Practical No. 06

Aim: Perform pre-processing, classification and visualisation techniques on agricultural datasets.

BHAGWATI
PAGE NO.:
DATE:

Practical No. 06

Name of Practical

Aim: Perform pre-processing, classification and visualisation techniques on agricultural datasets.

Theory:

Data preprocessing is a data mining technique that involves transforming raw data into an understandable format. Real-world data is often incomplete, inconsistent and/or lacking in certain behaviors or trends, and is likely to contain many errors. Data preprocessing is a proven method of resolving such issues. Data preprocessing prepares raw data for further processing. Data preprocessing includes cleaning, instance ~~extraction~~ selection, normalisation, transformation, feature extraction and selection, etc.

The sample data set used in this experiment is "student" data available at arff format. This document presumes that appropriate data preprocessing has been performed.

Steps involved in this experiment:

Step-I: We begin the experiment by loading the data into Weka.

Step 2: Now, we select the "classify" tab and click "choose" button to select the "j48" classifier.

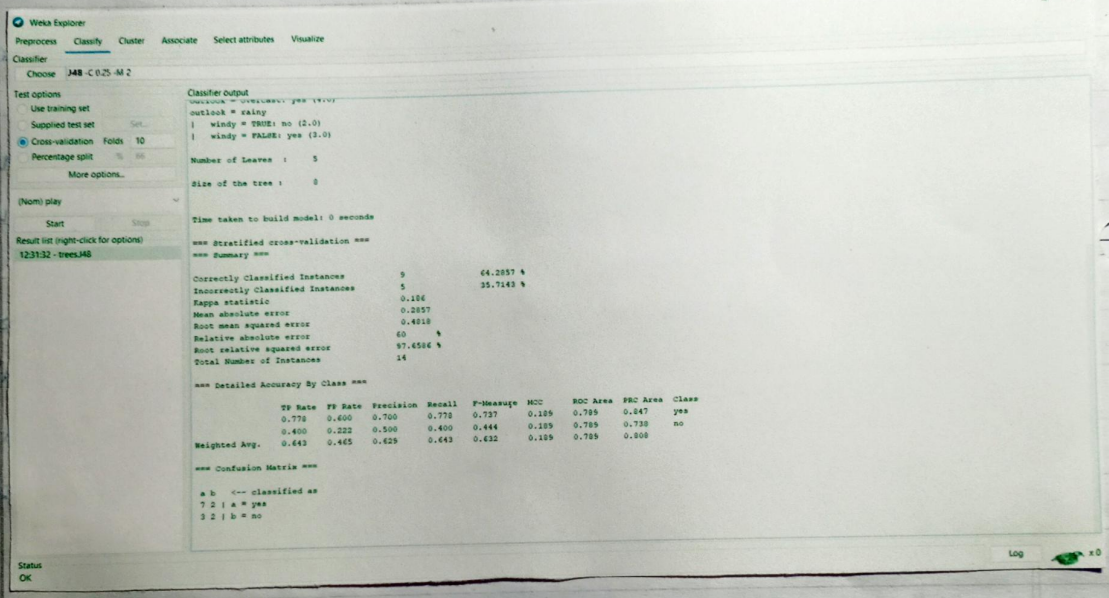
Step 3: Now, we specify the various parameters. This can be specified by clicking in the textbox to the right of the choose button. In this example we accept the default values. The default version does perform some pruning but does not perform error pruning.

Teacher's Signature.....

Practical No. 05

Aim : Perform pre-processing, classification and Visualization techniques on agricultural datasets.

Theory :



Steps involved in this experiment:

Step-1 : We begin the experiment by loading the data into Weka.

Step 2 : Now, we select the "classify" tab and click "choose" button to select the "J48" classifier.

Step 3 : Now, we specify the various parameters. This 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 some pruning but does not perform error pruning.

BHAGWATI
PAGE NO. :
DATE :

Name of Practical

Step 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 seperate evaluation data set, this is necessary to get a reasonable idea of accuracy of generated model.

Step 5: We now click "Start" to generated 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.

Step 6: Note 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.)

Step 7: Now, Weka also lets us view a graphical version of the classification tree. This can be done by the right clicking the last result set and selecting "visualisinge tree" from the pop-up menu.

Step 8: We will use our model to classify the new instances.

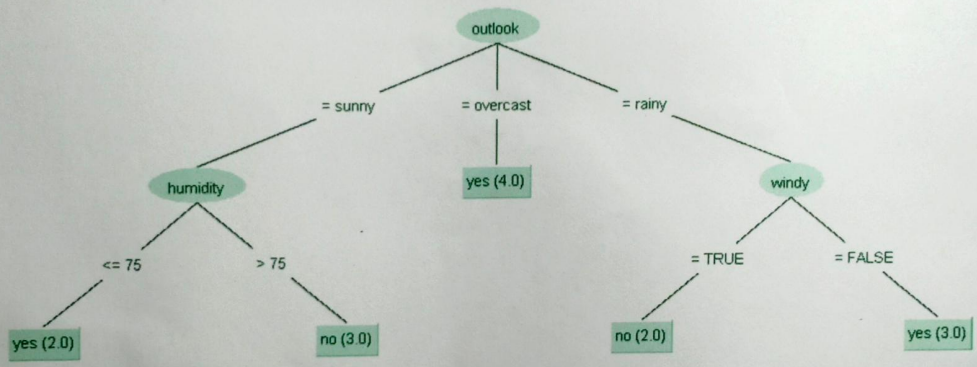
Step 9: In the main panel under "text" options click the "Supplied test set" radio button and then click the "set" button. This wills pop-up a window which will allow you to open the file containing test instances.

These are various ways of manipulating the visualizations available from the visualize panel in Weka.

Teacher's Signature.....

Weka Classifier Tree Visualizer: 12:31:32 - trees.J48 (weather)

Tree View



Ex

Conclusion :

Through Weka explore options, I performed preprocessing, classification and visualisation on Agriculture datasets.

BHAGWATI
PAGE NO. :
DATE :

Name of Practical

- Select an area of the graphs to zoom in on. To do this choose an option from lower right-hand drop-down box on the visualisation window (underneath the Y-axis drop down.) You can choose to draw onto the canvas using a polygon or polyline. Once you've selected an area click Submit button to zoom in.

- It is also possible to save the visualization data to an ARFF file. just use the "Save" button. You can then load it back into the Explorer and use any of Weka's filters to manipulate it.

Conclusion:

Through Weka explorers options of Pre-processing, Classification and Visualization on Agriculture dataset.

Teacher's Signature.....