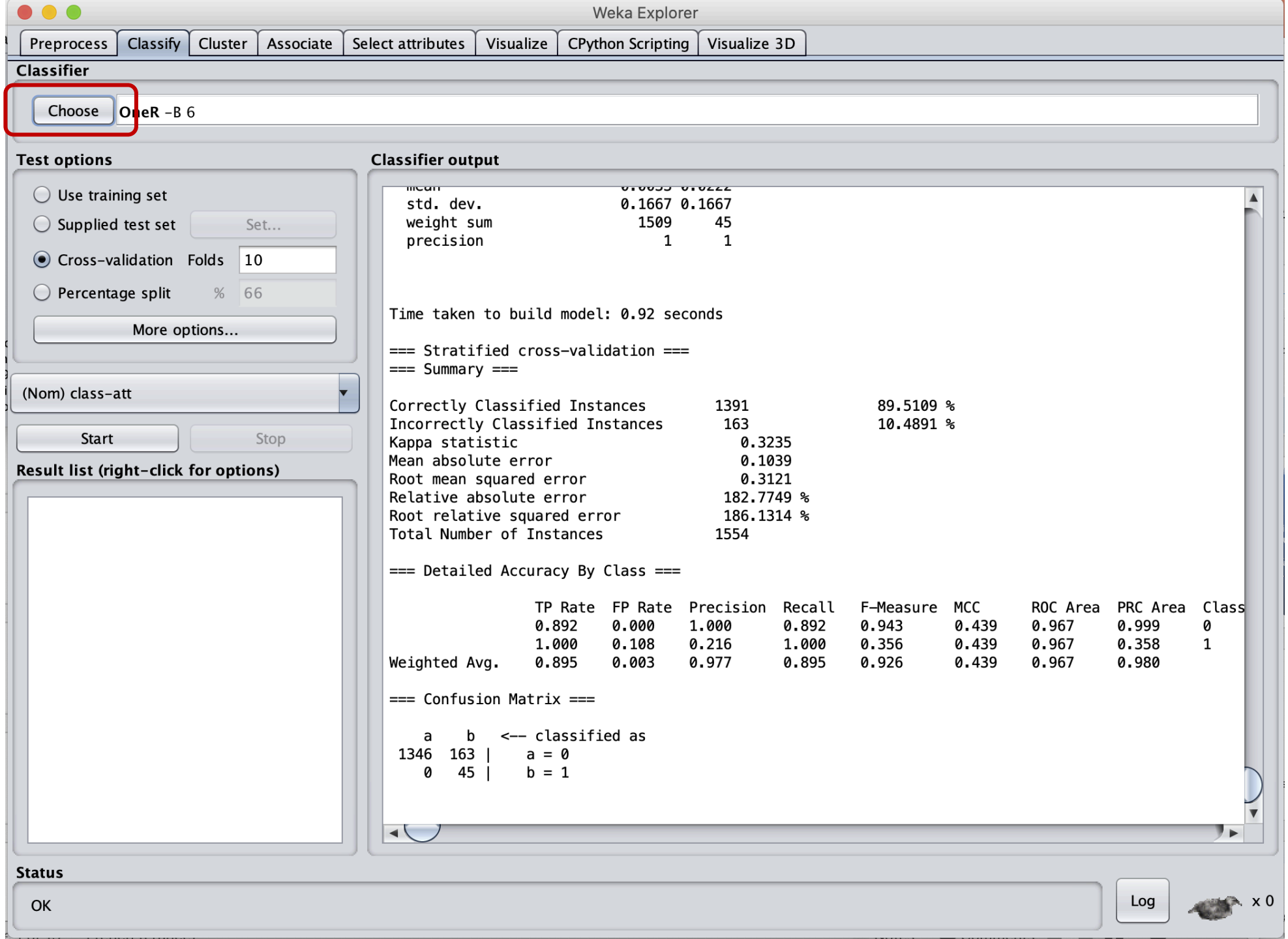
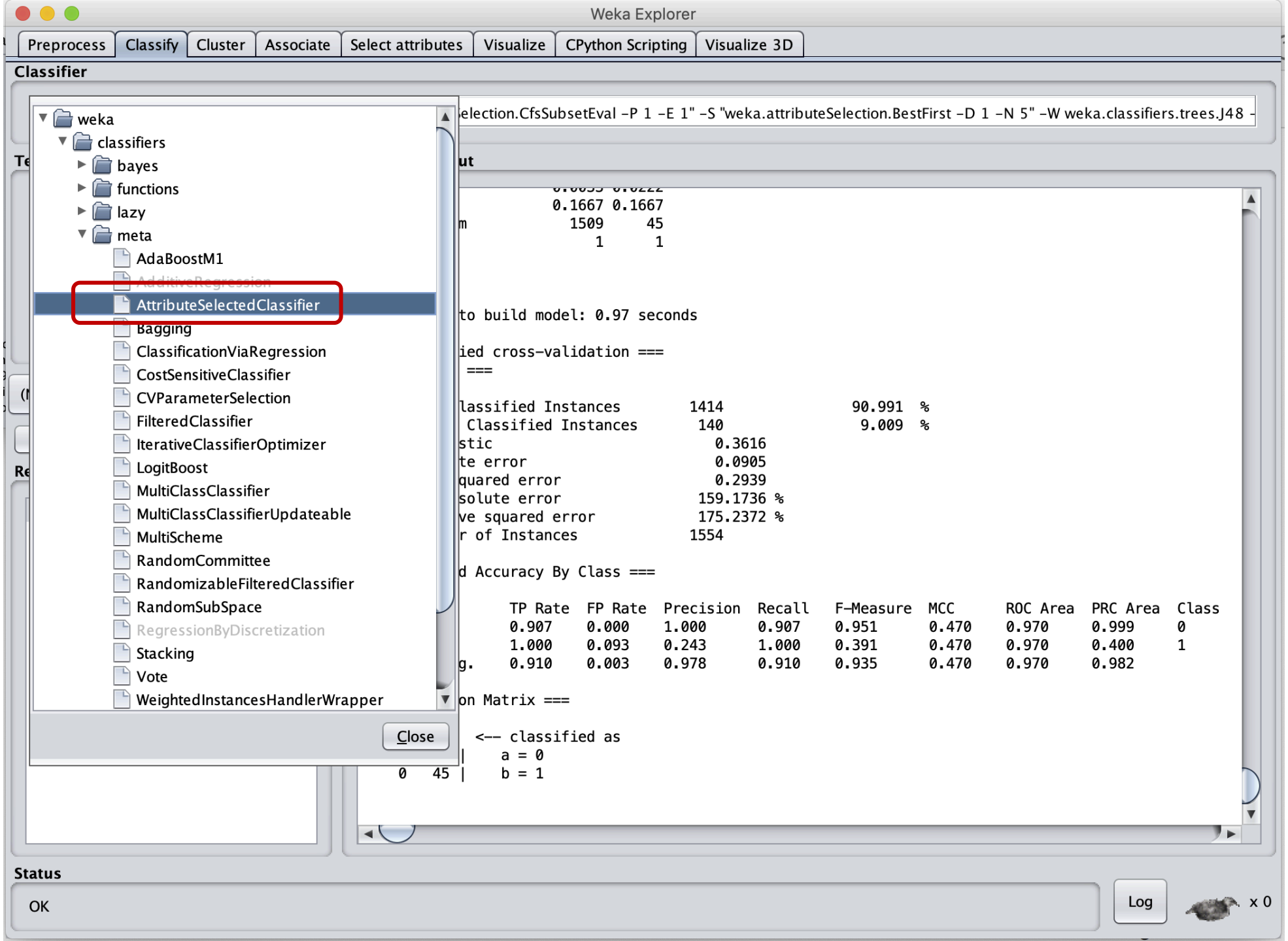
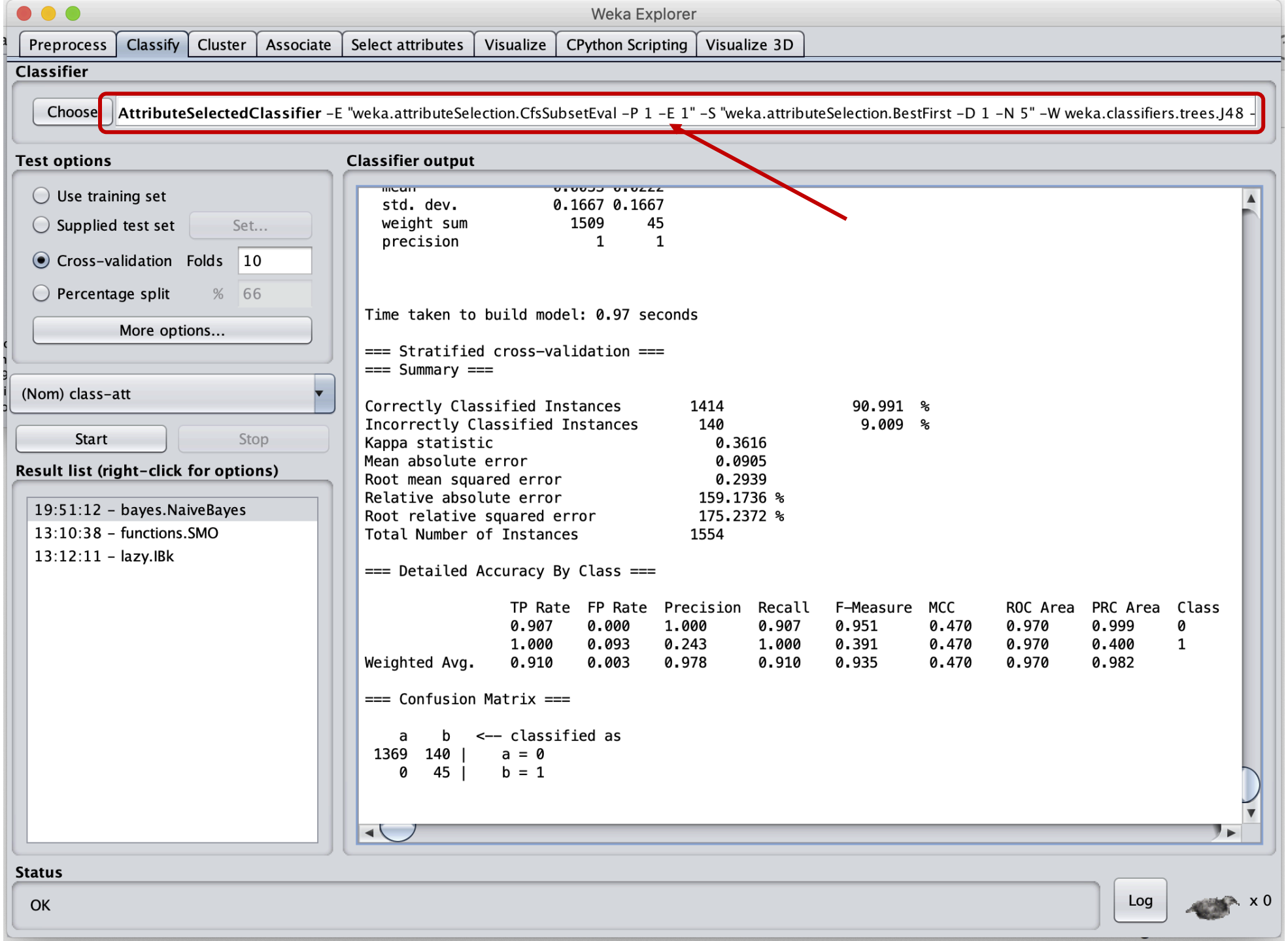
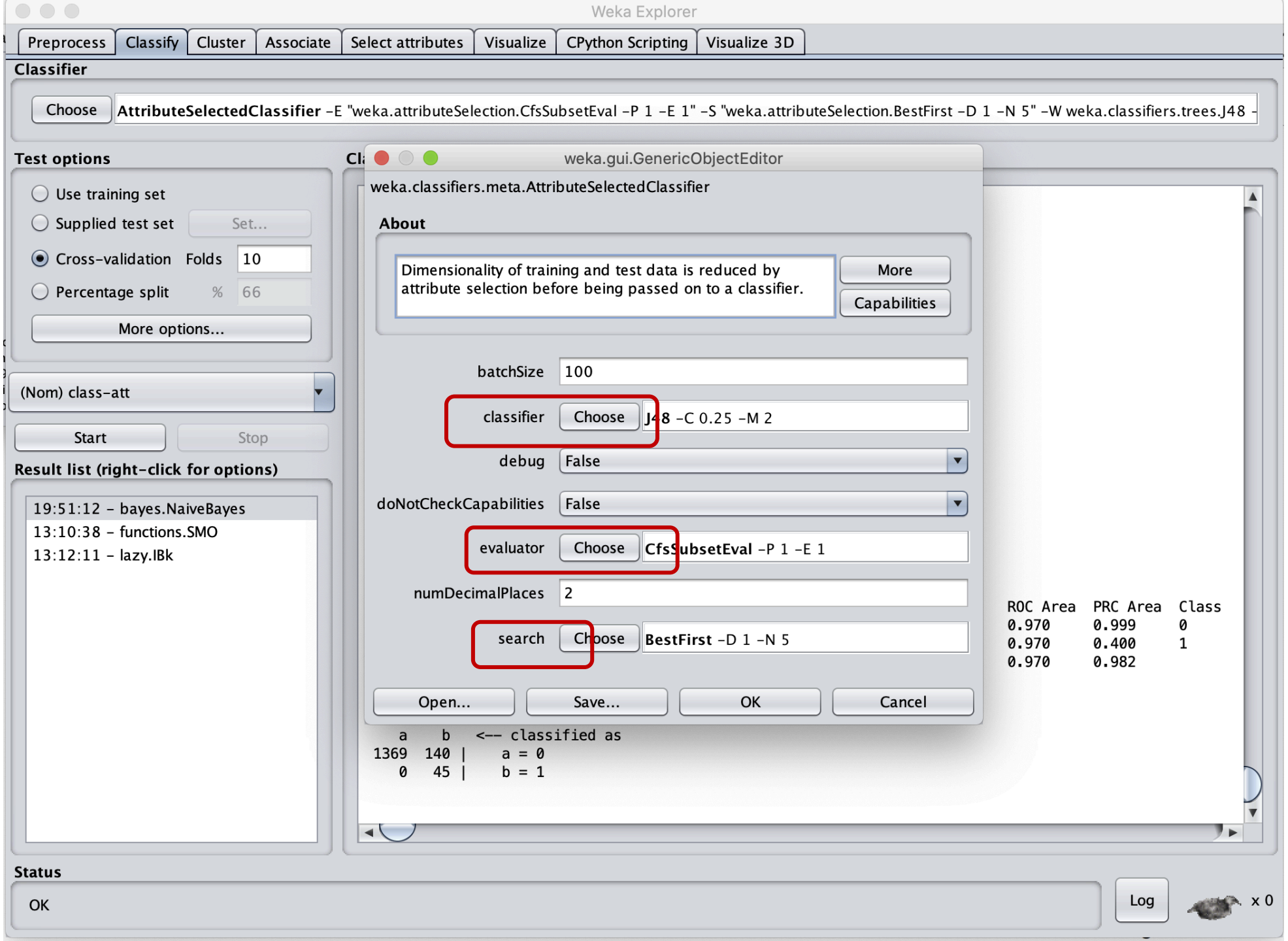
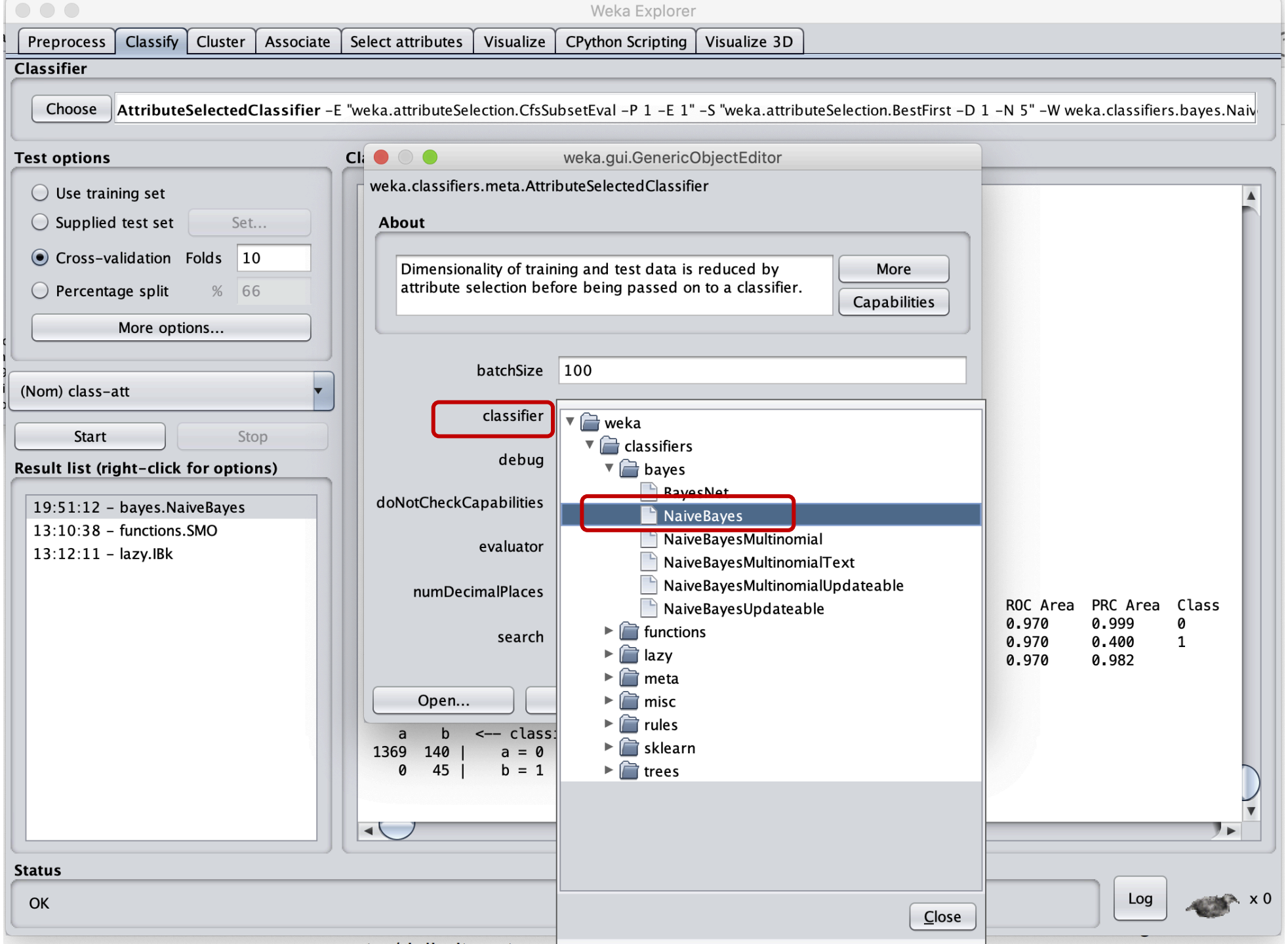
Task 05: Creating the classifiers WITH feature selection

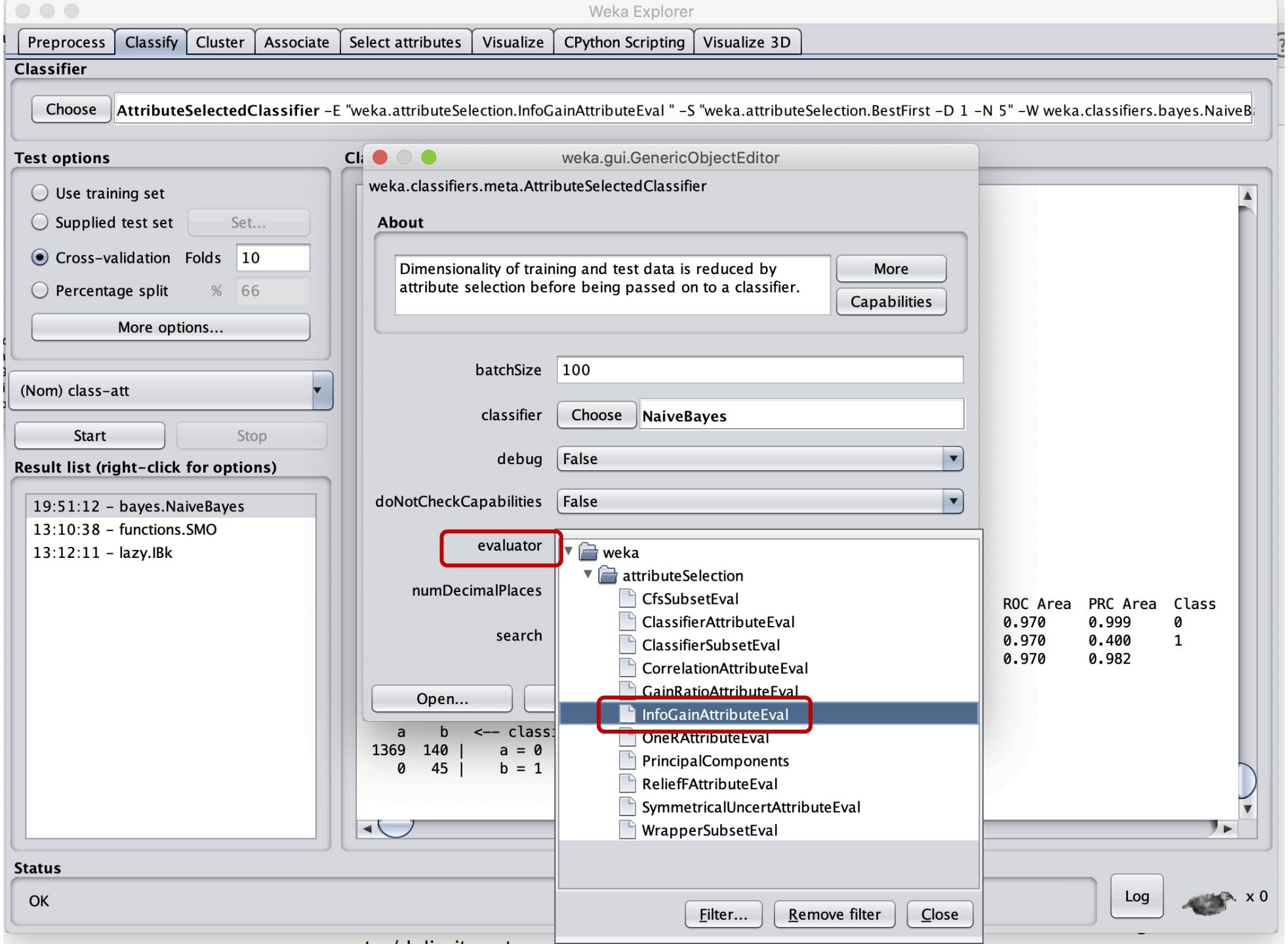

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
CPython Scripting
Visualize 3D
Classifier
Choose
OneR -B 6
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) class-att
Start
Stop
Result list (right-click for options)
Classifier output
std. dev. 0.1667 0.1667
weight sum 1509 45
precision 1 1
Time taken to build model: 0.92 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 1391 89.5109 %
Incorrectly Classified Instances 163 10.4891 %
Kappa statistic 0.3235
Mean absolute error 0.1039
Root mean squared error 0.3121
Relative absolute error 182.7749 %
Root relative squared error 186.1314 %
Total Number of Instances 1554
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.892 0.000 1.000 0.892 0.943 0.439 0.967 0.999 0
1.000 0.108 0.216 1.000 0.356 0.439 0.967 0.358 1
Weighted Avg. 0.895 0.003 0.977 0.895 0.926 0.439 0.967 0.980
=== Confusion Matrix ===
a b <-- classified as
1346 163 | a = 0
0 45 | b = 1
Status
OK
Log
x 0

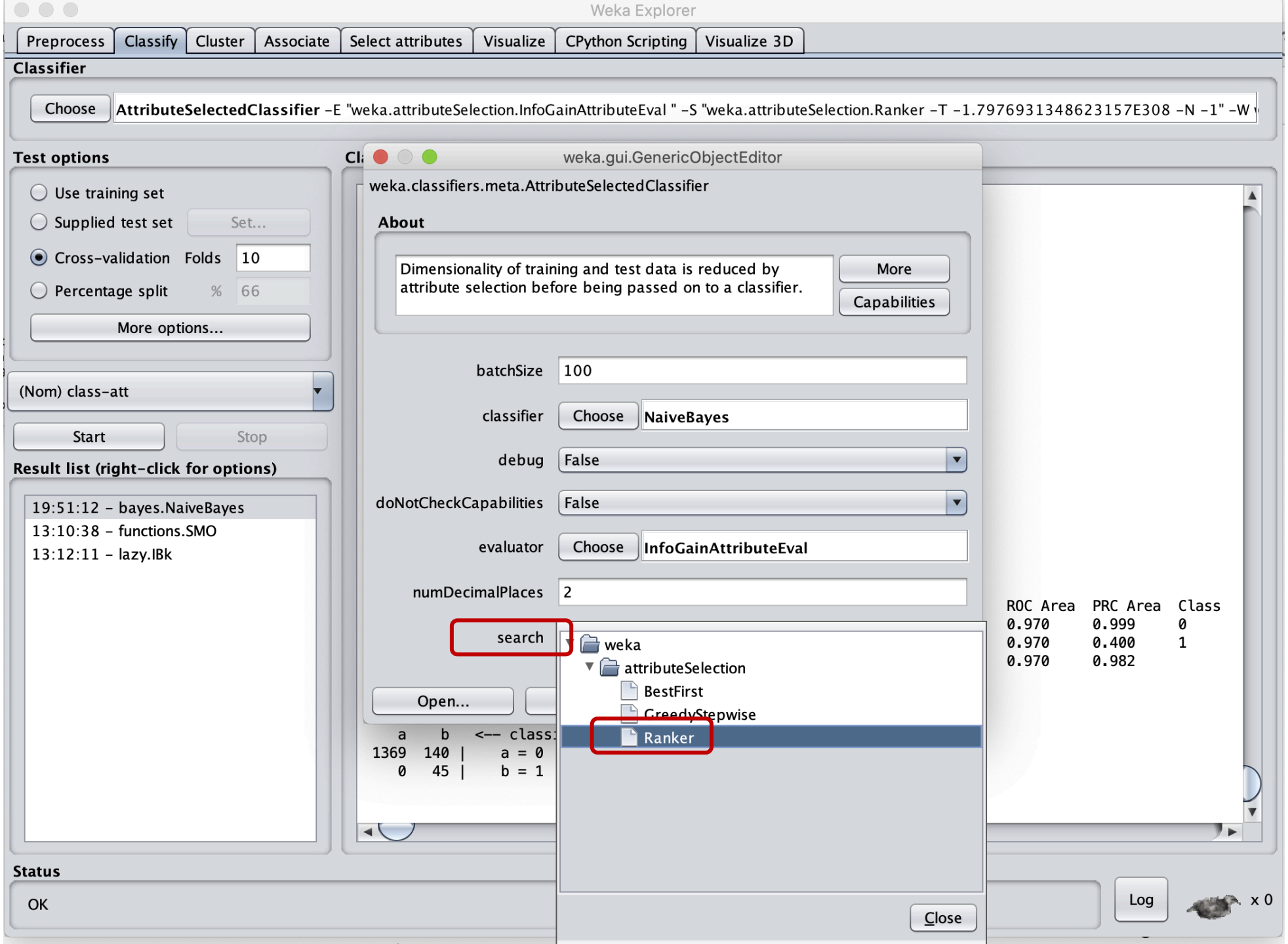

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
CPython Scripting
Visualize 3D
Classifier
election.CfsSubsetEval -P 1 -E 1" -S "weka.attributeSelection.BestFirst -D 1 -N 5" -W weka.classifiers.trees.J48 -
weka
classifiers
bayes
functions
lazy
meta
AdaBoostM1
AdditiveRegression
AttributeSelectedClassifier
Bagging
ClassificationViaRegression
CostSensitiveClassifier
CVParameterSelection
FilteredClassifier
IterativeClassifierOptimizer
LogitBoost
MultiClassClassifier
MultiClassClassifierUpdateable
MultiScheme
RandomCommittee
RandomizableFilteredClassifier
RandomSubSpace
RegressionByDiscretization
Stacking
Vote
WeightedInstancesHandlerWrapper
Close
0.1667 0.1667
1509 45
1 1
to build model: 0.97 seconds
ied cross-validation ===
===
lassified Instances 1414 90.991 %
Classified Instances 140 9.009 %
stic 0.3616
te error 0.0905
quared error 0.2939
solute error 159.1736 %
ve squared error 175.2372 %
r of Instances 1554
d Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.907 0.000 1.000 0.907 0.951 0.470 0.970 0.999 0
1.000 0.093 0.243 1.000 0.391 0.470 0.970 0.400 1
g. 0.910 0.003 0.978 0.910 0.935 0.470 0.970 0.982
on Matrix ===
<-- classified as
| a = 0
0 45 | b = 1
Status
OK
Log
x 0

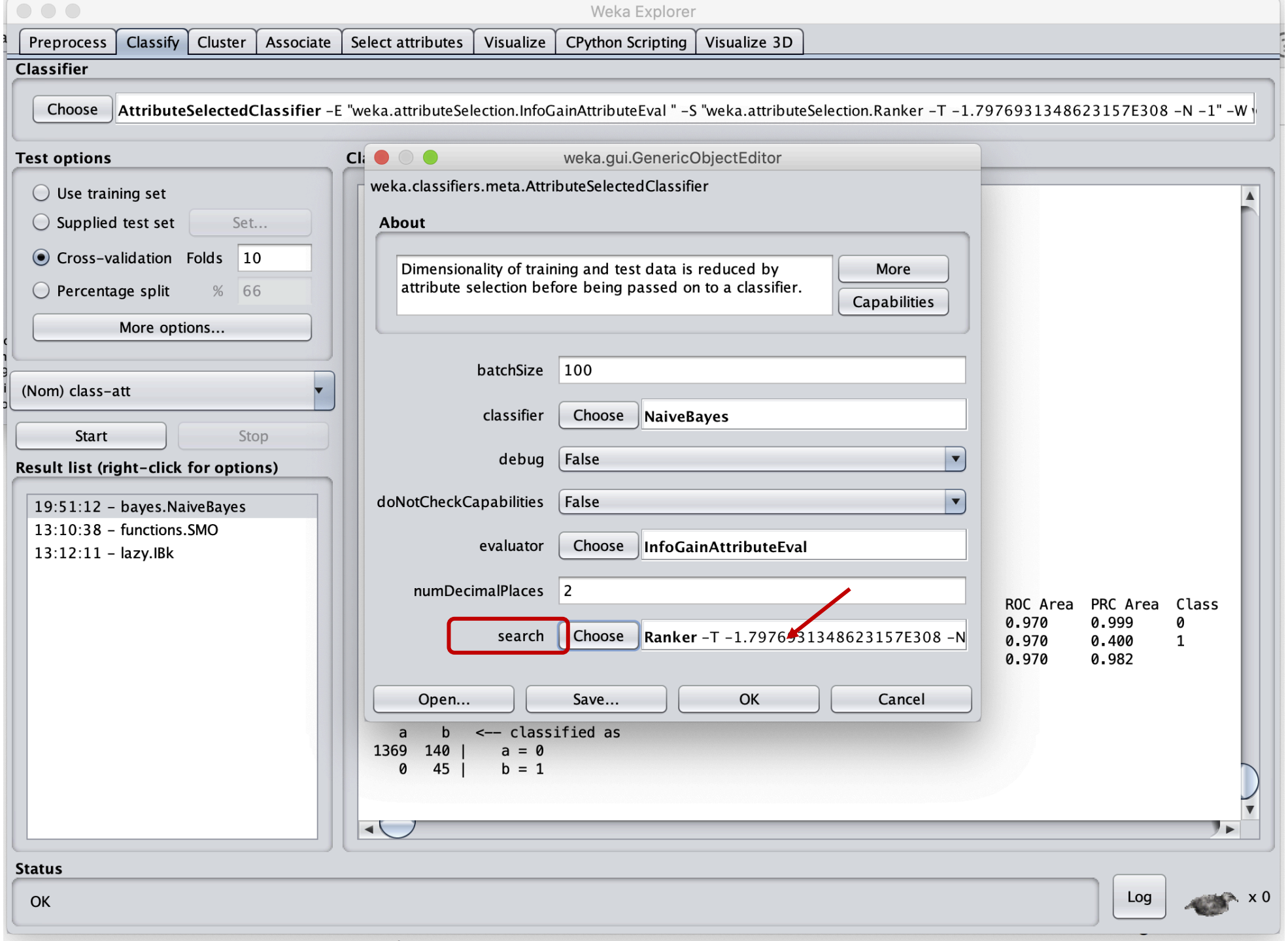

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
CPython Scripting
Visualize 3D
Classifier
Choose
AttributeSelectedClassifier -E "weka.attributeSelection.CfsSubsetEval -P 1 -E 1" -S "weka.attributeSelection.BestFirst -D 1 -N 5" -W weka.classifiers.trees.J48 -
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) class-att
Start
Stop
Result list (right-click for options)
19:51:12 - bayes.NaiveBayes
13:10:38 - functions.SMO
13:12:11 - lazy.IBk
Classifier output
std. dev. 0.1667 0.1667
weight sum 1509 45
precision 1 1
Time taken to build model: 0.97 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 1414 90.991 %
Incorrectly Classified Instances 140 9.009 %
Kappa statistic 0.3616
Mean absolute error 0.0905
Root mean squared error 0.2939
Relative absolute error 159.1736 %
Root relative squared error 175.2372 %
Total Number of Instances 1554
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.907 0.000 1.000 0.907 0.951 0.470 0.970 0.999 0
1.000 0.093 0.243 1.000 0.391 0.470 0.970 0.400 1
Weighted Avg. 0.910 0.003 0.978 0.910 0.935 0.470 0.970 0.982
=== Confusion Matrix ===
a b <-- classified as
1369 140 | a = 0
0 45 | b = 1
Status
OK
Log
x 0

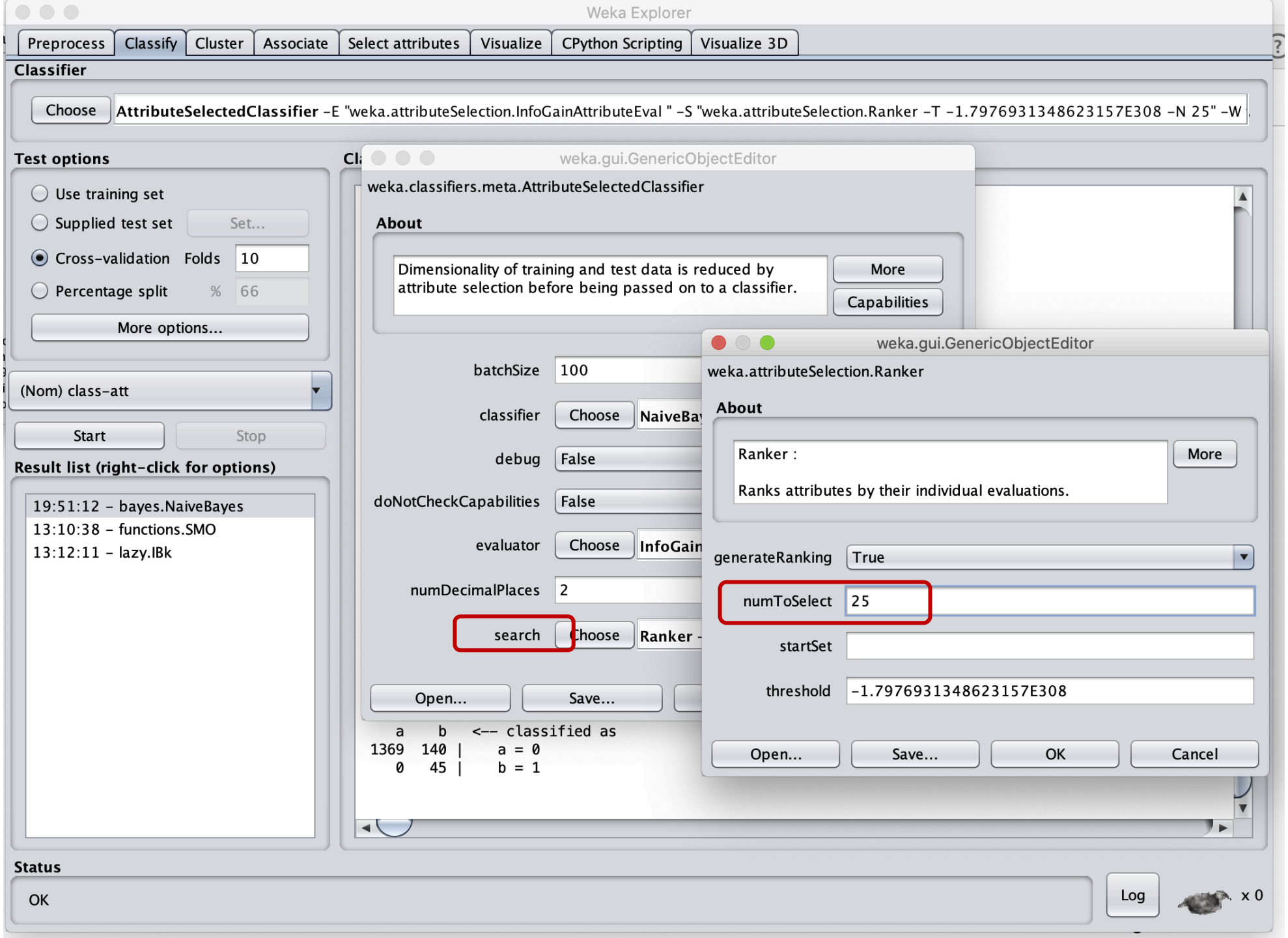

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
CPython Scripting
Visualize 3D
Classifier
Choose
AttributeSelectedClassifier -E "weka.attributeSelection.CfsSubsetEval -P 1 -E 1" -S "weka.attributeSelection.BestFirst -D 1 -N 5" -W weka.classifiers.trees.J48 -
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) class-att
Start
Stop
Result list (right-click for options)
19:51:12 - bayes.NaiveBayes
13:10:38 - functions.SMO
13:12:11 - lazy.IBk
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
ROC Area PRC Area Class
0.970 0.999 0
0.970 0.400 1
0.970 0.982
a b <-- classified as
1369 140 | a = 0
0 45 | b = 1
Status
OK
Log
x 0


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
CPython Scripting
Visualize 3D
Classifier
Choose
AttributeSelectedClassifier -E "weka.attributeSelection.CfsSubsetEval -P 1 -E 1" -S "weka.attributeSelection.BestFirst -D 1 -N 5" -W weka.classifiers.bayes.Naiv
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) class-att
Start
Stop
Result list (right-click for options)
19:51:12 - bayes.NaiveBayes
13:10:38 - functions.SMO
13:12:11 - lazy.IBk
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
debug
doNotCheckCapabilities
evaluator
numDecimalPlaces
search
Open...
weka
classifiers
bayes
BayesNet
NaiveBayes
NaiveBayesMultinomial
NaiveBayesMultinomialText
NaiveBayesMultinomialUpdateable
NaiveBayesUpdateable
functions
lazy
meta
misc
rules
sklearn
trees
Close
a b <-- classi
1369 140 | a = 0
0 45 | b = 1
ROC Area PRC Area Class
0.970 0.999 0
0.970 0.400 1
0.970 0.982
Status
OK
Log
x 0


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
CPython Scripting
Visualize 3D
Classifier
Choose
AttributeSelectedClassifier -E "weka.attributeSelection.InfoGainAttributeEval " -S "weka.attributeSelection.BestFirst -D 1 -N 5" -W weka.classifiers.bayes.NaiveB
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) class-att
Start
Stop
Result list (right-click for options)
19:51:12 - bayes.NaiveBayes
13:10:38 - functions.SMO
13:12:11 - lazy.IBk
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
numDecimalPlaces
search
Open...
a b <-- class
1369 140 | a = 0
0 45 | b = 1
weka
attributeSelection
CfsSubsetEval
ClassifierAttributeEval
ClassifierSubsetEval
CorrelationAttributeEval
GainRatioAttributeEval
InfoGainAttributeEval
OneRAttributeEval
PrincipalComponents
ReliefFAttributeEval
SymmetricalUncertAttributeEval
WrapperSubsetEval
Filter...
Remove filter
Close
ROC Area PRC Area Class
0.970 0.999 0
0.970 0.400 1
0.970 0.982
Status
OK
Log
x 0

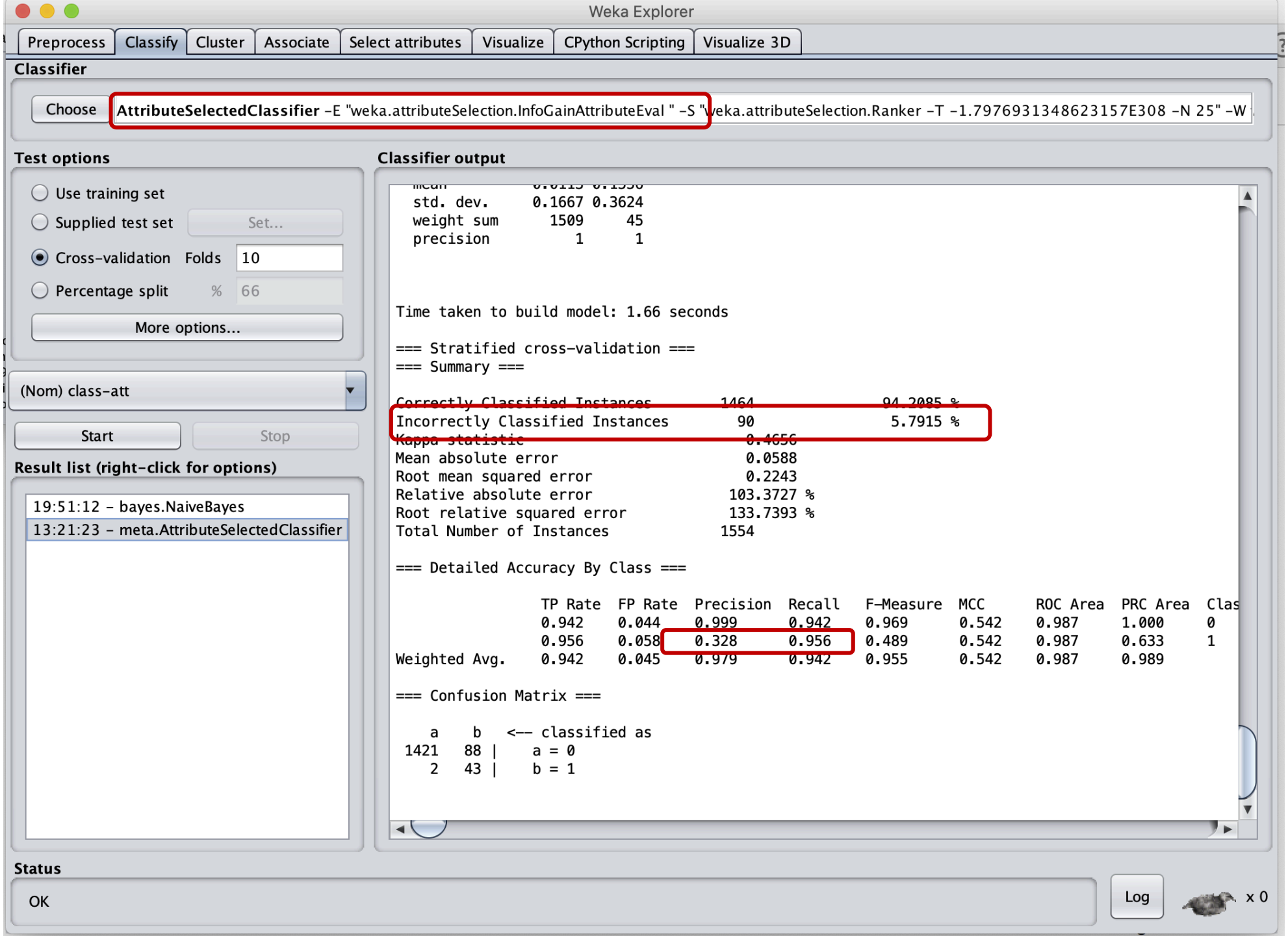

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
CPython Scripting
Visualize 3D
Classifier
Choose
AttributeSelectedClassifier -E "weka.attributeSelection.InfoGainAttributeEval " -S "weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N -1" -W
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) class-att
Start
Stop
Result list (right-click for options)
19:51:12 - bayes.NaiveBayes
13:10:38 - functions.SMO
13:12:11 - lazy.IBk
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
InfoGainAttributeEval
numDecimalPlaces
2
search
Open...
a b <-- classi
1369 140 | a = 0
0 45 | b = 1
ROC Area PRC Area Class
0.970 0.999 0
0.970 0.400 1
0.970 0.982
weka
attributeSelection
BestFirst
GreedyStepwise
Ranker
Close
Status
OK
Log
x 0


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
CPython Scripting
Visualize 3D
Classifier
Choose
AttributeSelectedClassifier -E "weka.attributeSelection.InfoGainAttributeEval " -S "weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N -1" -W
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) class-att
Start
Stop
Result list (right-click for options)
19:51:12 - bayes.NaiveBayes
13:10:38 - functions.SMO
13:12:11 - lazy.IBk
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
InfoGainAttributeEval
numDecimalPlaces
2
search
Choose
Ranker -T -1.7976931348623157E308 -N
Open...
Save...
OK
Cancel
ROC Area PRC Area Class
0.970 0.999 0
0.970 0.400 1
0.970 0.982
a b <-- classified as
1369 140 | a = 0
0 45 | b = 1
Status
OK
Log
x 0


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
CPython Scripting
Visualize 3D
Classifier
Choose
AttributeSelectedClassifier -E "weka.attributeSelection.InfoGainAttributeEval " -S "weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N 25" -W
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) class-att
Start
Stop
Result list (right-click for options)
19:51:12 - bayes.NaiveBayes
13:10:38 - functions.SMO
13:12:11 - lazy.IBk
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
NaiveBay
debug
False
doNotCheckCapabilities
False
evaluator
Choose
InfoGain
numDecimalPlaces
2
search
Choose
Ranker -
Open...
Save...
a b <-- classified as
1369 140 | a = 0
0 45 | b = 1
weka.gui.GenericObjectEditor
weka.attributeSelection.Ranker
About
Ranker :
Ranks attributes by their individual evaluations.
More
generateRanking
True
numToSelect
25
startSet
threshold
-1.7976931348623157E308
Open...
Save...
OK
Cancel
Status
OK
Log
x 0

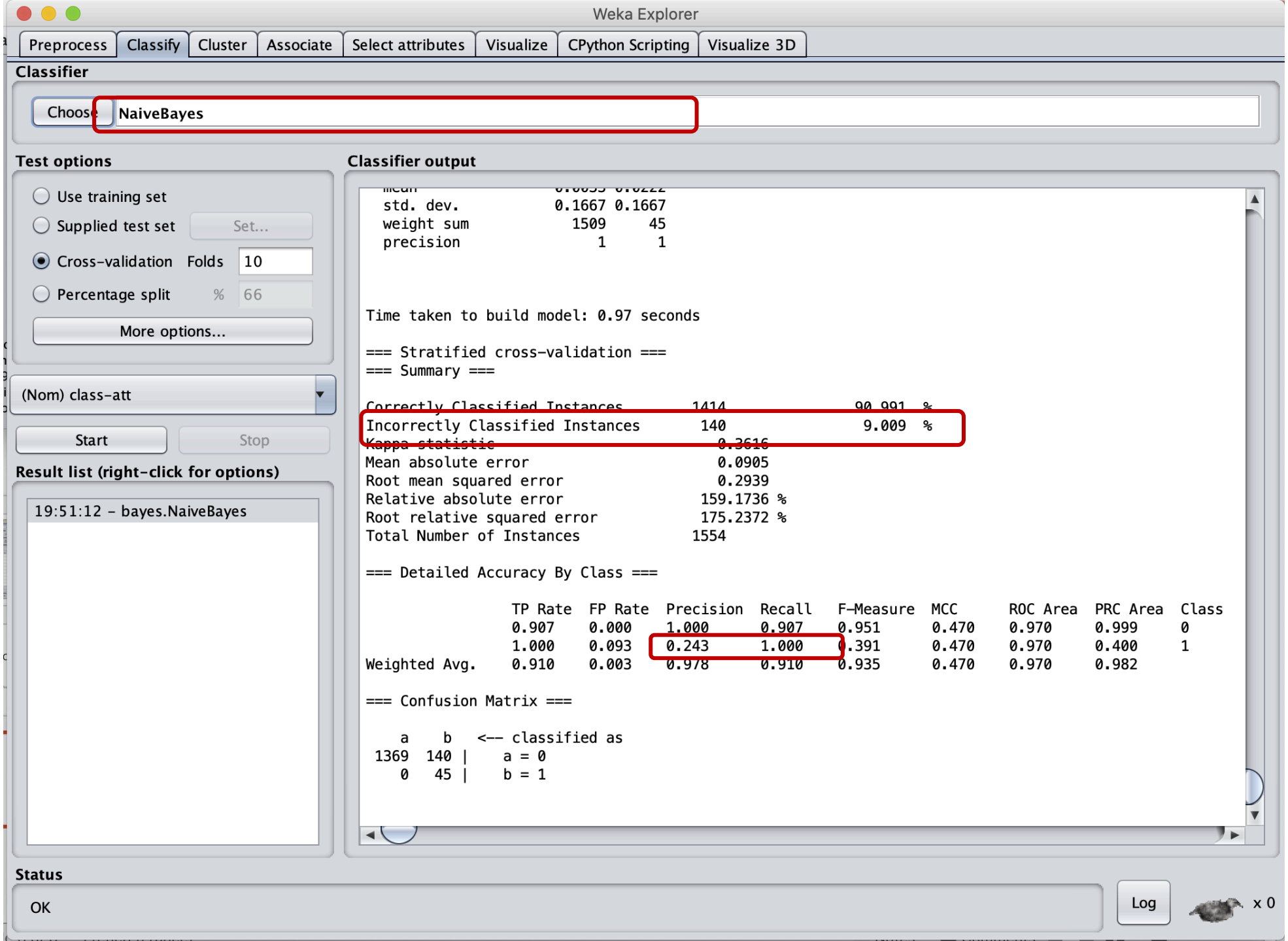
Task 05: Creating the classifiers WITH feature selection

Naive Bayes classifier: results


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
CPython Scripting
Visualize 3D
Classifier
Choose
AttributeSelectedClassifier -E "weka.attributeSelection.InfoGainAttributeEval " -S "weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N 25" -W
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) class-att
Start
Stop
Result list (right-click for options)
19:51:12 - bayes.NaiveBayes
13:21:23 - meta.AttributeSelectedClassifier
Classifier output
std. dev. 0.1667 0.3624
weight sum 1509 45
precision 1 1
Time taken to build model: 1.66 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 1464 94.2085 %
Incorrectly Classified Instances 90 5.7915 %
Kappa statistic 0.4656
Mean absolute error 0.0588
Root mean squared error 0.2243
Relative absolute error 103.3727 %
Root relative squared error 133.7393 %
Total Number of Instances 1554
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Clas
0.942 0.044 0.999 0.942 0.969 0.542 0.987 1.000 0
0.956 0.058 0.328 0.956 0.489 0.542 0.987 0.633 1
Weighted Avg. 0.942 0.045 0.979 0.942 0.955 0.542 0.987 0.989
=== Confusion Matrix ===
a b <-- classified as
1421 88 | a = 0
2 43 | b = 1
Status
OK
Log
x 0

Task 04: Creating the classifiers WITHOUT feature selection

Naive Bayes classifier: results


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
CPython Scripting
Visualize 3D
Classifier
Choose
NaiveBayes
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) class-att
Start
Stop
Result list (right-click for options)
19:51:12 - bayes.NaiveBayes
Classifier output
std. dev. 0.1667 0.1667
weight sum 1509 45
precision 1 1
Time taken to build model: 0.97 seconds
=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances 1414 90.991 %
Incorrectly Classified Instances 140 9.009 %
Kappa statistic 0.3616
Mean absolute error 0.0905
Root mean squared error 0.2939
Relative absolute error 159.1736 %
Root relative squared error 175.2372 %
Total Number of Instances 1554
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.907 0.000 1.000 0.907 0.951 0.470 0.970 0.999 0
1.000 0.093 0.243 1.000 0.391 0.470 0.970 0.400 1
Weighted Avg. 0.910 0.003 0.978 0.910 0.935 0.470 0.970 0.982
=== Confusion Matrix ===
a b <-- classified as
1369 140 | a = 0
0 45 | b = 1
Status
OK
Log
x 0