Experiment No:4

**Aim:** To study & implement following classifiers (Weka Tool).

1. Naive Bayes Classifier
2. Decision Tree

**Part A**

**Naive Bayes Classifier**

**Theory:**

This experiment illustrates the use of naïve bayes classifier in weka. The sample data set used in this experiment is “employee and Student ”data available at arff format. This document assumes that appropriate data pre processing has been performed.

***Steps involved in this experiment:***

1. We begin the experiment by loading the data (employee.arff) into weka.
2. Next we select the “classify” tab and click “choose” button to select the “Naive Bayes”classifier.
3. Now we specify the various parameters. These can be specified by clicking in the text box to the right of the choose button. In this example, we accept the default values his default version does perform some pruning but does not perform error pruning.
4. Under the “text “options in the main panel. We select the 10-fold cross validation as our evaluation approach. Since we don’t have separate evaluation data set, this is necessary to get a reasonable idea of accuracy of generated model.
5. We now click”start”to generate the model .the ASCII version of the tree as well as evaluation statistic will appear in the right panel when the model construction is complete.
6. Note that the classification accuracy of model is about 69%.this indicates that we may find more work. (Either in preprocessing or in selecting current parameters for the classification)
7. Now weka also lets us a view a graphical version of the classification tree. This can be done by right clicking the last result set and selecting “visualize tree” from the pop-up menu.
8. We will use our model to classify the new instances.
9. In the main panel under “text “options click the “supplied test set” radio button and then click the “set” button. This will show pop-up window which will allow you to open the file containing test instances.

***Data set*** nb.arff***:***

@relation 'weather.symbolic-weka.filters.unsupervised.attribute.Remove-R1,4'

@attribute temperature {hot,mild,cool}

@attribute humidity {high,normal}

@attribute play {yes,no}

@data

hot,high,no

hot,high,no

hot,high,yes

mild,high,yes

cool,normal,yes

cool,high,no

cool,normal,yes

mild,high,no

cool,normal,yes

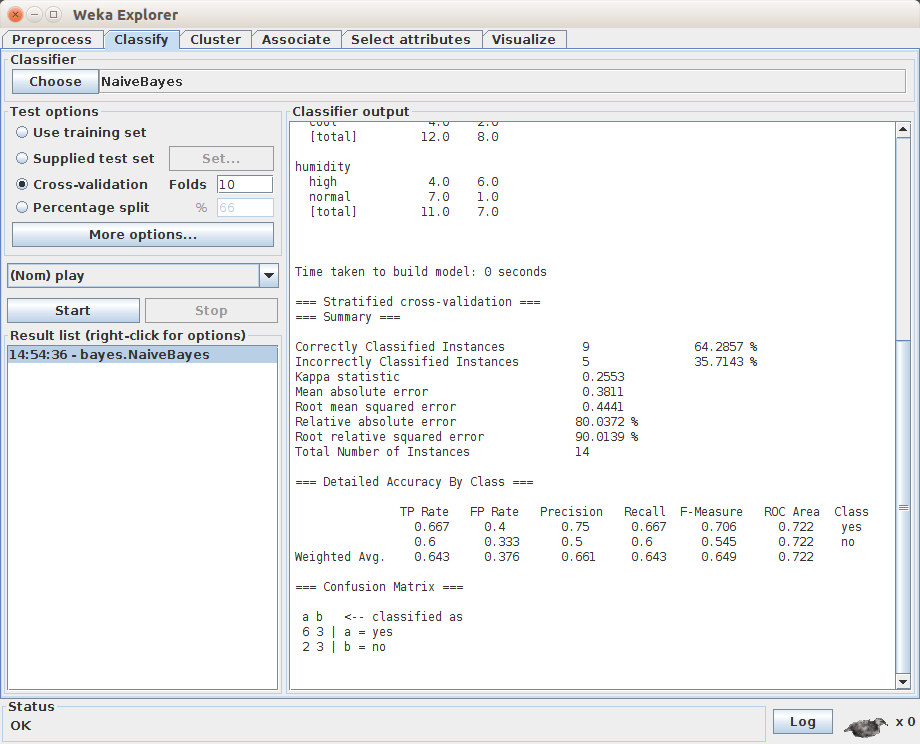
mild,normal,yes

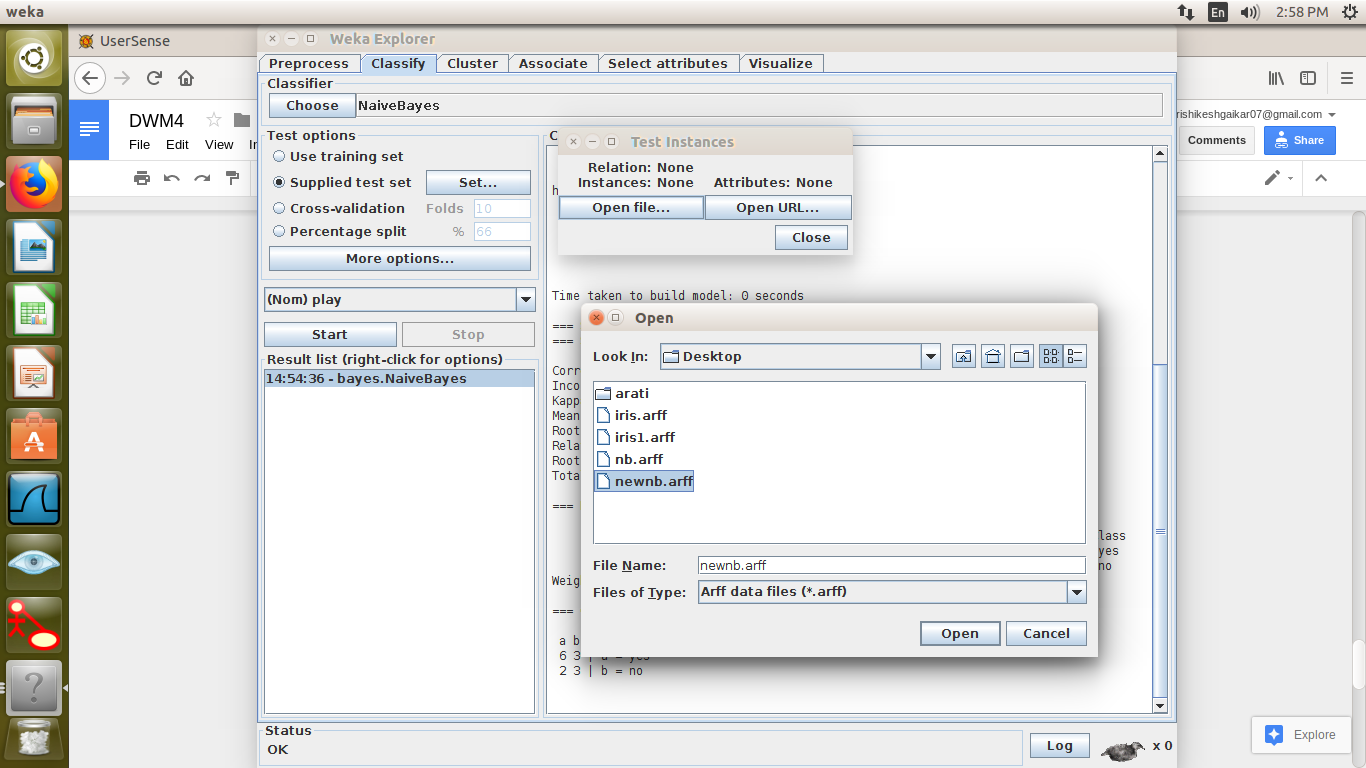
mild,normal,yes

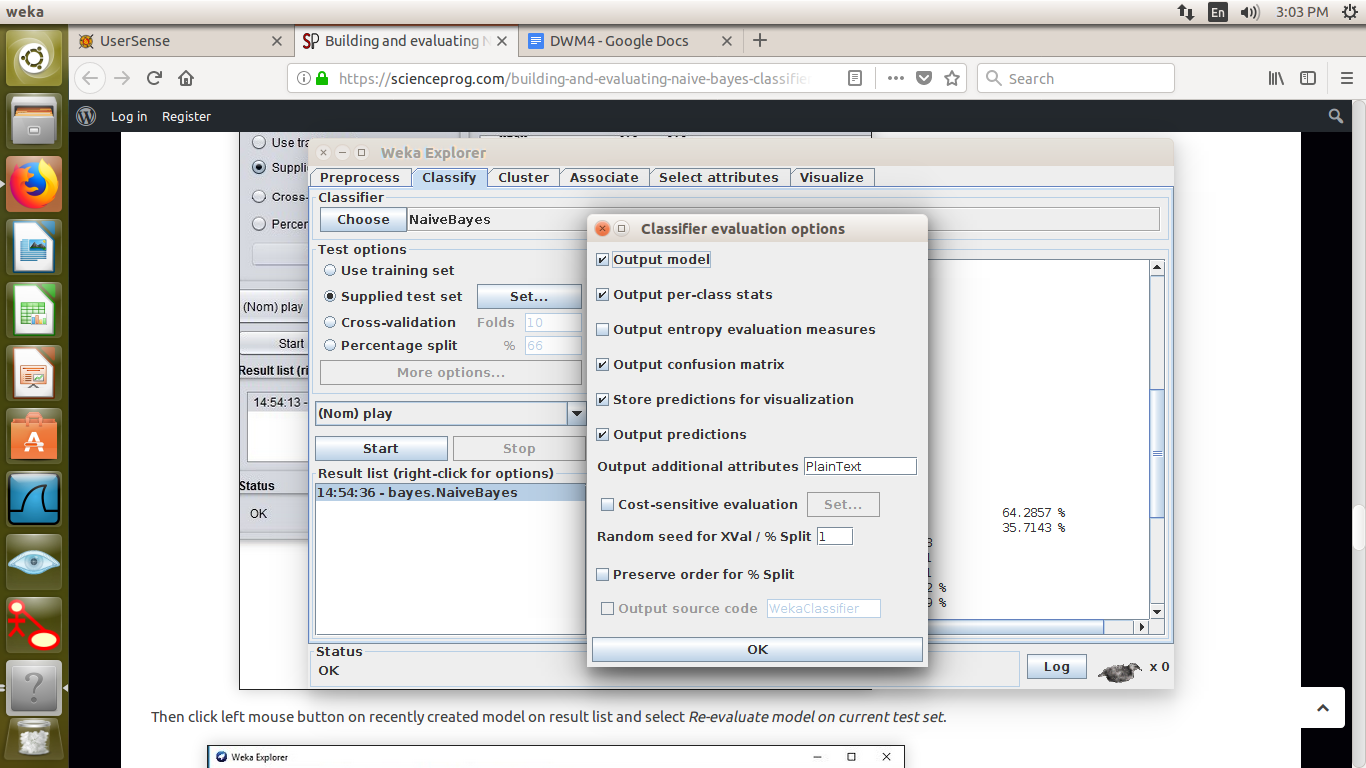
mild,high,yes

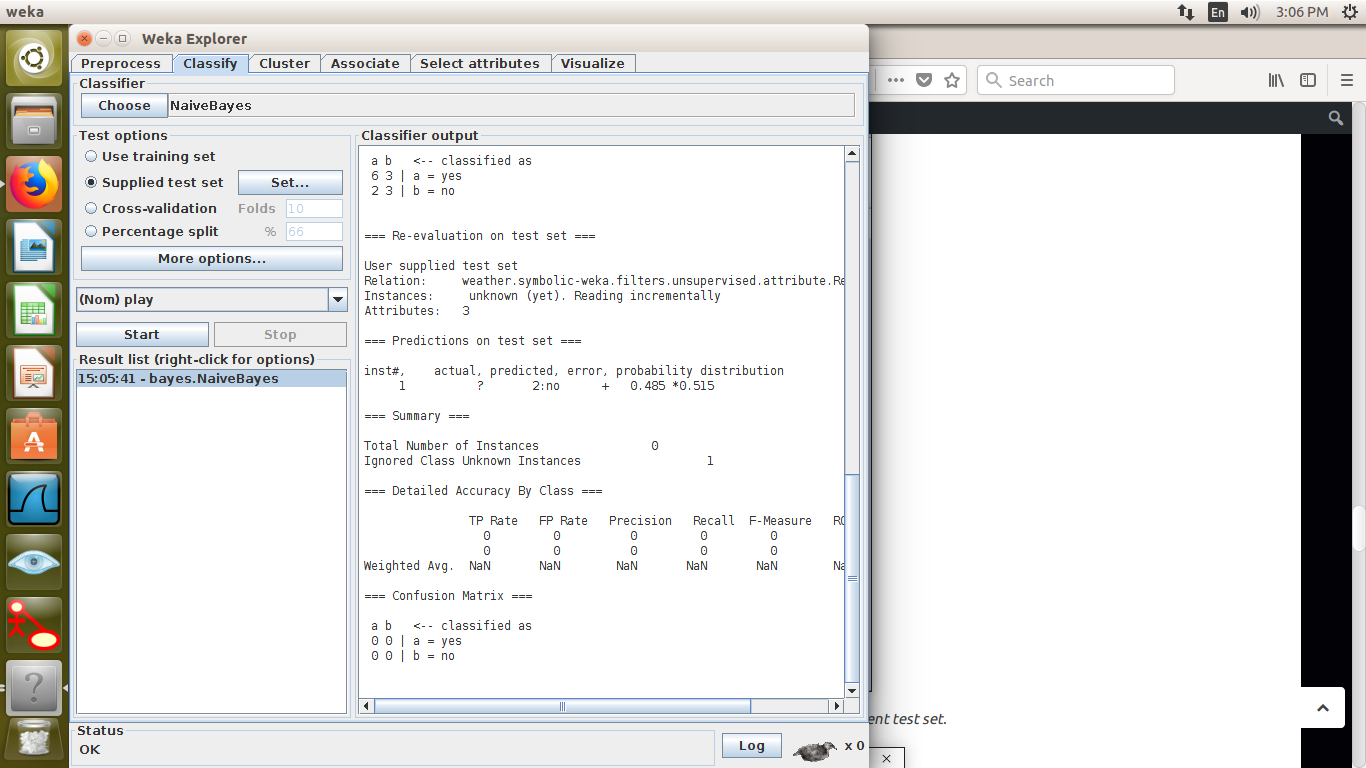
hot,normal,yes

mild,high,no









**Part B**

**Decision Tree**

**Theory:**

Decision tree learning is a method for assessing the most likely outcome value by taking into account the known values of the stored data instances. This learning method is among the most popular of inductive inference algorithms and has been successfully applied in broad range of tasks such as assessing the credit risk of applicants and improving loyalty of regular customers.

This experiment illustrates the use of j-48 classifier in weka. The sample data set used in this experiment is “tennis” data available at arff format. This document assumes that appropriate data preprocessing has been performed.

**Steps involved in this experiment:**

1. We begin the experiment by loading the data (PlayTennis.arff) into weka.
2. Next we select the classify tab and choose to select the “j48” classifier.
3. Now we specify the various parameters. These can be specified by clicking in the textbox to the right of the choose button. In this example,we accept the default values. The default version does perform soke pruning but does not perform error pruning.
4. Under the “text” option in the main panel. We select the 10-fold cross validation as our evaluation approach. Since we don't have a separate evaluation data set, this is necessary to get a reasonable idea of accuracy of the generated model
5. We now click “Start” to generate model. The ascii version of the tree as well as evaluation statistic will appear in the right panel when the model construction is complete.
6. Note that the classification accuracy model is about 69%. This indicates that we may find more work.
7. Now weka also lets us view a graphical version of the classification tree. This can done by right clicking the last result set and selecting “Visualize tree” from the popup menu.
8. We will use our own model to classify new instances.
9. In the main panel under “text” option select the “supplied test set” radio button and then click the “set” button. This will popup a window which will allow you to open the file containing the test instances

***Dataset*** *PlayTennis.arff****:***

*@relation PlayTennis*

*@attribute day numeric*

*@attribute outlook {Sunny, Overcast, Rain}*

*@attribute temperature {Hot, Mild, Cool}*

*@attribute humidity {High, Normal}*

*@attribute wind {Weak, Strong}*

*@attribute playTennis {Yes, No}*

*@data*

*1,Sunny,Hot,High,Weak,No,?*

*2,Sunny,Hot,High,Strong,No,?*

*3,Overcast,Hot,High,Weak,Yes,?*

*4,Rain,Mild,High,Weak,Yes,?*

*5,Rain,Cool,Normal,Weak,Yes,?*

*6,Rain,Cool,Normal,Strong,No,?*

*7,Overcast,Cool,Normal,Strong,Yes,?*

*8,Sunny,Mild,High,Weak,No,?*

*9,Sunny,Cool,Normal,Weak,Yes,?*

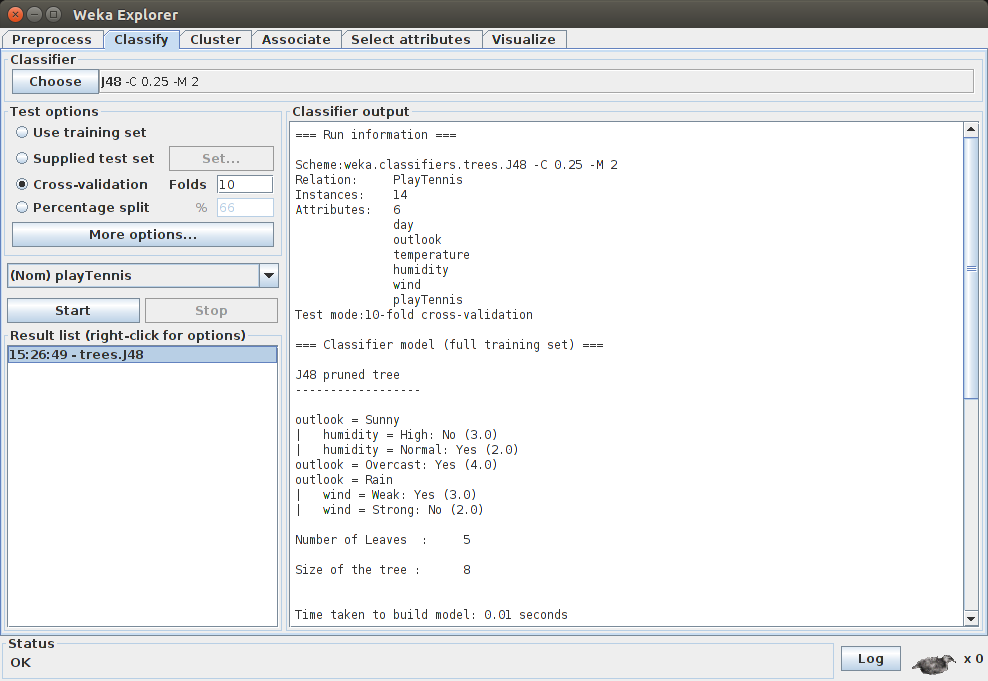
*10,Rain,Mild,Normal,Weak,Yes,?*

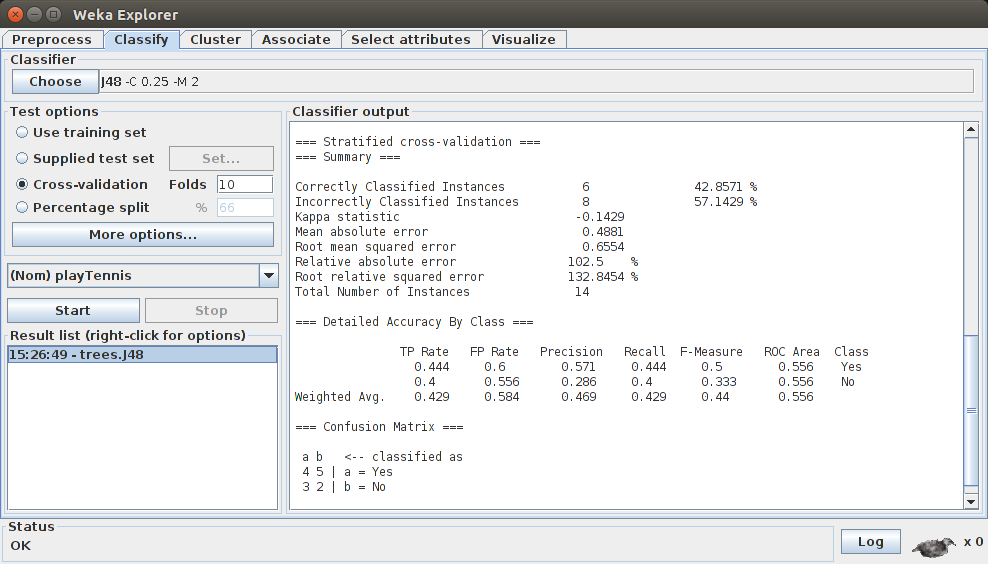
*11,Sunny,Mild,Normal,Strong,Yes,?*

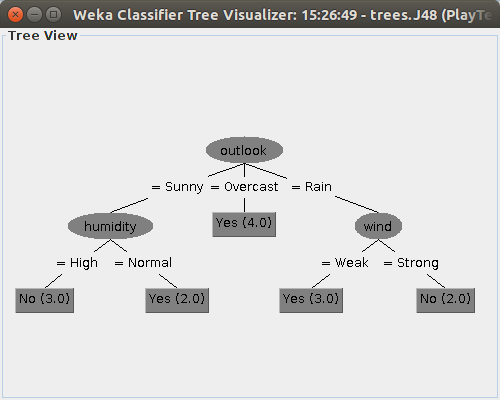
*12,Overcast,Mild,High,Strong,Yes,?*

*13,Overcast,Hot,Normal,Weak,Yes,?*

*14,Rain,Mild,High,Strong,No,?*







**Conclusion:**

Naive Bayes classifier and Decision Tree have been studied and implemented successfully using WEKA tool.