



Academic Year: 2024-2025

Semester: V

Class / Branch: TE/CSE-DS

Subject: DWM Lab

Name of Instructor: Prof. Sarala Mary

Name of Student: Gauri Iyer

Student ID: 22107030

Date Of Performance: 22/8/24

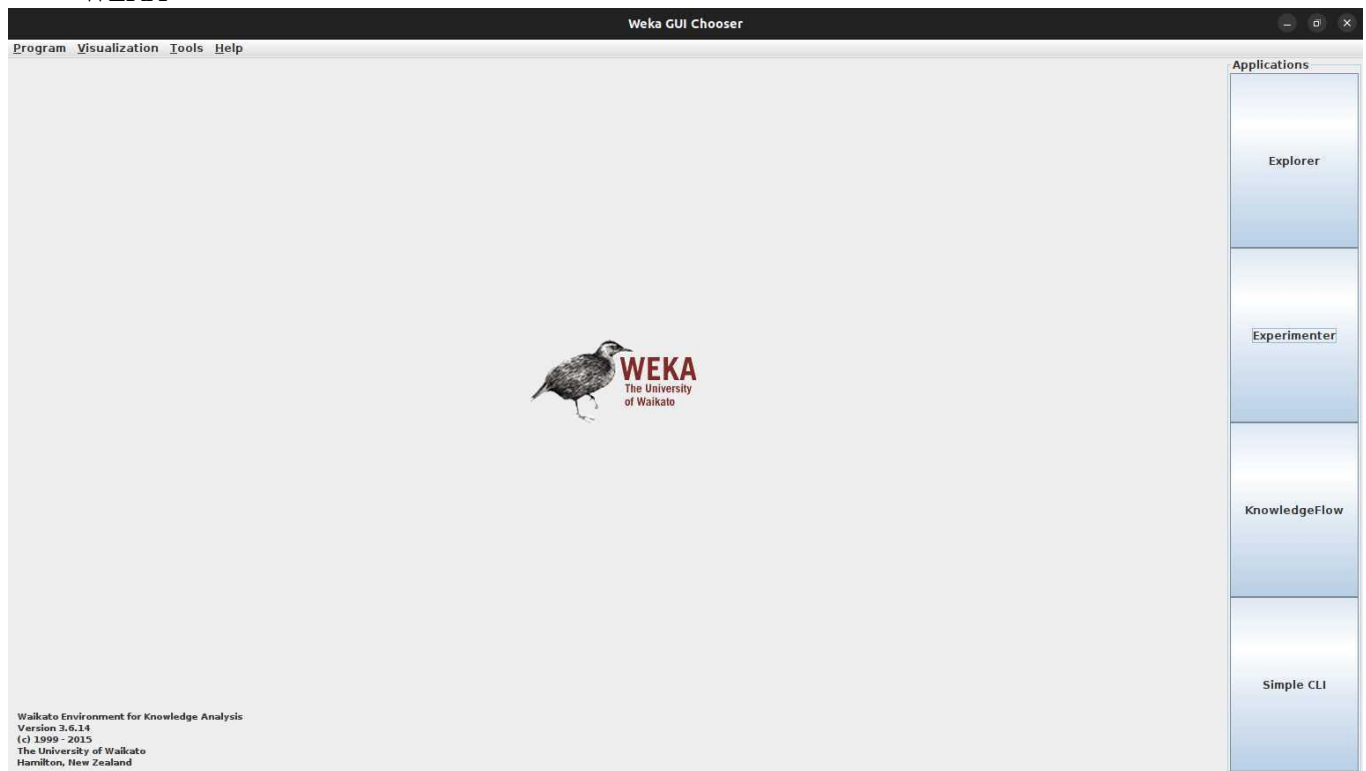
Date Of Submission: 22/8/24

Experiment No. 4

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

Outputs :-

- WEKA





- Preprocessing (Before removing missing values)

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

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

Filter: Choose RemoveWithValues -S 0.0 -C 1 -L first-last -M Apply

Current relation
Relation: labor-neg-data
Instances: 57
Attributes: 17

Attributes

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

Remove

Selected attribute
Name: duration
Missing: 1 (2%)
Distinct: 3
Type: Numeric
Unique: 0 (0%)

Statistic	Value
Minimum	1
Maximum	3
Mean	2.161
StdDev	0.708

Class: class (Nom) Visualize All

Status
OK Log

weka.gui.GenericObjectEditor

weka.filters.unsupervised.instance.RemoveWithValues

About
Filters instances according to the value of an attribute.

More
Capabilities

attributeIndex 1

dontFilterAfterFirstBatch False

invertSelection False

matchMissingValues True

modifyHeader False

nominalIndices first-last

splitPoint 0.0

Open... Save... OK Cancel



- After removing missing values

The screenshot shows the Weka Preprocess tab with the 'RemoveWithValues' filter selected. The filter is configured to remove instances where the 'duration' attribute has values less than or equal to 0.0, less than 10, or less than -1. The 'Current relation' is 'labor-neg-data-weka.filters.unsupervised.instance.RemoveWithValues-S0.0-C1-1-first-last-M'. The 'Selected attribute' is 'duration', which is numeric and has 0 missing values. The 'Class' is 'class (Nom)'. The 'Visualize All' button is visible at the bottom right.

- Classify

The screenshot shows the Weka Explorer interface with the 'J48' classifier selected. The 'Test options' are set to 'Cross-validation' with 'Folds' set to 10. The 'Result list' shows '10:40:00 - trees.J48'. The 'Classifier output' tab is active, displaying the 'J48 pruned tree' and various performance metrics.

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 : 8
Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===
=== Summary ===

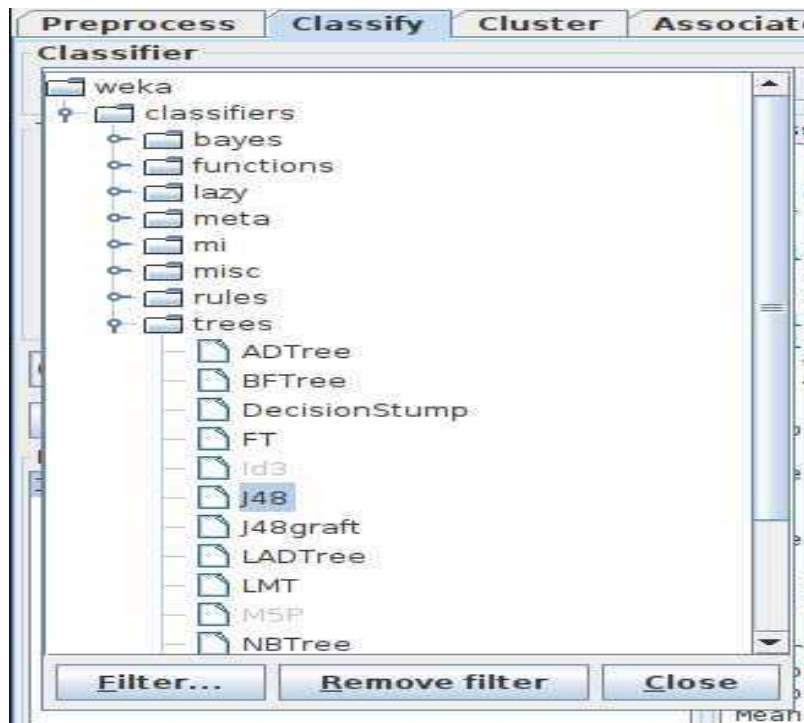
Metric	Value	Percentage
Correctly Classified Instances	9	64.2857 %
Incorrectly Classified Instances	5	35.7143 %
Kappa statistic	0.196	
Mean absolute error	0.2857	
Root mean squared error	0.4818	
Relative absolute error	60 %	
Root relative squared error	97.6586 %	
Total Number of Instances	14	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
Weighted Avg.	0.778	0.643	0.222	0.629	0.444	0.632	yes
	0.4	0.222	0.5	0.4	0.444	0.789	no

=== Confusion Matrix ===

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



- After

