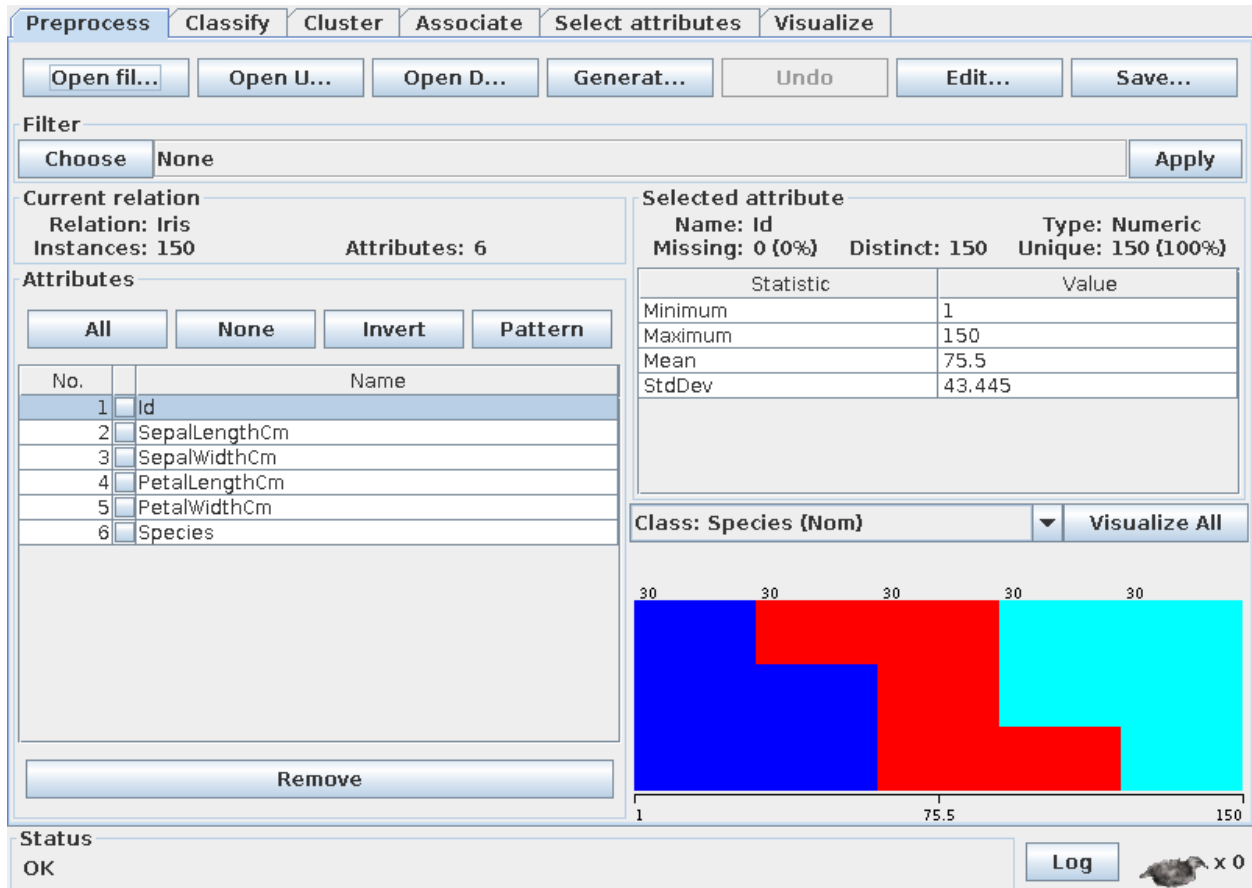


## Experiment – 1 (WEKA Tool)

WEKA tool – Implement the following concepts using WEKA tool and visualize the results.

Visualization of Iris Dataset.



### 1. Linear and Logistic Regression

Preprocess **Classify** Cluster Associate Select attributes Visualize

Classifier

Choose Logistic -R 1.0E-8 -M 1

Test options

☐ Use training set

☐ Supplied test set Set...

☐ Cross-validation Folds 10

☒ Percentage split % 70

More options...

(Nom) Species

Start Stop

Result list (right-click for options)

19:19:17 - functions.Logistic

Classifier output

Time taken to build model: 0.05 seconds

=== Evaluation on test split ===

=== Summary ===

Correctly Classified Instances	45	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0.0008		
Root mean squared error	0.007		
Relative absolute error	0.1911 %		
Root relative squared error	1.4795 %		
Total Number of Instances	45		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
1	0	1	1	1	1	Iris-setosa
1	0	1	1	1	1	Iris-versicolor
1	0	1	1	1	1	Iris-virginica
Weighted Avg.	1	0	1	1	1	

=== Confusion Matrix ===

```

a b c <-- classified as
14 0 0 | a = Iris-setosa
0 16 0 | b = Iris-versicolor
0 0 15 | c = Iris-virginica

```

Status OK Log x 0

Preprocess **Classify** Cluster Associate Select attributes Visualize

Classifier

Choose LinearRegression -S 0 -R 1.0E-8

Test options

☐ Use training set

☐ Supplied test set Set...

☐ Cross-validation Folds 10

☒ Percentage split % 80

More options...

(Nom) Species

Start Stop

Result list (right-click for options)

19:19:17 - functions.Logistic

19:19:52 - functions.Logistic

Classifier output

Time taken to build model: 0.03 seconds

=== Evaluation on test split ===

=== Summary ===

Correctly Classified Instances	30	100	%
Incorrectly Classified Instances	0	0	%
Kappa statistic	1		
Mean absolute error	0		
Root mean squared error	0		
Relative absolute error	0.0001 %		
Root relative squared error	0.0004 %		
Total Number of Instances	30		

=== Detailed Accuracy By Class ===

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
1	0	1	1	1	1	Iris-setosa
1	0	1	1	1	1	Iris-versicolor
1	0	1	1	1	1	Iris-virginica
Weighted Avg.	1	0	1	1	1	

=== Confusion Matrix ===

```

a b c <-- classified as
11 0 0 | a = Iris-setosa
0 10 0 | b = Iris-versicolor
0 0 9 | c = Iris-virginica

```

Status OK Log x 0

## 2. Decision Trees

## Dataset for Decision Trees

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

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

Filter: Choose None Apply

Current relation  
Relation: exp3\_train  
Instances: 14

Attributes: 5

Attributes: All None Invert Pattern

No.	Name
1	Outlook
2	Temperature
3	Humidity
4	Wind
5	Play Tennis

Remove

Status: OK

Selected attribute  
Name: Outlook  
Missing: 0 (0%)  
Distinct: 3  
Type: Nominal  
Unique: 0 (0%)

No.	Label	Count
1	Sunny	5
2	Overcast	4
3	Rain	5

Class: Play Tennis (Nom) Visualize All

Log x 0

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

Classifier: Choose Id3

Test options:  
☐ Use training set  
☐ Supplied test set (Set...)  
☐ Cross-validation (Folds: 10)  
☒ Percentage split (%: 70) More options...

(Nom) Play Tennis Start Stop

Result list (right-click for options):  
19:19:17 - functions.Logistic  
19:19:52 - functions.Logistic  
19:24:35 - trees.Id3

Classifier output

Outlook = Sunny  
| Humidity = High: No  
| Humidity = Normal: Yes  
Outlook = Overcast: Yes  
Outlook = Rain  
| Wind = Weak: Yes  
| Wind = Strong: No

Time taken to build model: 0 seconds

=== Evaluation on test split ===  
=== Summary ===

Metric	Value	%
Correctly Classified Instances	2	50
Incorrectly Classified Instances	2	50
Kappa statistic	0	
Mean absolute error	0.5	
Root mean squared error	0.7071	
Relative absolute error	100	%
Root relative squared error	134.1641	%
Total Number of Instances	4	

=== Detailed Accuracy By Class ===

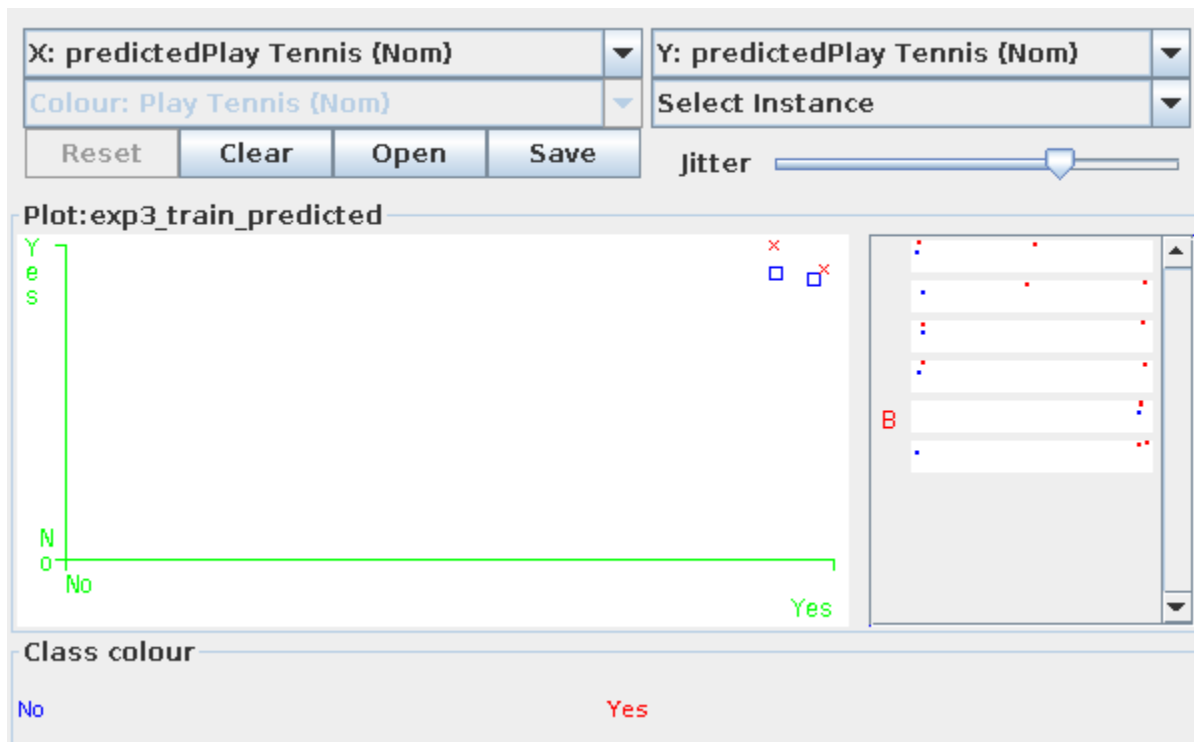
	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0	0	0	0	0	0.5	No
	1	1	0.5	1	0.667	0.5	Yes
Weighted Avg.	0.5	0.5	0.25	0.5	0.333	0.5	

=== Confusion Matrix ===

a b <- classified as  
0 2 | a = No  
0 2 | b = Yes

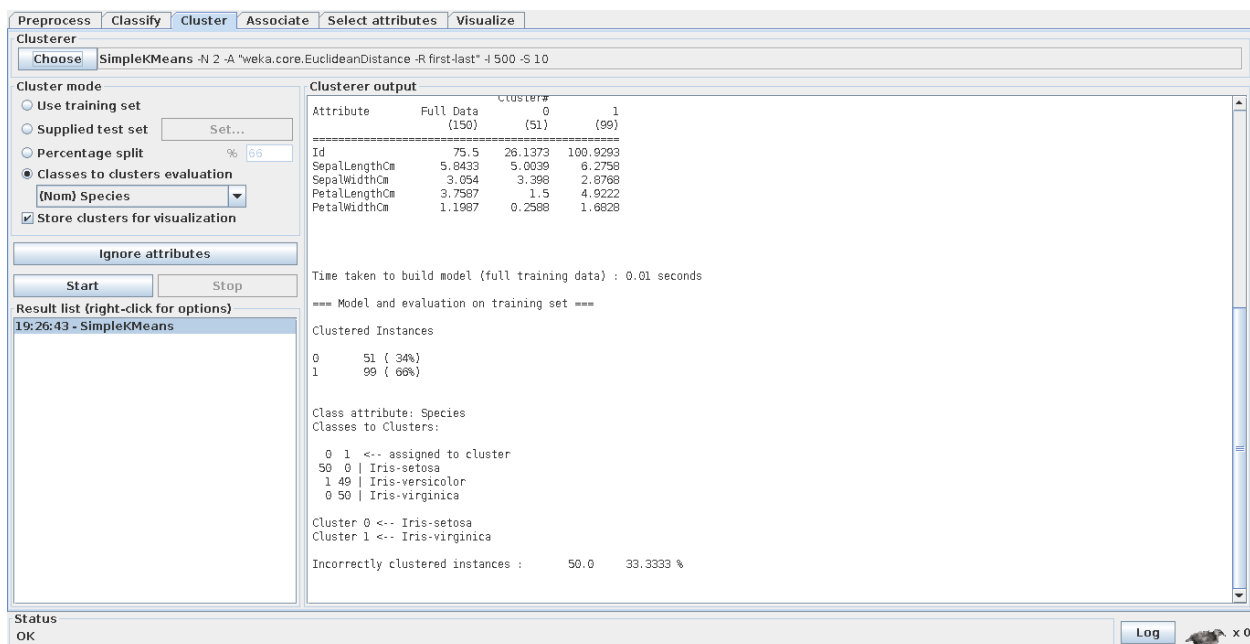
Status: OK

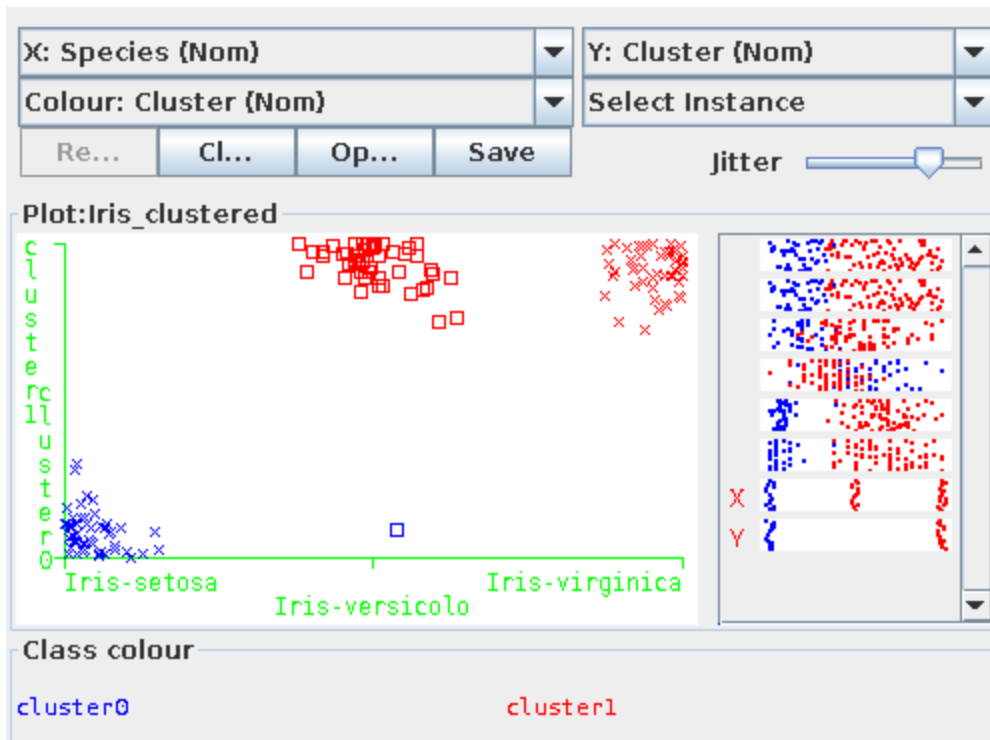
Log x 0



### 3. K-Means Clustering

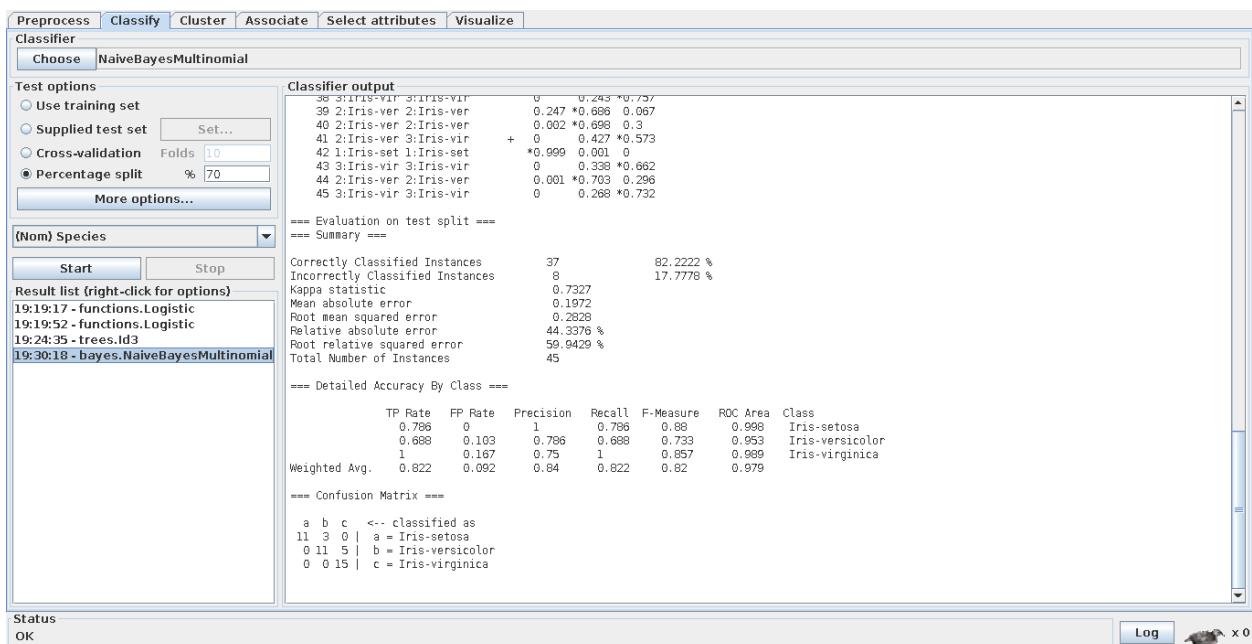
#### Dataset for K-Means Clustering





## 4. Naïve Bayes Classifier

### Dataset for Naïve Bayes Classifier



X: Species {Nom}

Y: predictedSpecies {Nom}

Colour: Species {Nom}

Select Instance

Reset

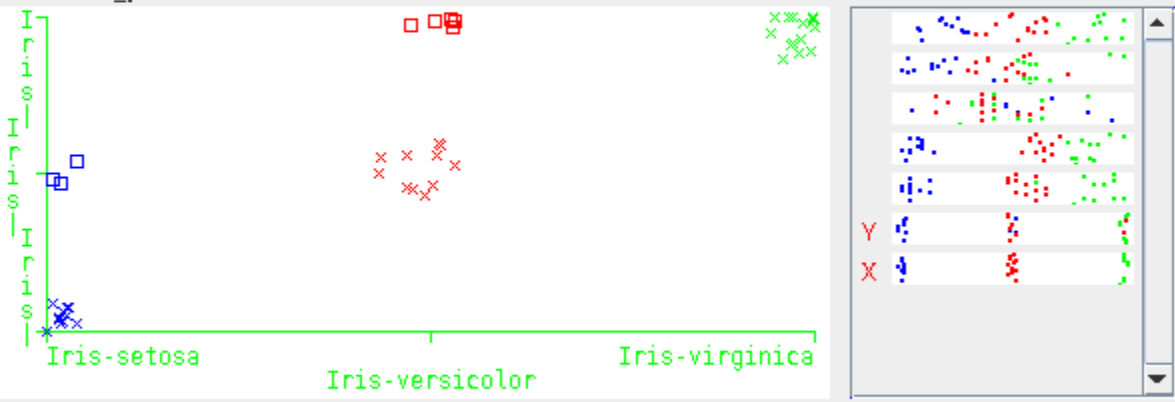
Clear

Open

Save

Jitter

Plot:Iris\_predicted



Class colour

Iris-setosa	Iris-versicolor	Iris-virginica
-------------	-----------------	----------------