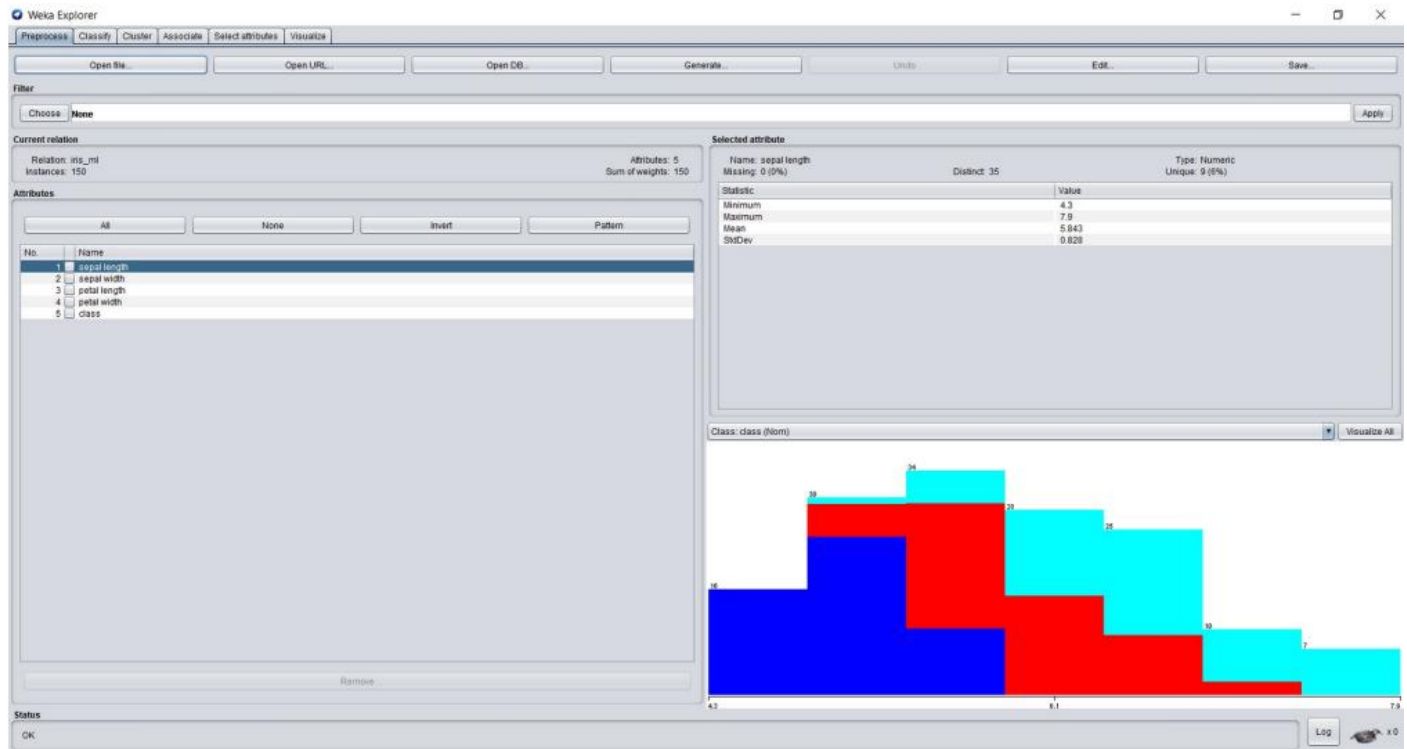


13/09/2021

ML ONSPOT 2

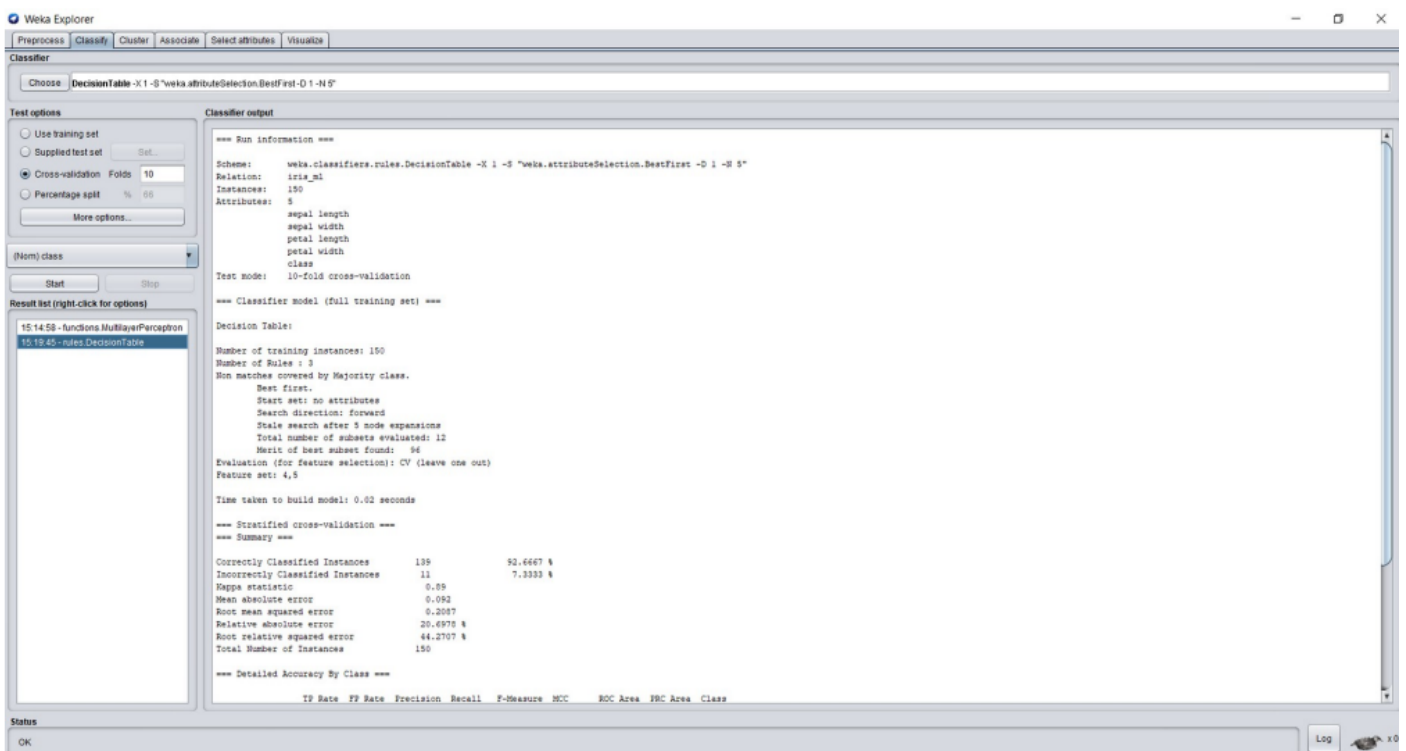
2019103573

Changed the data set so that matrix is generated



Classify table:

Decision table:



=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Iris-setosa
	0.880	0.050	0.898	0.880	0.889	0.834	0.946	0.861	Iris-versicolor
	0.900	0.060	0.882	0.900	0.891	0.836	0.947	0.869	Iris-virginica
Weighted Avg.	0.927	0.037	0.927	0.927	0.927	0.890	0.964	0.910	

=== Confusion Matrix ===

```

a b c <-- classified as
50 0 0 | a = Iris-setosa
0 44 6 | b = Iris-versicolor
0 5 45 | c = Iris-virginica

```

weka.gui.GenericObjectEditor



weka.classifiers.rules.DecisionTable

About

Class for building and using a simple decision table majority classifier.

More

Capabilities

batchSize 100

crossVal 1

debug False

displayRules False

doNotCheckCapabilities False

evaluationMeasure Default: accuracy (discrete class); RMSE (numeric class)

numDecimalPlaces 2

search Choose BestFirst -D 1 -N 5

useIBk False

Open...

Save...

OK

Cancel

weka.classifiers.rules.DecisionTable

About

Class for building and using a simple decision table majority classifier.

[More](#)[Capabilities](#)

batchSize 100

crossVal 2

debug False

displayRules False

doNotCheckCapabilities False

evaluationMeasure Default: accuracy (discrete class); RMSE (numeric class)

numDecimalPlaces 2

search Choose BestFirst -D 1 -N 5

useIBk False

[Open...](#)[Save...](#)[OK](#)[Cancel](#)

Classifier output

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

Decision Table:

Number of training instances: 150

Number of Rules : 3

Non matches covered by Majority class.

Best first.

Start set: no attributes

Search direction: forward

Stale search after 5 node expansions

Total number of subsets evaluated: 13

Merit of best subset found: 96

Evaluation (for feature selection): CV (2 fold)

Feature set: 4,5

Time taken to build model: 0.01 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	139	92.6667 %
Incorrectly Classified Instances	11	7.3333 %
Kappa statistic	0.89	
Mean absolute error	0.0897	
Root mean squared error	0.2987	
Relative absolute error	20.1838 %	
Root relative squared error	44.2717 %	
Total Number of Instances	150	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	1.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	Iris-setosa
	0.880	0.050	0.898	0.880	0.889	0.834	0.948	0.883	Iris-versicolor
	0.900	0.060	0.882	0.900	0.891	0.836	0.948	0.863	Iris-virginica
Weighted Avg.	0.927	0.037	0.927	0.927	0.927	0.890	0.966	0.915	

=== Confusion Matrix ===

```
a b c <-- classified as
50 0 0 | a = Iris-setosa
0 44 6 | b = Iris-versicolor
0 5 45 | c = Iris-virginica
```

So the following changes are observed

Cluster Table:

EM algorithm:

Initial observed values

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Clusterer

Choose EM -I 103 -N 1 -K 10 -max-iter 100 -M 10E-6 -K 10 -num-slots 1 -S 100

Cluster mode

☒ Use training set
☐ Supplied test set (Set...)
☐ Percentage split % 66
☐ Classes to clusters evaluation (Norm) class
☒ Store clusters for visualization

Ignore attributes

Start Stop

Result list (right-click for options)

15:24:22 - EM

Clusterer output

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

Attribute	Cluster			
	0 (0.92)	1 (0.33)	2 (0.2)	3 (0.14)
sepal length				
mean	5.897	5.806	6.9426	6.1304
std. dev.	0.6275	0.3489	0.499	0.2943
sepal width				
mean	2.7519	3.418	3.1103	2.6098
std. dev.	0.3183	0.3772	0.2952	0.2361
petal length				
mean	4.2267	1.464	5.8559	5.0993
std. dev.	0.445	0.1718	0.4626	0.2462
petal width				
mean	1.3134	0.244	2.1495	1.8254
std. dev.	0.1844	0.1061	0.232	0.2152
class				
Iris-setosa	1	51	1	1
Iris-versicolour	48.1125	1	1.0182	3.8693
Iris-virginica	2.0983	1	31.0375	19.8641
[total]	51.2108	53	33.0557	24.7335

Time taken to build model (full training data) : 0.36 seconds

=== Model and evaluation on training set ===

Clustered Instances

0	48 (32%)
1	50 (33%)
2	29 (19%)
3	23 (15%)

Log likelihood: -2.03504

weka.clusterers.EM

About

Simple EM (expectation maximisation) class.

More

Capabilities

debug False

displayModelInOldFormat False

doNotCheckCapabilities False

maxIterations 100

maximumNumberOfClusters -1

minLogLikelihoodImprovementCV 1.0E-6

minLogLikelihoodImprovementIterating 1.0E-6

minStdDev 1.0E-6

numClusters -1

numExecutionSlots 1

numFolds 10

numKMeansRuns 10

seed 100

Open...

Save...

OK

Cancel

Changing :

weka.gui.GenericObjectEditor

weka.clusterers.EM

About

Simple EM (expectation maximisation) class.

More

Capabilities

debug False

displayModelInOldFormat False

doNotCheckCapabilities False

maxIterations 100

maximumNumberOfClusters -1

minLogLikelihoodImprovementCV 1.0E-6

minLogLikelihoodImprovementIterating 1.0E-6

minStdDev 1.0E-6

numClusters -1

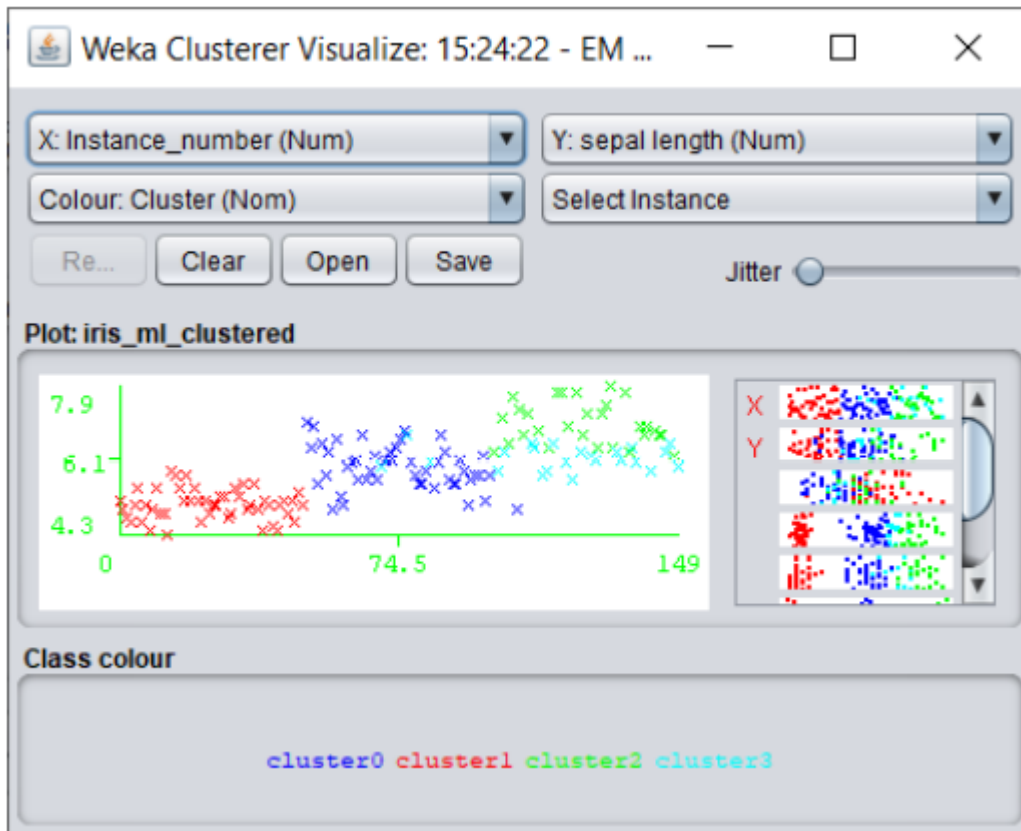
numExecutionSlots 2

numFolds 10

numKMeansRuns 10

seed 100

Open... Save... OK Cancel



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
std. dev.	0.445	0.1718	0.4626	0.2462
petal width				
mean	1.3134	0.244	2.1495	1.8254
std. dev.	0.1864	0.1061	0.232	0.2152
class				
Iris-setosa	1	51	1	1
Iris-versicolor	48.1125	1	1.0182	3.8693
Iris-virginica	2.0983	1	31.0375	19.8641
[total]	51.2108	53	33.0557	24.7335

Time taken to build model (full training data) : 0.4 seconds

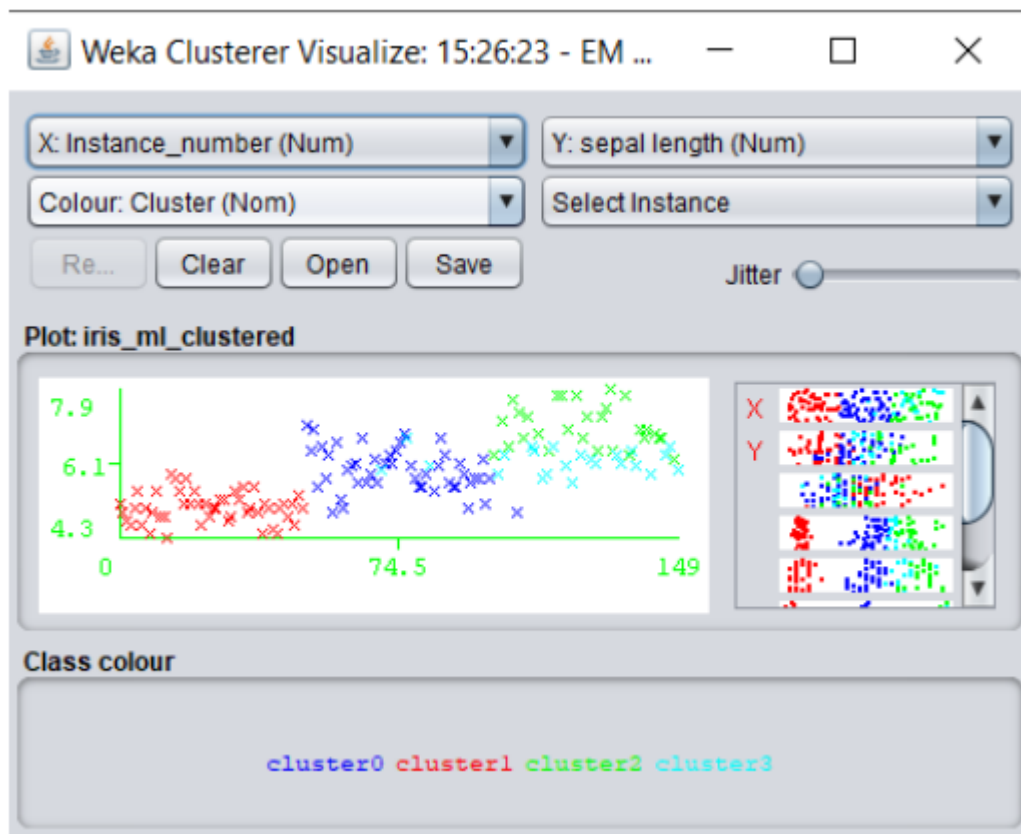
=== Model and evaluation on training set ===

Clustered Instances

0	48 (32%)
1	50 (33%)
2	29 (19%)
3	23 (15%)

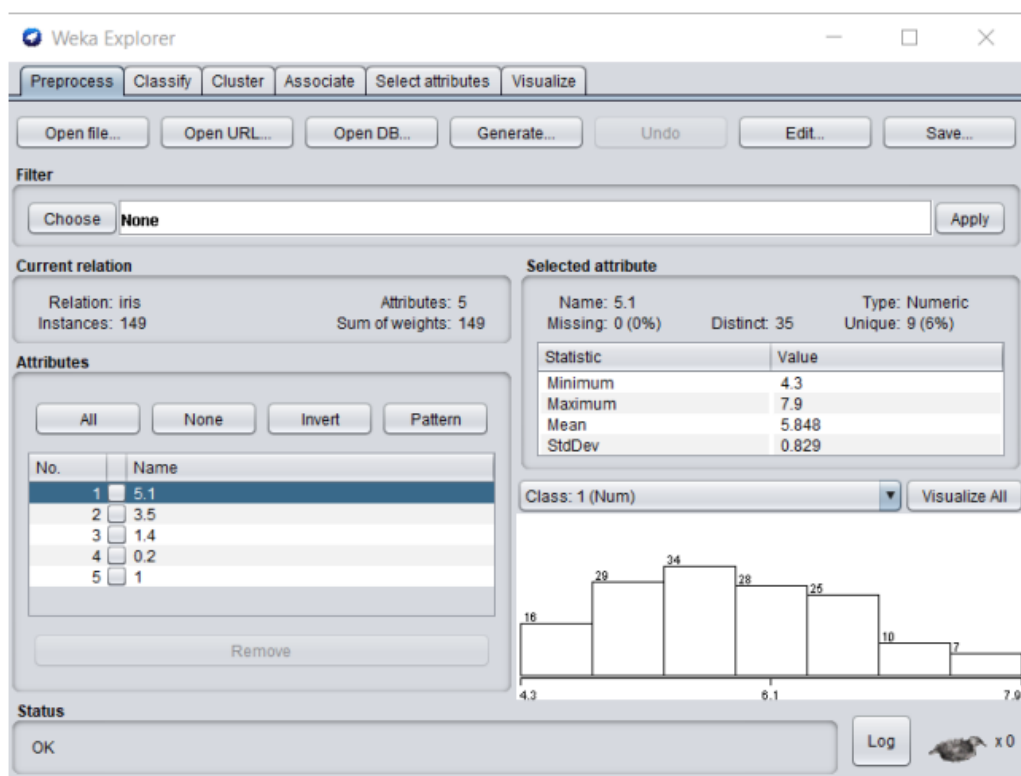
Log likelihood: -2.03504

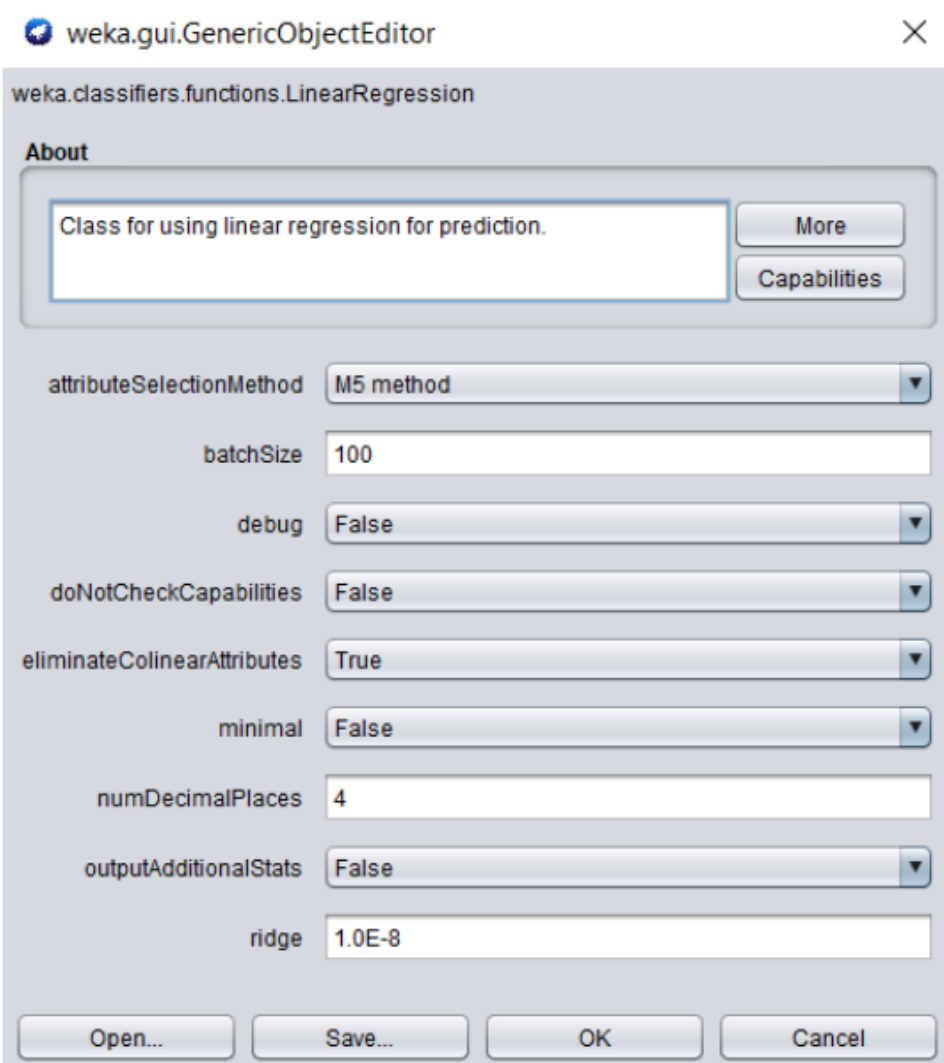
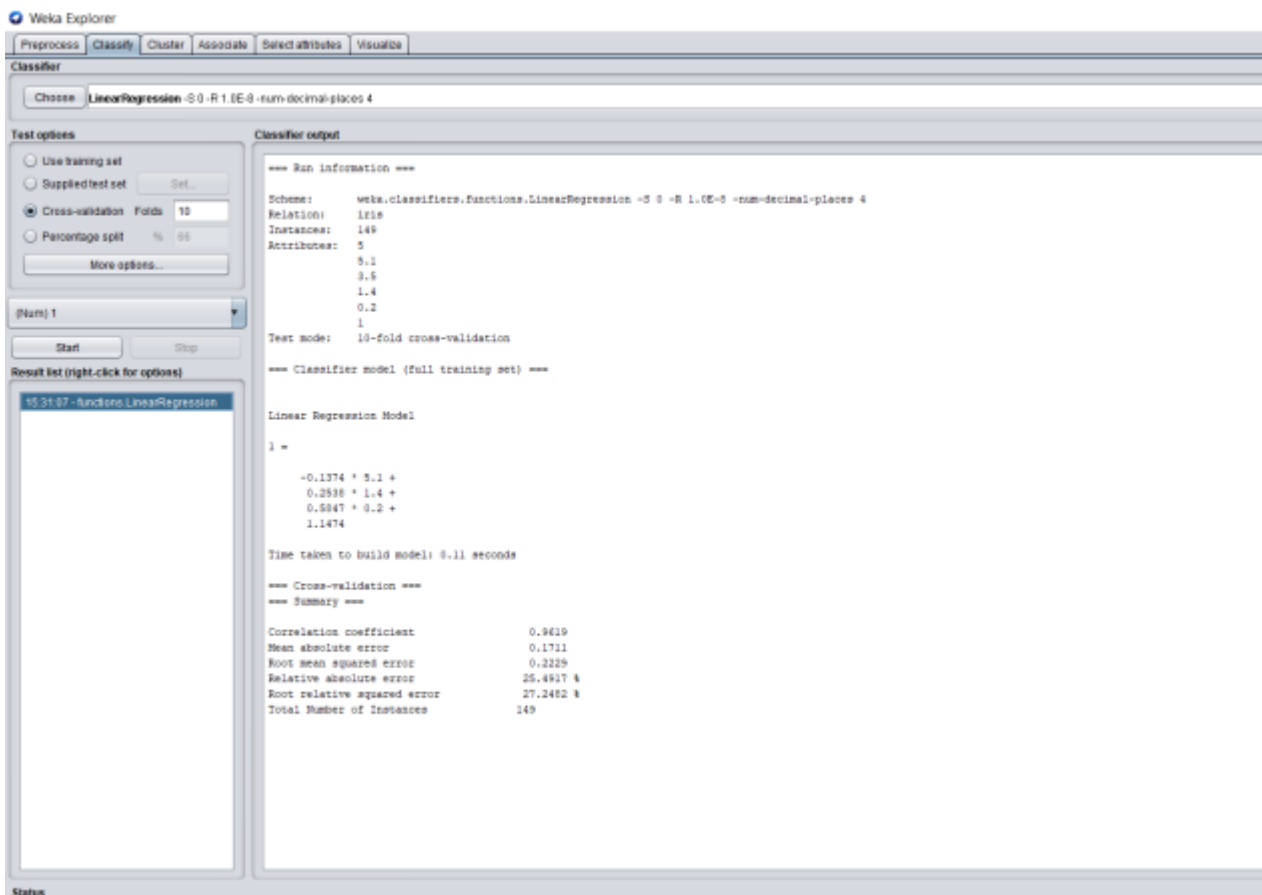
Time Changed



Linear Regression :

Using old dataset:





Changing

weka.gui.GenericObjectEditor ✕

weka.classifiers.functions.LinearRegression

About

Class for using linear regression for prediction.

More

Capabilities

attributeSelectionMethod M5 method ▼

batchSize 100

debug False ▼

doNotCheckCapabilities False ▼

eliminateColinearAttributes True ▼

minimal False ▼

numDecimalPlaces 2

outputAdditionalStats False ▼

ridge 1.0E-8

Open... Save... OK Cancel

Classifier

Choose LinearRegression -S 0 -R 1.0E-8

Test options

- ☐ Use training set
☐ Supplied test set
☒ Cross-validation Folds
☐ Percentage split %

(Num) 1

Start

Stop

Result list (right-click for options)

15:31:07 - functions.LinearRegression
15:32:15 - functions.LinearRegression

Classifier output

```
=== Run information ===

Scheme:      weka.classifiers.functions.LinearRegression -S 0 -R 1.0E-8
Relation:    iris
Instances:   149
Attributes:  5
              5.1
              3.5
              1.4
              0.2
              1
Test mode:   10-fold cross-validation

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

Linear Regression Model

1 =

      -0.14 * 5.1 +
       0.25 * 1.4 +
       0.58 * 0.2 +
       1.15

Time taken to build model: 0 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient      0.9619
Mean absolute error         0.1711
Root mean squared error     0.2229
Relative absolute error     25.4917 %
Root relative squared error 27.2482 %
Total Number of Instances   149
```