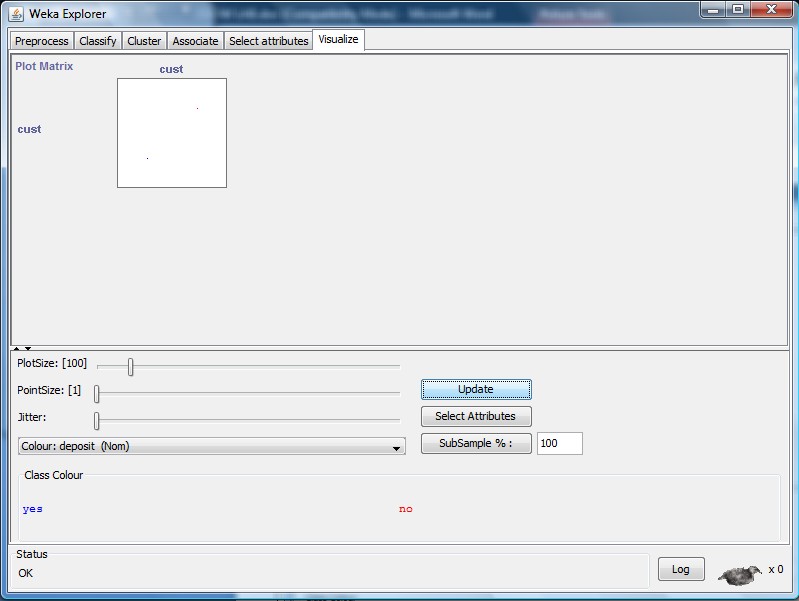
1. After that click on the **Visualize tab** on the top of the Menu bar.
2. When we select **Visualize tab** then **Plot Matrix** is displayed on the screen.

# 2-D Plot Matrix:



1. After that we select the **Select Attribute button**, then select **Cust attribute** and clock OK.
2. Click on the **Update button** to display the output.

# Output:



1. After that select the **Select Attribute button** and select **Accno attribute** and then click OK.
2. **Increase** the **Plot Size** and **Point Size**.
3. Click on the **Update button** to display the output.

# Output:

1. After that we select the **Select Attribute button**, then select **Bankname attribute** and clock OK.
2. Click on the **Update button** to display the output.

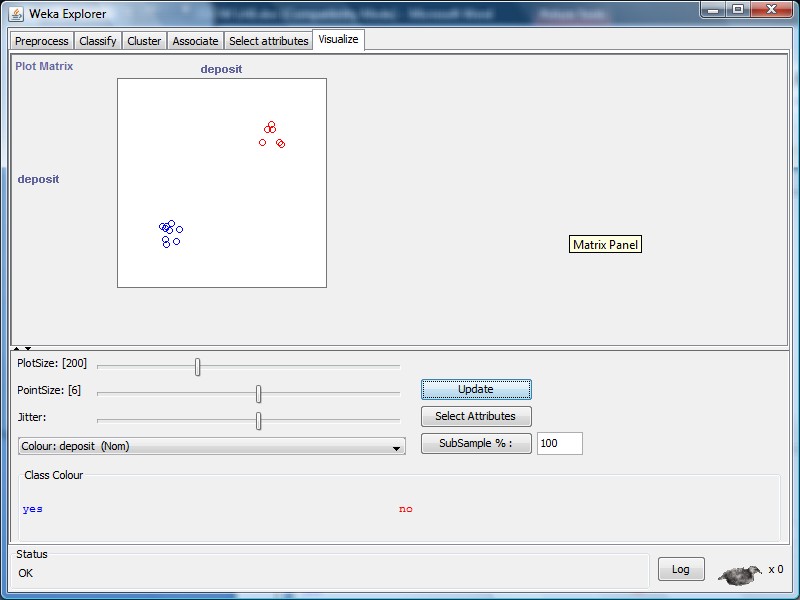
# Output:

1. After that select the **Select Attribute button** and select **location attribute** and then click OK.
2. **Increase** the **Jitter Size**.
3. Click on the **Update button** to display the output.

# Output:

1. After that we select the **Select Attribute button**, then select **Deposit attribute** and clock OK.
2. Click on the **Update button** to display the output.

# Output:



**Result:**

This program has been successfully executed.

**EXPERIMENT NO:15**

**Aim:**

Write a procedure for cross-validation using J48 Algorithm for weather table.

# Description:

**Cross-validation**, sometimes called **rotation estimation**, is a technique for assessing how the results of a [statistical](http://en.wikipedia.org/wiki/Statistics) analysis will generalize to an independent data set. It is mainly used in settings where the goal is prediction, and one wants to estimate how [accurately](http://en.wikipedia.org/wiki/Accuracy) a predictive model will perform in practice. One round of cross-validation involves [partitioning](http://en.wikipedia.org/wiki/Partition_of_a_set) a [sample](http://en.wikipedia.org/wiki/Statistical_sample) of [data](http://en.wikipedia.org/wiki/Data) into [complementary](http://en.wikipedia.org/wiki/Complement_%28set_theory%29) subsets, performing the analysis on one subset (called the *training set*), and validating the analysis on the other subset (called the *validation set* or *testing set*).

# Creation of Weather Table:

**Procedure:**

* 1. Open Start  Programs  Accessories  Notepad
  2. Type the following training data set with the help of Notepad for Weather Table. @relation weather

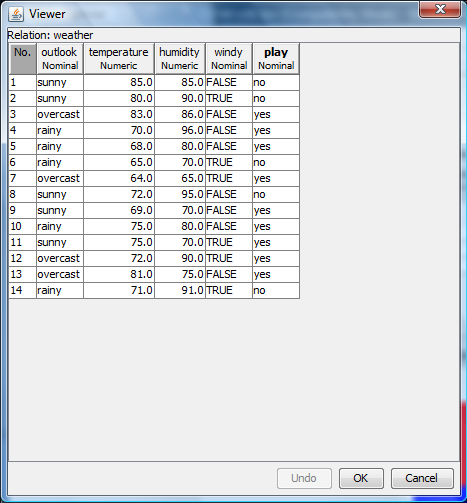
@attribute outlook {sunny, rainy, overcast} @attribute temperature numeric

@attribute humidity numeric @attribute windy {TRUE, FALSE} @attribute play {yes, no}

@data sunny,85,85,FALSE,no sunny,80,90,TRUE,no overcast,83,86,FALSE,yes rainy,70,96,FALSE,yes rainy,68,80,FALSE,yes rainy,65,70,TRUE,no overcast,64,65,TRUE,yes sunny,72,95,FALSE,no sunny,69,70,FALSE,yes rainy,75,80,FALSE,yes sunny,75,70,TRUE,yes overcast,72,90,TRUE,yes overcast,81,75,FALSE,yes rainy,71,91,TRUE,no

* 1. After that the file is saved with **.arff** file format.
  2. Minimize the arff file and then open Start  Programs  weka-3-4.
  3. Click on **weka-3-4**, then Weka dialog box is displayed on the screen.
  4. In that dialog box there are four modes, click on **explorer**.
  5. Explorer shows many options. In that click on **‘open file’** and select the arff file
  6. Click on **edit button** which shows weather table on weka.

# Training Data Set  Weather Table



**Procedure:**

1. **Start** -> **Programs** -> **Weka 3.4**
2. Open **Knowledge Flow**.
3. Select **Data Source** tab & choose **Arff Loader**.
4. Place **Arff Loader** component on the **layout area** by clicking on that component.
5. Specify an Arff file to load by **right clicking on Arff Loader** icon, and then a pop-up menu will appear. In that select **Configure** & browse to the location of **weather.arff**
6. Click on the **Evaluation** tab & choose **Class Assigner** & place it on the layout.
7. Now **connect** the **Arff Loader** to the **Class Assigner** by right clicking on Arff Loader, and then select

**Data Set** option, now a link will be established.

1. Right click on **Class Assigner** & choose **Configure** option, and then a new window will appear & specify a class to our data.
2. Select **Evaluation** tab & select **Cross-Validation Fold Maker** & place it on the layout.
3. Now **connect** the **Class Assigner** to the **Cross-Validation Fold Maker**.
4. Select **Classifiers** tab & select **J48** component & place it on the layout.
5. Now **connect Cross-Validation Fold Maker** to **J48 twice**; **first** choose **Training Data Set** option and

**then Test Data Set** option.

1. Select **Evaluation Tab** & select **Classifier Performance Evaluator** component & place it on the layout.
2. Connect **J48** to **Classifier Performance Evaluator** component by right clicking on J48 & selecting

**Batch Classifier**.

1. Select **Visualization** tab & select **Text Viewer** component & place it on the layout.
2. Connect **Text Viewer** to **Classifier Performance Evaluator** by right clicking on Text Viewer & by selecting **Text** option.
3. Start the flow of execution by selecting **Start Loading** from **Arff Loader**.
4. For viewing **result**, **right click** on **Text Viewer** & select the **Show Results**, and then the result will be displayed on the new window.

# Output:

**Result:**

The program has been successfully executed.

**EXPERIMENT NO:16**

**Aim:** Write a procedure for Clustering Buying data using Cobweb Algorithm.

# Description:

**Cluster analysis** or **clustering** is the task of assigning a set of objects into groups (called **clusters**) so that the objects in the same cluster are more similar (in some sense or another) to each other than to those in other clusters. Clustering is a main task of explorative [data mining,](http://en.wikipedia.org/wiki/Data_mining) and a common technique for [statistical](http://en.wikipedia.org/wiki/Statistics) [data analysis](http://en.wikipedia.org/wiki/Data_analysis) used in many fields, including [machine learning,](http://en.wikipedia.org/wiki/Machine_learning) [pattern recognition,](http://en.wikipedia.org/wiki/Pattern_recognition) [image analysis,](http://en.wikipedia.org/wiki/Image_analysis) [information retrieval](http://en.wikipedia.org/wiki/Information_retrieval), and [bioinformatics.](http://en.wikipedia.org/wiki/Bioinformatics)

# Creation of Buying Table:

**Procedure:**

1. Open Start  Programs  Accessories  Notepad
2. Type the following training data set with the help of Notepad for Buying Table. @relation buying

@attribute age {L20,20-40,G40} @attribute income {high,medium,low} @attribute stud {yes,no}

@attribute creditrate {fair,excellent} @attribute buyscomp {yes,no} @data

L20,high,no,fair,yes 20-40,low,yes,fair,yes

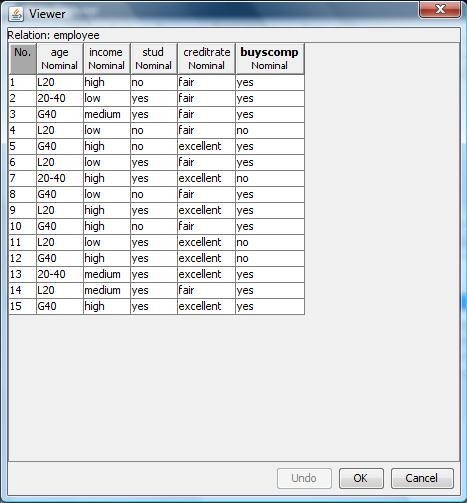
G40,medium,yes,fair,yes L20,low,no,fair,no G40,high,no,excellent,yes L20,low,yes,fair,yes

20-40,high,yes,excellent,no G40,low,no,fair,yes L20,high,yes,excellent,yes G40,high,no,fair,yes L20,low,yes,excellent,no G40,high,yes,excellent,no

20-40,medium,yes,excellent,yes L20,medium,yes,fair,yes G40,high,yes,excellent,yes

1. After that the file is saved with **.arff** file format.
2. Minimize the arff file and then open Start  Programs  weka-3-4.
3. Click on **weka-3-4**, then Weka dialog box is displayed on the screen.
4. In that dialog box there are four modes, click on **explorer**.
5. Explorer shows many options. In that click on **‘open file’** and select the arff file
6. Click on **edit button** which shows buying table on weka.

# Training Data Set  Buying Table



**Procedure:**

* 1. Click **Start** -> **Programs** -> **Weka 3.4**
  2. Click on **Explorer**.
  3. Click on **open file** & then select **Buying.arff** file.
  4. Click on **Cluster menu**. In this there are different algorithms are there.
  5. Click on **Choose button** and then select **cobweb** algorithm.
  6. Click on **Start button** and then **output** will be displayed on the screen.

# Output:

**Result:**

The program has been successfully executed.

**EXPERIMENT NO:17**

**Aim:** Write a procedure for Clustering Weather data using EM Algorithm.

# Description:

**Cluster analysis** or **clustering** is the task of assigning a set of objects into groups (called **clusters**) so that the objects in the same cluster are more similar (in some sense or another) to each other than to those in other clusters. Clustering is a main task of explorative [data mining,](http://en.wikipedia.org/wiki/Data_mining) and a common technique for [statistical](http://en.wikipedia.org/wiki/Statistics) [data analysis](http://en.wikipedia.org/wiki/Data_analysis) used in many fields, including [machine learning,](http://en.wikipedia.org/wiki/Machine_learning) [pattern recognition,](http://en.wikipedia.org/wiki/Pattern_recognition) [image analysis,](http://en.wikipedia.org/wiki/Image_analysis) [information retrieval](http://en.wikipedia.org/wiki/Information_retrieval), and [bioinformatics.](http://en.wikipedia.org/wiki/Bioinformatics)

# Creation of Weather Table:

**Procedure:**

1. Open Start  Programs  Accessories  Notepad
2. Type the following training data set with the help of Notepad for Weather Table. @relation weather

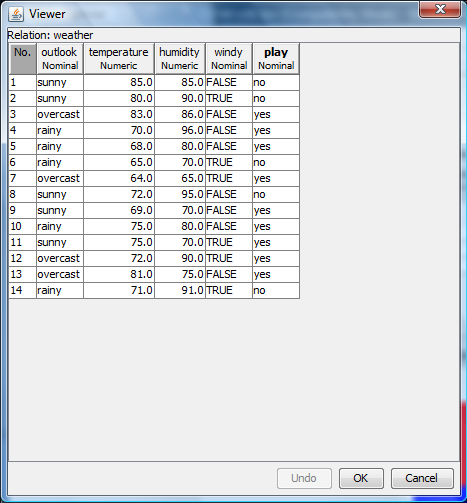
@attribute outlook {sunny, rainy, overcast} @attribute temperature numeric

@attribute humidity numeric @attribute windy {TRUE, FALSE} @attribute play {yes, no}

@data sunny,85,85,FALSE,no sunny,80,90,TRUE,no overcast,83,86,FALSE,yes rainy,70,96,FALSE,yes rainy,68,80,FALSE,yes rainy,65,70,TRUE,no overcast,64,65,TRUE,yes sunny,72,95,FALSE,no sunny,69,70,FALSE,yes rainy,75,80,FALSE,yes sunny,75,70,TRUE,yes overcast,72,90,TRUE,yes overcast,81,75,FALSE,yes rainy,71,91,TRUE,no

1. After that the file is saved with **.arff** file format.
2. Minimize the arff file and then open Start  Programs  weka-3-4.
3. Click on **weka-3-4**, then Weka dialog box is displayed on the screen.
4. In that dialog box there are four modes, click on **explorer**.
5. Explorer shows many options. In that click on **‘open file’** and select the arff file
6. Click on **edit button** which shows weather table on weka.

# Training Data Set  Weather Table



**Procedure:**

1. Click **Start** -> **Programs** -> **Weka 3.4**
2. Click on **Explorer**.
3. Click on **open file** & then select **Weather.arff** file.
4. Click on **Cluster menu**. In this there are different algorithms are there.
5. Click on **Choose button** and then select **EM** algorithm.
6. Click on **Start button** and then **output** will be displayed on the screen.

# Output:

**Result:**

The program has been successfully executed.

**EXPERIMENT NO:18**

**Aim:** Write a procedure for Banking data using Farthest First Algorithm.

# Description:

**Cluster analysis** or **clustering** is the task of assigning a set of objects into groups (called **clusters**) so that the objects in the same cluster are more similar (in some sense or another) to each other than to those in other clusters. Clustering is a main task of explorative [data mining,](http://en.wikipedia.org/wiki/Data_mining) and a common technique for [statistical](http://en.wikipedia.org/wiki/Statistics) [data analysis](http://en.wikipedia.org/wiki/Data_analysis) used in many fields, including [machine learning,](http://en.wikipedia.org/wiki/Machine_learning) [pattern recognition,](http://en.wikipedia.org/wiki/Pattern_recognition) [image analysis,](http://en.wikipedia.org/wiki/Image_analysis) [information retrieval](http://en.wikipedia.org/wiki/Information_retrieval), and [bioinformatics.](http://en.wikipedia.org/wiki/Bioinformatics)

# Creation of Banking Table:

**Procedure:**

1. Open Start  Programs  Accessories  Notepad
2. Type the following training data set with the help of Notepad for Banking Table. @relation bank

@attribute cust {male,female} @attribute accno

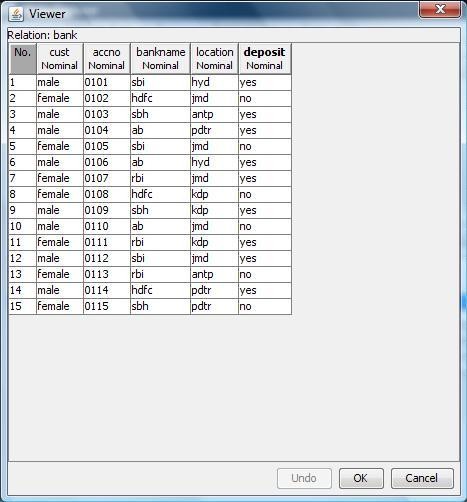
{0101,0102,0103,0104,0105,0106,0107,0108,0109,0110,0111,0112,0113,0114,0115}

@attribute bankname {sbi,hdfc,sbh,ab,rbi} @attribute location {hyd,jmd,antp,pdtr,kdp} @attribute deposit {yes,no}

@data male,0101,sbi,hyd,yes female,0102,hdfc,jmd,no male,0103,sbh,antp,yes male,0104,ab,pdtr,yes female,0105,sbi,jmd,no male,0106,ab,hyd,yes female,0107,rbi,jmd,yes female,0108,hdfc,kdp,no male,0109,sbh,kdp,yes male,0110,ab,jmd,no female,0111,rbi,kdp,yes male,0112,sbi,jmd,yes female,0113,rbi,antp,no male,0114,hdfc,pdtr,yes female,0115,sbh,pdtr,no

1. After that the file is saved with **.arff** file format.
2. Minimize the arff file and then open Start  Programs  weka-3-4.
3. Click on **weka-3-4**, then Weka dialog box is displayed on the screen.
4. In that dialog box there are four modes, click on **explorer**.
5. Explorer shows many options. In that click on **‘open file’** and select the arff file
6. Click on **edit button** which shows banking table on weka.

# Training Data Set  Banking Table



**Procedure:**

* 1. Click **Start** -> **Programs** -> **Weka 3.4**
  2. Click on **Explorer**.
  3. Click on **open file** & then select **Banking.arff** file.
  4. Click on **Cluster menu**. In this there are different algorithms are there.
  5. Click on **Choose button** and then select **FarthestFirst** algorithm.
  6. Click on **Start button** and then **output** will be displayed on the screen.

# Output:

**Result:**

The program has been successfully executed.

**EXPERIMENT NO:19**

**Aim:** Write a procedure for Employee data using Make Density Based Cluster Algorithm.

# Description:

**Cluster analysis** or **clustering** is the task of assigning a set of objects into groups (called **clusters**) so that the objects in the same cluster are more similar (in some sense or another) to each other than to those in other clusters. Clustering is a main task of explorative [data mining,](http://en.wikipedia.org/wiki/Data_mining) and a common technique for [statistical](http://en.wikipedia.org/wiki/Statistics) [data analysis](http://en.wikipedia.org/wiki/Data_analysis) used in many fields, including [machine learning,](http://en.wikipedia.org/wiki/Machine_learning) [pattern recognition,](http://en.wikipedia.org/wiki/Pattern_recognition) [image analysis,](http://en.wikipedia.org/wiki/Image_analysis) [information retrieval](http://en.wikipedia.org/wiki/Information_retrieval), and [bioinformatics.](http://en.wikipedia.org/wiki/Bioinformatics)

# Creation of Employee Table:

**Procedure:**

1. Open Start  Programs  Accessories  Notepad
2. Type the following training data set with the help of Notepad for Employee Table. @relation employee

@attribute eid numeric

@attribute ename {raj,ramu,anil,sunil,rajiv,sunitha,kavitha,suresh,ravi,ramana,ram,kavya,navya} @attribute salary numeric

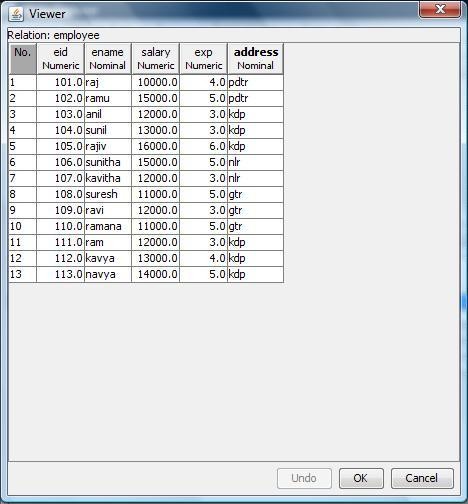
@attribute exp numeric

@attribute address {pdtr,kdp,nlr,gtr}

@data 101,raj,10000,4,pdtr 102,ramu,15000,5,pdtr 103,anil,12000,3,kdp 104,sunil,13000,3,kdp 105,rajiv,16000,6,kdp 106,sunitha,15000,5,nlr 107,kavitha,12000,3,nlr 108,suresh,11000,5,gtr 109,ravi,12000,3,gtr 110,ramana,11000,5,gtr 111,ram,12000,3,kdp 112,kavya,13000,4,kdp 113,navya,14000,5,kdp

1. After that the file is saved with **.arff** file format.
2. Minimize the arff file and then open Start  Programs  weka-3-4.
3. Click on **weka-3-4**, then Weka dialog box is displayed on the screen.
4. In that dialog box there are four modes, click on **explorer**.
5. Explorer shows many options. In that click on **‘open file’** and select the arff file
6. Click on **edit button** which shows employee table on weka.

# Training Data Set  Employee Table



**Procedure:**

1. Click **Start** -> **Programs** -> **Weka 3.4**
2. Click on **Explorer**.
3. Click on **open file** & then select **Employee.arff** file.
4. Click on **Cluster menu**. In this there are different algorithms are there.
5. Click on **Choose button** and then select **MakeDensityBasedClusterer** algorithm.
6. Click on **Start button** and then **output** will be displayed on the screen.

# Output:

**Result:**

The program has been successfully executed.

**EXPERIMENT NO:20**

**Aim:** Write a procedure for Clustering Customer data using Simple KMeans Algorithm.

# Description:

**Cluster analysis** or **clustering** is the task of assigning a set of objects into groups (called **clusters**) so that the objects in the same cluster are more similar (in some sense or another) to each other than to those in other clusters. Clustering is a main task of explorative [data mining,](http://en.wikipedia.org/wiki/Data_mining) and a common technique for [statistical](http://en.wikipedia.org/wiki/Statistics) [data analysis](http://en.wikipedia.org/wiki/Data_analysis) used in many fields, including [machine learning,](http://en.wikipedia.org/wiki/Machine_learning) [pattern recognition,](http://en.wikipedia.org/wiki/Pattern_recognition) [image analysis,](http://en.wikipedia.org/wiki/Image_analysis) [information retrieval](http://en.wikipedia.org/wiki/Information_retrieval), and [bioinformatics.](http://en.wikipedia.org/wiki/Bioinformatics)

# Creation of Customer Table:

**Procedure:**

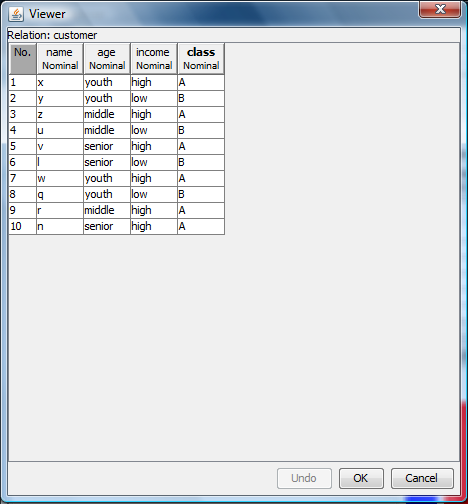
* 1. Open Start  Programs  Accessories  Notepad
  2. Type the following training data set with the help of Notepad for Buying Table. @relation customer

@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

* 1. After that the file is saved with **.arff** file format.
  2. Minimize the arff file and then open Start  Programs  weka-3-4.
  3. Click on **weka-3-4**, then Weka dialog box is displayed on the screen.
  4. In that dialog box there are four modes, click on **explorer**.
  5. Explorer shows many options. In that click on **‘open file’** and select the arff file
  6. Click on **edit button** which shows buying table on weka.

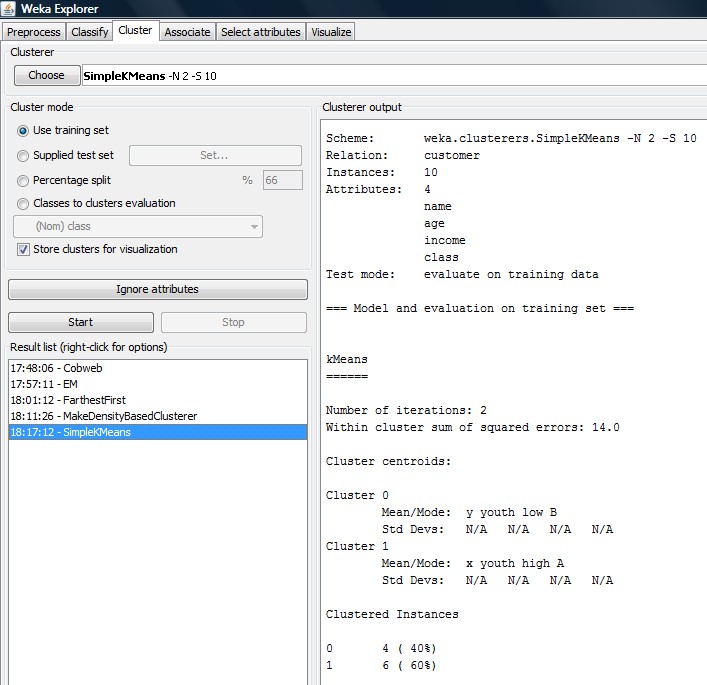
# Training Data Set  Customer Table



**Procedure:**

1. Click **Start** -> **Programs** -> **Weka 3.4**
2. Click on **Explorer**.
3. Click on **open file** & then select **Customer.arff** file.
4. Click on **Cluster menu**. In this there are different algorithms are there.
5. Click on **Choose button** and then select **SimpleKMeans** algorithm.
6. Click on **Start button** and then **output** will be displayed on the screen.

# Output:



**Result:**

The program has been successfully executed.