



Daffodil International University

Lab Report

Course Code: CSE 322

Course Title: Data Mining and Machine Learning Lab.

Submitted to:

Nasima Islam Bithi

Lecturer

Department of CSE, DIU.

Submitted by:

Md Ashikur Rahman

ID:201-15-3768

Section: PC-R

Department of CSE

Daffodil International University.

First, I downloaded the “vote” dataset from website. In this dataset :17 attribute with class attribute. In this dataset the output attribute called class, has two values: ‘democrat’, and ‘republican’ showing the vote of a voter.

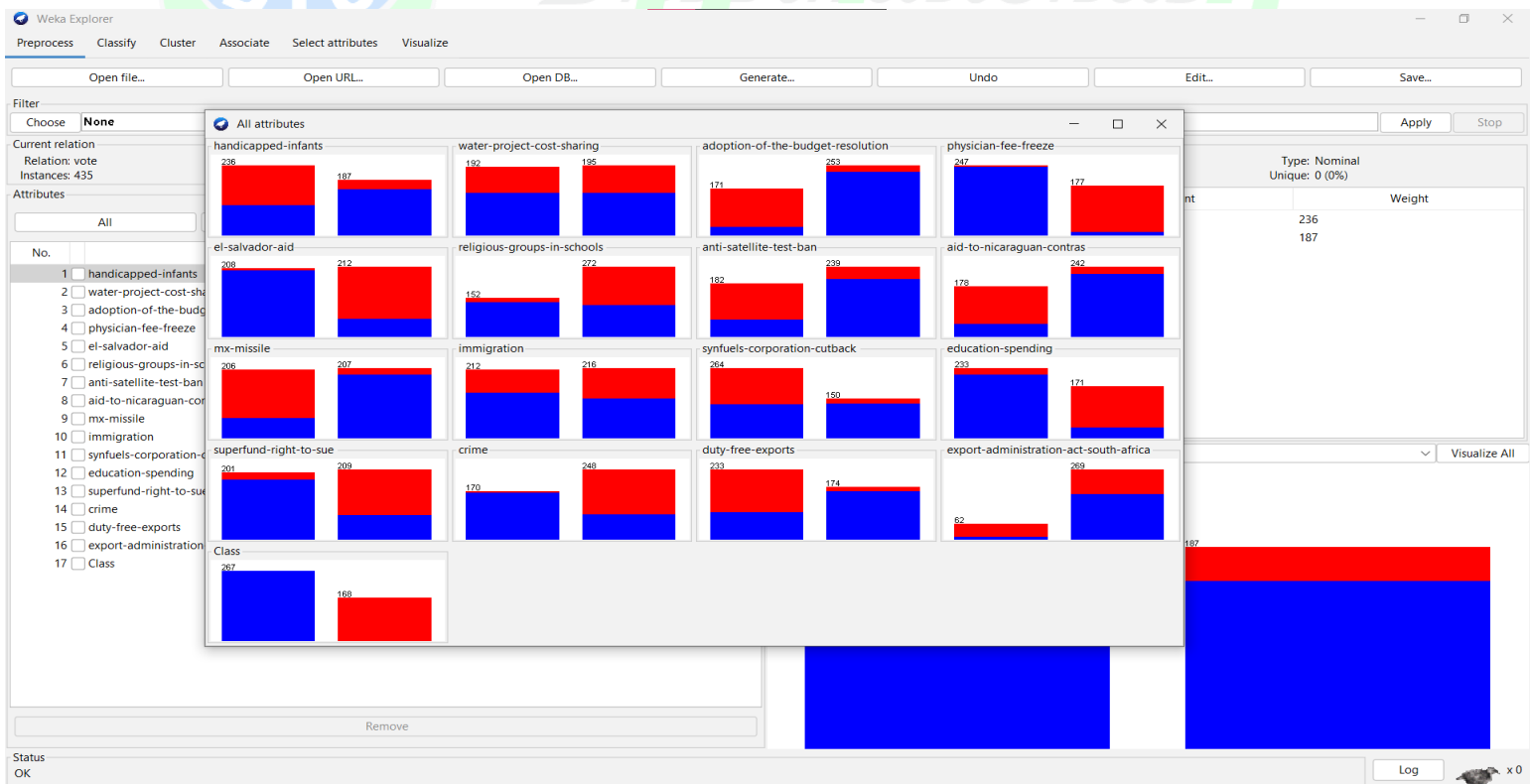
I use Weka Data Mining software.

Step:1

Data Pre-processing:

I open the “vote” dataset in Weka data mining software. After downloading the dataset of the congressional voting record of 1984, I load the dataset (arff format) on Weka explorer.

*Data Visualization:



*Missing Value:

In this dataset there are some missing values.

The screenshot shows the Weka Explorer interface. The 'Preprocess' tab is active. The 'Filter' dropdown is set to 'None'. The 'Current relation' is 'vote' with 435 instances. The 'Attributes' list on the left includes: 1. handicapped-infants, 2. water-project-cost-sharing, 3. adoption-of-the-budget-resolution, 4. physician-fee-freeze, 5. el-salvador-aid, 6. religious-groups-in-schools, 7. anti-satellite-test-ban, 8. aid-to-nicaraguan-contras, 9. mx-missile, 10. immigration, 11. synfuels-corporation-cutbs, 12. education-spending, 13. superfund-right-to-sue, 14. crime, 15. duty-free-exports, 16. export-administration-act, and 17. Class. The 'Viewer' window is open, showing a table with 23 instances and 7 attributes. The 'Class' attribute is highlighted in the left sidebar. The table shows missing values (greyed out) for several instances. The 'Status' bar at the bottom indicates 'OK'.

No.	1: handicapped-infants	2: water-project-cost-sharing	3: adoption-of-the-budget-resolution	4: physician-fee-freeze	5: el-salvador-aid	6: religious-groups-in-schools	7: anti-satellite-test-ban
1	n	y	n	y	y	y	n
2	n	y	n	y	y	y	n
3		y	y		y	y	n
4	n	y	y	n		y	n
5	y	y	y	n	y	y	n
6	n	y	y	n	y	y	n
7	n	y	n	y	y	y	n
8	n	y	n	y	y	y	n
9	n	y	n	y	y	y	n
10	y	y	y	n	n	n	y
11	n	y	n	y	y	n	n
12	n	y	n	y	y	y	n
13	n	y	y	n	n	n	y
14	y	y	y	n	n	y	y
15	n	y	n	y	y	y	n
16	n	y	n	y	y	y	n
17	y	n	y	n	n	y	n
18	y		y	n	n	n	y
19	n	y	n	y	y	y	n
20	y	y	y	n	n	n	y
21	y	y	y	n	n		y
22	y	y	y	n	n	n	y
23	y		y	n	n	n	y

*Handle Missing Value:

We can handle missing value with some ways:

- *ReplaceMissingValues
- * ReplaceMissingValuesWithUserConstant
- * ReplaceWithMissingValues

Hare I use **ReplaceMissingValues Algorithm** for remove missing values.

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

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

Filter

Choose **ReplaceMissingValues** Apply Stop

Current relation: vote-weka.filters.unsupervised.attribute.ReplaceMissingValues
Instances: 435

Attributes

All N

No. 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

☒ 1 handicapped-infants ☐ 2 water-project-cost-sharing ☐ 3 adoption-of-the-budget-resolution ☐ 4 physician-fee-freeze ☐ 5 el-salvador-aid ☐ 6 religious-groups-in-schools ☐ 7 anti-satellite-test-ban ☐ 8 aid-to-nicaraguan-contras ☐ 9 mx-missile ☐ 10 immigration ☐ 11 synfuels-corporation-cutback ☐ 12 education-spending ☐ 13 superfund-right-to-sue ☐ 14 crime ☐ 15 duty-free-exports ☐ 16 export-administration-act-south-africa ☐ 17 Class

Viewer

Relation: vote-weka.filters.unsupervised.attribute.ReplaceMissingValues

No.	1: handicapped-infants	2: water-project-cost-sharing	3: adoption-of-the-budget-resolution	4: physician-fee-freeze	5: el-salvador-aid	6: religious-groups-in-schools	7: anti-satellite-test-ban	8: aid-to-nicaraguan-contras	9: mx-missile	10: immigration	11: synfuels-corporation-cutback	12: education-spending	13: superfund-right-to-sue	14: crime	15: duty-free-exports	16: export-administration-act-south-africa	17: Class
1	n	y	n	y	y	y	n	n	n	n	n	n	n	n	n	n	n
2	n	y	n	y	y	y	n	n	n	n	n	n	n	n	n	n	n
3	n	y	y	n	y	y	n	n	n	n	n	n	n	n	n	n	n
4	n	y	y	n	y	y	n	n	n	n	n	n	n	n	n	n	n
5	y	y	y	n	y	y	n	n	n	n	n	n	n	n	n	n	n
6	n	y	y	n	y	y	n	n	n	n	n	n	n	n	n	n	n
7	n	y	n	y	y	y	n	n	n	n	n	n	n	n	n	n	n
8	n	y	n	y	y	y	n	n	n	n	n	n	n	n	n	n	n
9	n	y	n	y	y	y	n	n	n	n	n	n	n	n	n	n	n
10	y	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n
11	n	y	n	y	y	n	n	n	n	n	n	n	n	n	n	n	n
12	n	y	n	y	y	y	n	n	n	n	n	n	n	n	n	n	n
13	n	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n
14	y	y	y	n	n	y	n	n	n	n	n	n	n	n	n	n	n
15	n	y	n	y	y	y	n	n	n	n	n	n	n	n	n	n	n
16	n	y	n	y	y	y	n	n	n	n	n	n	n	n	n	n	n
17	y	n	y	n	n	y	n	n	n	n	n	n	n	n	n	n	n
18	y	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n
19	n	y	n	y	y	y	n	n	n	n	n	n	n	n	n	n	n
20	y	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n
21	y	y	y	n	n	y	n	n	n	n	n	n	n	n	n	n	n
22	y	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n
23	y	y	y	n	n	n	n	n	n	n	n	n	n	n	n	n	n

Type: Nominal
Unique: 0 (0%)

248
187

Visualize All

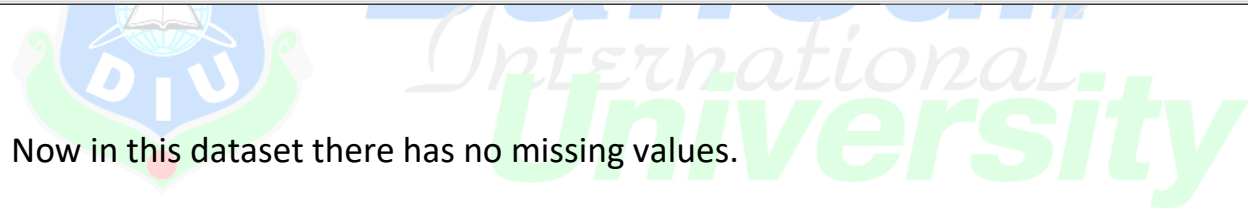
Add instance Undo OK Cancel

Remove

Status
OK

Log x 0

Now in this dataset there has no missing values.



*Duplicate Value:

After load dataset, I saw that it has some Duplicates values. In this time, I solve these things by using the **RemoveDuplicate Algorithm**.

Before Remove Duplicate Values:

The screenshot shows the Weka Explorer interface with the 'ReplaceMissingValues' filter applied to the 'handicapped-infants' dataset. The interface is divided into several sections:

- Filter:** Shows the 'ReplaceMissingValues' filter applied to the 'handicapped-infants' dataset. The current relation is 'vote-weka.filters.unsupervised.attribute.ReplaceMissingValues' with 435 instances.
- Attributes:** A list of 17 attributes is shown, including 'handicapped-infants', 'water-project-cost-sharing', 'adoption-of-the-budget-resolution', 'physician-fee-freeze', 'el-salvador-aid', 'religious-groups-in-schools', 'anti-satellite-test-ban', 'aid-to-nicaraguan-contras', 'mx-missile', 'immigration', 'synfuels-corporation-cutback', 'education-spending', 'superfund-right-to-sue', 'crime', 'duty-free-exports', 'export-administration-act-south-africa', and 'Class'.
- Selected attribute:** A table showing the selected attribute 'handicapped-infants' with 2 distinct values (n and y) and 0 missing values (0%).
- Class: Class (Nom):** A bar chart visualization showing the distribution of the 'Class' attribute. The chart has two bars: one for 'n' (248 instances) and one for 'y' (187 instances). The 'n' bar is red and the 'y' bar is blue.

No.	Label	Count	Weight
1	n	248	248
2	y	187	187

After Remove Duplicate Values:

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

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

Filter

Choose **RemoveDuplicates** Apply Stop

Current relation

Relation: vote-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.instance.RemoveD... Attributes: 17
Instances: 281 Sum of weights: 281

Attributes

All None Invert Pattern

No.	Name
1	<input checked="" type="checkbox"/> handicapped-infants
2	<input type="checkbox"/> water-project-cost-sharing
3	<input type="checkbox"/> adoption-of-the-budget-resolution
4	<input type="checkbox"/> physician-fee-freeze
5	<input type="checkbox"/> el-salvador-aid
6	<input type="checkbox"/> religious-groups-in-schools
7	<input type="checkbox"/> anti-satellite-test-ban
8	<input type="checkbox"/> aid-to-nicaraguan-contras
9	<input type="checkbox"/> mx-missile
10	<input type="checkbox"/> immigration
11	<input type="checkbox"/> synfuels-corporation-cutback
12	<input type="checkbox"/> education-spending
13	<input type="checkbox"/> superfund-right-to-sue
14	<input type="checkbox"/> crime
15	<input type="checkbox"/> duty-free-exports
16	<input type="checkbox"/> export-administration-act-south-africa
17	<input type="checkbox"/> Class

Remove

Selected attribute

Name: handicapped-infants
Missing: 0 (0%)
Distinct: 2
Type: Nominal
Unique: 0 (0%)

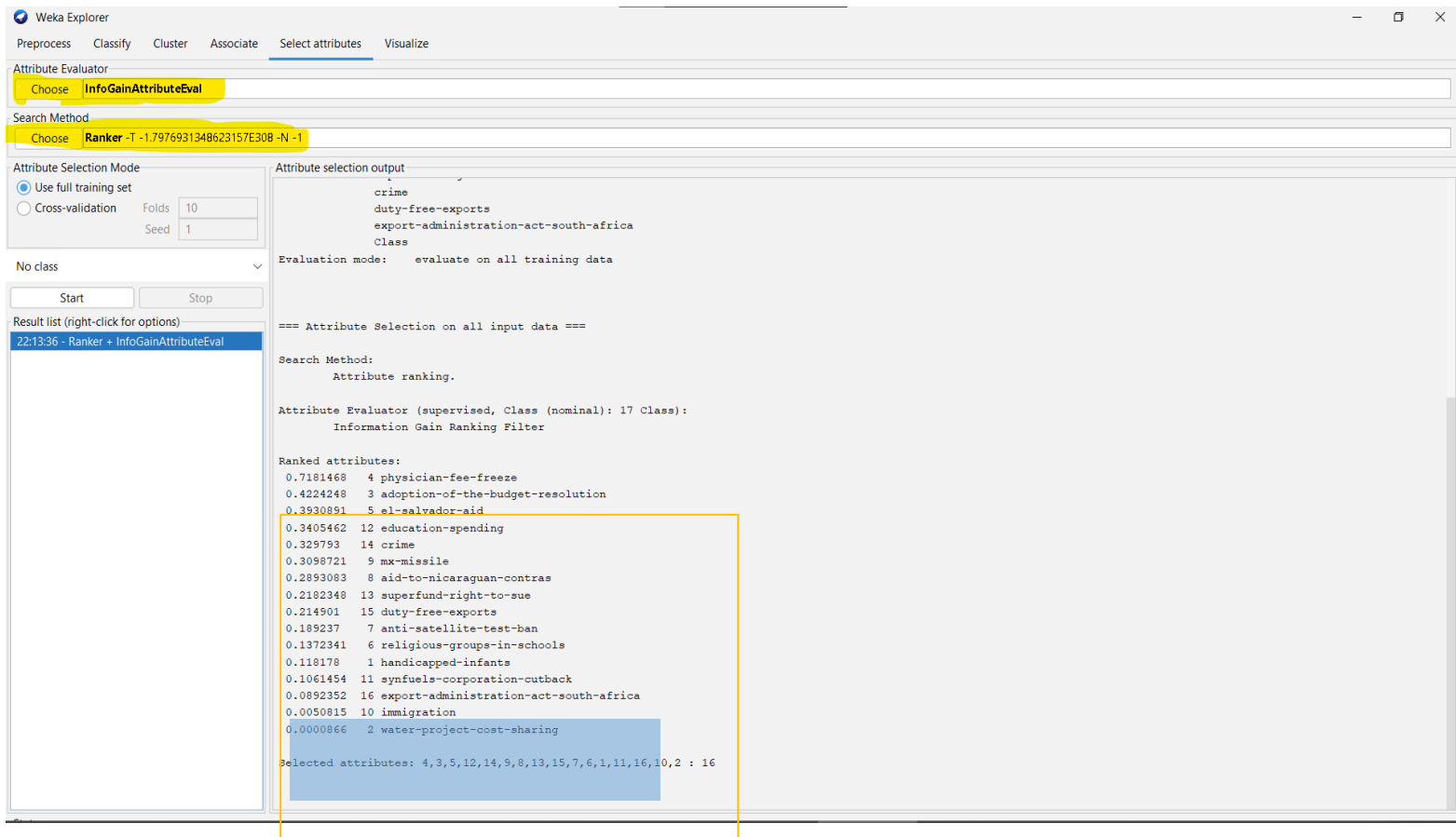
No.	Label	Count	Weight
1	n	159	159
2	y	122	122

Class: Class (Nom) Visualize All

Status OK Log x 0

*Select the 12 best Attributes:

We can find our best attributes from “Select attributes” option.



Here I use **Algorithm: InfoGainAttributeEval** and **SelectMethod: Ranker** for select best attribute.

We can remove:

1. handicapped-infants
11. synfuels-corporation-cutback
16. export-administration-act-south-africa
10. immigration
2. water-project-cost-sharing

These attribute from dataset. Because they are less priorities.

***Remove Attributes:**

After applying the methods, we find out best 12 attributes. They are 4,3,5,14,12,9,8,15,11,13,16,7 no attribute. Now I am going to remove other attributes form the dataset. After remove the low-value attribute the data set will be like this:

The screenshot shows the Weka Explorer application window. The 'Preprocess' tab is active, and the 'ReplaceMissingValues' filter is applied. The 'Attributes' list on the left shows 17 attributes, with 'water-project-cost-sharing' selected. The 'Selected attribute' panel on the right displays the attribute's details: 'Name: water-project-cost-sharing', 'Type: Nominal', 'Missing: 0 (0%)', and 'Distinct: 2'. Below this, a bar chart visualizes the distribution of the attribute, showing two categories: 'n' (192 instances) and 'y' (243 instances). The bars are colored red and blue respectively.

Weka Explorer

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

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

Filter: Choose **ReplaceMissingValues** [Apply] [Stop]

Current relation
Relation: vote-weka.filters.unsupervised.attribute.ReplaceMissingValues
Instances: 435

Attributes: 17
Sum of weights: 435

Attributes: All | None | Invert | Pattern

No.	Name
1	<input checked="" type="checkbox"/> handicapped-infants
2	<input checked="" type="checkbox"/> water-project-cost-sharing
3	<input type="checkbox"/> adoption-of-the-budget-resolution
4	<input type="checkbox"/> physician-fee-freeze
5	<input type="checkbox"/> el-salvador-aid
6	<input type="checkbox"/> religious-groups-in-schools
7	<input type="checkbox"/> anti-satellite-test-ban
8	<input type="checkbox"/> aid-to-nicaraguan-contras
9	<input type="checkbox"/> mx-missile
10	<input checked="" type="checkbox"/> immigration
11	<input checked="" type="checkbox"/> synfuels-corporation-cutback
12	<input type="checkbox"/> education-spending
13	<input type="checkbox"/> superfund-right-to-sue
14	<input type="checkbox"/> crime
15	<input type="checkbox"/> duty-free-exports
16	<input checked="" type="checkbox"/> export-administration-act-south-africa
17	<input type="checkbox"/> Class

Remove

Selected attribute
Name: water-project-cost-sharing
Missing: 0 (0%)
Distinct: 2
Type: Nominal
Unique: 0 (0%)

No.	Label	Count	Weight
1	n	192	192
2	y	243	243

Class: Class (Nom) [Visualize All]

192 243

Status OK Log x 0

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

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

Filter
Choose: **ReplaceMissingValues** Apply Stop

Current relation
Relation: vote-weka.filters.unsupervised.attribute.ReplaceMissingValues-weka.filters.unsupervised.attribute.Remove... Attributes: 12
Instances: 435 Sum of weights: 435

Attributes
All None Invert Pattern

No.	Name
1	<input checked="" type="checkbox"/> adoption-of-the-budget-resolution
2	<input type="checkbox"/> physician-fee-freeze
3	<input type="checkbox"/> el-salvador-aid
4	<input type="checkbox"/> religious-groups-in-schools
5	<input type="checkbox"/> anti-satellite-test-ban
6	<input type="checkbox"/> aid-to-nicaraguan-contras
7	<input type="checkbox"/> mx-missile
8	<input type="checkbox"/> education-spending
9	<input type="checkbox"/> superfund-right-to-sue
10	<input type="checkbox"/> crime
11	<input type="checkbox"/> duty-free-exports
12	<input type="checkbox"/> Class

Remove

Selected attribute
Name: adoption-of-the-budget-resolution
Missing: 0 (0%) Distinct: 2 Type: Nominal
Unique: 0 (0%)

No.	Label	Count	Weight
1	n	171	171
2	y	264	264

Class: Class (Nom) Visualize All

Status
OK Log x 0

In this dataset there are 12 best Attributes.

Step:2

Apply precisely 4 different classification Algorithms

*Apply 1st Algorithm: J48 (Without Changing Parameter)

In classification we apply some algorithm for find our best model. We need high accuracy model.

The screenshot shows the Weka Explorer interface with the J48 classifier selected. The 'Test options' section shows 'Percentage split' at 80%. The 'Classifier output' section displays the following results:

```
Classifier output
| | | | water-project-cost-sharing = n: republican (2.0)
| | | | water-project-cost-sharing = y: democrat (4.0)
| | mx-missile = y: democrat (5.0/1.0)

Number of Leaves : 6
Size of the tree : 11

Time taken to build model: 0.06 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

=== Summary ===

Correctly Classified Instances 82 94.2529 %
Incorrectly Classified Instances 5 5.7471 %
Kappa statistic 0.8821
Mean absolute error 0.0926
Root mean squared error 0.23
Relative absolute error 19.4124 %
Root relative squared error 46.8597 %
Total Number of Instances 87

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.923 0.029 0.980 0.923 0.950 0.884 0.947 0.950 democrat
0.971 0.077 0.895 0.971 0.932 0.884 0.947 0.881 republican
Weighted Avg. 0.943 0.048 0.945 0.943 0.943 0.884 0.947 0.922

=== Confusion Matrix ===

a b <-- classified as
48 4 | a = democrat
1 34 | b = republican
```

The 'Result list' on the left shows '00:55:58 - trees.J48' selected. The status bar at the bottom indicates 'Status OK'.

First, I apply the **Tree algorithm: J84**. I use 80% data for training and 20% data for testing.

***Algorithm: J48 (tree)**

***Accuracy: 94.2529 %**

Without changing any parameter.

***J48 (With Changing Parameter)**

Now I change some parameter:

Before:

confidenceFactor: 0.25

unpruned: False

After:

confidenceFactor: 0.5

unpruned: True

The screenshot shows the Weka Explorer interface with the J48 classifier selected. The 'Test options' section shows 'Percentage split' at 80%. The 'Classifier output' section shows the evaluation results for the J48 classifier. The 'Summary' section shows the following results:

Area	PRC Area	Class
47	0.950	democrat
47	0.881	republican
47	0.922	

The 'Detailed Accuracy by Class' section shows the following results:

Class	Actual \ Predicted	democrat	republican
democrat	48	4	1
republican	34	1	3

The 'Confusion Matrix' section shows the following results:

Class	Actual \ Predicted	democrat	republican
democrat	48	4	1
republican	34	1	3

The 'weka.gui.GenericObjectEditor' dialog box is open, showing the configuration for the J48 classifier. The 'confidenceFactor' is set to 0.5 and 'unpruned' is set to True. The 'About' section of the dialog box shows the following text:

Class for generating a pruned or unpruned C4.

The 'More' button is highlighted.

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier: Choose J48 -U -M 2

Test options:

- ☐ Use training set
- ☐ Supplied test set Set...
- ☐ Cross-validation Folds 10
- ☒ Percentage split % 80

More options...

(Nom) Class

Start Stop

Result list (right-click for options)

- 00:55:58 - trees.J48
- 01:02:14 - trees.J48

Classifier output

```

| | | | water-project-cost-sharing = n: republican (2.0)
| | | | water-project-cost-sharing = y: democrat (4.0)
| | mx-missile = y: democrat (5.0/1.0)

Number of Leaves : 13
Size of the tree : 25

Time taken to build model: 0.02 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

=== Summary ===

Correctly Classified Instances      83      95.4023 %
Incorrectly Classified Instances    4      4.5977 %
Kappa statistic                    0.9053
Mean absolute error                 0.0705
Root mean squared error             0.2005
Relative absolute error             14.7732 %
Root relative squared error         40.8628 %
Total Number of Instances          87

=== Detailed Accuracy By Class ===

      TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
      0.942    0.029    0.980     0.942    0.961      0.906    0.981     0.976    democrat
      0.971    0.058    0.919     0.971    0.944      0.906    0.981     0.975    republican
Weighted Avg.   0.954    0.040    0.955     0.954    0.954      0.906    0.981     0.976

=== Confusion Matrix ===

  a  b  <-- classified as
49  3  | a = democrat
 1 34 | b = republican

```

Status OK

Log x 0

After changing the value of parameters our accuracy is increases.

***Accuracy: 95.4023 %**

Apply 2nd Algorithm: IBk (lazy)

The screenshot shows the Weka Explorer interface with the IBk classifier selected. The 'Test options' section on the left shows 'Percentage split' set to 80%. The 'Result list' on the left shows three entries: '00:55:58 - trees.J48', '01:08:14 - trees.J48', and '01:16:45 - lazy.IBk', with the last one selected. The 'Classifier output' pane on the right displays the following information:

Classifier: IBk-K 1 -W 0 -A "weka.core.neighboursearch.LinearNNSearch -A "weka.core.EuclideanDistance -R first-last"

Test mode: split 80.0% train, remainder test

=== Classifier model (full training set) ===

IBk1 instance-based classifier
using 1 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0.02 seconds

=== Summary ===

Correctly Classified Instances	80	91.954 %
Incorrectly Classified Instances	7	8.046 %
Kappa statistic	0.838	
Mean absolute error	0.0907	
Root mean squared error	0.2605	
Relative absolute error	19.0096 %	
Root relative squared error	53.0884 %	
Total Number of Instances	87	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.865	0.000	1.000	0.865	0.928	0.849	0.973	0.983	democrat
	1.000	0.135	0.833	1.000	0.909	0.849	0.973	0.929	republican
Weighted Avg.	0.920	0.054	0.933	0.920	0.920	0.849	0.973	0.961	

=== Confusion Matrix ===

```
a b <-- classified as
45 7 | a = democrat
0 35 | b = republican
```

Status: OK

Log x 0

***Algorithm: IBk (lazy)**

***Accuracy: 91.954 %**

Without changing any parameter.

Now if I make some changes:

The screenshot shows the Weka Explorer interface with the 'Classify' tab selected. The classifier chosen is 'IBk-K 1 -W 0 -X -A "weka.core.neighboursearch.LinearNNSearch -A "weka.core.EuclideanDistance -R first-last"' with 3 decimal places. The test options are set to 'Percentage split' at 80%. The classifier output shows the following results:

```
Classifier output
Test mode: split 80.0% train, remainder test
=== Classifier model (full training set) ===
IBk instance-based classifier
using 1 nearest neighbour(s) for classification

Time taken to build model: 0 seconds

=== Evaluation on test split ===
Time taken to test model on test split: 0.03 seconds

=== Summary ===
Correctly Classified Instances      80      91.954 %
Incorrectly Classified Instances    7       8.046 %
Kappa statistic                    0.838
Mean absolute error                 0.0907
Root mean squared error             0.2605
Relative absolute error             19.0096 %
Root relative squared error         53.0884 %
Total Number of Instances          87

=== Detailed Accuracy By Class ===
               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
               0.865    0.000    1.000     0.865    0.928      0.849    0.973     0.983    democrat
               1.000    0.135    0.833     1.000    0.909      0.849    0.973     0.929    republican
Weighted Avg.   0.920    0.054    0.933     0.920    0.920      0.849    0.973     0.961

=== Confusion Matrix ===
  a  b  <-- classified as
45  7  |  a = democrat
 0 35 |  b = republican
```

The result list on the left shows three entries: '00:55:58 - trees.J48', '01:08:14 - trees.J48', and '01:23:53 - lazy.IBk', with the last one selected.

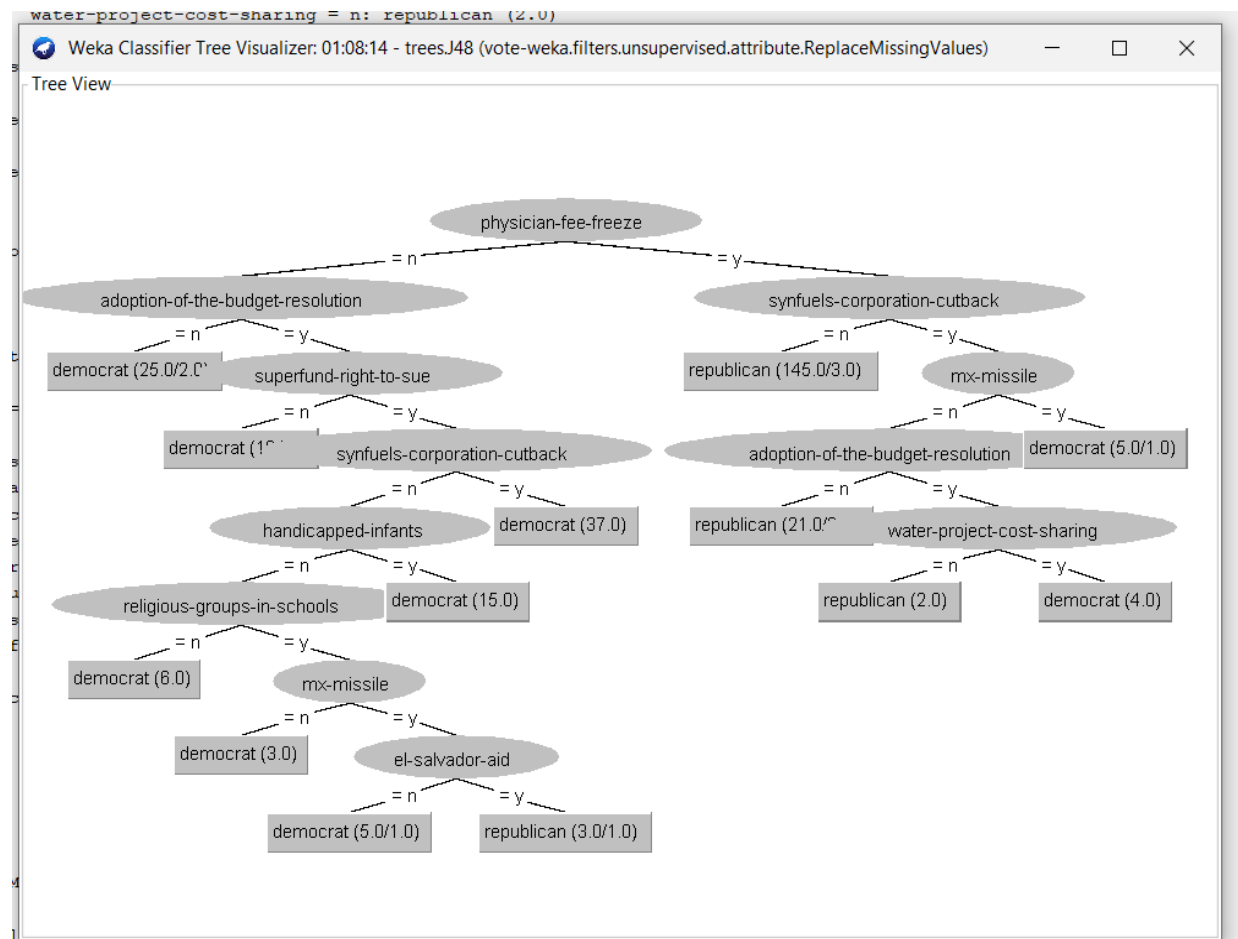
After changes parameters:

crossValidate: True

numDecimalPlaces: 3

There is no change in accuracy.

*Decision tree for j48(tree) Algorithm:



Step:03

Calculate for each model the accuracy, precision, recall, F1-score for the class 'democrat':

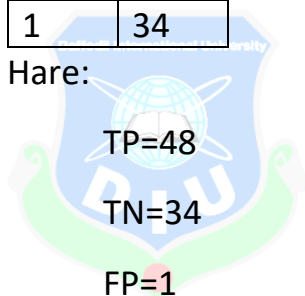
For 1st Algorithm:

***J48 (Without Changing Parameter):**

***Confusion matrix:**

48	4
1	34

Hare:



TP=48

TN=34

FP=1

FN=4

Daffodil
International
University

1.Accuracy: $(48+34) / (48+34+1+4) = 94.2529\%$

2.Precision: $(48) / (48+1) = 0.975$

3.Recall: $(48) / (48+4) = 0.9230$

4.F1-Score: $(2 * 0.975 * 0.9230) / (0.975 + 0.9230) = 0.9482$

***J48 (With Changing Parameter):**

***Confusion matrix:**

49	3
1	34

Hare:

$$TP=49$$

$$TN=34$$

$$FP=1$$

$$FN=3$$

1.Accuracy: $(49+34)/ (49+34+1+3) = 95.4023\%$

2.Precision: $(49)/ (49+1) = 0.98$

3.Recall: $(49)/ (49+3) = 0.9423$

4.F1-Score: $(2*0.98*0.9423)/(0.98+0.9423)=0.9607$

For 2nd Algorithm:

***IBk (Without changing any parameter)**

***Confusion matrix:**

45	7
0	35

Hare:

$$TP=45$$

$$TN=35$$

$$FP=0$$

$$FN=7$$

1.Accuracy: $(45+35)/ (45+35+0+7) = 91.954\%$

2.Precision: $(45)/ (45+0) = 1$

3.Recall: $(45)/ (45+7) = 0.8653$

4.F1-Score: $(2*1*0.8653)/ (1+0.8653) = 0.9277$

After change parameters there are no change in accuracy, precision, recall, F1-score for the class 'democrat'.

Step:4

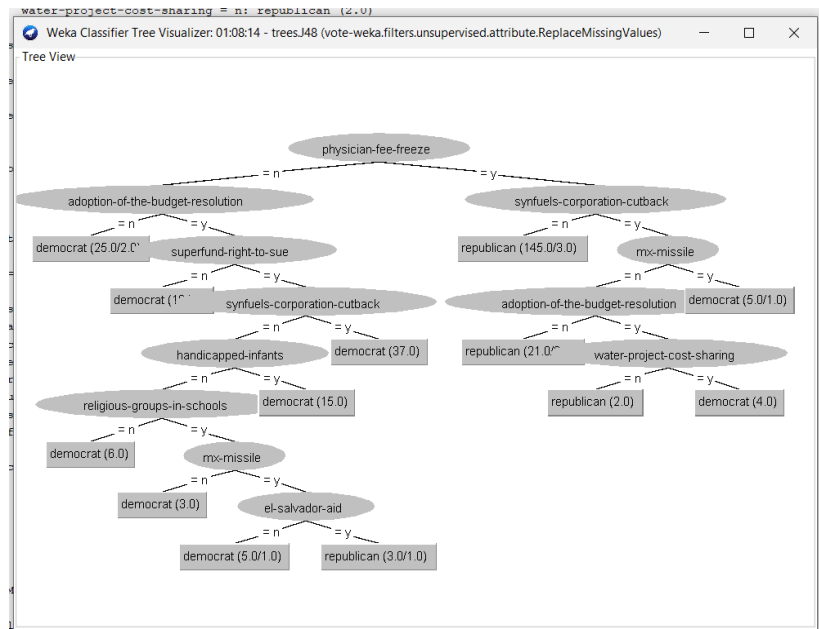
Choose the best, the second best and the third best model from step (2):

Best Model: J48 (With Changing Parameter) because this model has high accuracy between 4 model

Second Best Model: J48 (With Changing Parameter) because this model has second high accuracy between 4 model.

Third Best Model: IBk because this model has low accuracy between 4 model.

Step:5



If I analysis from the decision tree generated by J48 algorithm, three characteristic of democrat voter are **physician-fee-freeze**, **mx-missile** and **adoption-of-the-budget-resolution**.

The association rules that I have used are given below:

1. physician-fee-freeze=n export-administration-act-south-africa=y 172 ==> Class=democrat 168.
2. physician-fee-freeze=n 179 ==> Class=democrat 175.
3. physician-fee-freeze=n 179 ==> export-administration-act-south-africa=y Class=democrat 16
4. adoption-of-the-budget-resolution=y 183 ==> export-administration-act-south-africa=y 178

Thank You



Daffodil
International
University