

Experiment No. 4

Aim:- To study WEKA and implement the decision tree algorithm using WEKA.

Outputs :-

- WEKA

- **Preprocessing (Before removing missing values)**

The screenshot shows the WEKA Preprocess window. The 'Filter' tab is active, and the 'RemoveWithValues' filter is selected. The filter's configuration is: 'RemoveWithValues < 0.0 < 1 -1 first-last -M'. The 'Current relation' is 'labor-neg-data' with 57 instances and 17 attributes. The 'Selected attribute' is 'duration', which is numeric with 3 distinct values and 1 missing value (2%). The 'Statistics' table for the 'duration' attribute is as follows:

Statistic	Value
Minimum	1
Maximum	3
Mean	2.161
StdDev	0.708

The 'Class' is 'class (Nom)'. A visualization of the data is shown at the bottom right, with red and blue bars representing different classes. The status bar at the bottom indicates 'OK'.

The screenshot shows the 'weka.gui.GenericObjectEditor' window for the 'weka.filters.unsupervised.instance.RemoveWithValues' filter. The 'About' section states: 'Filters instances according to the value of an attribute.' The configuration parameters are:

- attributeIndex: 1
- dontFilterAfterFirstBatch: False
- invertSelection: False
- matchMissingValues: True
- modifyHeader: False
- nominalIndices: first-last
- splitPoint: 0.0

The window includes buttons for 'Open...', 'Save...', 'OK', and 'Cancel'.

- After removing missing values

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Open file... | Open URL... | Open DB... | Generate... | Undo | Edit... | Save...

Filter: Choose RemoveWithValues -5 0.0 <10 <1 first last M Apply

Current relation: labor-neg-data-weka.filters.unsupervised.instance.RemoveWithValues-5,0,0,10,1,first,last,M
Instances: 0 Attributes: 17

Attributes: All None Insert Pattern

No.	Name
1	duration
2	wage-increase-first-year
3	wage-increase-second-year
4	wage-increase-third-year
5	cost-of-living-adjustment
6	working-hours
7	pension
8	standby-pay
9	shift-differential
10	education-allowance
11	statutory-holidays
12	vacation
13	longterm-disability-assistance
14	contribution-to-dental-plan
15	bereavement-assistance
16	contribution-to-health-plan
17	class

Selected attribute: Name: duration Missing: 0 (0%) Distinct: 0 Type: Numeric Unique: 0 (0%)

Statistic	Value
Minimum	NaN
Maximum	NaN
Mean	NaN
StdDev	NaN

Class: class (Nom) Visualize All

- Classify

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize

Classifier: Choose J48 < 0.25 -M 2

Test options: Use training set, Supplied test set, Cross-validation (Folds: 10), Percentage split (%: 0.5), More options...

(Nom) play Start Stop

Result list (right-click for options): 10-10-00-trees.j48

Classifier output: J48 pruned tree

```

outlook = sunny
| humidity >= 75: yes (2.0)
| humidity > 75: no (3.0)
outlook = overcast: yes (4.0)
outlook = rainy
| windy = TRUE: no (2.0)
| windy = FALSE: yes (3.0)
  
```

Number of Leaves : 5
Size of the tree : 9
Time taken to build model: 0.01 seconds

Stratified cross-validation Summary

	Correctly Classified Instances	Incorrectly Classified Instances	Kappa statistic	Mean absolute error	Root mean squared error	Relative absolute error	Root relative squared error	Total Number of Instances
	9	5	0.186	0.2857	0.4819	60 %	97.6586 %	14

Detailed Accuracy By Class

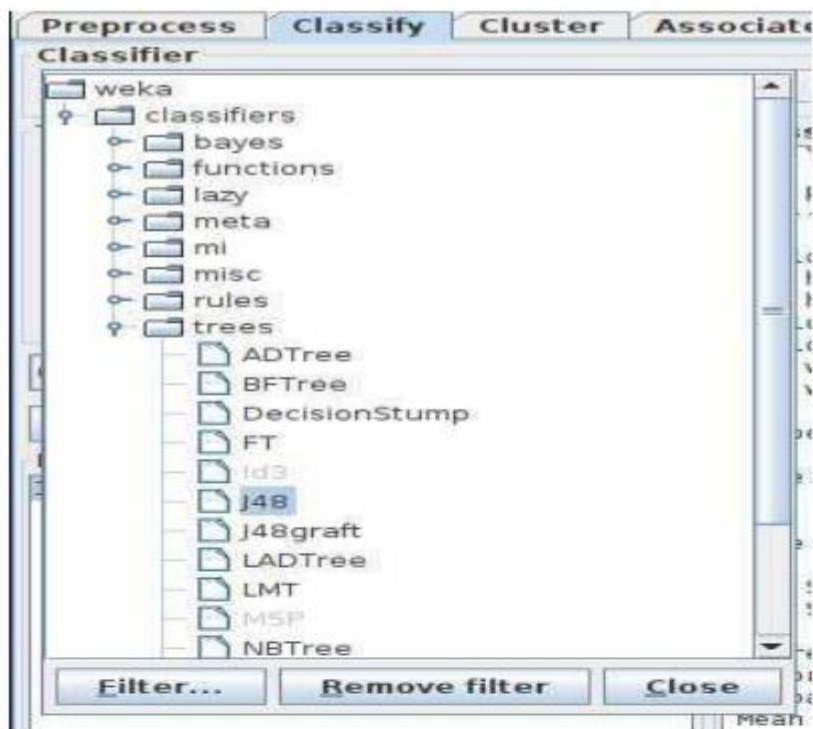
	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
Weighted Avg.	0.643	0.405	0.629	0.643	0.632	0.709	yes
							no

Confusion Matrix

```

a b  <-- classified as
7 2 | a = yes
3 2 | b = no
  
```

Status: OK Log



- After

