Terna Engineering College
Computer Engineering Department

Program: Sem VII

Course: Artificial Intelligence & Soft Computing (AI&SC)

Experiment No. 05

PART B

(PART B: TO BE COMPLETED BY STUDENTS)

(Students must submit the soft copy as per the following segments within two hours of the practical. The soft copy must be uploaded on the Blackboard or emailed to the concerned lab in charge faculties at the end of the practical in case there is no Blackboard access available)

Roll No. 50	Name: AMEY THAKUR
Class: BE-COMPS-50	Batch: B3
Date of Experiment: 24-08-2021	Date of Submission: 24-08-2021
Grade :	

Aim: Identify the Classification problem and create a Knowledge database for that problem and apply appropriate search methods for optimization.

B.1 Software Code written by student:

(Paste your Problem Statement for Classification and Data set Used as a knowledge Database for Given Classification Problem)

Problem Statement

To categorise the flower dataset in order to determine if it is in stock or not, using a combination of classifiers, evaluators, and search algorithms to improve accuracy and compare findings.

B.2 Input and Output:

(Paste your screenshot of Analysis of Data, Relevant Attributes Selection by using at least Three methods)

Flowers.arff

@relation Flowers

@attribute Type {Lilies, Orchids, Roses, Tulips}

@attribute Color {Red,Yellow,Blue}

@attribute Feature {Fresh,Luster,Vibrance,Strong}

@attribute InStock {Yes,No}

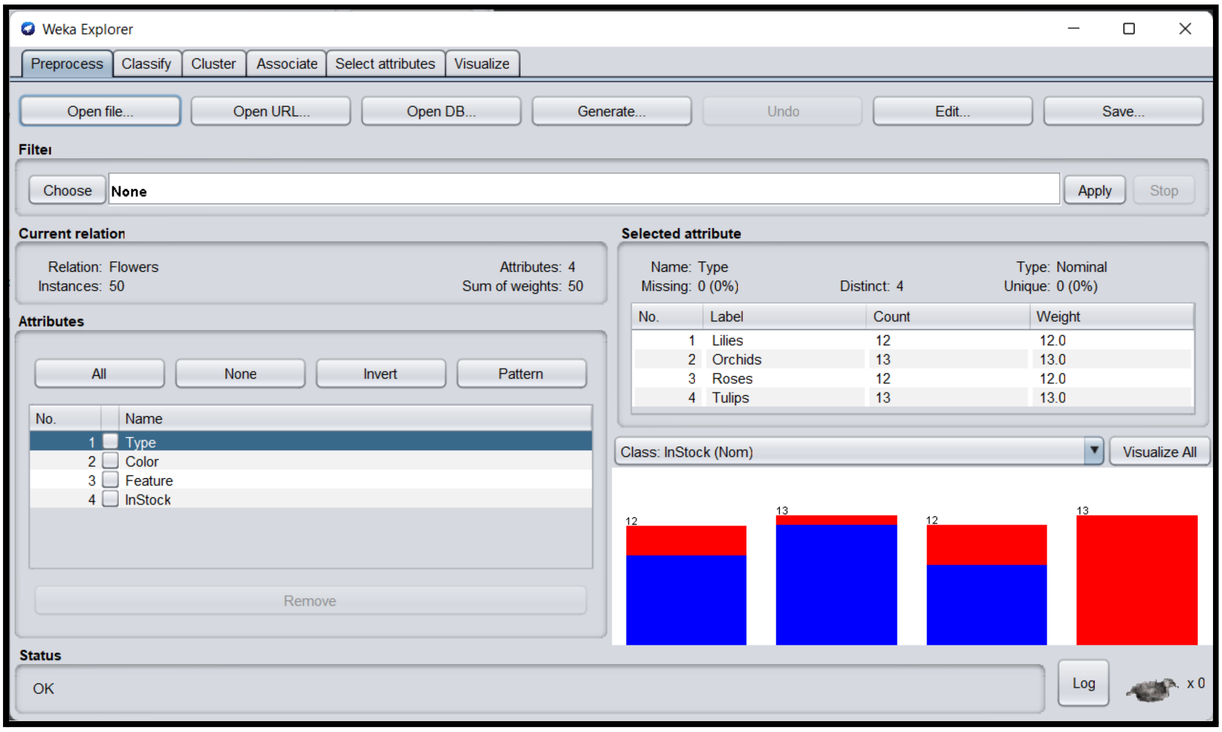
@data

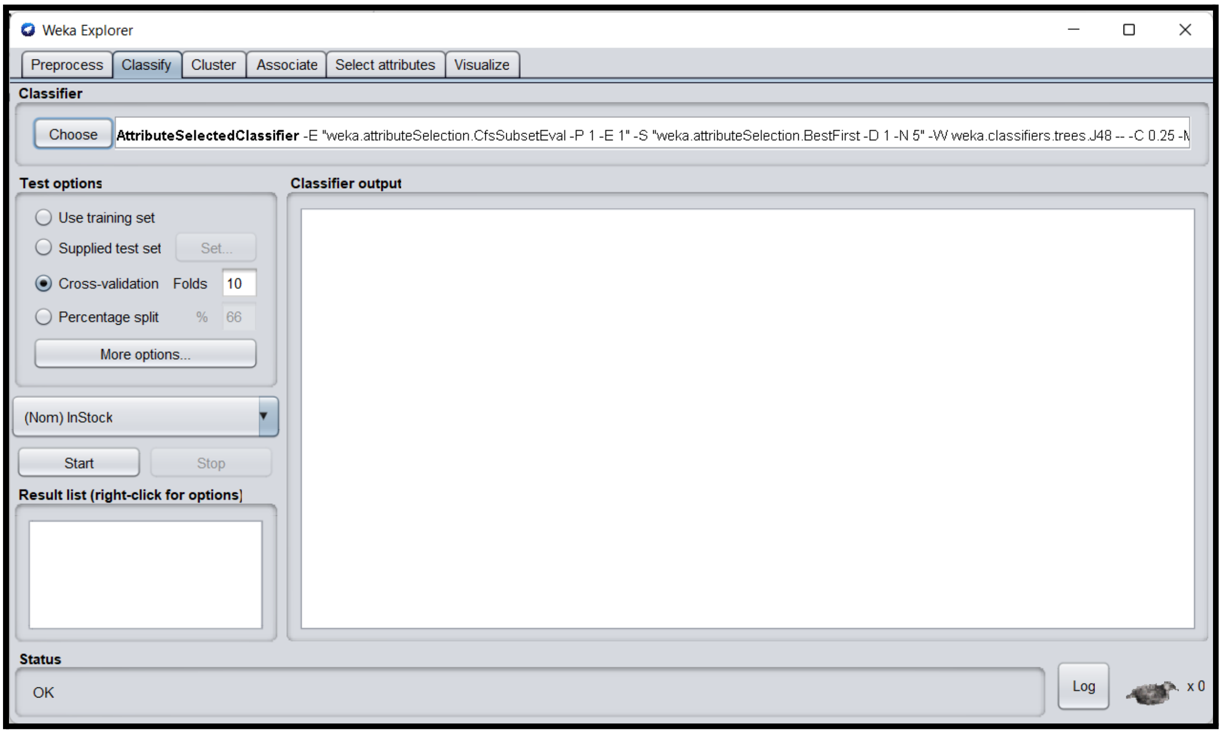
Lilies	Red	Fresh	Yes
Orchids	Blue	Luster	Yes
Orchids	Yellow	Fresh	Yes
Tulips	Red	Strong	No
Lilies	Yellow	Vibrance	Yes
Tulips	Red	Fresh	No
Roses	Yellow	Strong	No
Roses	Blue	Luster	Yes
Lilies	Blue	Strong	Yes
Orchids	Red	Vibrance	Yes
Roses	Yellow	Fresh	Yes
Tulips	Red	Luster	No
Tulips	Yellow	Strong	No
Roses	Blue	Vibrance	Yes
Orchids	Blue	Luster	Yes
Orchids	Red	Strong	No
Lilies	Blue	Fresh	Yes
Tulips	Yellow	Vibrance	No
Roses	Red	Luster	Yes
Roses	Yellow	Strong	No
Lilies	Red	Vibrance	Yes
Orchids	Blue	Fresh	Yes
Lilies	Red	Fresh	Yes

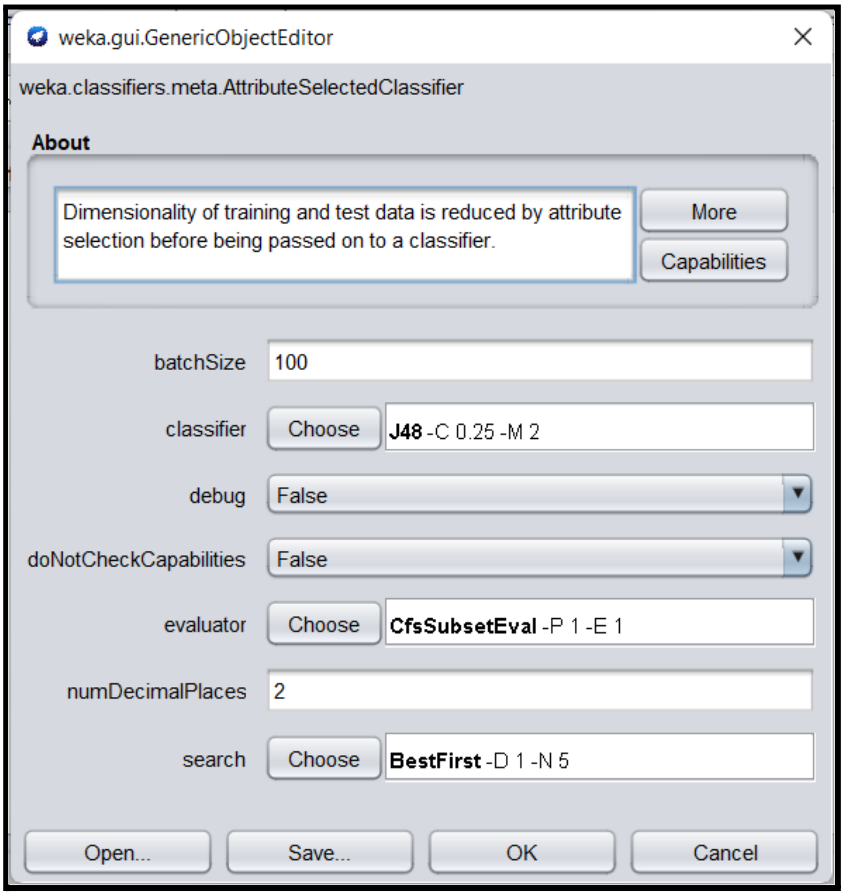
Tulips	Yellow	Vibrance	No
Roses	Blue	Luster	Yes
Roses	Red	Strong	No
Lilies	Red	Luster	Yes
Orchids	Yellow	Fresh	Yes
Tulips	Red	Luster	No
Roses	Blue	Fresh	Yes
Lilies	Blue	Fresh	Yes
Lilies	Red	Strong	No
Orchids	Red	Fresh	Yes
Tulips	Yellow	Vibrance	No
Orchids	Blue	Luster	Yes
Tulips	Yellow	Strong	No
Lilies	Red	Vibrance	Yes
Tulips	Red	Luster	No
Roses	Blue	Strong	No
Orchids	Blue	Vibrance	Yes
Orchids	Red	Vibrance	Yes
Roses	Yellow	Fresh	Yes
Tulips	Blue	Luster	No
Lilies	Yellow	Strong	No
Roses	Red	Vibrance	Yes
Orchids	Blue	Luster	Yes
Tulips	Red	Fresh	No
Lilies	Blue	Strong	No
Tulips	Yellow	Luster	No

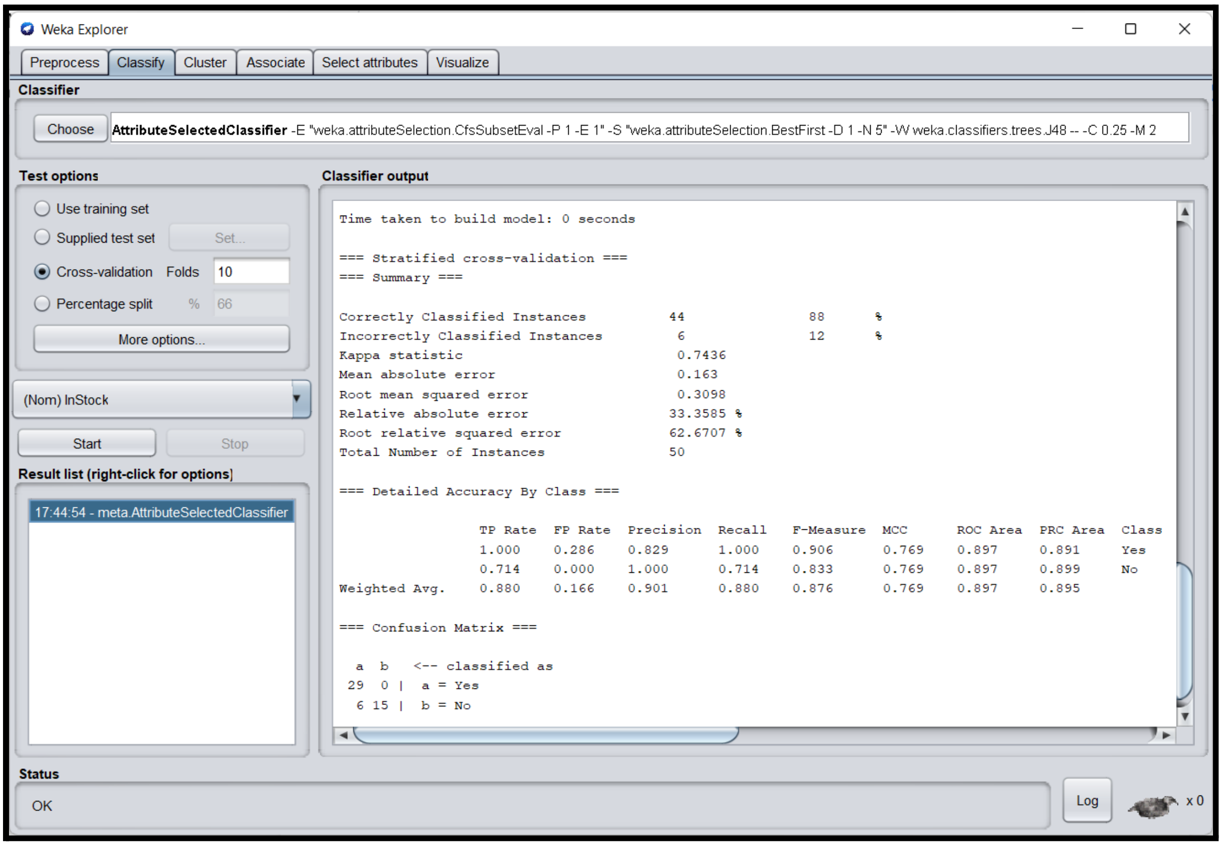
Orchids	Red	Luster	Yes

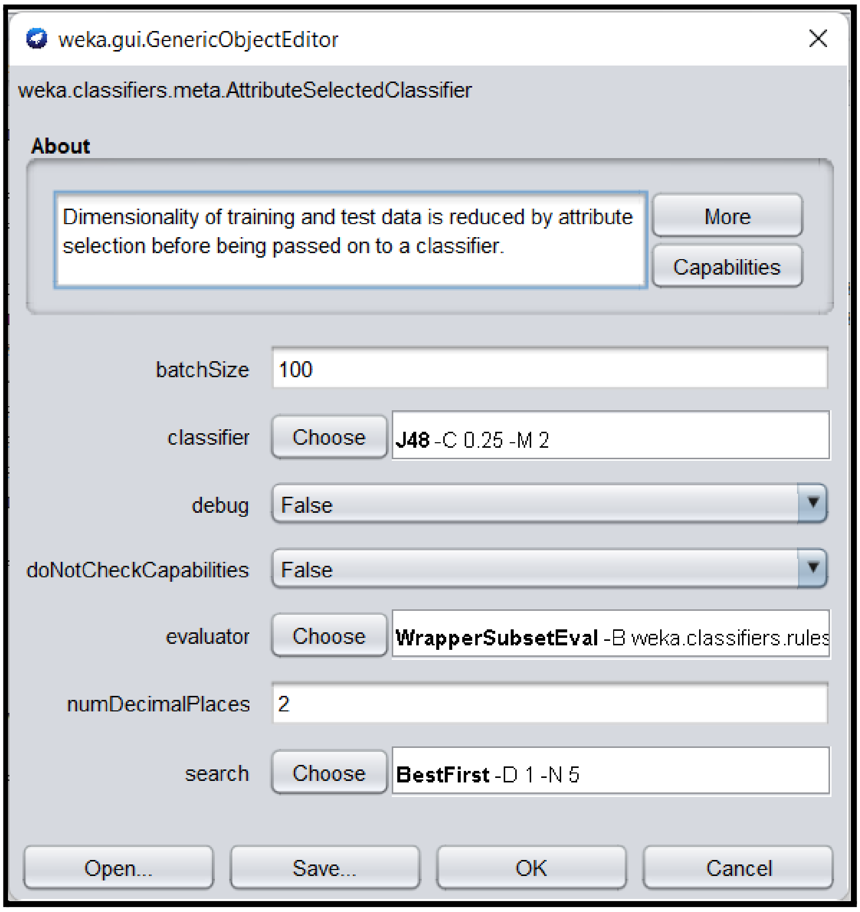
Weka

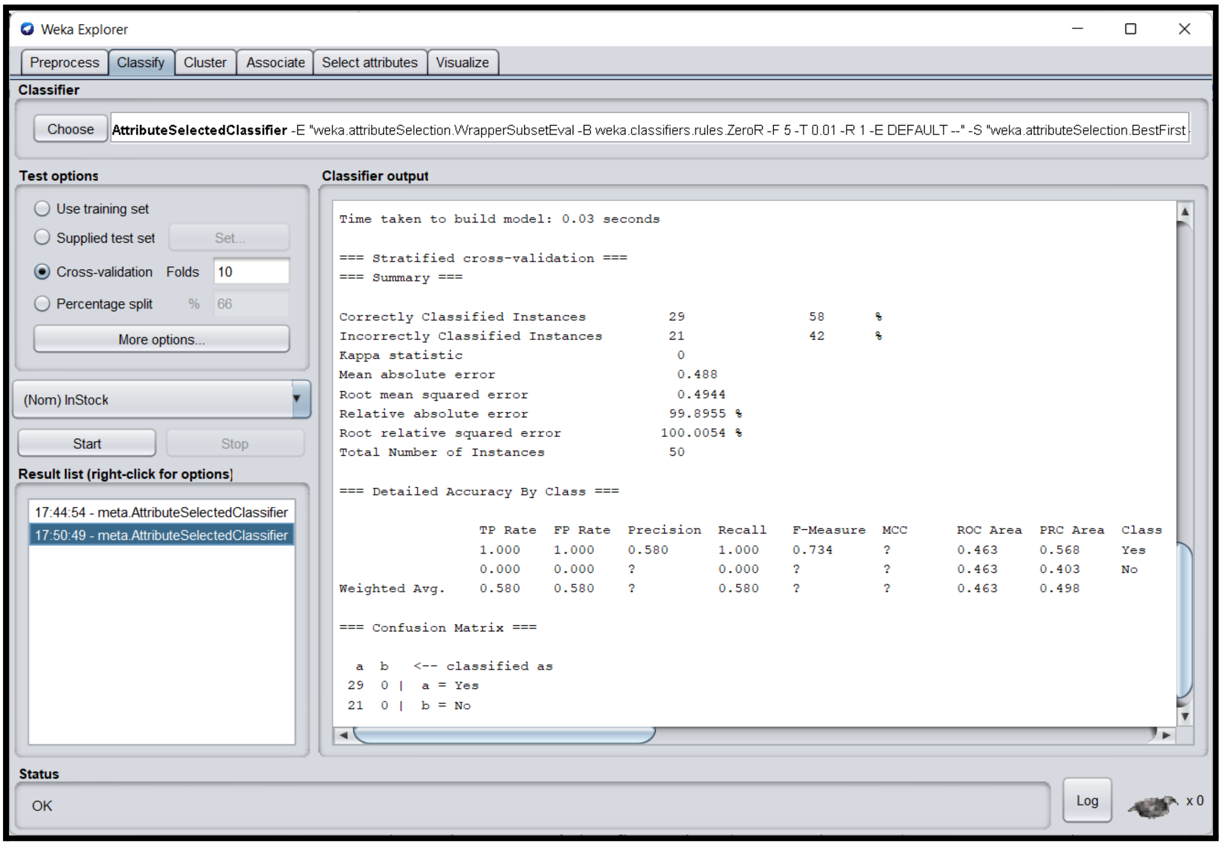


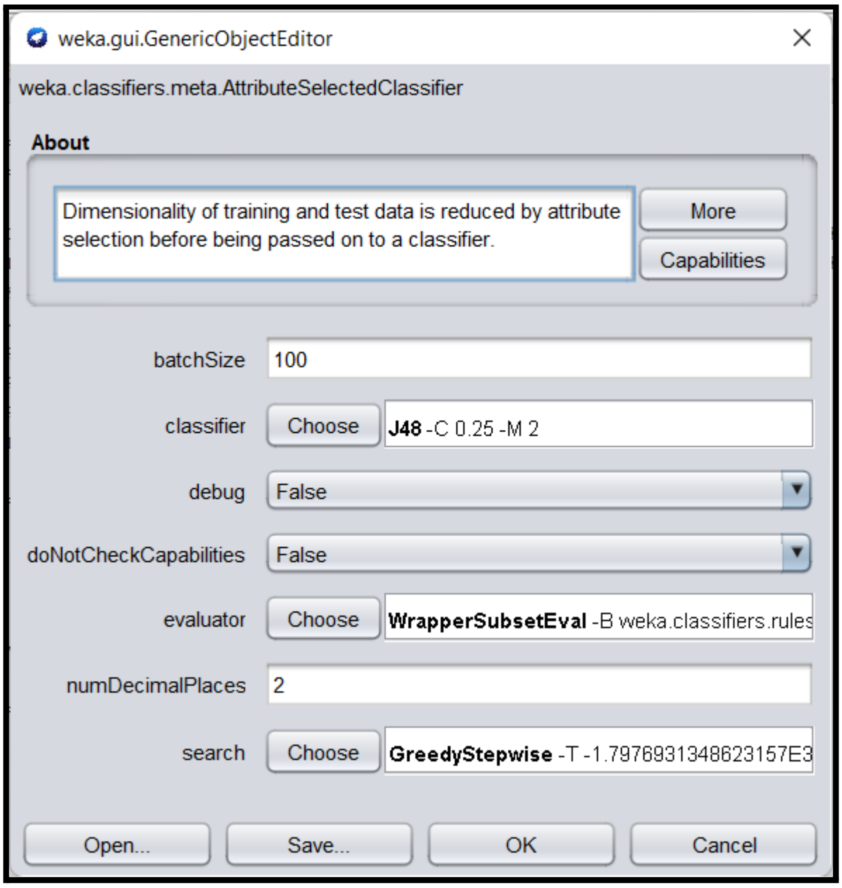


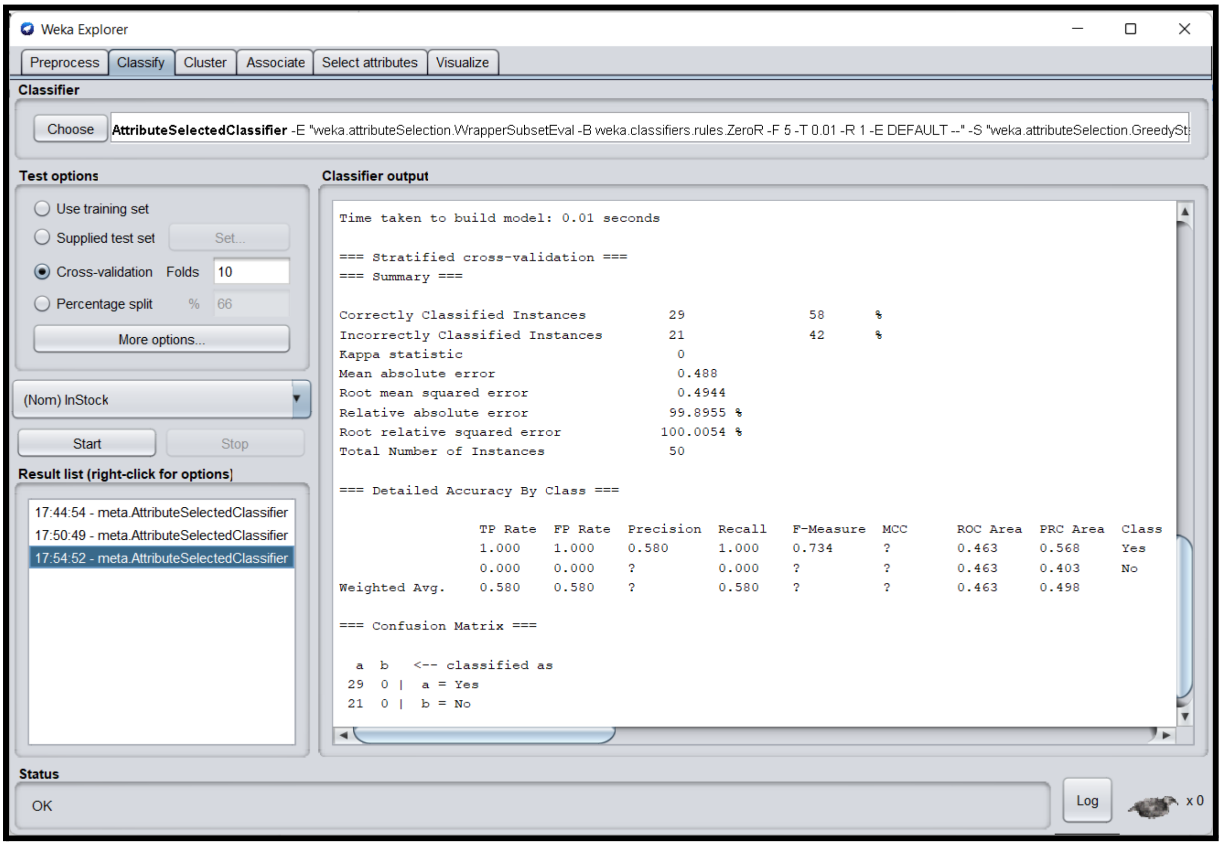

weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize 100
classifier Choose J48 -C 0.25 -M 2
debug False
doNotCheckCapabilities False
evaluator Choose CfsSubsetEval -P 1 -E 1
numDecimalPlaces 2
search Choose BestFirst -D 1 -N 5
Open...
Save...
OK
Cancel

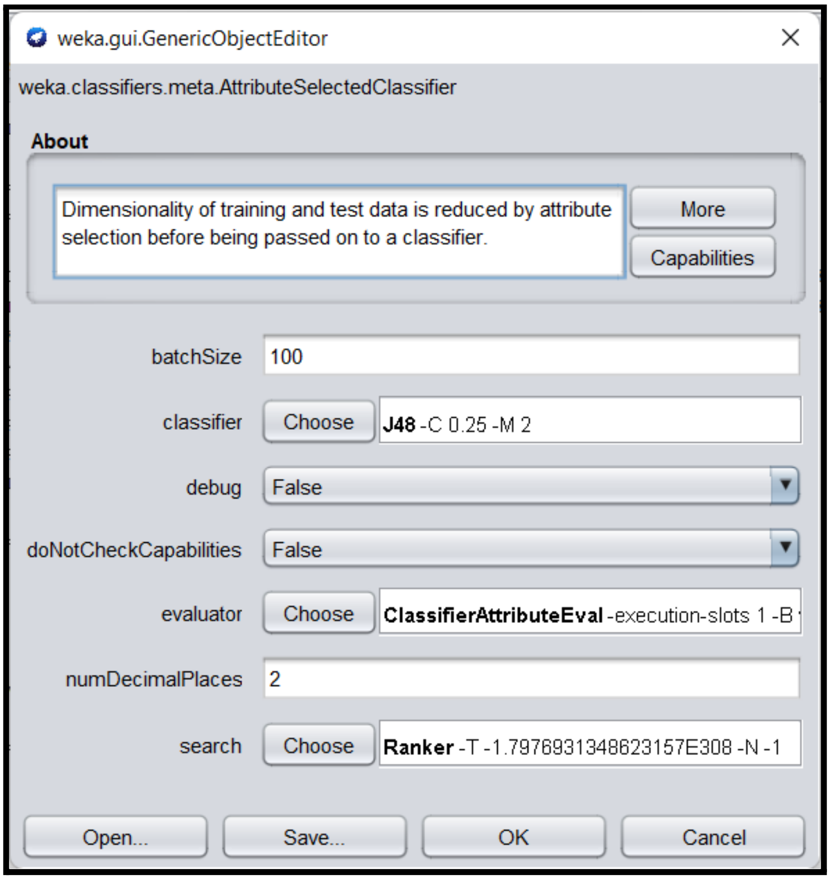

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose AttributeSelectedClassifier -E "weka.attributeSelection.CfsSubsetEval -P 1 -E 1" -S "weka.attributeSelection.BestFirst -D 1 -N 5" -W weka.classifiers.trees.J48 -- -C 0.25 -M 2
Test options
Use training set
Supplied test set
Set...
Cross-validation Folds 10
Percentage split % 66
More options...
(Nom) InStock
Start
Stop
Result list (right-click for options)
17:44:54 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 44 88 %
Incorrectly Classified Instances 6 12 %
Kappa statistic 0.7436
Mean absolute error 0.163
Root mean squared error 0.3098
Relative absolute error 33.3585 %
Root relative squared error 62.6707 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
1.000 0.286 0.829 1.000 0.906 0.769 0.897 0.891 Yes
0.714 0.000 1.000 0.714 0.833 0.769 0.897 0.899 No
Weighted Avg. 0.880 0.166 0.901 0.880 0.876 0.769 0.897 0.895
=== Confusion Matrix ===
a b <-- classified as
29 0 | a = Yes
6 15 | b = No
Status
OK
Log
x 0

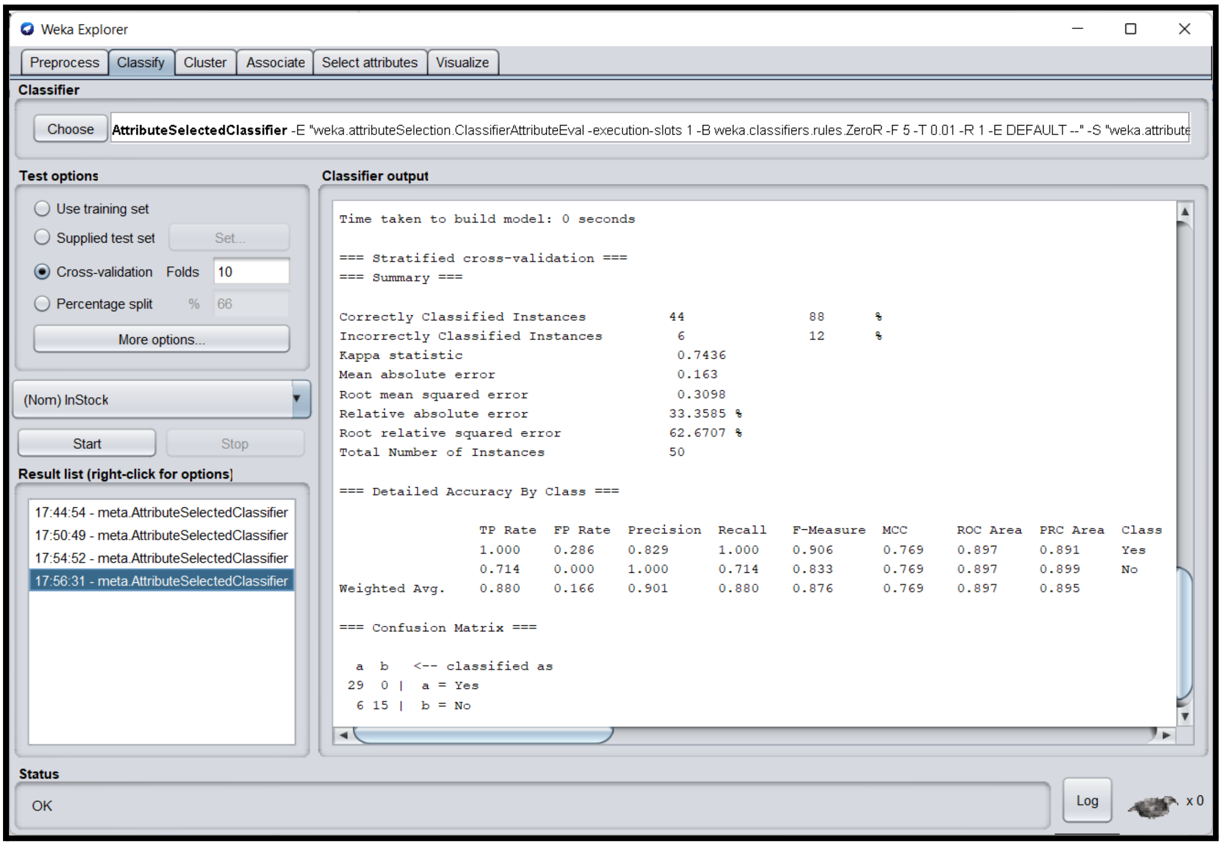

weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize
100
classifier
Choose
J48 -C 0.25 -M 2
debug
False
doNotCheckCapabilities
False
evaluator
Choose
WrapperSubsetEval -B weka.classifiers.rules
numDecimalPlaces
2
search
Choose
BestFirst -D 1 -N 5
Open...
Save...
OK
Cancel

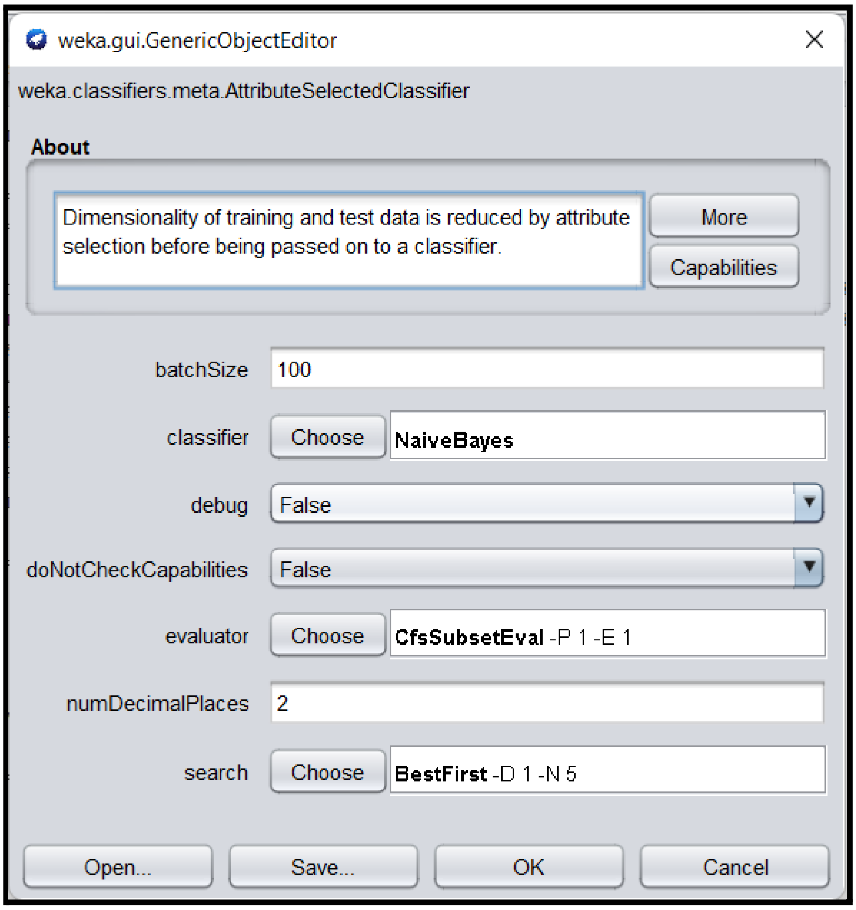

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
AttributeSelectedClassifier -E "weka.attributeSelection.WrapperSubsetEval -B weka.classifiers.rules.ZeroR -F 5 -T 0.01 -R 1 -E DEFAULT --" -S "weka.attributeSelection.BestFirst
Test options
Use training set
Supplied test set
Set...
Cross-validation Folds 10
Percentage split % 66
More options...
(Nom) InStock
Start
Stop
Result list (right-click for options)
17:44:54 - meta.AttributeSelectedClassifier
17:50:49 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0.03 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 29 58 %
Incorrectly Classified Instances 21 42 %
Kappa statistic 0
Mean absolute error 0.488
Root mean squared error 0.4944
Relative absolute error 99.8955 %
Root relative squared error 100.0054 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
1.000 1.000 0.580 1.000 0.734 ? 0.463 0.568 Yes
0.000 0.000 ? 0.000 ? ? 0.463 0.403 No
Weighted Avg. 0.580 0.580 ? 0.580 ? ? 0.463 0.498
=== Confusion Matrix ===
a b <-- classified as
29 0 | a = Yes
21 0 | b = No
Status
OK
Log
x 0

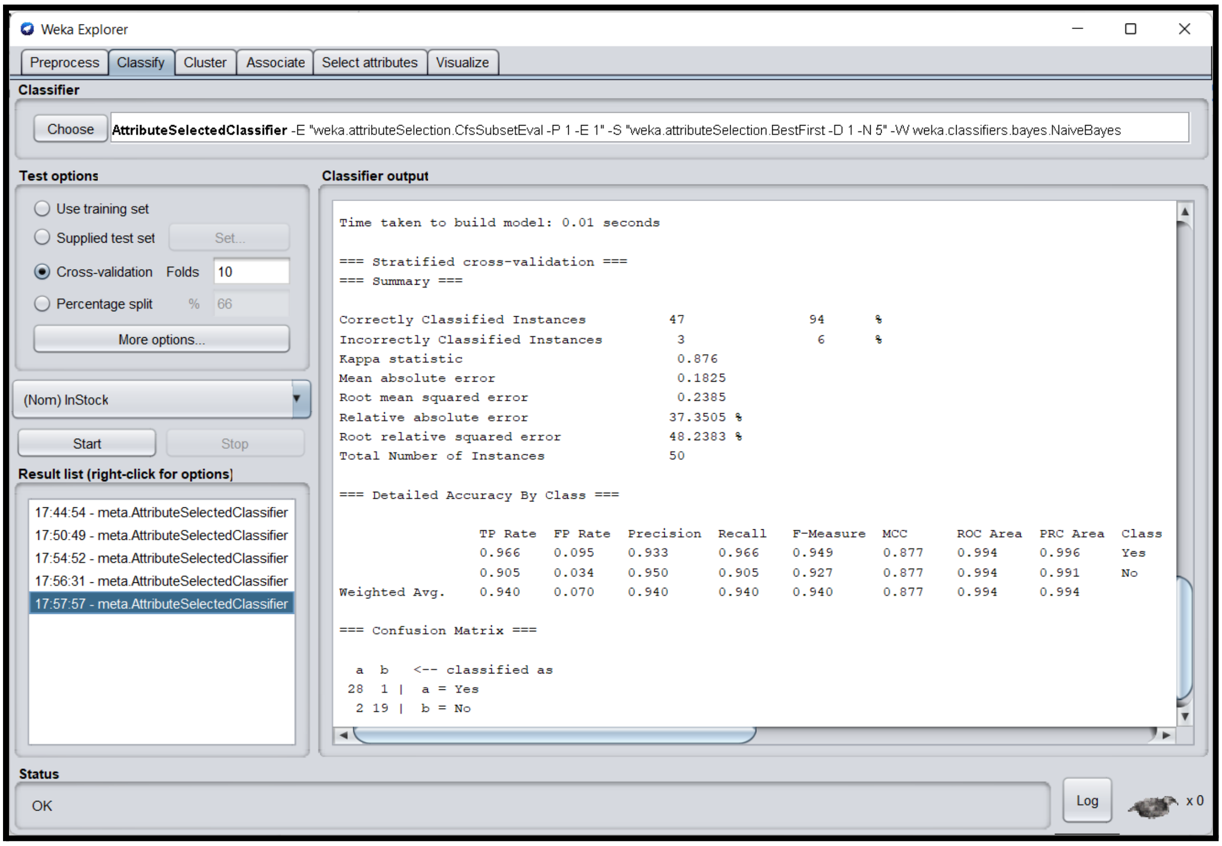

weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize 100
classifier Choose J48 -C 0.25 -M 2
debug False
doNotCheckCapabilities False
evaluator Choose WrapperSubsetEval -B weka.classifiers.rules
numDecimalPlaces 2
search Choose GreedyStepwise -T -1.7976931348623157E3
Open...
Save...
OK
Cancel

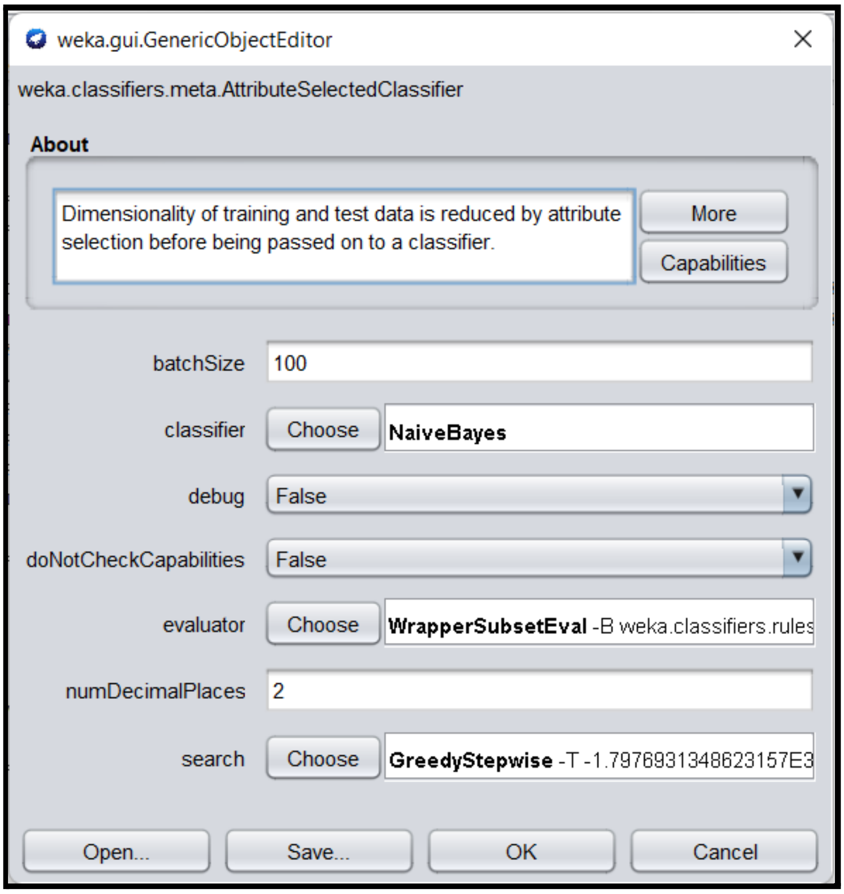

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
AttributeSelectedClassifier -E "weka.attributeSelection.WrapperSubsetEval -B weka.classifiers.rules.ZeroR -F 5 -T 0.01 -R 1 -E DEFAULT --" -S "weka.attributeSelection.GreedySt
Test options
Use training set
Supplied test set
Set...
Cross-validation Folds 10
Percentage split % 66
More options...
(Nom) InStock
Start
Stop
Result list (right-click for options)
17:44:54 - meta.AttributeSelectedClassifier
17:50:49 - meta.AttributeSelectedClassifier
17:54:52 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0.01 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 29 58 %
Incorrectly Classified Instances 21 42 %
Kappa statistic 0
Mean absolute error 0.488
Root mean squared error 0.4944
Relative absolute error 99.8955 %
Root relative squared error 100.0054 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
1.000 1.000 0.580 1.000 0.734 ? 0.463 0.568 Yes
0.000 0.000 ? 0.000 ? ? 0.463 0.403 No
Weighted Avg. 0.580 0.580 ? 0.580 ? ? 0.463 0.498
=== Confusion Matrix ===
a b <-- classified as
29 0 | a = Yes
21 0 | b = No
Status
OK
Log
x 0

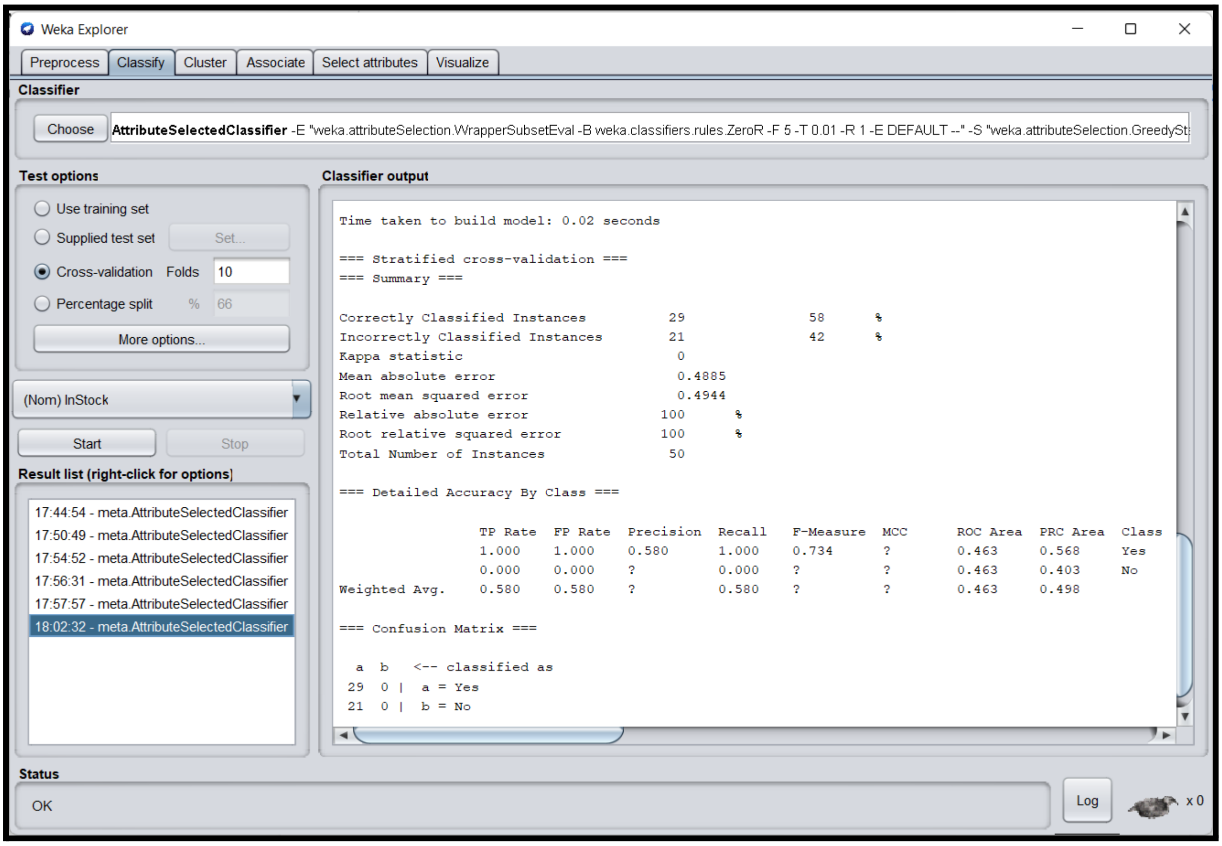

weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize
100
classifier
Choose
J48 -C 0.25 -M 2
debug
False
doNotCheckCapabilities
False
evaluator
Choose
ClassifierAttributeEval -execution-slots 1 -B
numDecimalPlaces
2
search
Choose
Ranker -T -1.7976931348623157E308 -N -1
Open...
Save...
OK
Cancel

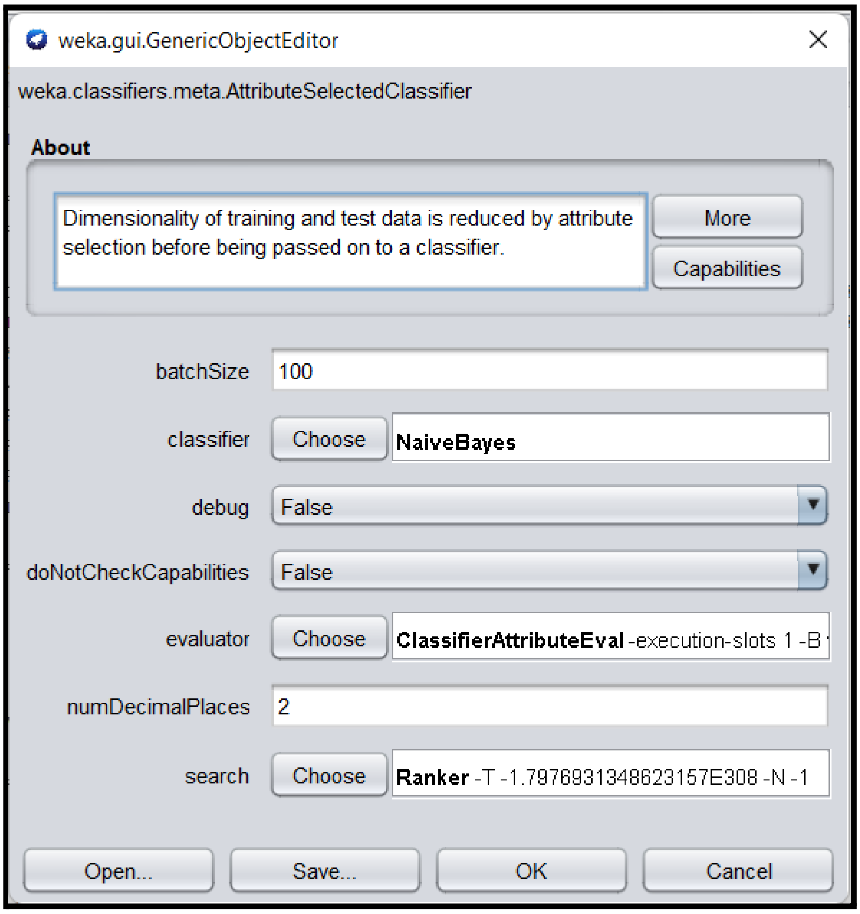

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
AttributeSelectedClassifier -E "weka.attributeSelection.ClassifierAttributeEval -execution-slots 1 -B weka.classifiers.rules.ZeroR -F 5 -T 0.01 -R 1 -E DEFAULT --" -S "weka.attribute
Test options
Use training set
Supplied test set
Set...
Cross-validation Folds 10
Percentage split % 66
More options...
(Nom) InStock
Start
Stop
Result list (right-click for options)
17:44:54 - meta.AttributeSelectedClassifier
17:50:49 - meta.AttributeSelectedClassifier
17:54:52 - meta.AttributeSelectedClassifier
17:56:31 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 44 88 %
Incorrectly Classified Instances 6 12 %
Kappa statistic 0.7436
Mean absolute error 0.163
Root mean squared error 0.3098
Relative absolute error 33.3585 %
Root relative squared error 62.6707 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
1.000 0.286 0.829 1.000 0.906 0.769 0.897 0.891 Yes
0.714 0.000 1.000 0.714 0.833 0.769 0.897 0.899 No
Weighted Avg. 0.880 0.166 0.901 0.880 0.876 0.769 0.897 0.895
=== Confusion Matrix ===
a b <-- classified as
29 0 | a = Yes
6 15 | b = No
Status
OK
Log
x 0

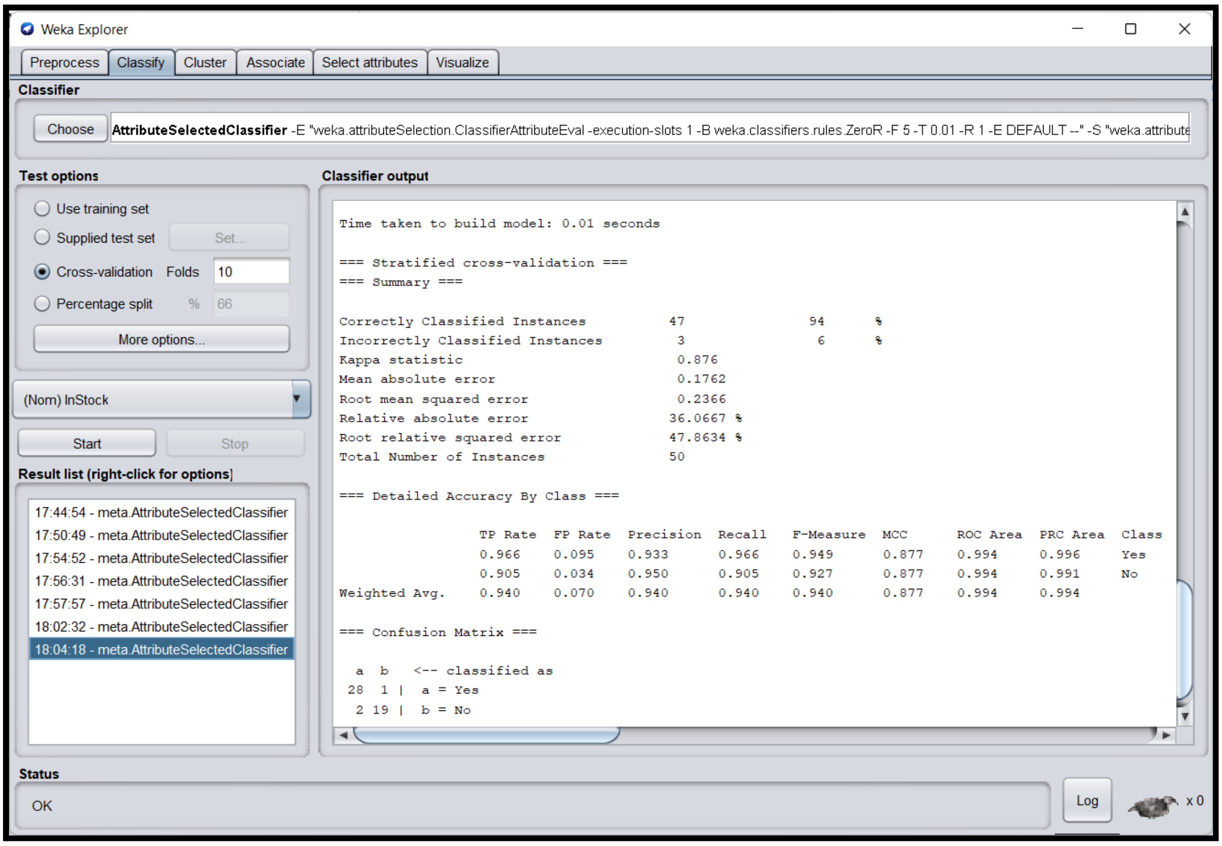

weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize 100
classifier Choose NaiveBayes
debug False
doNotCheckCapabilities False
evaluator Choose CfsSubsetEval -P 1 -E 1
numDecimalPlaces 2
search Choose BestFirst -D 1 -N 5
Open...
Save...
OK
Cancel

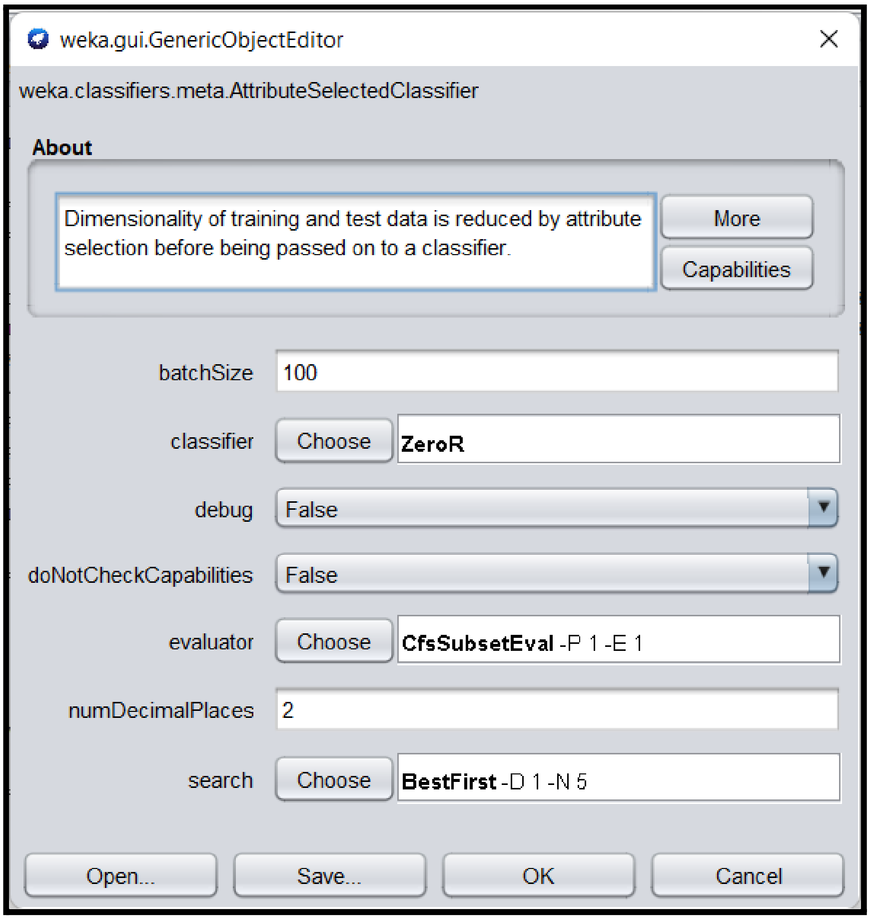

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose AttributeSelectedClassifier -E "weka.attributeSelection.CfsSubsetEval -P 1 -E 1" -S "weka.attributeSelection.BestFirst -D 1 -N 5" -W weka.classifiers.bayes.NaiveBayes
Test options
Use training set
Supplied test set Set...
Cross-validation Folds 10
Percentage split % 66
More options...
(Nom) InStock
Start
Stop
Result list (right-click for options)
17:44:54 - meta.AttributeSelectedClassifier
17:50:49 - meta.AttributeSelectedClassifier
17:54:52 - meta.AttributeSelectedClassifier
17:56:31 - meta.AttributeSelectedClassifier
17:57:57 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0.01 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 47 94 %
Incorrectly Classified Instances 3 6 %
Kappa statistic 0.876
Mean absolute error 0.1825
Root mean squared error 0.2385
Relative absolute error 37.3505 %
Root relative squared error 48.2383 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.966 0.095 0.933 0.966 0.949 0.877 0.994 0.996 Yes
0.905 0.034 0.950 0.905 0.927 0.877 0.994 0.991 No
Weighted Avg. 0.940 0.070 0.940 0.940 0.940 0.877 0.994 0.994
=== Confusion Matrix ===
a b <-- classified as
28 1 | a = Yes
2 19 | b = No
Status
OK
Log
x 0

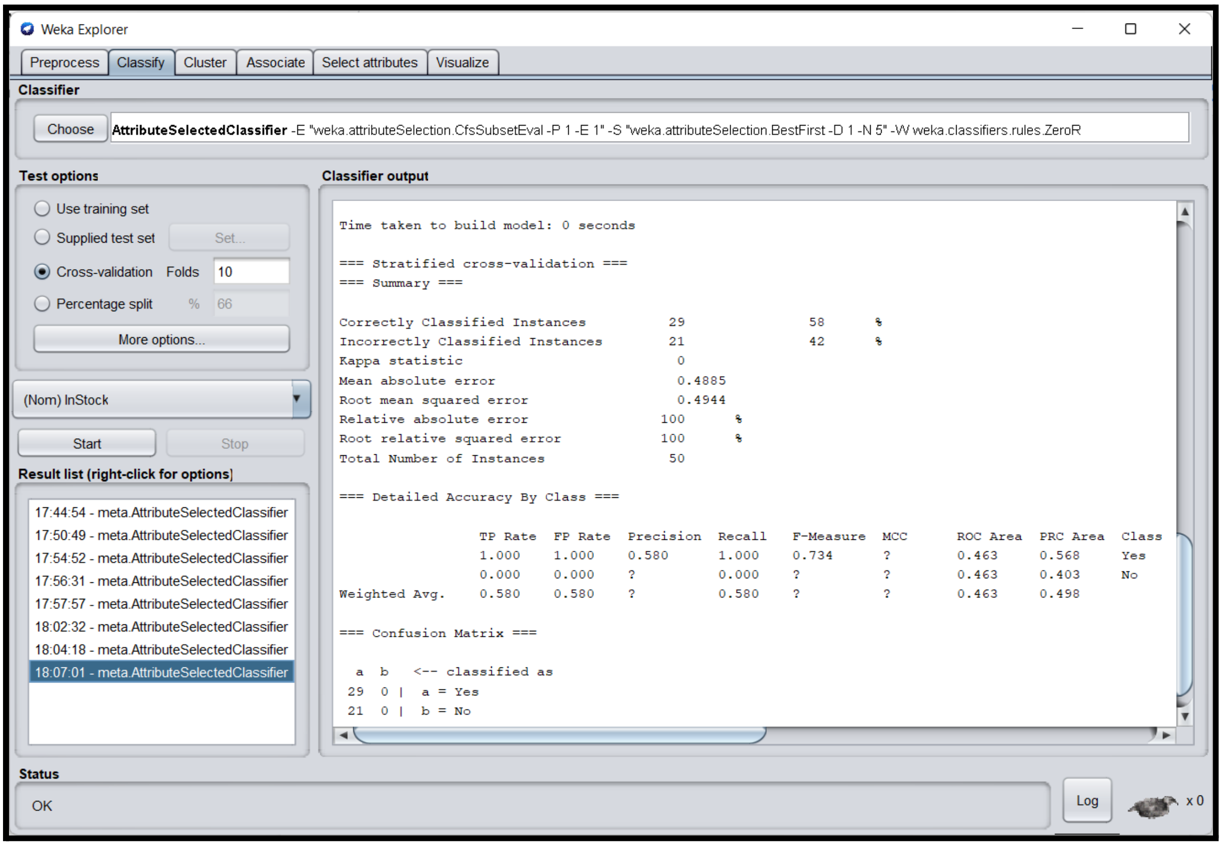

weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize 100
classifier Choose NaiveBayes
debug False
doNotCheckCapabilities False
evaluator Choose WrapperSubsetEval -B weka.classifiers.rules
numDecimalPlaces 2
search Choose GreedyStepwise -T -1.7976931348623157E3
Open... Save... OK Cancel

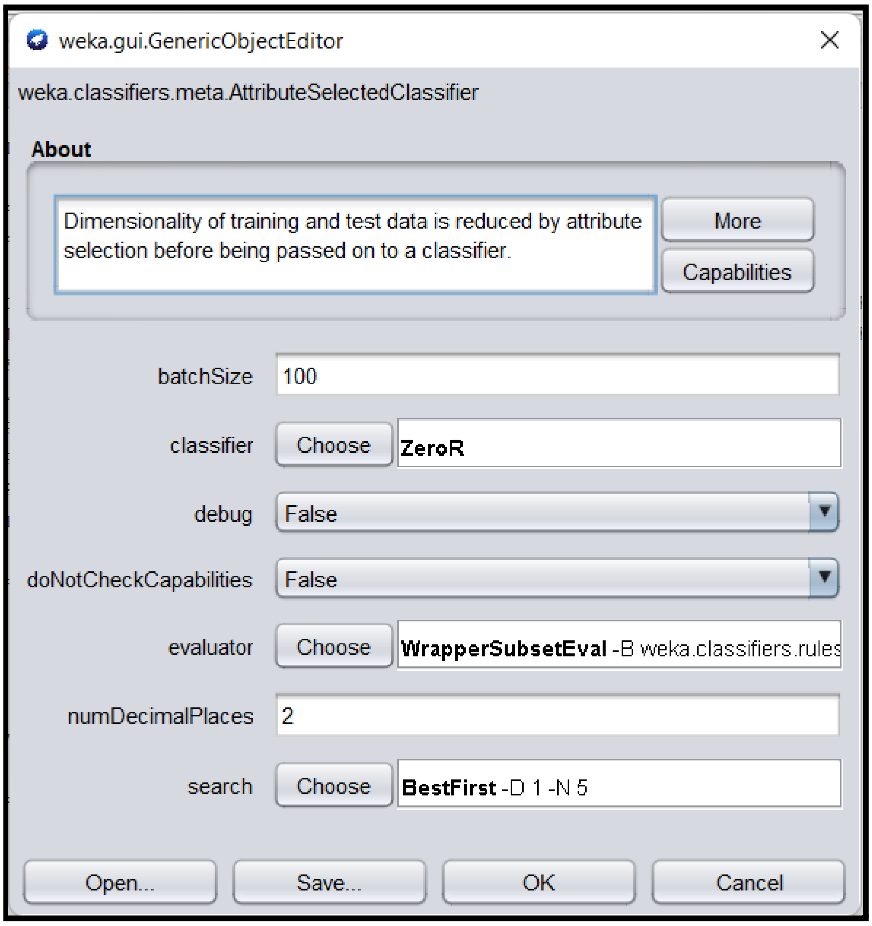

Weka Explorer
Preprocess Classify Cluster Associate Select attributes Visualize
Classifier
Choose AttributeSelectedClassifier -E "weka.attributeSelection.WrapperSubsetEval -B weka.classifiers.rules.ZeroR -F 5 -T 0.01 -R 1 -E DEFAULT --" -S "weka.attributeSelection.GreedySt
Test options
Use training set
Supplied test set Set...
Cross-validation Folds 10
Percentage split % 66
More options...
(Nom) InStock
Start Stop
Result list (right-click for options)
17:44:54 - meta.AttributeSelectedClassifier
17:50:49 - meta.AttributeSelectedClassifier
17:54:52 - meta.AttributeSelectedClassifier
17:56:31 - meta.AttributeSelectedClassifier
17:57:57 - meta.AttributeSelectedClassifier
18:02:32 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0.02 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 29 58 %
Incorrectly Classified Instances 21 42 %
Kappa statistic 0
Mean absolute error 0.4885
Root mean squared error 0.4944
Relative absolute error 100 %
Root relative squared error 100 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
1.000 1.000 0.580 1.000 0.734 ? 0.463 0.568 Yes
0.000 0.000 ? 0.000 ? ? 0.463 0.403 No
Weighted Avg. 0.580 0.580 ? 0.580 ? ? 0.463 0.498
=== Confusion Matrix ===
a b <-- classified as
29 0 | a = Yes
21 0 | b = No
Status
OK
Log
x 0

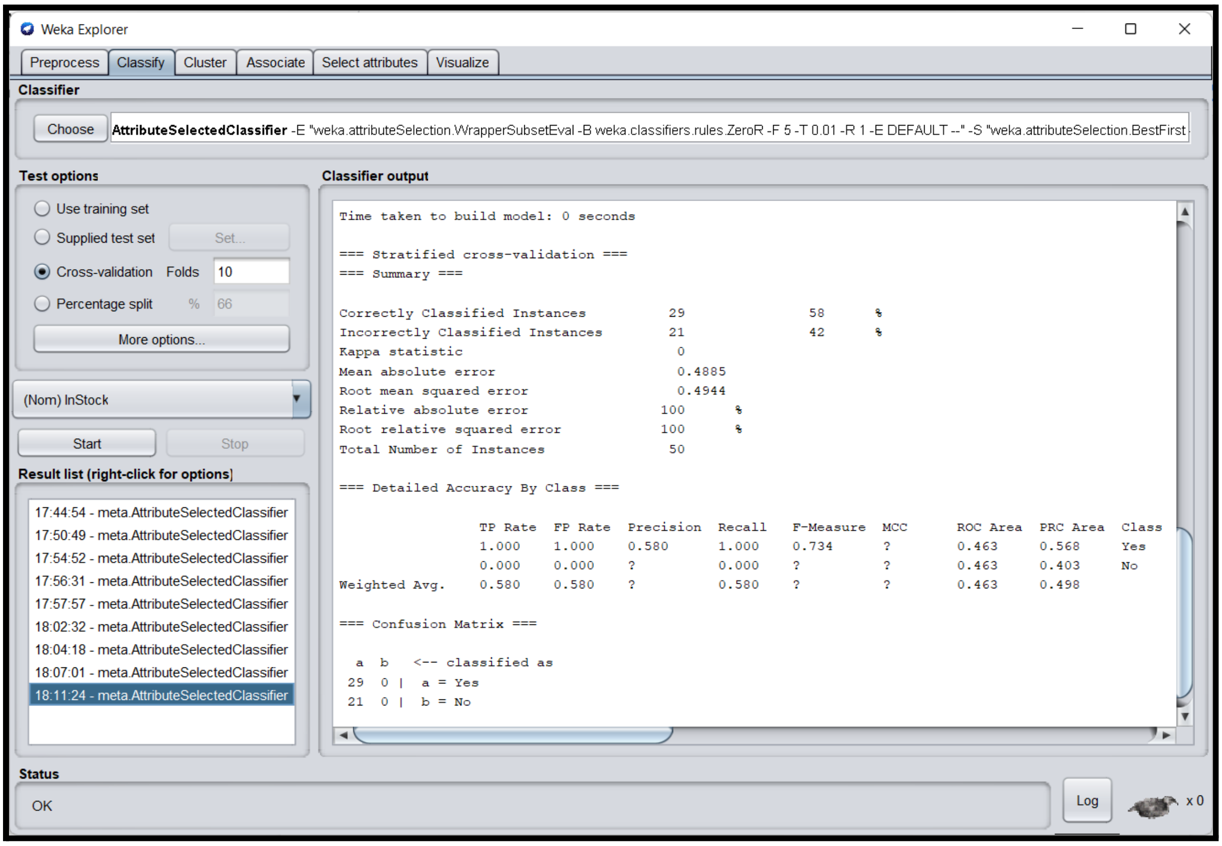

weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize 100
classifier Choose NaiveBayes
debug False
doNotCheckCapabilities False
evaluator Choose ClassifierAttributeEval -execution-slots 1 -B
numDecimalPlaces 2
search Choose Ranker -T -1.7976931348623157E308 -N -1
Open... Save... OK Cancel

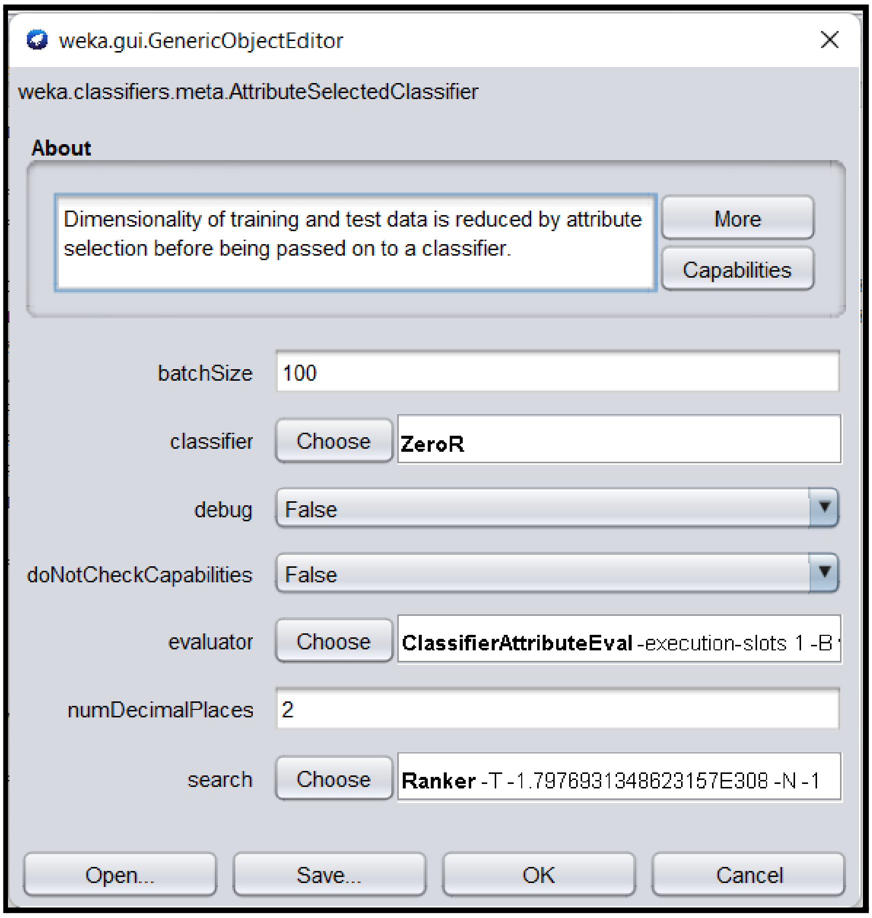

Weka Explorer
Preprocess Classify Cluster Associate Select attributes Visualize
Classifier
Choose AttributeSelectedClassifier -E "weka.attributeSelection.ClassifierAttributeEval -execution-slots 1 -B weka.classifiers.rules.ZeroR -F 5 -T 0.01 -R 1 -E DEFAULT --" -S "weka.attribute
Test options
Use training set
Supplied test set Set...
Cross-validation Folds 10
Percentage split % 66
More options...
(Nom) InStock
Start Stop
Result list (right-click for options)
17:44:54 - meta.AttributeSelectedClassifier
17:50:49 - meta.AttributeSelectedClassifier
17:54:52 - meta.AttributeSelectedClassifier
17:56:31 - meta.AttributeSelectedClassifier
17:57:57 - meta.AttributeSelectedClassifier
18:02:32 - meta.AttributeSelectedClassifier
18:04:18 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0.01 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 47 94 %
Incorrectly Classified Instances 3 6 %
Kappa statistic 0.876
Mean absolute error 0.1762
Root mean squared error 0.2366
Relative absolute error 36.0667 %
Root relative squared error 47.8634 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.966 0.095 0.933 0.966 0.949 0.877 0.994 0.996 Yes
0.905 0.034 0.950 0.905 0.927 0.877 0.994 0.991 No
Weighted Avg. 0.940 0.070 0.940 0.940 0.940 0.877 0.994 0.994
=== Confusion Matrix ===
a b <-- classified as
28 1 | a = Yes
2 19 | b = No
Status
OK
Log
x 0

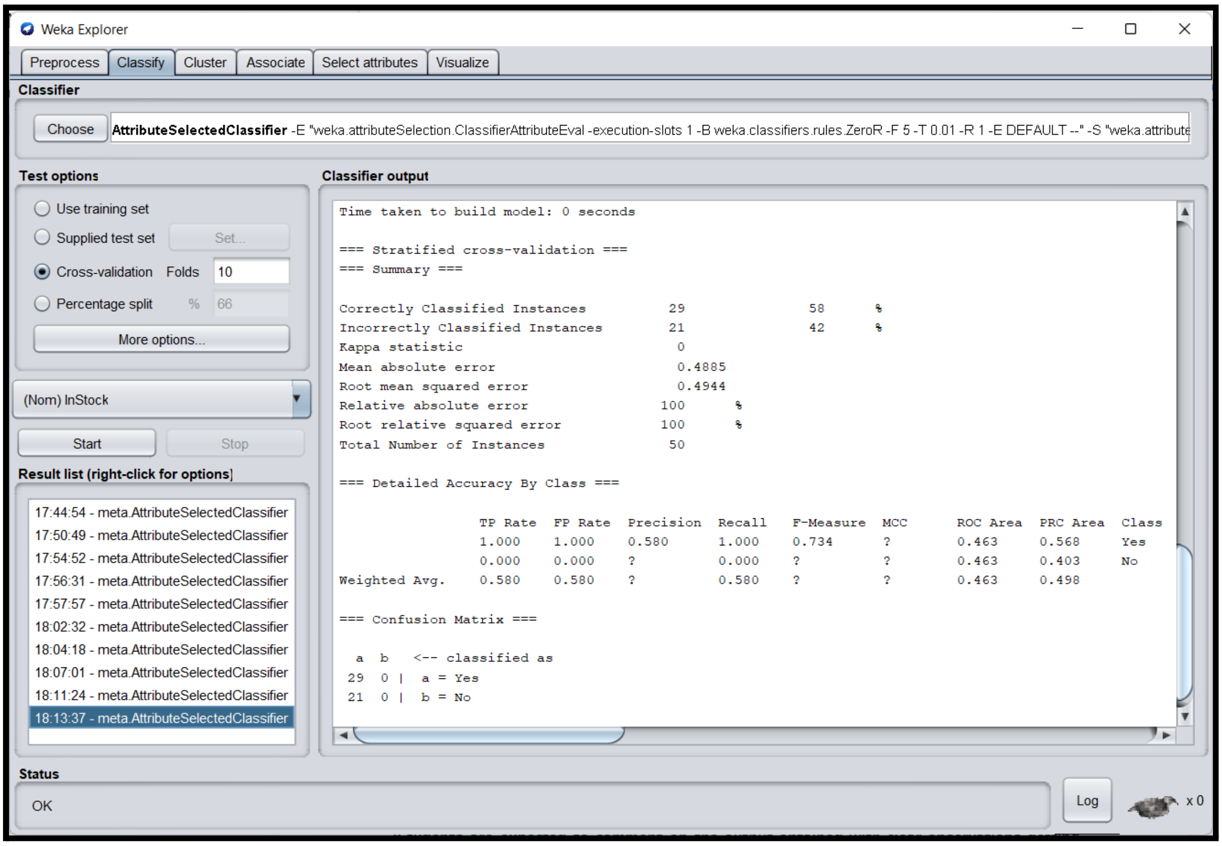

weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize 100
classifier Choose ZeroR
debug False
doNotCheckCapabilities False
evaluator Choose CfsSubsetEval -P 1 -E 1
numDecimalPlaces 2
search Choose BestFirst -D 1 -N 5
Open...
Save...
OK
Cancel

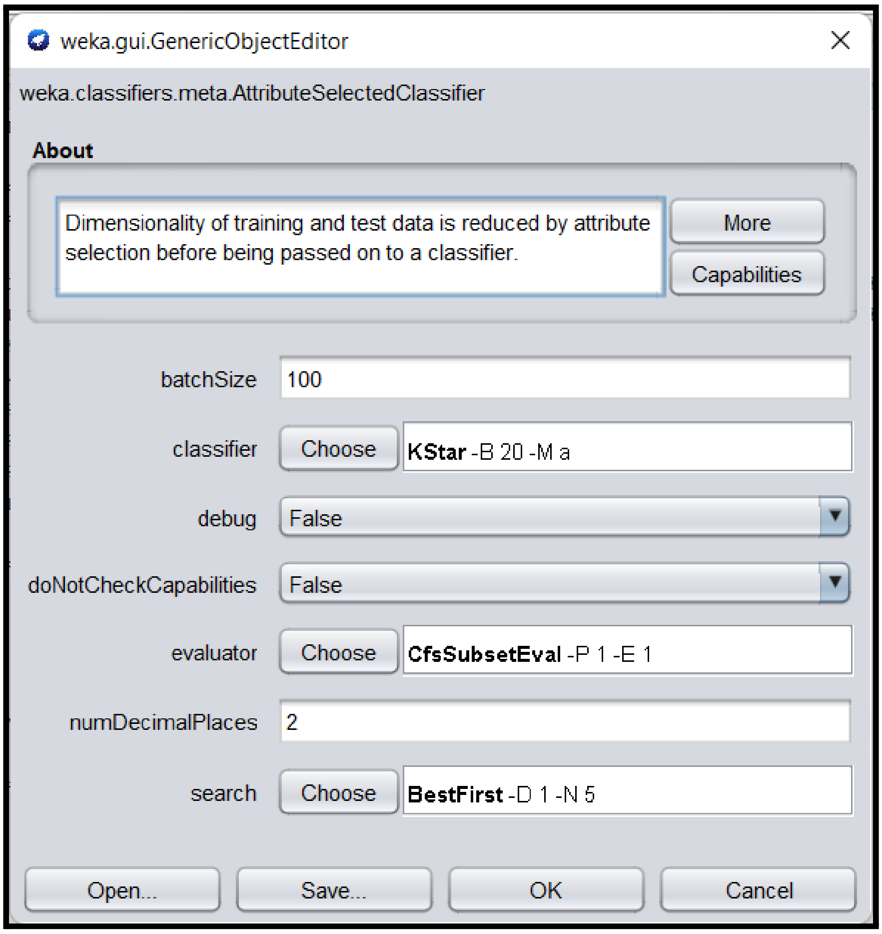

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose AttributeSelectedClassifier -E "weka.attributeSelection.CfsSubsetEval -P 1 -E 1" -S "weka.attributeSelection.BestFirst -D 1 -N 5" -W weka.classifiers.rules.ZeroR
Test options
Use training set
Supplied test set Set...
Cross-validation Folds 10
Percentage split % 66
More options...
(Nom) InStock
Start
Stop
Result list (right-click for options)
17:44:54 - meta.AttributeSelectedClassifier
17:50:49 - meta.AttributeSelectedClassifier
17:54:52 - meta.AttributeSelectedClassifier
17:56:31 - meta.AttributeSelectedClassifier
17:57:57 - meta.AttributeSelectedClassifier
18:02:32 - meta.AttributeSelectedClassifier
18:04:18 - meta.AttributeSelectedClassifier
18:07:01 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 29 58 %
Incorrectly Classified Instances 21 42 %
Kappa statistic 0
Mean absolute error 0.4885
Root mean squared error 0.4944
Relative absolute error 100 %
Root relative squared error 100 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
1.000 1.000 0.580 1.000 0.734 ? 0.463 0.568 Yes
0.000 0.000 ? 0.000 ? ? 0.463 0.403 No
Weighted Avg. 0.580 0.580 ? 0.580 ? ? 0.463 0.498
=== Confusion Matrix ===
a b <-- classified as
29 0 | a = Yes
21 0 | b = No
Status
OK
Log
x 0

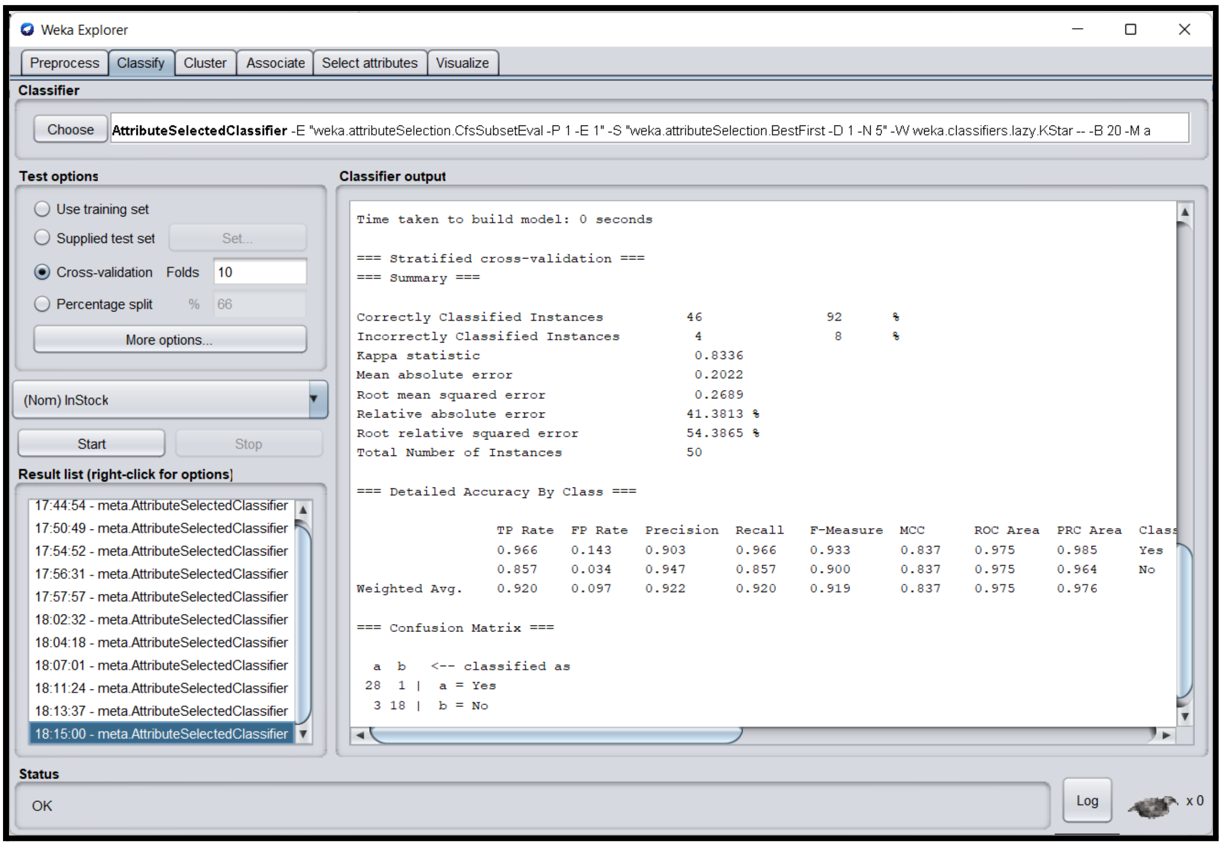

weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize
100
classifier
Choose
ZeroR
debug
False
doNotCheckCapabilities
False
evaluator
Choose
WrapperSubsetEval -B weka.classifiers.rules
numDecimalPlaces
2
search
Choose
BestFirst -D 1 -N 5
Open...
Save...
OK
Cancel

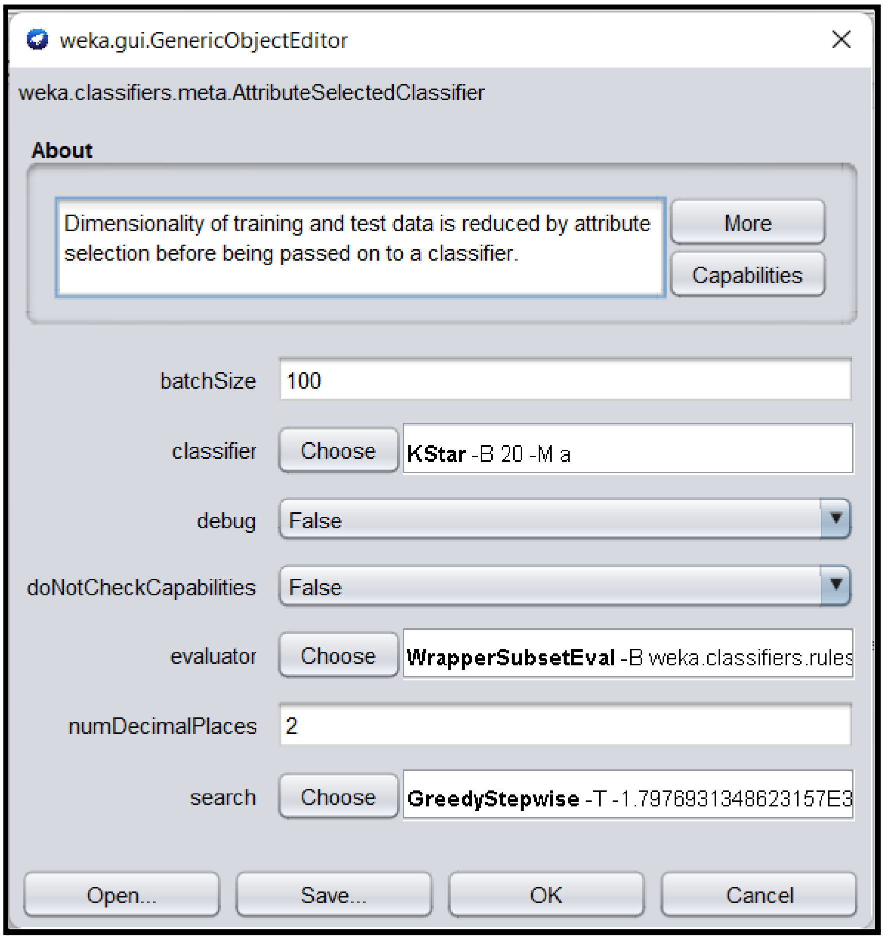

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
AttributeSelectedClassifier -E "weka.attributeSelection.WrapperSubsetEval -B weka.classifiers.rules.ZeroR -F 5 -T 0.01 -R 1 -E DEFAULT --" -S "weka.attributeSelection.BestFirst
Test options
Use training set
Supplied test set
Set...
Cross-validation Folds 10
Percentage split % 66
More options...
(Nom) InStock
Start
Stop
Result list (right-click for options)
17:44:54 - meta.AttributeSelectedClassifier
17:50:49 - meta.AttributeSelectedClassifier
17:54:52 - meta.AttributeSelectedClassifier
17:56:31 - meta.AttributeSelectedClassifier
17:57:57 - meta.AttributeSelectedClassifier
18:02:32 - meta.AttributeSelectedClassifier
18:04:18 - meta.AttributeSelectedClassifier
18:07:01 - meta.AttributeSelectedClassifier
18:11:24 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 29 58 %
Incorrectly Classified Instances 21 42 %
Kappa statistic 0
Mean absolute error 0.4885
Root mean squared error 0.4944
Relative absolute error 100 %
Root relative squared error 100 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
1.000 1.000 0.580 1.000 0.734 ? 0.463 0.568 Yes
0.000 0.000 ? 0.000 ? ? 0.463 0.403 No
Weighted Avg. 0.580 0.580 ? 0.580 ? ? 0.463 0.498
=== Confusion Matrix ===
a b <-- classified as
29 0 | a = Yes
21 0 | b = No
Status
OK
Log
x 0

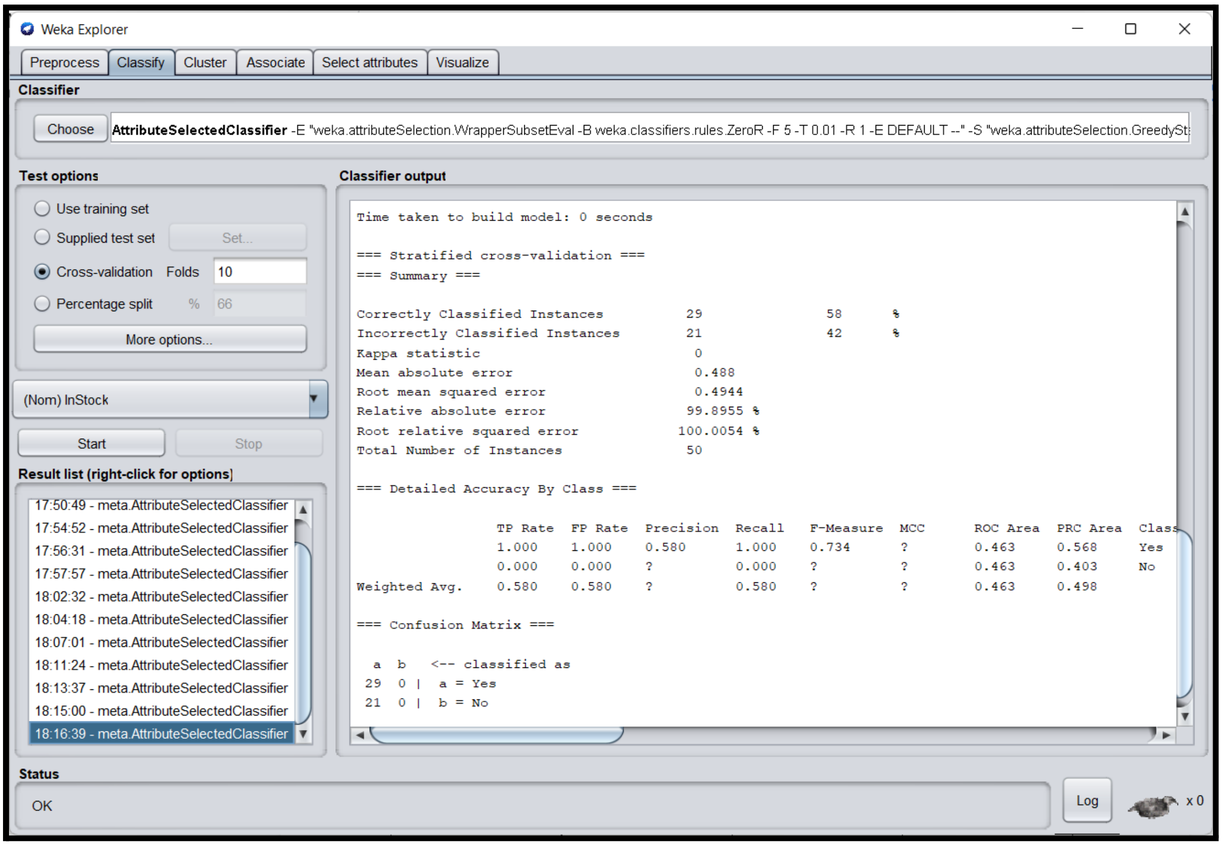

weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize
100
classifier
Choose
ZeroR
debug
False
doNotCheckCapabilities
False
evaluator
Choose
ClassifierAttributeEval -execution-slots 1 -B
numDecimalPlaces
2
search
Choose
Ranker -T -1.7976931348623157E308 -N -1
Open...
Save...
OK
Cancel

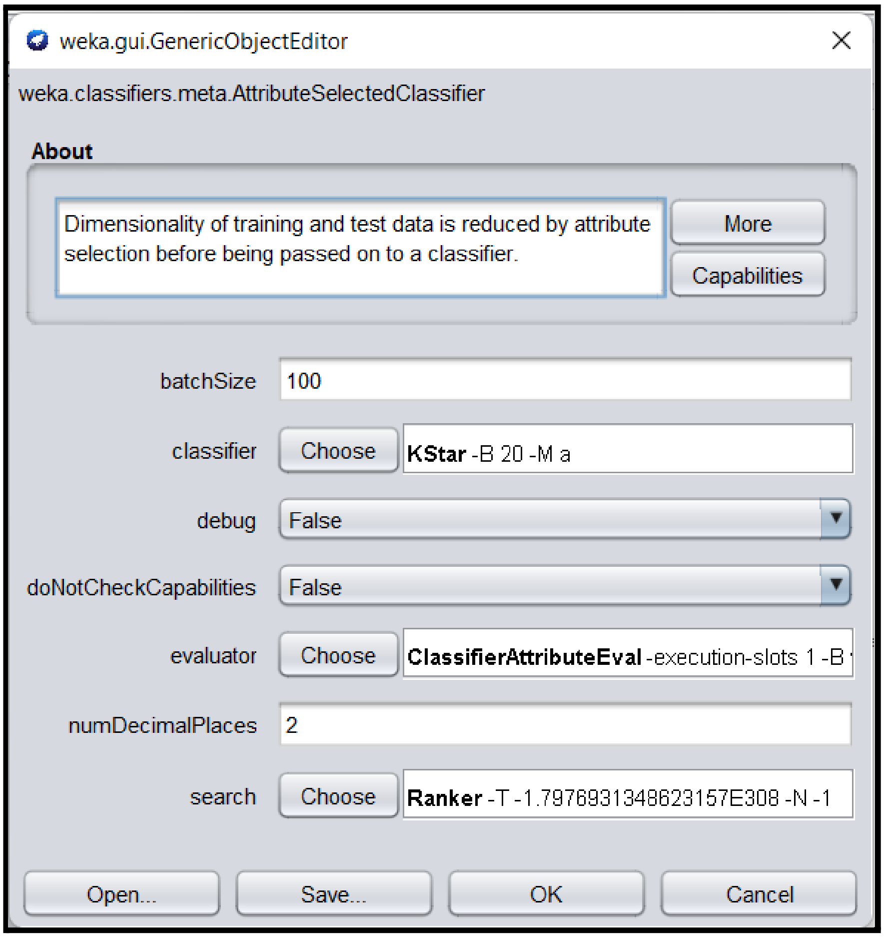

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
AttributeSelectedClassifier -E "weka.attributeSelection.ClassifierAttributeEval -execution-slots 1 -B weka.classifiers.rules.ZeroR -F 5 -T 0.01 -R 1 -E DEFAULT --" -S "weka.attribute
Test options
Use training set
Supplied test set
Set...
Cross-validation Folds 10
Percentage split % 66
More options...
(Nom) InStock
Start
Stop
Result list (right-click for options)
17:44:54 - meta.AttributeSelectedClassifier
17:50:49 - meta.AttributeSelectedClassifier
17:54:52 - meta.AttributeSelectedClassifier
17:56:31 - meta.AttributeSelectedClassifier
17:57:57 - meta.AttributeSelectedClassifier
18:02:32 - meta.AttributeSelectedClassifier
18:04:18 - meta.AttributeSelectedClassifier
18:07:01 - meta.AttributeSelectedClassifier
18:11:24 - meta.AttributeSelectedClassifier
18:13:37 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 29 58 %
Incorrectly Classified Instances 21 42 %
Kappa statistic 0
Mean absolute error 0.4885
Root mean squared error 0.4944
Relative absolute error 100 %
Root relative squared error 100 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
1.000 1.000 0.580 1.000 0.734 ? 0.463 0.568 Yes
0.000 0.000 ? 0.000 ? ? 0.463 0.403 No
Weighted Avg. 0.580 0.580 ? 0.580 ? ? 0.463 0.498
=== Confusion Matrix ===
a b <-- classified as
29 0 | a = Yes
21 0 | b = No
Status
OK
Log
x 0

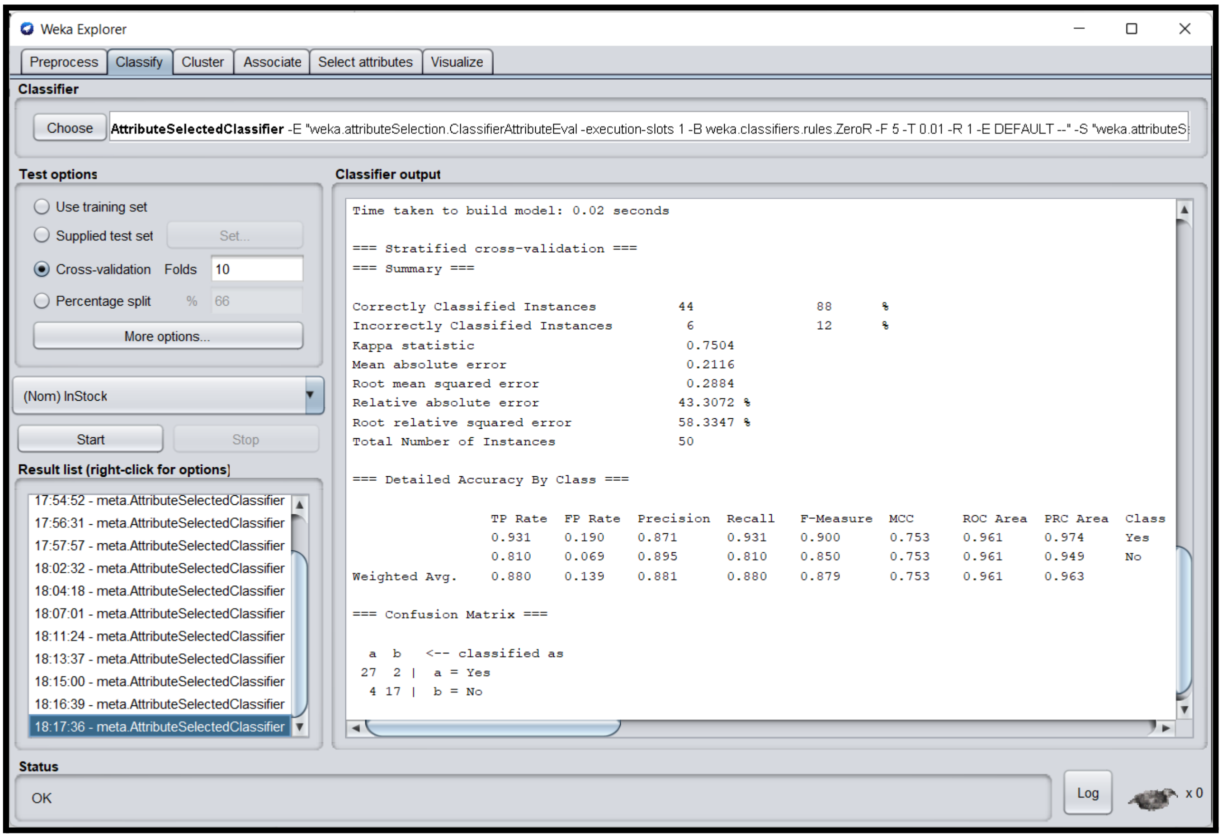

weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize
100
classifier
Choose
KStar -B 20 -M a
debug
False
doNotCheckCapabilities
False
evaluator
Choose
CfsSubsetEval -P 1 -E 1
numDecimalPlaces
2
search
Choose
BestFirst -D 1 -N 5
Open...
Save...
OK
Cancel

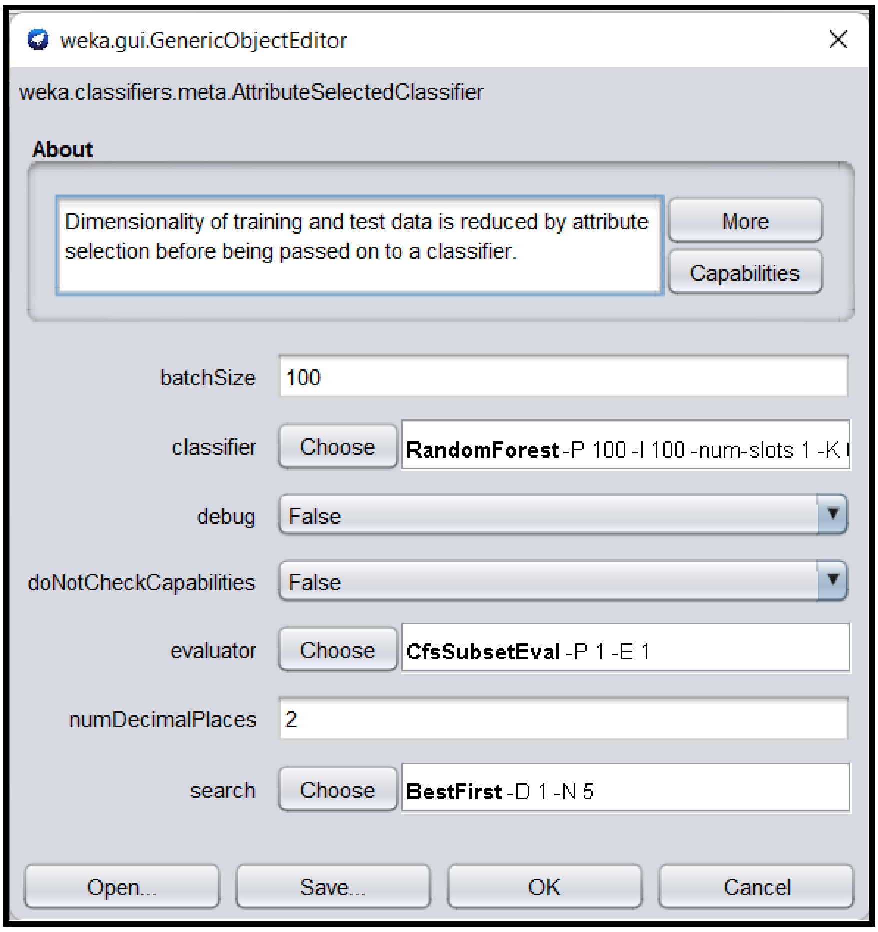

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
AttributeSelectedClassifier -E "weka.attributeSelection.CfsSubsetEval -P 1 -E 1" -S "weka.attributeSelection.BestFirst -D 1 -N 5" -W weka.classifiers.lazy.KStar -- -B 20 -M a
Test options
Use training set
Supplied test set
Set...
Cross-validation Folds 10
Percentage split % 66
More options...
(Nom) InStock
Start
Stop
Result list (right-click for options)
17:44:54 - meta.AttributeSelectedClassifier
17:50:49 - meta.AttributeSelectedClassifier
17:54:52 - meta.AttributeSelectedClassifier
17:56:31 - meta.AttributeSelectedClassifier
17:57:57 - meta.AttributeSelectedClassifier
18:02:32 - meta.AttributeSelectedClassifier
18:04:18 - meta.AttributeSelectedClassifier
18:07:01 - meta.AttributeSelectedClassifier
18:11:24 - meta.AttributeSelectedClassifier
18:13:37 - meta.AttributeSelectedClassifier
18:15:00 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 46 92 %
Incorrectly Classified Instances 4 8 %
Kappa statistic 0.8336
Mean absolute error 0.2022
Root mean squared error 0.2689
Relative absolute error 41.3813 %
Root relative squared error 54.3865 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.966 0.143 0.903 0.966 0.933 0.837 0.975 0.985 Yes
0.857 0.034 0.947 0.857 0.900 0.837 0.975 0.964 No
Weighted Avg. 0.920 0.097 0.922 0.920 0.919 0.837 0.975 0.976
=== Confusion Matrix ===
a b <-- classified as
28 1 | a = Yes
3 18 | b = No
Status
OK
Log
x 0

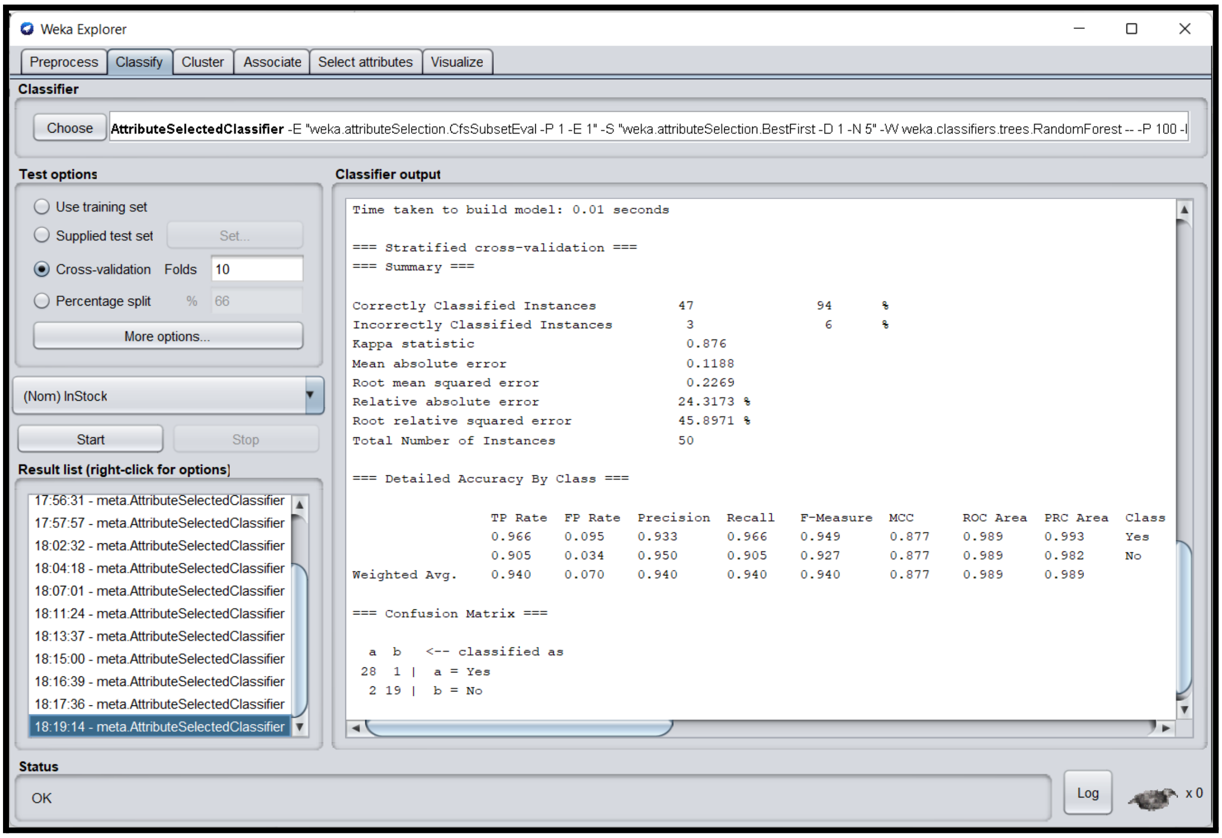

weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize 100
classifier Choose KStar -B 20 -M a
debug False
doNotCheckCapabilities False
evaluator Choose WrapperSubsetEval -B weka.classifiers.rules
numDecimalPlaces 2
search Choose GreedyStepwise -T -1.7976931348623157E3
Open... Save... OK Cancel

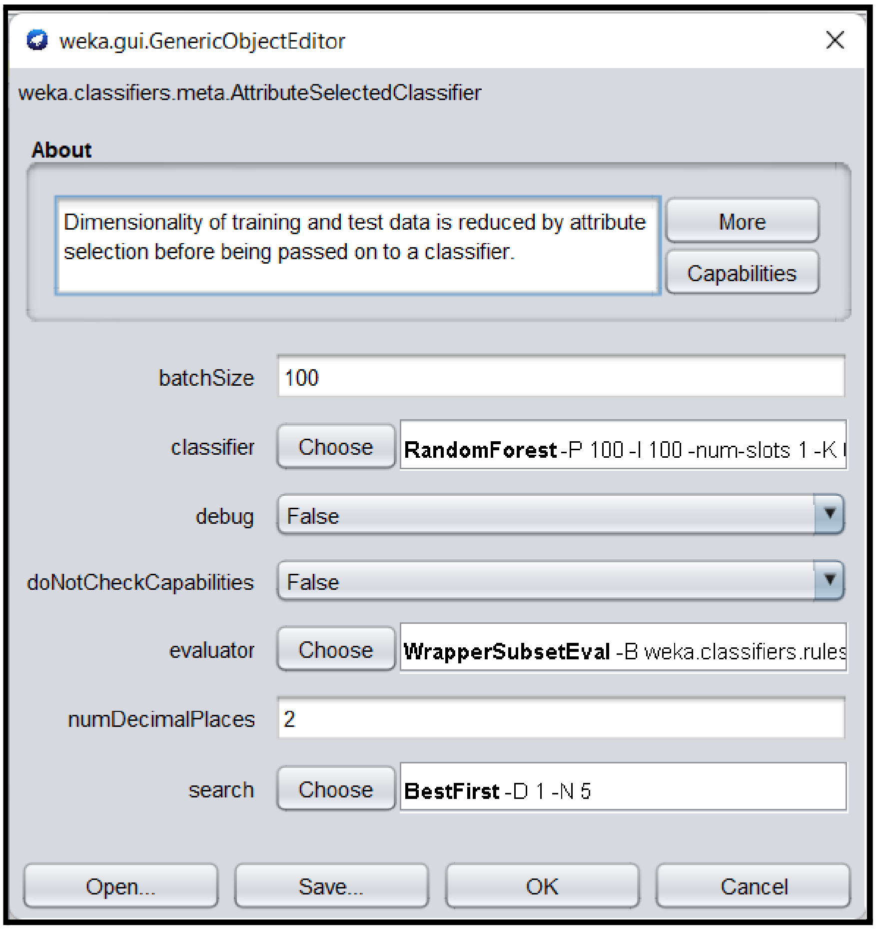

Weka Explorer
Preprocess Classify Cluster Associate Select attributes Visualize
Classifier
Choose AttributeSelectedClassifier -E "weka.attributeSelection.WrapperSubsetEval -B weka.classifiers.rules.ZeroR -F 5 -T 0.01 -R 1 -E DEFAULT --" -S "weka.attributeSelection.GreedySt
Test options
Use training set
Supplied test set Set...
Cross-validation Folds 10
Percentage split % 66
More options...
(Nom) InStock
Start Stop
Result list (right-click for options)
17:50:49 - meta.AttributeSelectedClassifier
17:54:52 - meta.AttributeSelectedClassifier
17:56:31 - meta.AttributeSelectedClassifier
17:57:57 - meta.AttributeSelectedClassifier
18:02:32 - meta.AttributeSelectedClassifier
18:04:18 - meta.AttributeSelectedClassifier
18:07:01 - meta.AttributeSelectedClassifier
18:11:24 - meta.AttributeSelectedClassifier
18:13:37 - meta.AttributeSelectedClassifier
18:15:00 - meta.AttributeSelectedClassifier
18:16:39 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 29 58 %
Incorrectly Classified Instances 21 42 %
Kappa statistic 0
Mean absolute error 0.488
Root mean squared error 0.4944
Relative absolute error 99.8955 %
Root relative squared error 100.0054 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
1.000 1.000 0.580 1.000 0.734 ? 0.463 0.568 Yes
0.000 0.000 ? 0.000 ? ? 0.463 0.403 No
Weighted Avg. 0.580 0.580 ? 0.580 ? ? 0.463 0.498
=== Confusion Matrix ===
a b <-- classified as
29 0 | a = Yes
21 0 | b = No
Status
OK
Log
x 0

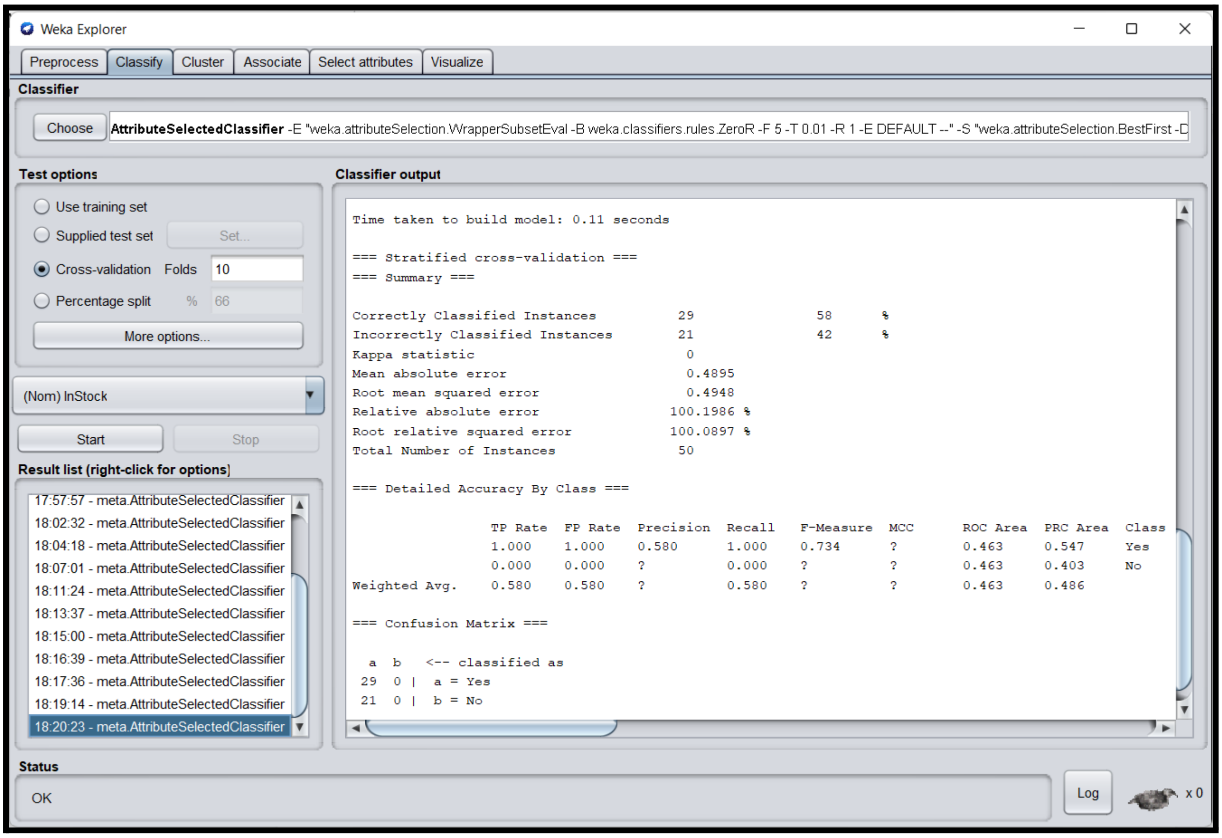

weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize 100
classifier Choose KStar -B 20 -M a
debug False
doNotCheckCapabilities False
evaluator Choose ClassifierAttributeEval -execution-slots 1 -B
numDecimalPlaces 2
search Choose Ranker -T -1.7976931348623157E308 -N -1
Open... Save... OK Cancel

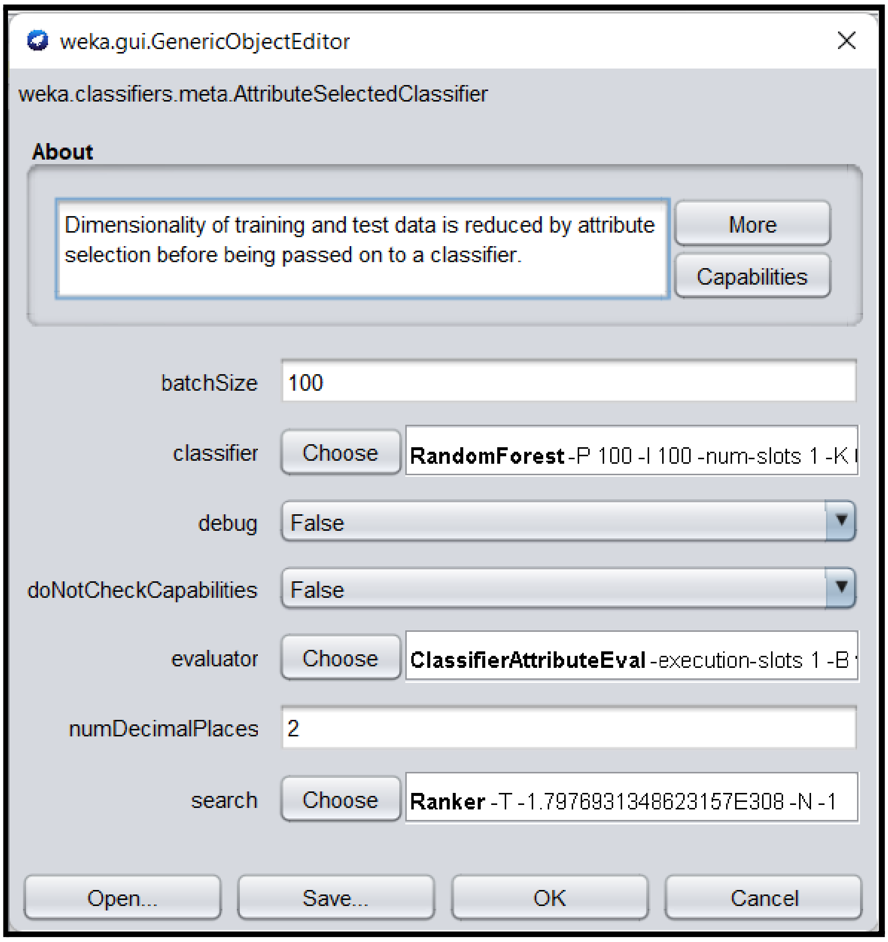

Weka Explorer
Preprocess Classify Cluster Associate Select attributes Visualize
Classifier
Choose AttributeSelectedClassifier -E "weka.attributeSelection.ClassifierAttributeEval -execution-slots 1 -B weka.classifiers.rules.ZeroR -F 5 -T 0.01 -R 1 -E DEFAULT --" -S "weka.attributeS
Test options
Use training set
Supplied test set Set...
Cross-validation Folds 10
Percentage split % 66
More options...
(Nom) InStock
Start Stop
Result list (right-click for options)
17:54:52 - meta.AttributeSelectedClassifier
17:56:31 - meta.AttributeSelectedClassifier
17:57:57 - meta.AttributeSelectedClassifier
18:02:32 - meta.AttributeSelectedClassifier
18:04:18 - meta.AttributeSelectedClassifier
18:07:01 - meta.AttributeSelectedClassifier
18:11:24 - meta.AttributeSelectedClassifier
18:13:37 - meta.AttributeSelectedClassifier
18:15:00 - meta.AttributeSelectedClassifier
18:16:39 - meta.AttributeSelectedClassifier
18:17:36 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0.02 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 44 88 %
Incorrectly Classified Instances 6 12 %
Kappa statistic 0.7504
Mean absolute error 0.2116
Root mean squared error 0.2884
Relative absolute error 43.3072 %
Root relative squared error 58.3347 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.931 0.190 0.871 0.931 0.900 0.753 0.961 0.974 Yes
0.810 0.069 0.895 0.810 0.850 0.753 0.961 0.949 No
Weighted Avg. 0.880 0.139 0.881 0.880 0.879 0.753 0.961 0.963
=== Confusion Matrix ===
a b <-- classified as
27 2 | a = Yes
4 17 | b = No
Status
OK
Log
x 0

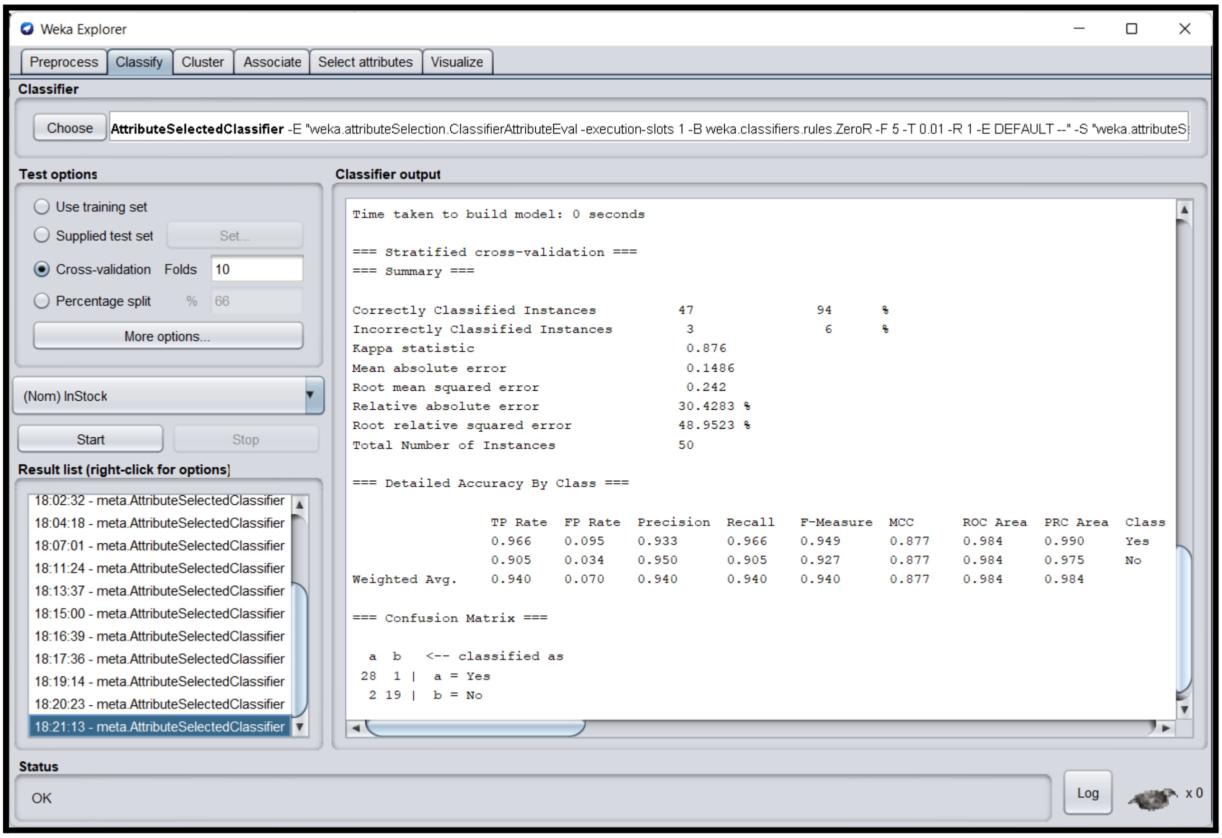

weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize
100
classifier
Choose
RandomForest -P 100 -I 100 -num-slots 1 -K
debug
False
doNotCheckCapabilities
False
evaluator
Choose
CfsSubsetEval -P 1 -E 1
numDecimalPlaces
2
search
Choose
BestFirst -D 1 -N 5
Open...
Save...
OK
Cancel


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
AttributeSelectedClassifier -E "weka.attributeSelection.CfsSubsetEval -P 1 -E 1" -S "weka.attributeSelection.BestFirst -D 1 -N 5" -W weka.classifiers.trees.RandomForest -- -P 100 -I
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) InStock
Start
Stop
Result list (right-click for options)
17:56:31 - meta.AttributeSelectedClassifier
17:57:57 - meta.AttributeSelectedClassifier
18:02:32 - meta.AttributeSelectedClassifier
18:04:18 - meta.AttributeSelectedClassifier
18:07:01 - meta.AttributeSelectedClassifier
18:11:24 - meta.AttributeSelectedClassifier
18:13:37 - meta.AttributeSelectedClassifier
18:15:00 - meta.AttributeSelectedClassifier
18:16:39 - meta.AttributeSelectedClassifier
18:17:36 - meta.AttributeSelectedClassifier
18:19:14 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0.01 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 47 94 %
Incorrectly Classified Instances 3 6 %
Kappa statistic 0.876
Mean absolute error 0.1188
Root mean squared error 0.2269
Relative absolute error 24.3173 %
Root relative squared error 45.8971 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.966 0.095 0.933 0.966 0.949 0.877 0.989 0.993 Yes
0.905 0.034 0.950 0.905 0.927 0.877 0.989 0.982 No
Weighted Avg. 0.940 0.070 0.940 0.940 0.940 0.877 0.989 0.989
=== Confusion Matrix ===
a b <-- classified as
28 1 | a = Yes
2 19 | b = No
Status
OK
Log
x 0


weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize 100
classifier Choose RandomForest -P 100 -I 100 -num-slots 1 -K
debug False
doNotCheckCapabilities False
evaluator Choose WrapperSubsetEval -B weka.classifiers.rules
numDecimalPlaces 2
search Choose BestFirst -D 1 -N 5
Open...
Save...
OK
Cancel


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose AttributeSelectedClassifier -E "weka.attributeSelection.WrapperSubsetEval -B weka.classifiers.rules.ZeroR -F 5 -T 0.01 -R 1 -E DEFAULT --" -S "weka.attributeSelection.BestFirst -D
Test options
Use training set
Supplied test set Set...
Cross-validation Folds 10
Percentage split % 66
More options...
(Nom) InStock
Start
Stop
Result list (right-click for options)
17:57:57 - meta.AttributeSelectedClassifier
18:02:32 - meta.AttributeSelectedClassifier
18:04:18 - meta.AttributeSelectedClassifier
18:07:01 - meta.AttributeSelectedClassifier
18:11:24 - meta.AttributeSelectedClassifier
18:13:37 - meta.AttributeSelectedClassifier
18:15:00 - meta.AttributeSelectedClassifier
18:16:39 - meta.AttributeSelectedClassifier
18:17:36 - meta.AttributeSelectedClassifier
18:19:14 - meta.AttributeSelectedClassifier
18:20:23 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0.11 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 29 58 %
Incorrectly Classified Instances 21 42 %
Kappa statistic 0
Mean absolute error 0.4895
Root mean squared error 0.4948
Relative absolute error 100.1986 %
Root relative squared error 100.0897 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
1.000 1.000 0.580 1.000 0.734 ? 0.463 0.547 Yes
0.000 0.000 ? 0.000 ? ? 0.463 0.403 No
Weighted Avg. 0.580 0.580 ? 0.580 ? ? 0.463 0.486
=== Confusion Matrix ===
a b <-- classified as
29 0 | a = Yes
21 0 | b = No
Status
OK
Log
x 0


weka.gui.GenericObjectEditor
weka.classifiers.meta.AttributeSelectedClassifier
About
Dimensionality of training and test data is reduced by attribute selection before being passed on to a classifier.
More
Capabilities
batchSize 100
classifier Choose RandomForest -P 100 -I 100 -num-slots 1 -K
debug False
doNotCheckCapabilities False
evaluator Choose ClassifierAttributeEval -execution-slots 1 -B
numDecimalPlaces 2
search Choose Ranker -T -1.7976931348623157E308 -N -1
Open...
Save...
OK
Cancel


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose AttributeSelectedClassifier -E "weka.attributeSelection.ClassifierAttributeEval -execution-slots 1 -B weka.classifiers.rules.ZeroR -F 5 -T 0.01 -R 1 -E DEFAULT --" -S "weka.attributeS
Test options
Use training set
Supplied test set
Set...
Cross-validation Folds 10
Percentage split % 66
More options...
(Nom) InStock
Start
Stop
Result list (right-click for options)
18:02:32 - meta.AttributeSelectedClassifier
18:04:18 - meta.AttributeSelectedClassifier
18:07:01 - meta.AttributeSelectedClassifier
18:11:24 - meta.AttributeSelectedClassifier
18:13:37 - meta.AttributeSelectedClassifier
18:15:00 - meta.AttributeSelectedClassifier
18:16:39 - meta.AttributeSelectedClassifier
18:17:36 - meta.AttributeSelectedClassifier
18:19:14 - meta.AttributeSelectedClassifier
18:20:23 - meta.AttributeSelectedClassifier
18:21:13 - meta.AttributeSelectedClassifier
Classifier output
Time taken to build model: 0 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 47 94 %
Incorrectly Classified Instances 3 6 %
Kappa statistic 0.876
Mean absolute error 0.1486
Root mean squared error 0.242
Relative absolute error 30.4283 %
Root relative squared error 48.9523 %
Total Number of Instances 50
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.966 0.095 0.933 0.966 0.949 0.877 0.984 0.990 Yes
0.905 0.034 0.950 0.905 0.927 0.877 0.984 0.975 No
Weighted Avg. 0.940 0.070 0.940 0.940 0.940 0.877 0.984 0.984
=== Confusion Matrix ===
a b <-- classified as
28 1 | a = Yes
2 19 | b = No
Status
OK
Log
x 0

B.3 Observations and learning: (Performance Evaluation)

(Students are expected to comment on the output obtained with clear observations getting from Performance Evaluation after analyzing the data and learning for each task assigned)

CLASSIFIER	EVALUATOR	SEARCH METHOD	ACCURACY
J48	CFS Subset	Best First	88%
J48	Wrapper Subset	Best First	58%
J48	Wrapper Subset	Greedy Stepwise	58%
J48	Classifier Attribute	Ranker	88%
Naive Bayes	CFS Subset	Best First	94%
Naive Bayes	Wrapper Subset	Greedy Stepwise	58%
Naive Bayes	Classifier Attribute	Ranker	94%
ZeroR	CFS Subset	Best First	58%
ZeroR	Wrapper Subset	Best First	58%
ZeroR	Classifier Attribute	Ranker	58%
KStar	CFS Subset	Best First	92%
KStar	Wrapper Subset	Greedy Stepwise	58%
KStar	Classifier Attribute	Ranker	88%
Random Forest	CFS Subset	Best First	94%
Random Forest	Wrapper Subset	Best First	58%
Random Forest	Classifier Attribute	Ranker	94%

We can deduce the following from the given comparison:

- ➔ The least accurate classifier is the ZeroR, which is useful for defining a baseline performance as a standard for other classification algorithms.
- ➔ Greedy Stepwise Search has a lower accuracy than Best First and Ranker Search Algorithms.
- ➔ The Classifier Attribute and the CFS Subset Evaluator decide the most accuracy, while the Wrapper Subset Evaluator is accountable for the least accuracy.

B.4 Conclusion:

(Students must write the conclusion as per the attainment of individual outcome listed above and learning/observation noted in section B.3)

As a result, we were able to successfully incorporate a variety of problem-solving techniques while also optimising the accuracy for the provided dataset.

B.5 Question of Curiosity

(To be answered by student based on the practical performed and learning/observations)

Q1) What are the different methods for Relevant Attribute Selection?

Ans:

In the data mining process, Relevant Attribute Selection is a strategy for data minimization. Data reduction decreases the size of data so that it can be used more efficiently for analysis.

Attribute Subset Selection Methods

1. Stepwise Forward Selection
2. Stepwise Backward Elimination
3. Combination of Forward Selection and Backward Elimination
4. Decision Tree Induction

Q2) Explain Performance Evaluation Parameters for Classification Problem.

Ans:

1. Confusion Matrix

   The counts of test records successfully and erroneously predicted by the model are used to evaluate the performance of a classification model. The confusion matrix offers a more detailed view of not just a predictive model's performance, but also which classes are being forecasted correctly and erroneously, as well as the kind of errors that are being produced.

Accuracy:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

Recall:

$$Recall = \frac{TP}{TP + FN}$$

Precision:

$$Precision = \frac{TP}{TP + FP}$$

F_1 score:

$$F_1 = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}}$$

2. Combining Precision and Recall — F1 Score

   We aim to maximise either recall or accuracy at the expense of the other measure in the three examples above. For example, we would like to reduce FN to improve recall in the case of a good or bad loan categorization. However, in instances when we wish to discover the best balance of accuracy and recall, we may use the F1 score to combine the two measures.

$$F_1 = 2 * \frac{precision * recall}{precision + recall}$$

3. Decision Threshold & Receiver Operating Characteristic (ROC) curve

   The ROC plot is a popular method of displaying the performance of a classification model. It describes the trade-off between the true positive rate (tpr) and the false positive rate (fpr) for different probability thresholds in a prediction model.

$$true\ positive\ rate = \frac{true\ positives}{true\ positives + false\ negatives} \qquad false\ positive\ rate = \frac{false\ positives}{false\ positives + true\ negatives}$$