

Name: Rampurawala Alefiya Abbas

Roll No: 43

Class: TE4

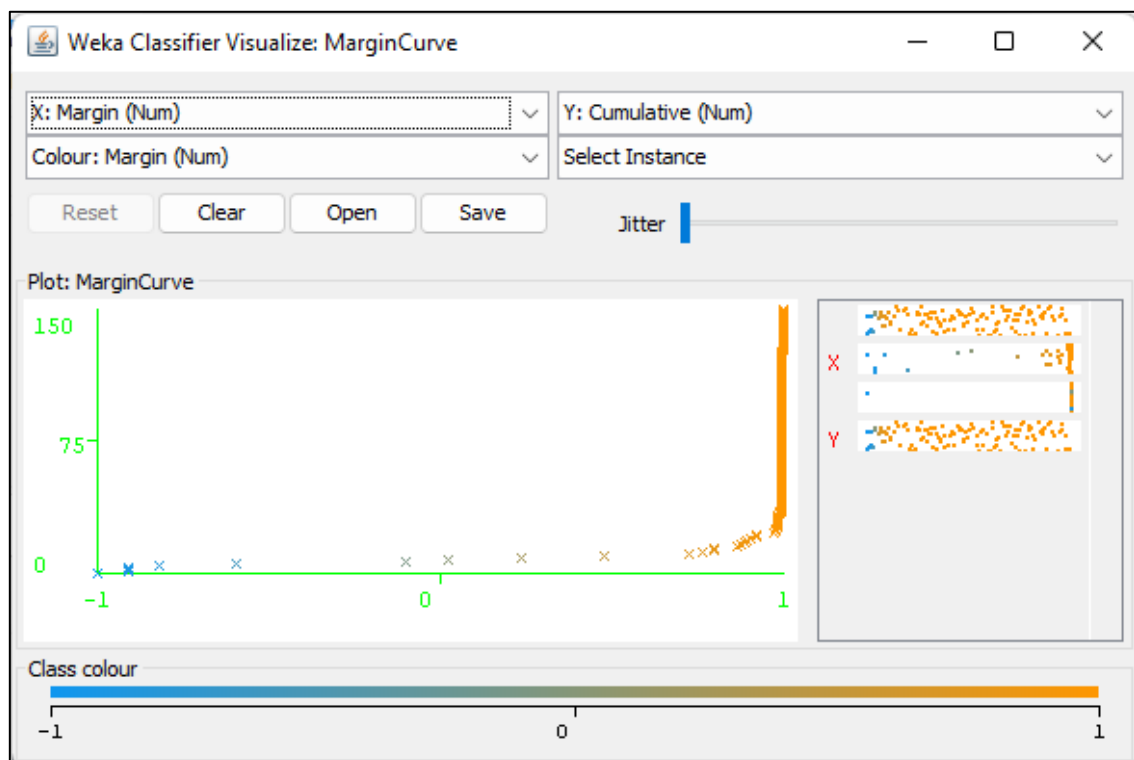
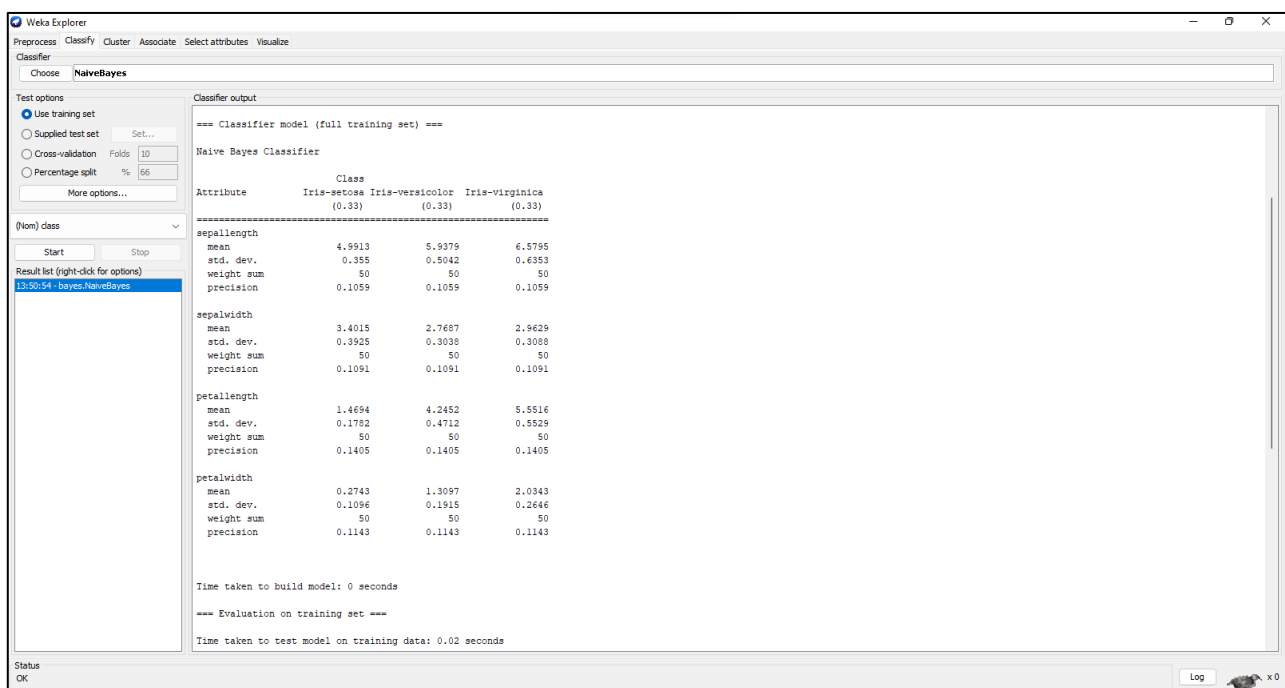
Batch: C

Subject: Data Warehousing & Mining

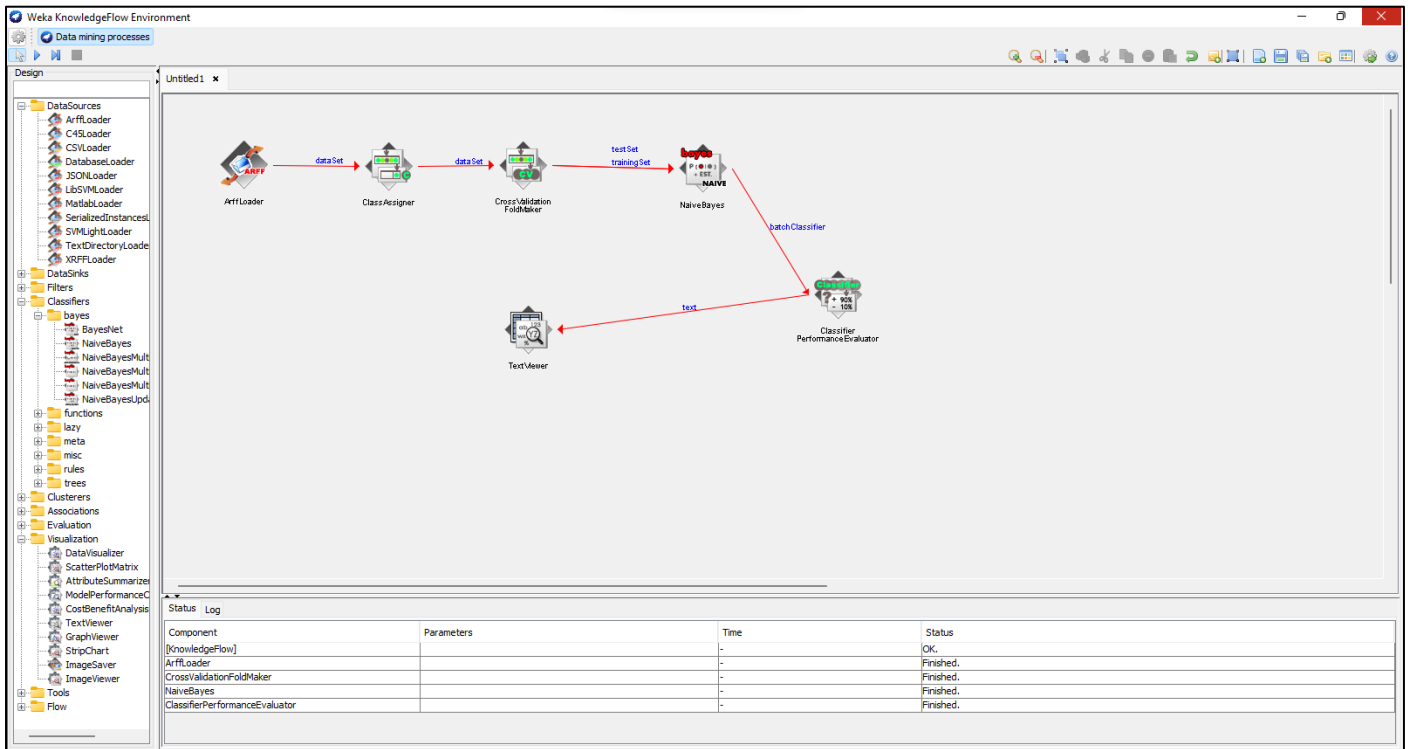
Experiment No.: 06

1. Weka Classification Using Naïve Bayesian:

a. Explorer



b. Knowledge Base



Text Viewer

Result list

14:32:14 - NaiveBayes

Text

=== Evaluation result ===

Scheme: NaiveBayes

Relation: breast-cancer

Metric	Value	Percentage
Correctly Classified Instances	205	71.6783 %
Incorrectly Classified Instances	81	28.3217 %
Kappa statistic	0.2857	
Mean absolute error	0.3272	
Root mean squared error	0.4534	
Relative absolute error	78.2036 %	
Root relative squared error	95.1872 %	
Coverage of cases (0.95 level)	97.5524 %	
Mean rel. region size (0.95 level)	91.2587 %	
Total Number of Instances	286	

=== Detailed Accuracy By Class ===

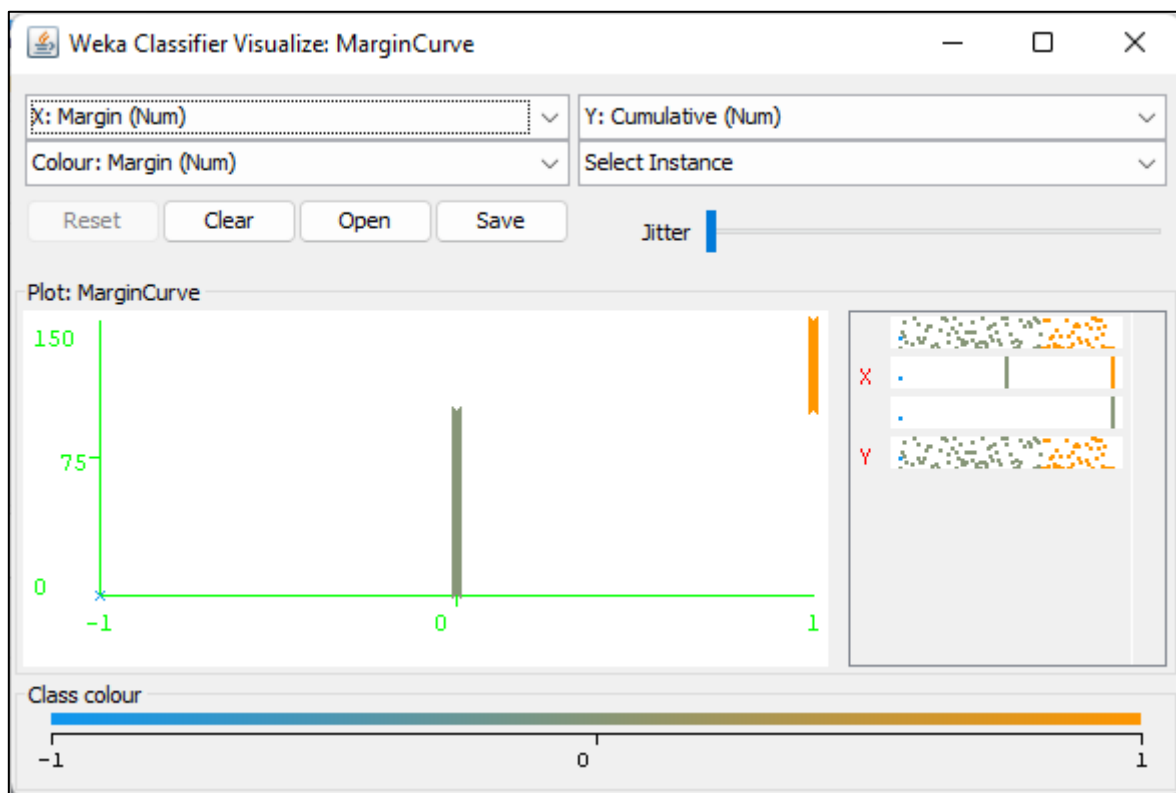
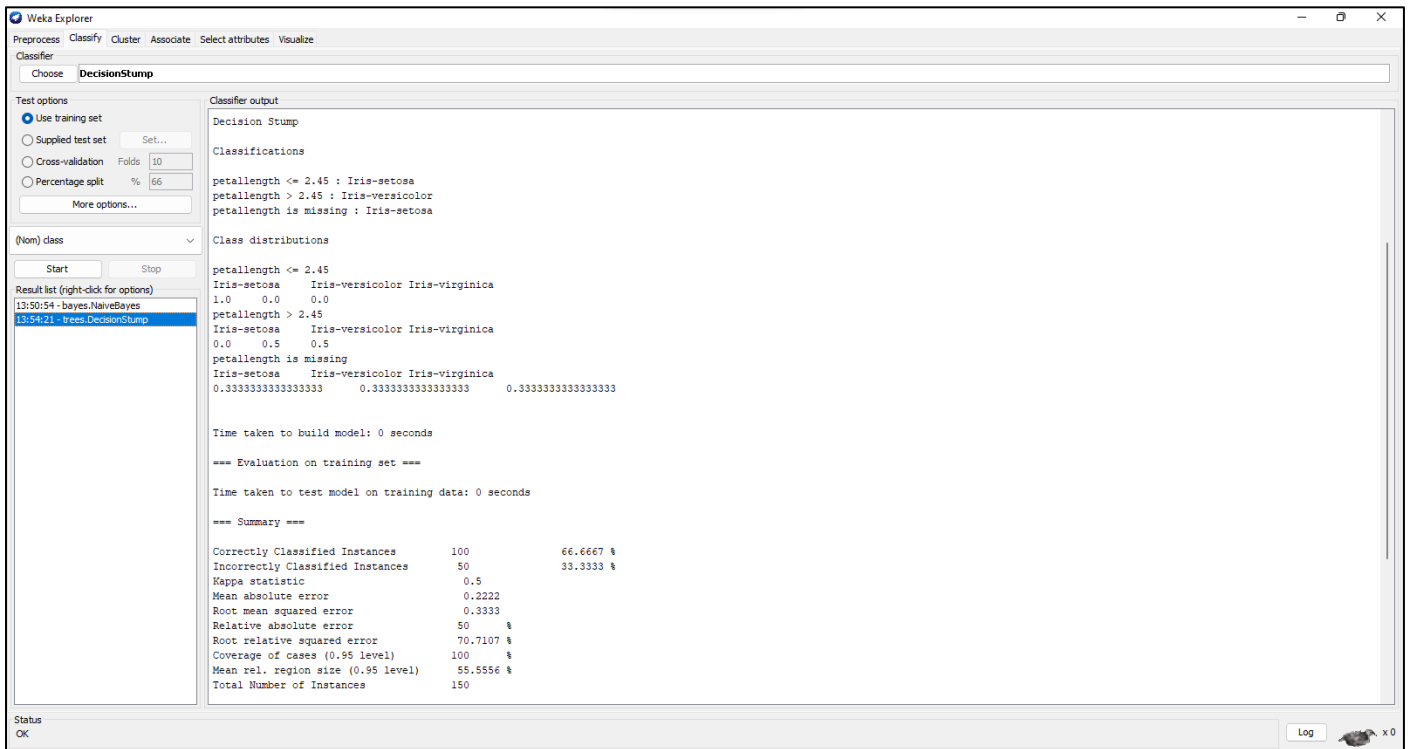
	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.836	0.565	0.778	0.836	0.806	0.288	0.701	0.837	no-recurrence-events
	0.435	0.164	0.529	0.435	0.477	0.288	0.701	0.514	recurrence-events
Weighted Avg.	0.717	0.446	0.704	0.717	0.708	0.288	0.701	0.741	

=== Confusion Matrix ===

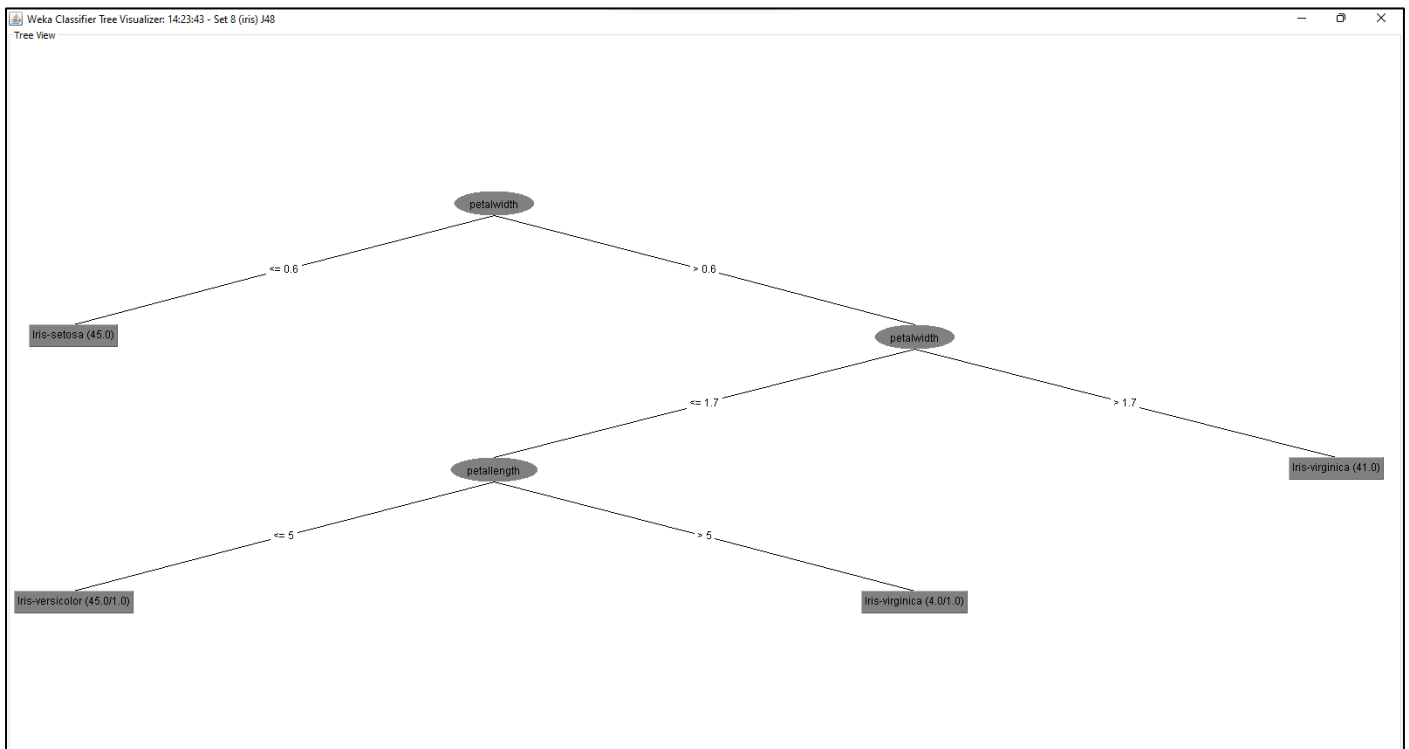
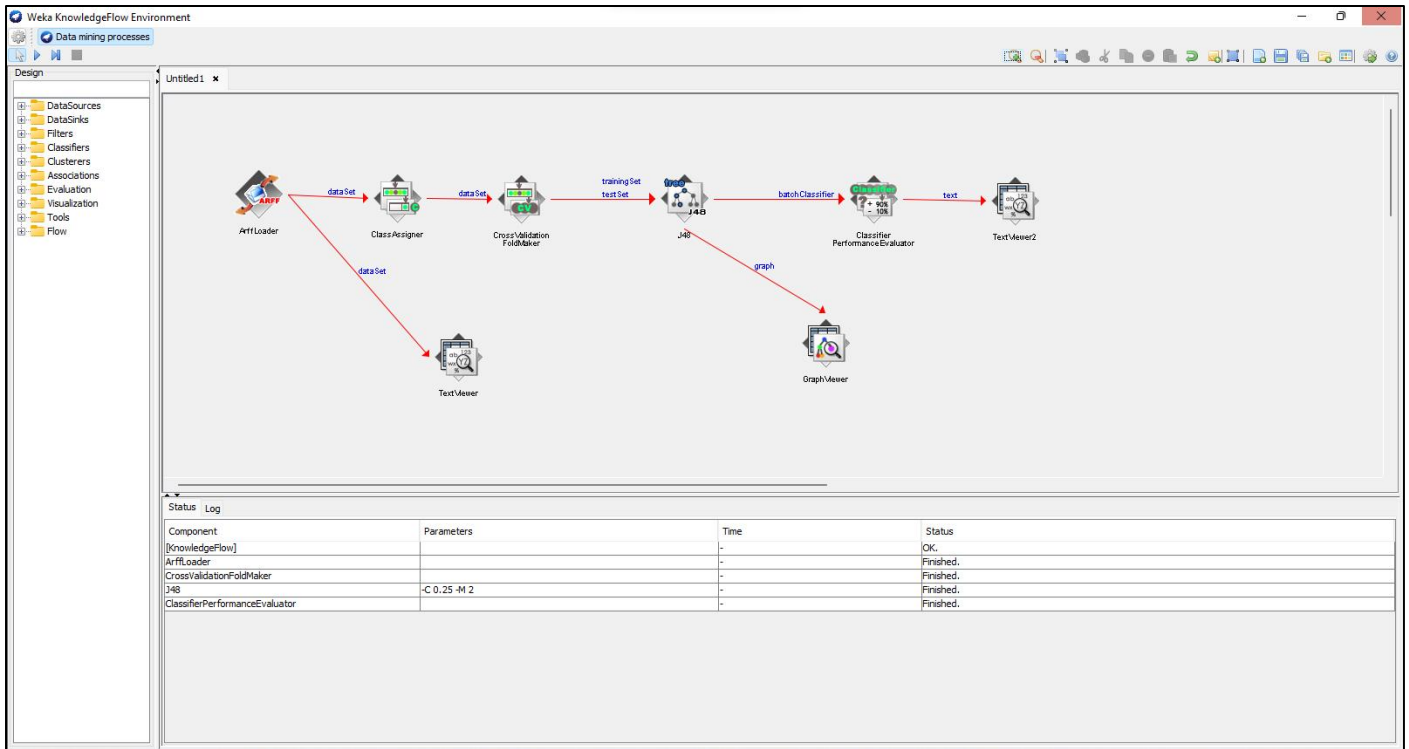
a \ b	no-recurrence-events	recurrence-events
no-recurrence-events	168	33
recurrence-events	48	37

2. Weka Classification Using Decision Tree

a. Explorer



b. Knowledge Base



```
Text Viewer
Result list
14:23:41 - #6
14:23:43 - #6

Text
@relation iris

@attribute sepal.length numeric
@attribute sepal.width numeric
@attribute petal.length numeric
@attribute petal.width numeric
@attribute class {Iris-setosa,Iris-versicolor,Iris-virginica}

@data
5.1,3.5,1.4,0.2,Iris-setosa
4.9,3.1,1.4,0.2,Iris-setosa
4.7,3.2,1.3,0.2,Iris-setosa
4.6,3.1,1.5,0.2,Iris-setosa
5.3,6.1,4.0,0.2,Iris-setosa
5.4,3.9,1.7,0.4,Iris-setosa
4.6,3.4,1.4,0.3,Iris-setosa
5.3,4.1,1.5,0.2,Iris-setosa
4.4,2.9,1.4,0.2,Iris-setosa
4.9,3.1,1.5,0.1,Iris-setosa
5.4,3.7,1.5,0.2,Iris-setosa
4.8,3.4,1.6,0.2,Iris-setosa
4.8,3.1,1.4,0.1,Iris-setosa
4.3,3.1,1.0,0.1,Iris-setosa
5.8,4.1,2.0,0.2,Iris-setosa
5.7,4.4,1.5,0.4,Iris-setosa
5.4,3.9,1.3,0.4,Iris-setosa
5.1,3.5,1.4,0.3,Iris-setosa
5.7,3.8,1.7,0.3,Iris-setosa
5.1,3.8,1.5,0.3,Iris-setosa
5.4,3.4,1.7,0.2,Iris-setosa
5.1,3.7,1.5,0.4,Iris-setosa
4.6,3.6,1.0,0.2,Iris-setosa
5.1,3.3,1.7,0.5,Iris-setosa
4.8,3.4,1.5,0.2,Iris-setosa
5.3,1.6,0.2,Iris-setosa
5.3,4.1,6.0,4,Iris-setosa
5.2,3.5,1.5,0.2,Iris-setosa
3.2,3.4,1.4,0.2,Iris-setosa
4.7,3.2,1.6,0.2,Iris-setosa
4.8,3.1,1.6,0.2,Iris-setosa
5.4,3.4,1.5,0.4,Iris-setosa
5.2,4.1,1.5,0.1,Iris-setosa
5.5,4.2,1.4,0.2,Iris-setosa
4.9,3.1,1.5,0.1,Iris-setosa
5.3,2.1,2.0,0.2,Iris-setosa
5.5,3.5,1.3,0.2,Iris-setosa
4.9,3.1,1.5,0.1,Iris-setosa
```

```
Graph Viewer
Graph list
14:18:57 - Set 2 (supermarket) J48
14:18:57 - Set 1 (supermarket) J48
14:18:57 - Set 4 (supermarket) J48
14:18:57 - Set 3 (supermarket) J48
14:18:58 - Set 5 (supermarket) J48
14:18:58 - Set 6 (supermarket) J48
14:18:58 - Set 7 (supermarket) J48
14:18:58 - Set 8 (supermarket) J48
14:18:58 - Set 9 (supermarket) J48
14:18:58 - Set 10 (supermarket) J48
14:23:43 - Set 1 (iris) J48
14:23:43 - Set 2 (iris) J48
14:23:43 - Set 4 (iris) J48
14:23:43 - Set 3 (iris) J48
14:23:43 - Set 5 (iris) J48
14:23:43 - Set 6 (iris) J48
14:23:43 - Set 8 (iris) J48
14:23:43 - Set 7 (iris) J48
14:23:43 - Set 10 (iris) J48
14:23:43 - Set 9 (iris) J48
```

```
Text Viewer
Result list
14:23:43 - J48

Text
=== Evaluation result ===

Scheme: J48
Options: -C 0.25 -M 2
Relation: supermarket

Correctly Classified Instances      2948           63.713 %
Incorrectly Classified Instances    1679           36.287 %
Kappa statistic                     0
Mean absolute error                 0.4624
Root mean squared error            0.4808
Relative absolute error             99.9961 %
Root relative squared error         100 %
Coverage of cases (0.95 level)     100 %
Mean rel. region size (0.95 level) 100 %
Total Number of Instances          4627

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
1.000    1.000    0.637    1.000    0.778    0.000    0.499    0.637    low
0.000    0.000    0.000    0.000    0.000    0.000    0.499    0.363    high
Weighted Avg.   0.637    0.637    0.406    0.637    0.496    0.000    0.499    0.537

=== Confusion Matrix ===
      a    b  <-- classified as
2948    0 |    a = low
1679    0 |    b = high
```

3. Weka clustering Using K-Means

a. Explorer

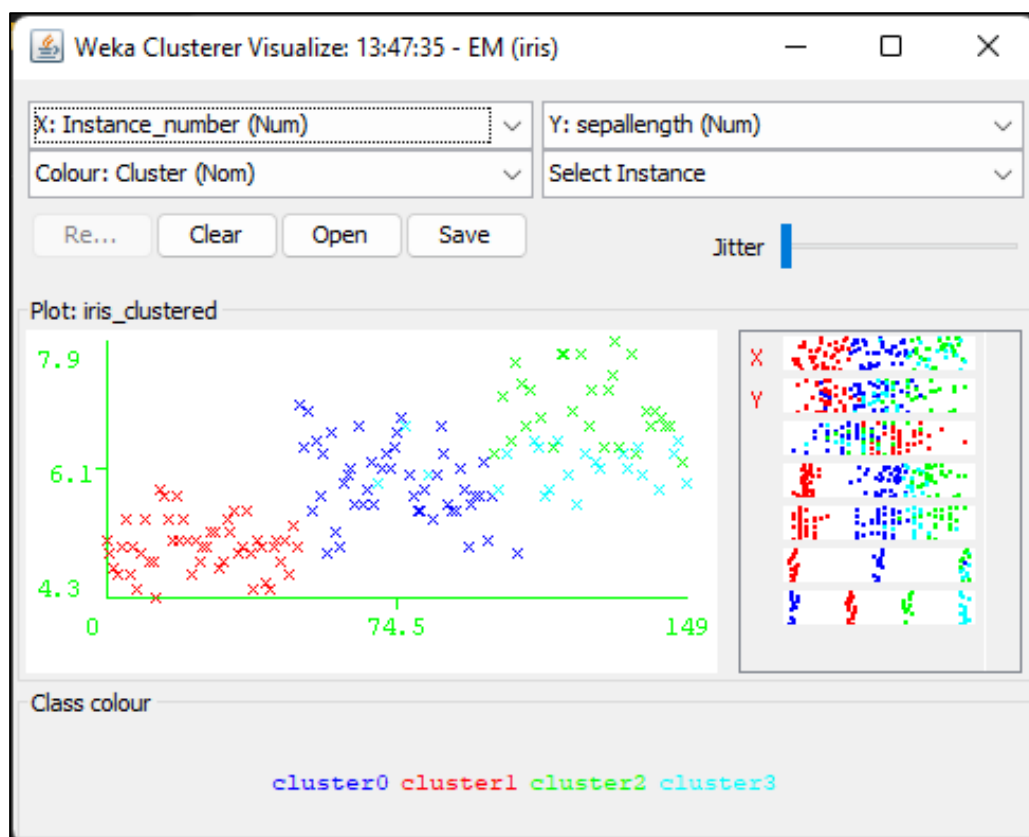
The screenshot shows the Weka Explorer window with the 'Cluster' tab selected. The 'Clusterer' dropdown is set to 'EM'. The 'Cluster mode' section has 'Use training set' selected. The 'Clusterer output' pane displays the following information:

EM
==

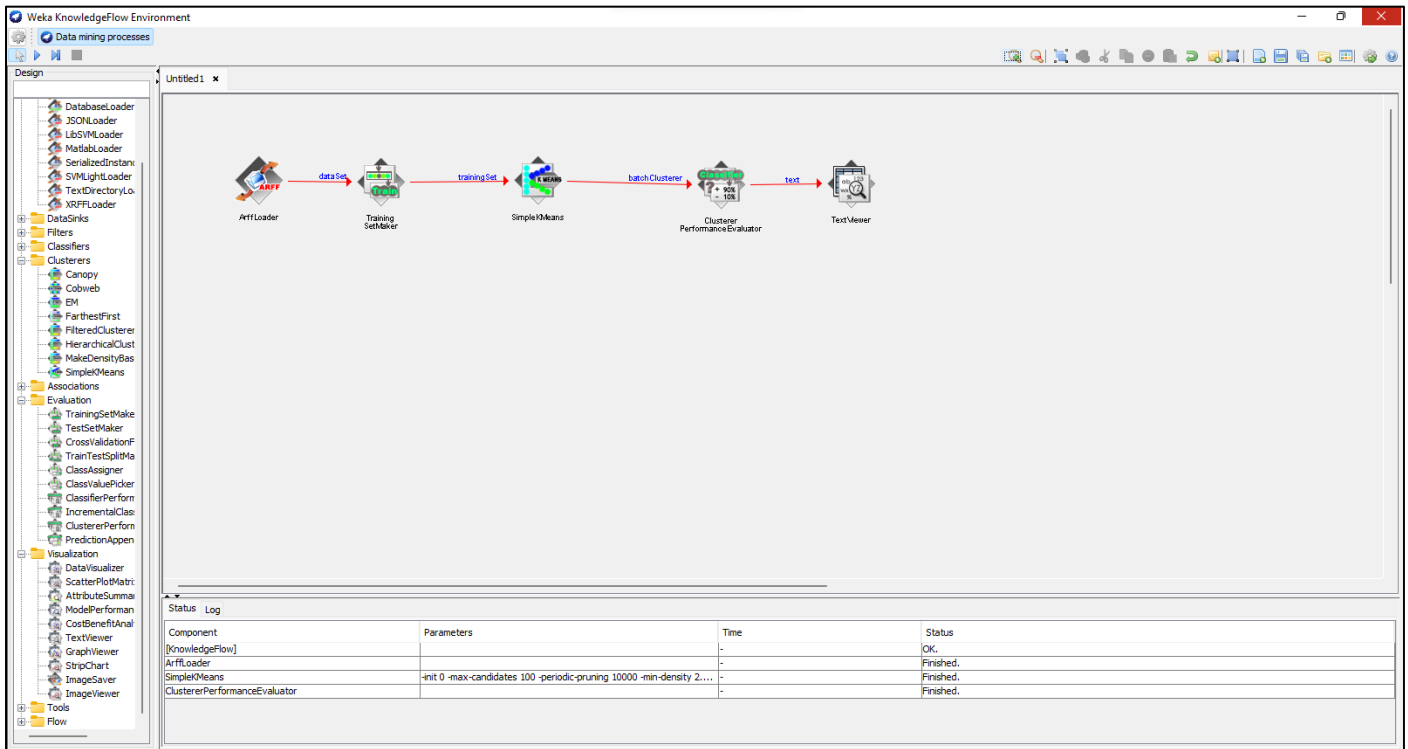
Number of clusters selected by cross validation: 4
Number of iterations performed: 16

Attribute	Cluster			
	0 (0.32)	1 (0.33)	2 (0.2)	3 (0.14)
=====				
sepal.length				
mean	5.897	5.006	6.9426	6.1304
std. dev.	0.5279	0.3489	0.498	0.2943
sepal.width				
mean	2.7519	3.418	3.1103	2.8088
std. dev.	0.3103	0.3772	0.2952	0.2361
petal.length				
mean	4.2267	1.464	5.8559	5.0993

The 'Result list' on the left shows '13:47:35 - EM' selected. The status bar at the bottom indicates 'OK'.



b. Knowledge Base



Text Viewer

Result list

14:39:22 - SimpleKMeans

Text

=== Evaluation result for training instances ===

Scheme: SimpleKMeans
Relation: breast-cancer

KMeans
=====

Number of iterations: 3
Within cluster sum of squared errors: 1177.0

Initial starting points (random):

Cluster 0: 50-59,premeno,10-14,0-2,no,2,right,left_up,no,no-recurrence-events
Cluster 1: 40-49,premeno,15-19,0-2,yes,3,right,left_up,no,recurrence-events

Missing values globally replaced with mean/mode

Final cluster centroids:

Attribute	Full Data (286.0)	Cluster# 0 (225.0)	1 (61.0)
age	50-59	50-59	40-49
menopause	premeno	premeno	premeno
tumor-size	30-34	25-29	30-34
inv-nodes	0-2	0-2	0-2
node-caps	no	no	yes
deg-malig	2	2	3
breast	left	left	left
breast-quad	left_low	left_low	left_low
irradiat	no	no	no
Class	no-recurrence-events	no-recurrence-events	recurrence-events

Clustered Instances

Cluster	Count	Percentage
0	225	(79%)
1	61	(21%)

4. Weka Association Using Apriori

a. Explorer

The screenshot shows the Weka Explorer interface with the Apriori algorithm selected. The 'Associate' tab is active, and the 'Apriori' model is displayed. The 'Choose' dropdown shows the dataset 'Apriori-N 10-T 0-C 0.9-D 0.05-U 1.0-M 0.1-S 1.0-c-1'. The 'Start' button is highlighted. The 'Result list (right-click...)' shows '13:55:08 - Apriori'. The 'Associator output' pane displays the following text:

```
node=caps
deg=mailg
breast
breast=quad
irradiat
Class
=== Associator model (full training set) ===

Apriori
=====

Minimum support: 0.5 (143 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 10

Generated sets of large itemsets:

Size of set of large itemsets L(1): 6
Size of set of large itemsets L(2): 6
Size of set of large itemsets L(3): 4
Size of set of large itemsets L(4): 1

Best rules found:

1. inv-nodes=0-2 irradiat=no Class=no-recurrence-events 147 ==> node=caps=no 145 <conf:(0.99)> lift:(1.27) lev:(0.11) [30] conv:(10.97)
2. inv-nodes=0-2 irradiat=no 183 ==> node=caps=no 177 <conf:(0.97)> lift:(1.25) lev:(0.12) [34] conv:(5.85)
3. node=caps=no irradiat=no Class=no-recurrence-events 151 ==> inv-nodes=0-2 145 <conf:(0.96)> lift:(1.25) lev:(0.11) [32] conv:(5.51)
4. inv-nodes=0-2 Class=no-recurrence-events 167 ==> node=caps=no 160 <conf:(0.96)> lift:(1.23) lev:(0.11) [30] conv:(4.67)
5. inv-nodes=0-2 213 ==> node=caps=no 201 <conf:(0.94)> lift:(1.22) lev:(0.12) [35] conv:(3.67)
6. node=caps=no irradiat=no 188 ==> inv-nodes=0-2 177 <conf:(0.94)> lift:(1.26) lev:(0.13) [36] conv:(4)
7. node=caps=no Class=no-recurrence-events 171 ==> inv-nodes=0-2 160 <conf:(0.94)> lift:(1.26) lev:(0.11) [32] conv:(3.64)
8. irradiat=no Class=no-recurrence-events 164 ==> node=caps=no 151 <conf:(0.92)> lift:(1.19) lev:(0.08) [23] conv:(2.62)
9. inv-nodes=0-2 node=caps=no Class=no-recurrence-events 160 ==> irradiat=no 145 <conf:(0.91)> lift:(1.19) lev:(0.08) [23] conv:(2.38)
10. node=caps=no 222 ==> inv-nodes=0-2 201 <conf:(0.91)> lift:(1.22) lev:(0.12) [35] conv:(2.58)
```

The 'Status' bar at the bottom shows 'OK'.

The screenshot shows the Weka Explorer interface with the Apriori algorithm selected. The 'Associate' tab is active, and the 'Apriori' model is displayed. The 'Choose' dropdown shows the dataset 'Apriori-N 10-T 0-C 0.9-D 0.05-U 1.0-M 0.1-S 1.0-c-1'. The 'Start' button is highlighted. The 'Result list (right-click...)' shows '14:57:28 - Apriori'. The 'Associator output' pane displays the following text:

```
Attributes: 217
[list of attributes omitted]
=== Associator model (full training set) ===

Apriori
=====

Minimum support: 0.15 (694 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 17

Generated sets of large itemsets:

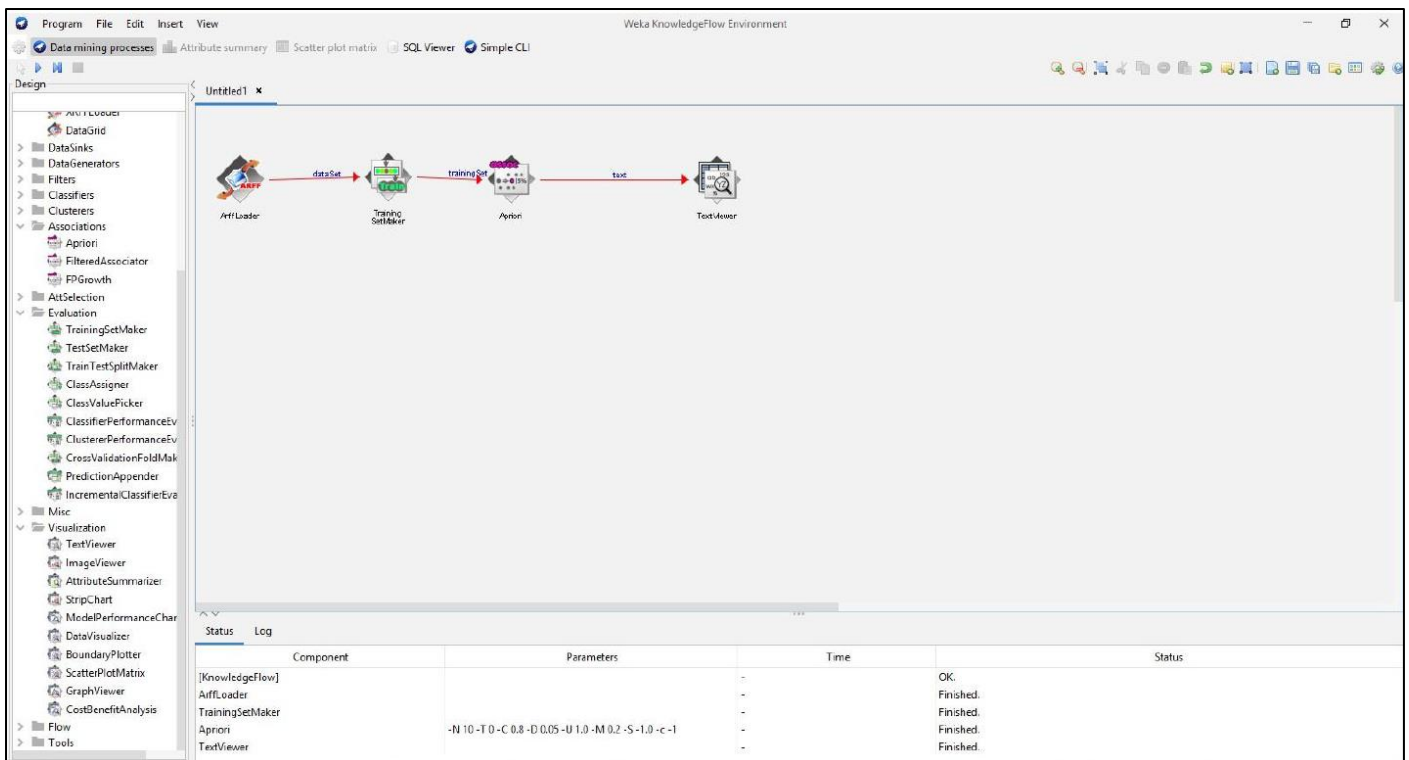
Size of set of large itemsets L(1): 44
Size of set of large itemsets L(2): 380
Size of set of large itemsets L(3): 910
Size of set of large itemsets L(4): 633
Size of set of large itemsets L(5): 105
Size of set of large itemsets L(6): 1

Best rules found:

1. biscuits=t frozen foods=t fruit=t total-high 788 ==> bread and cake=t 723 <conf:(0.92)> lift:(1.27) lev:(0.03) [155] conv:(3.35)
2. baking needs=t biscuits=t fruit=t total-high 760 ==> bread and cake=t 696 <conf:(0.92)> lift:(1.27) lev:(0.03) [149] conv:(3.28)
3. baking needs=t frozen foods=t fruit=t total-high 770 ==> bread and cake=t 705 <conf:(0.92)> lift:(1.27) lev:(0.03) [150] conv:(3.27)
4. biscuits=t fruit=t vegetables=t total-high 815 ==> bread and cake=t 746 <conf:(0.92)> lift:(1.27) lev:(0.03) [159] conv:(3.26)
5. party snack foods=t fruit=t total-high 854 ==> bread and cake=t 779 <conf:(0.91)> lift:(1.27) lev:(0.04) [164] conv:(3.15)
6. biscuits=t frozen foods=t vegetables=t total-high 797 ==> bread and cake=t 725 <conf:(0.91)> lift:(1.26) lev:(0.03) [151] conv:(3.06)
7. baking needs=t biscuits=t vegetables=t total-high 772 ==> bread and cake=t 701 <conf:(0.91)> lift:(1.26) lev:(0.03) [145] conv:(3.01)
8. biscuits=t fruit=t total-high 954 ==> bread and cake=t 866 <conf:(0.91)> lift:(1.26) lev:(0.04) [179] conv:(3)
9. frozen foods=t fruit=t vegetables=t total-high 834 ==> bread and cake=t 757 <conf:(0.91)> lift:(1.26) lev:(0.03) [156] conv:(3)
10. frozen foods=t fruit=t total-high 969 ==> bread and cake=t 877 <conf:(0.91)> lift:(1.26) lev:(0.04) [179] conv:(2.92)
```

The 'Status' bar at the bottom shows 'OK'.

b. Knowledge Base



Text Viewer

Result list:

- 15:05:08.082 - Model: Apriori
- 15:06:46.715 - Model: Apriori

Text

=== Associator model ===

Scheme: Apriori

Relation: supermarket

Apriori

=====

Minimum support: 0.3 (1389 instances)
Minimum metric (confidence): 0.8
Number of cycles performed: 14

Generated sets of large itemsets:

Size of set of large itemsets L(1): 25
Size of set of large itemsets L(2): 69
Size of set of large itemsets L(3): 20

Best rules found:

1. biscuits=vegetables=1764 ==> bread and cake=1497 <conf:(0.84)> lift:(1.17) lev:(0.05) [217] conv:(1.78)
2. cereal-high 1879 ==> bread and cake=1413 <conf:(0.84)> lift:(1.17) lev:(0.04) [204] conv:(1.75)
3. biscuits=milk-cream=1767 ==> bread and cake=1488 <conf:(0.84)> lift:(1.17) lev:(0.05) [213] conv:(1.75)
4. biscuits=fruit=1837 ==> bread and cake=1541 <conf:(0.84)> lift:(1.17) lev:(0.05) [218] conv:(1.73)
5. biscuits=frozen foods=1810 ==> bread and cake=1510 <conf:(0.83)> lift:(1.16) lev:(0.04) [207] conv:(1.69)
6. frozen foods=fruit=1861 ==> bread and cake=1548 <conf:(0.83)> lift:(1.16) lev:(0.05) [208] conv:(1.66)
7. frozen foods=milk-cream=1826 ==> bread and cake=1516 <conf:(0.83)> lift:(1.15) lev:(0.04) [201] conv:(1.65)
8. baking needs=milk-cream=1907 ==> bread and cake=1580 <conf:(0.83)> lift:(1.15) lev:(0.04) [207] conv:(1.63)
9. milk-cream=fruit=2038 ==> bread and cake=1684 <conf:(0.83)> lift:(1.15) lev:(0.05) [217] conv:(1.61)
10. baking needs=biscuits=1764 ==> bread and cake=1456 <conf:(0.83)> lift:(1.15) lev:(0.04) [186] conv:(1.6)

Close Settings Clear results