FEBRUARY 1, 2024

MACHINE LEARNING FOUNDATIONS

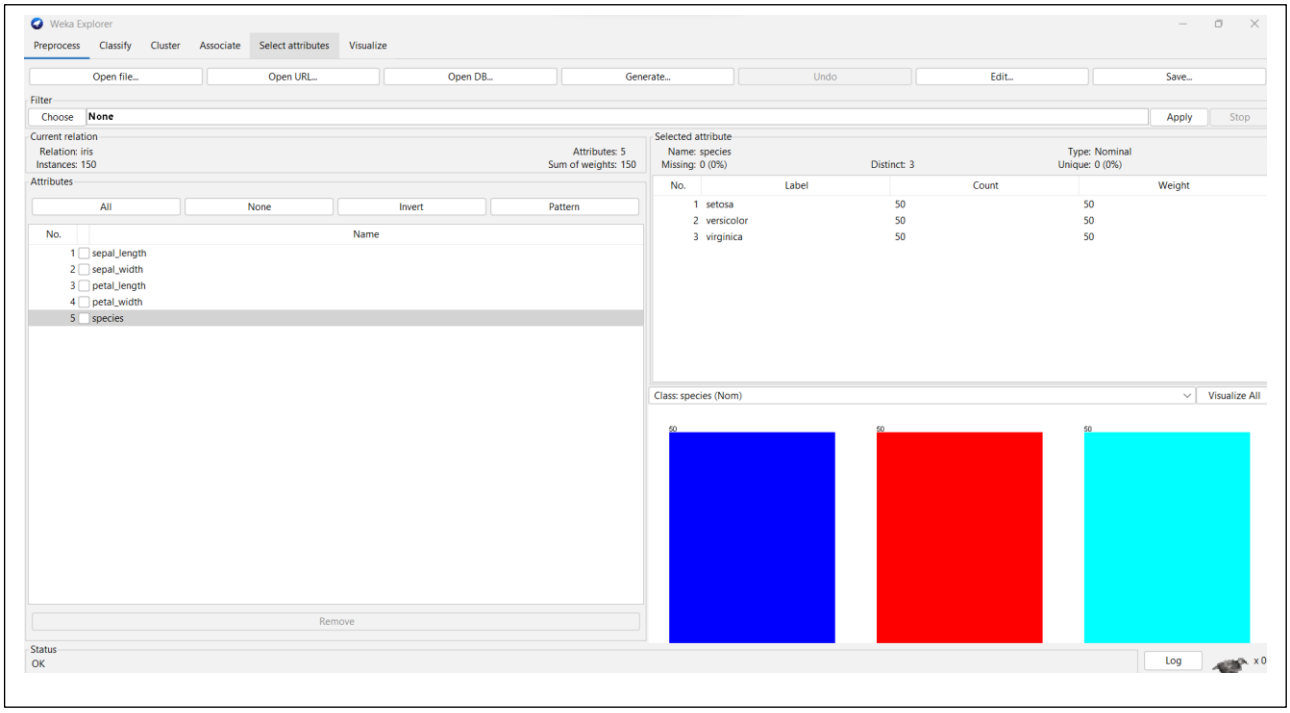
ASHFIRUNA ZAMAN

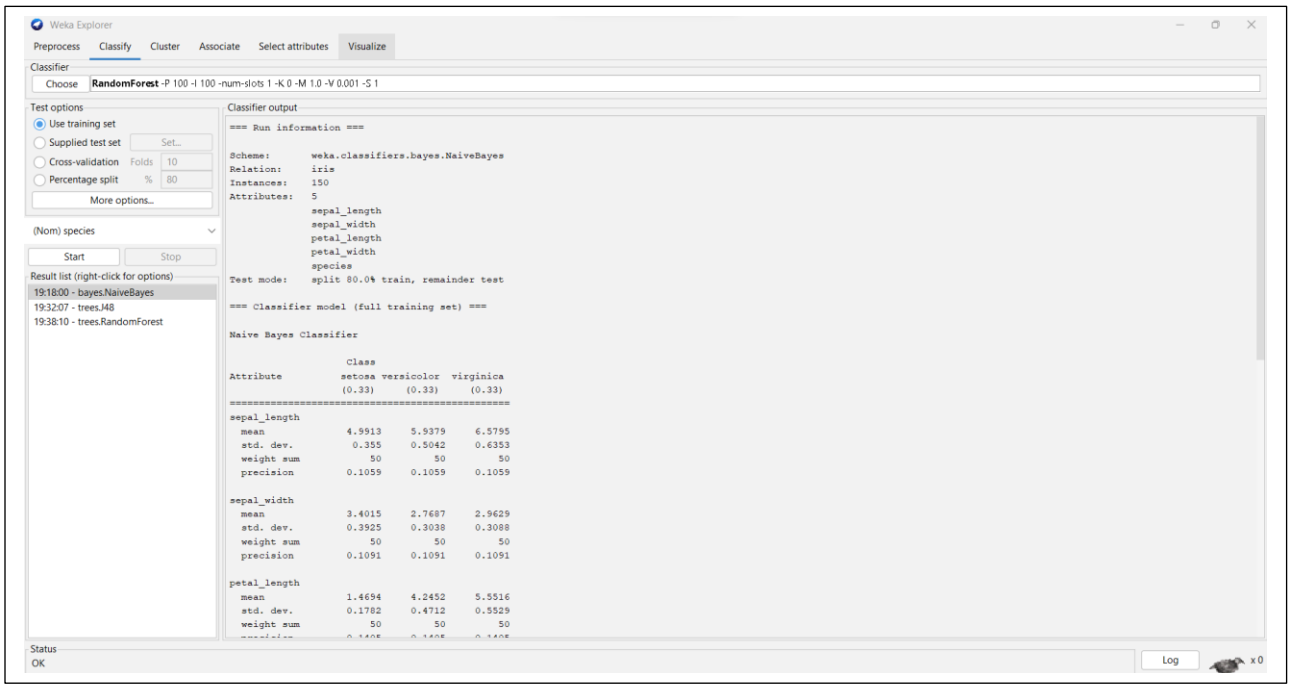
Classification

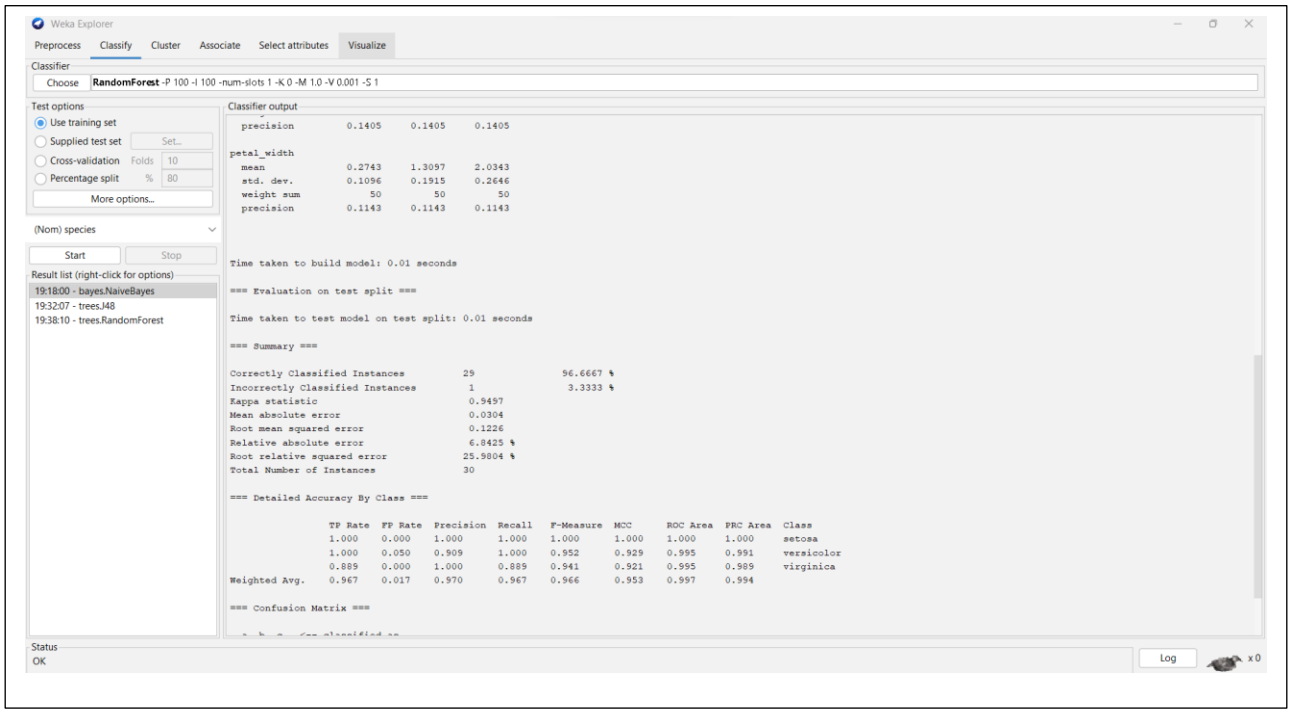
IRIS Dataset

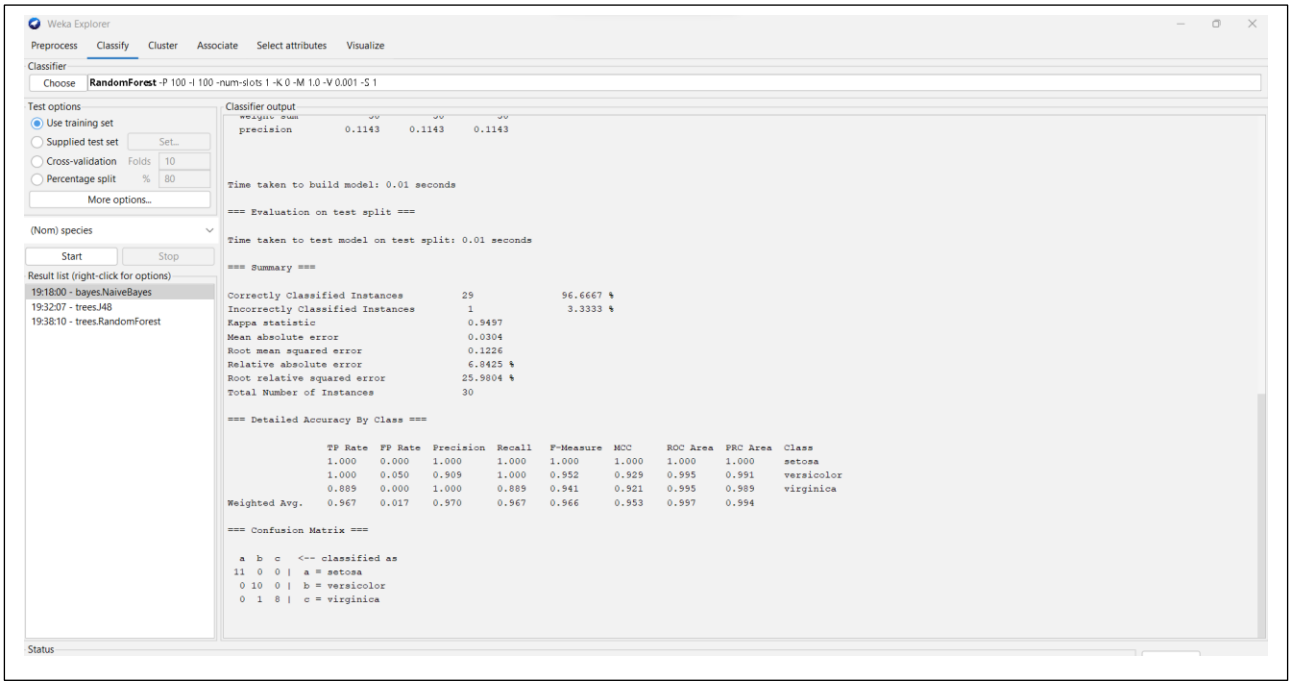
In comparing the performance of the J48 decision tree, Naive Bayes, and RandomForest classifiers on the Iris dataset using WEKA, several key insights emerge:

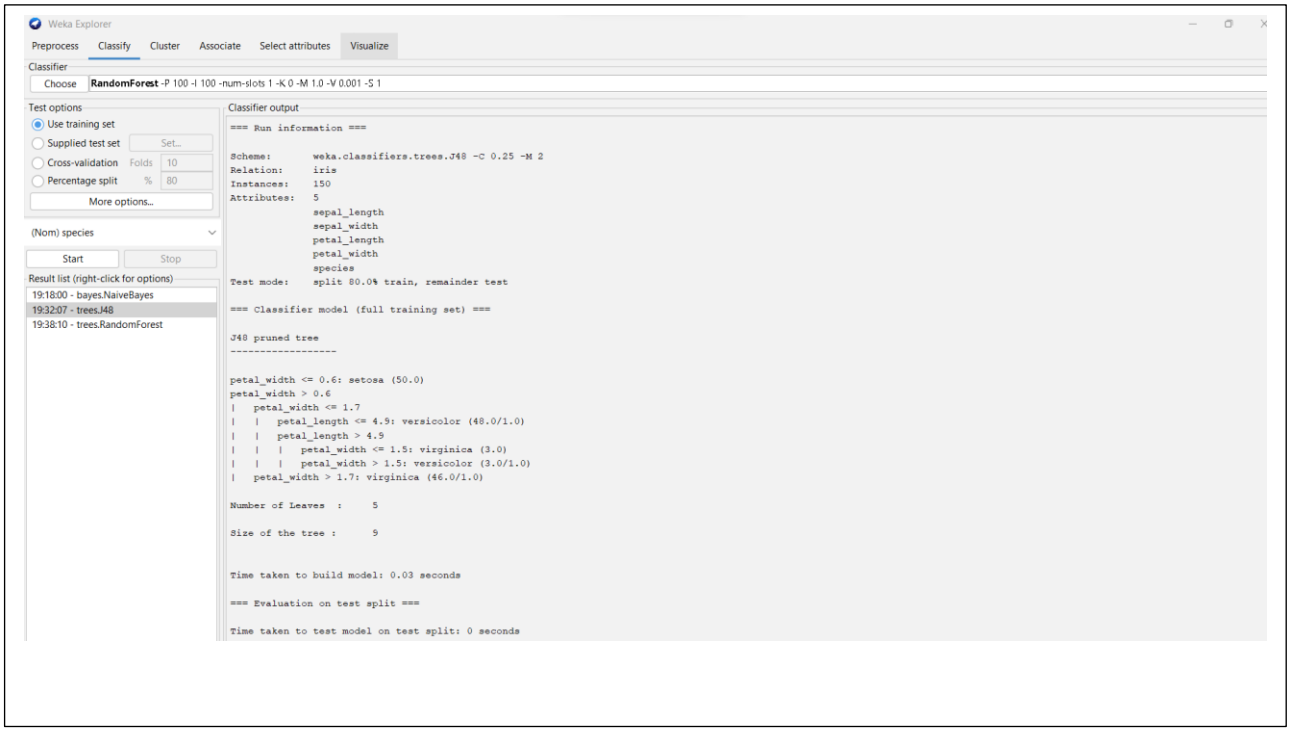
1. Performance Metrics:
    - J48 Decision Tree: Achieved a perfect classification rate of 100% with no errors. This suggests that the decision boundaries for the **Iris dataset** are well-defined and can be effectively captured by a simple decision tree.
    - Naive Bayes: Scored a high accuracy of 96.67%, with one instance of virginica misclassified as versicolor. Naive Bayes, which assumes independence between features, still performed admirably despite the potential correlation between attributes in the Iris dataset.
    - RandomForest: Also achieved a 96.67% accuracy, with the same misclassification issue as Naive Bayes. As an ensemble method, RandomForest typically provides a more robust generalization capability by combining multiple decision trees. However, in this case, it did not outperform the simpler J48 model.
2. Model Complexity and Interpretability:
    - The J48 model stands out for its simplicity and interpretability. The decision tree's structure provides clear insights into the decision-making process, which is valuable in applications where understanding the model's rationale is crucial.
    - Naive Bayes offers a good balance between simplicity and performance, especially in cases where the assumption of feature independence is approximately valid.
    - RandomForest, while potentially more robust to overfitting and capable of handling complex interactions between attributes, loses out on interpretability due to its ensemble nature.
3. Applicability and Generalization:
    - The success of the J48 model on this dataset might not generalize to more complex or less clearly separable datasets. Its performance, though outstanding here, could be an instance of the dataset's simplicity.
    - Naive Bayes and RandomForest are generally more adaptable to a variety of datasets. Their performance here, while slightly lower than J48, suggests they could potentially perform better on datasets with more complex or less distinct decision boundaries.

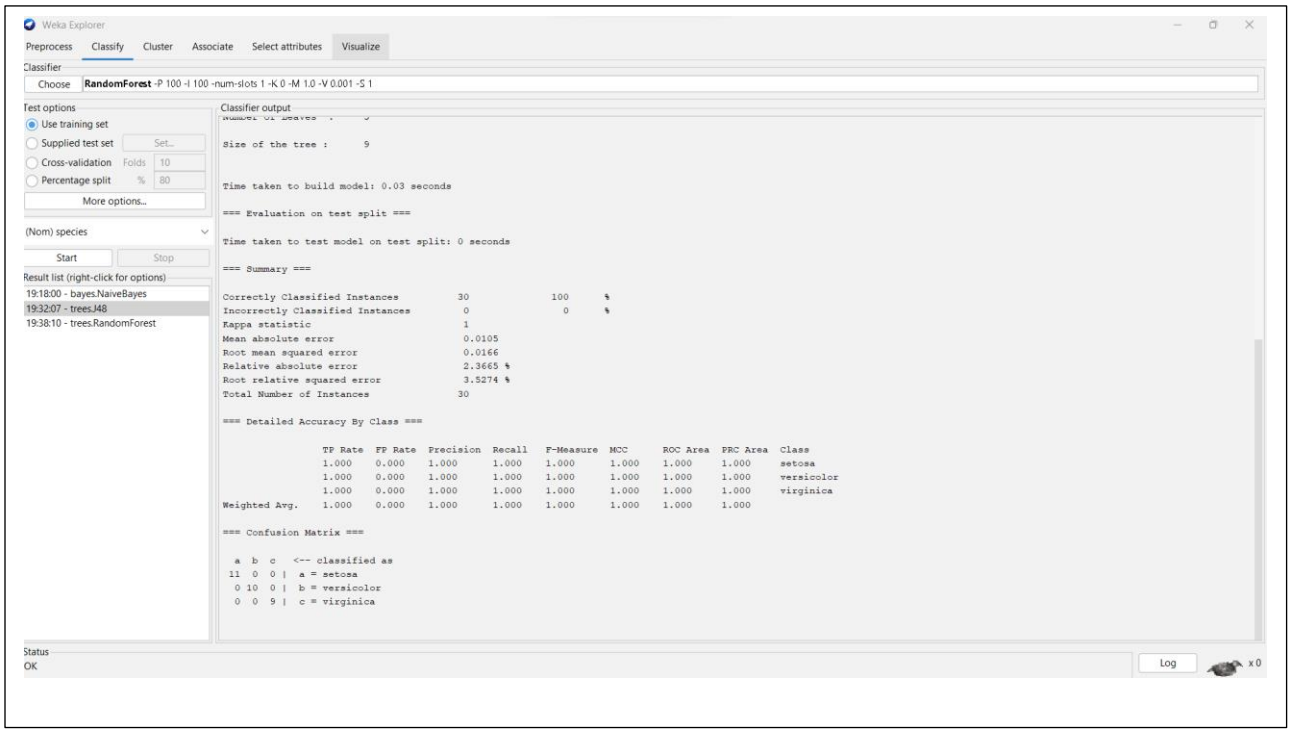

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Generate...
Undo
Edit...
Save...
Filter
Choose
None
Apply
Stop
Current relation
Relation: iris
Instances: 150
Attributes: 5
Sum of weights: 150
Selected attribute
Name: species
Missing: 0 (0%)
Distinct: 3
Type: Nominal
Unique: 0 (0%)
Attributes
All
None
Invert
Pattern
No.
Name
1 sepal_length
2 sepal_width
3 petal_length
4 petal_width
5 species
No.
Label
Count
Weight
1 setosa 50 50
2 versicolor 50 50
3 virginica 50 50
Class: species (Nom)
Visualize All
50
50
50
Remove
Status
OK
Log
x 0

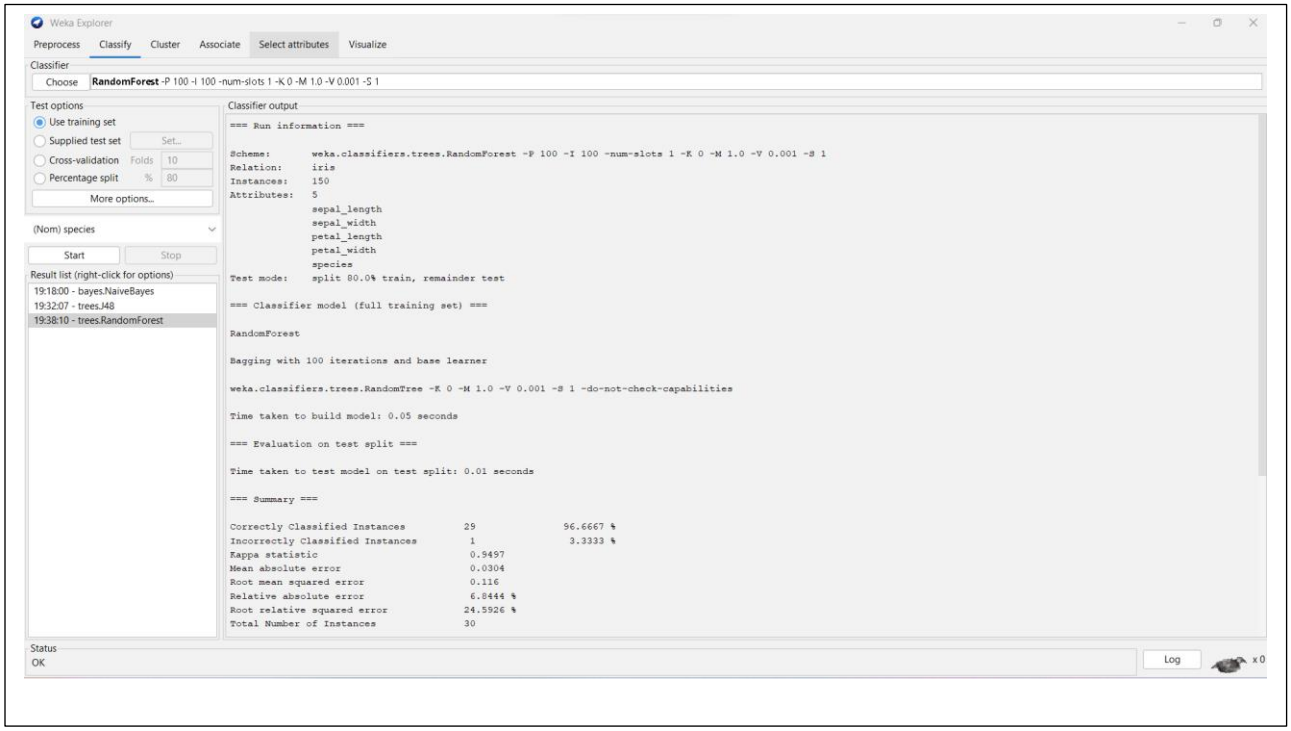

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
80
More options...
(Nom) species
Start
Stop
Result list (right-click for options)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
Classifier output
=== Run information ===
Scheme: weka.classifiers.bayes.NaiveBayes
Relation: iris
Instances: 150
Attributes: 5
sepal_length
sepal_width
petal_length
petal_width
species
Test mode: split 80.0% train, remainder test
=== Classifier model (full training set) ===
Naive Bayes Classifier
Class
Attribute setosa versicolor virginica
(0.33) (0.33) (0.33)
sepal_length
mean 4.9913 5.9379 6.5795
std. dev. 0.355 0.5042 0.6353
weight sum 50 50 50
precision 0.1059 0.1059 0.1059
sepal_width
mean 3.4015 2.7687 2.9629
std. dev. 0.3925 0.3038 0.3088
weight sum 50 50 50
precision 0.1091 0.1091 0.1091
petal_length
mean 1.4694 4.2452 5.5516
std. dev. 0.1782 0.4712 0.5529
weight sum 50 50 50
Status
OK
Log
x 0

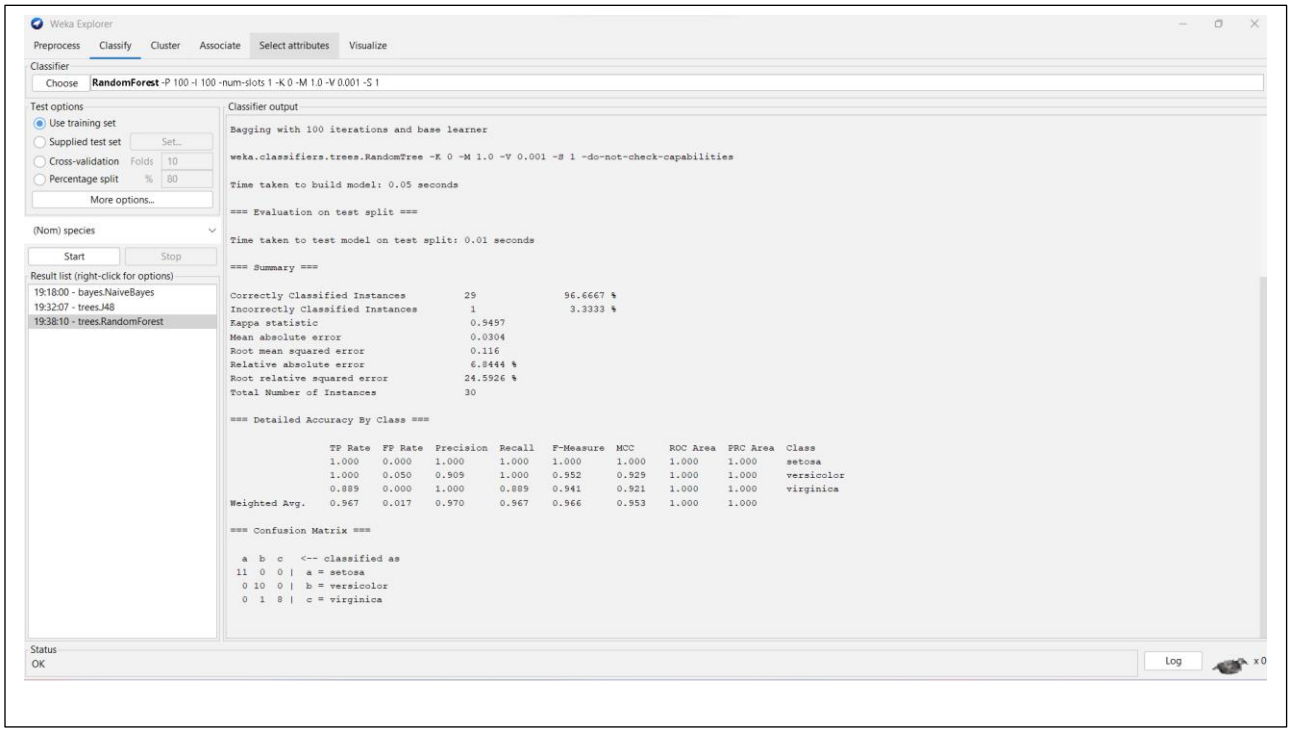

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
80
More options...
(Nom) species
Start
Stop
Result list (right-click for options)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
Classifier output
precision 0.1405 0.1405 0.1405
petal_width
mean 0.2743 1.3097 2.0343
std. dev. 0.1096 0.1915 0.2646
weight sum 50 50 50
precision 0.1143 0.1143 0.1143
Time taken to build model: 0.01 seconds
=== Evaluation on test split ===
Time taken to test model on test split: 0.01 seconds
=== Summary ===
Correctly Classified Instances 29 96.6667 %
Incorrectly Classified Instances 1 3.3333 %
Kappa statistic 0.9497
Mean absolute error 0.0304
Root mean squared error 0.1226
Relative absolute error 6.8425 %
Root relative squared error 25.9804 %
Total Number of Instances 30
=== 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 setosa
1.000 0.050 0.909 1.000 0.952 0.929 0.995 0.991 versicolor
0.889 0.000 1.000 0.889 0.941 0.921 0.995 0.989 virginica
Weighted Avg. 0.967 0.017 0.970 0.967 0.966 0.953 0.997 0.994
=== Confusion Matrix ===
Status
OK
Log
x 0


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
Classifier output
precision 0.1143 0.1143 0.1143
Time taken to build model: 0.01 seconds
=== Evaluation on test split ===
Time taken to test model on test split: 0.01 seconds
=== Summary ===
Correctly Classified Instances 29 96.6667 %
Incorrectly Classified Instances 1 3.3333 %
Kappa statistic 0.9497
Mean absolute error 0.0304
Root mean squared error 0.1226
Relative absolute error 6.8425 %
Root relative squared error 25.9804 %
Total Number of Instances 30
=== 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 setosa
1.000 0.050 0.909 1.000 0.952 0.929 0.995 0.991 versicolor
0.889 0.000 1.000 0.889 0.941 0.921 0.995 0.989 virginica
Weighted Avg. 0.967 0.017 0.970 0.967 0.966 0.953 0.997 0.994
=== Confusion Matrix ===
a b c <-- classified as
11 0 0 | a = setosa
0 10 0 | b = versicolor
0 1 8 | c = virginica
Status

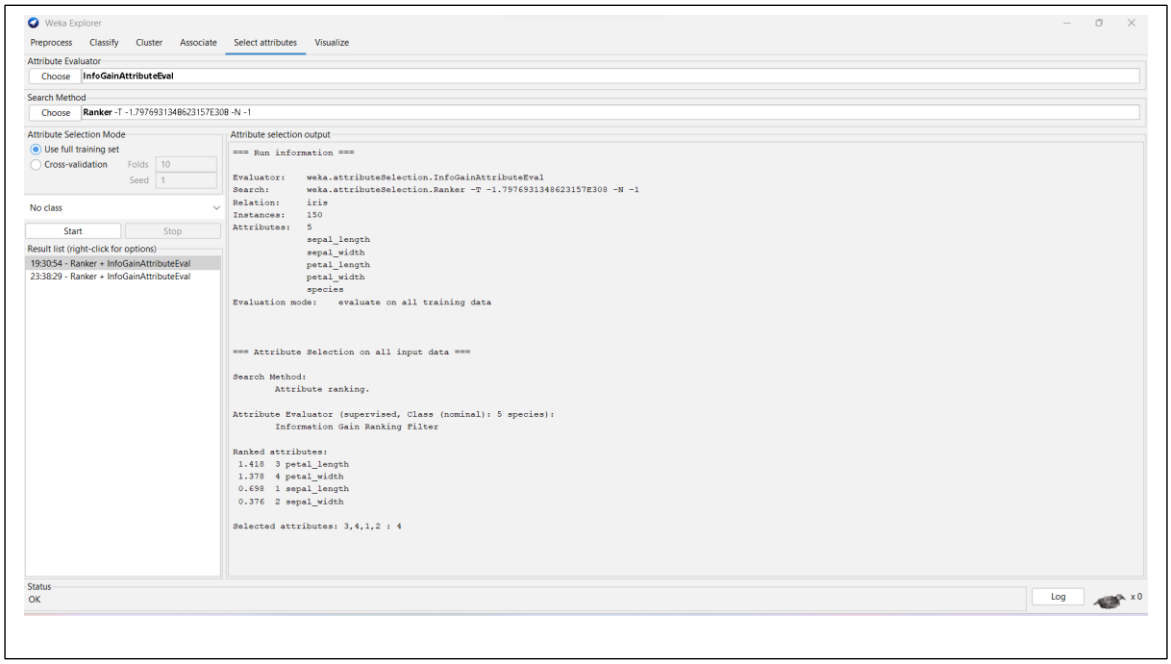

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
80
More options...
(Nom) species
Start
Stop
Result list (right-click for options)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
Classifier output
=== Run information ===
Scheme: weka.classifiers.trees.J48 -C 0.25 -M 2
Relation: iris
Instances: 150
Attributes: 5
sepal_length
sepal_width
petal_length
petal_width
species
Test mode: split 80.0% train, remainder test
=== Classifier model (full training set) ===
J48 pruned tree
petal_width <= 0.6: setosa (50.0)
petal_width > 0.6
| petal_width <= 1.7
| | petal_length <= 4.9: versicolor (48.0/1.0)
| | petal_length > 4.9
| | | petal_width <= 1.5: virginica (3.0)
| | | petal_width > 1.5: versicolor (3.0/1.0)
| petal_width > 1.7: virginica (46.0/1.0)
Number of Leaves : 5
Size of the tree : 9
Time taken to build model: 0.03 seconds
=== Evaluation on test split ===
Time taken to test model on test split: 0 seconds


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
Classifier output
Size of the tree : 9
Time taken to build model: 0.03 seconds
=== Evaluation on test split ===
Time taken to test model on test split: 0 seconds
=== Summary ===
Correctly Classified Instances 30 100 %
Incorrectly Classified Instances 0 0 %
Kappa statistic 1
Mean absolute error 0.0105
Root mean squared error 0.0166
Relative absolute error 2.3665 %
Root relative squared error 3.5274 %
Total Number of Instances 30
=== 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 setosa
1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 versicolor
1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000 virginica
Weighted Avg. 1.000 0.000 1.000 1.000 1.000 1.000 1.000 1.000
=== Confusion Matrix ===
a b c <-- classified as
11 0 0 | a = setosa
0 10 0 | b = versicolor
0 0 9 | c = virginica
Status
OK
Log
x 0


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
Classifier output
=== Run information ===
Scheme: weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Relation: iris
Instances: 150
Attributes: 5
sepal_length
sepal_width
petal_length
petal_width
species
Test mode: split 80.0% train, remainder test
=== Classifier model (full training set) ===
RandomForest
Bagging with 100 iterations and base learner
weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities
Time taken to build model: 0.05 seconds
=== Evaluation on test split ===
Time taken to test model on test split: 0.01 seconds
=== Summary ===
Correctly Classified Instances 29 96.6667 %
Incorrectly Classified Instances 1 3.3333 %
Kappa statistic 0.9497
Mean absolute error 0.0304
Root mean squared error 0.116
Relative absolute error 6.8444 %
Root relative squared error 24.5926 %
Total Number of Instances 30
Status
OK
Log
x 0


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
Classifier output
Bagging with 100 iterations and base learner
weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities
Time taken to build model: 0.05 seconds
=== Evaluation on test split ===
Time taken to test model on test split: 0.01 seconds
=== Summary ===
Correctly Classified Instances 29 96.6667 %
Incorrectly Classified Instances 1 3.3333 %
Kappa statistic 0.9497
Mean absolute error 0.0304
Root mean squared error 0.116
Relative absolute error 6.8444 %
Root relative squared error 24.5926 %
Total Number of Instances 30
=== 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 setosa
1.000 0.050 0.909 1.000 0.952 0.929 1.000 1.000 versicolor
0.889 0.000 1.000 0.889 0.941 0.921 1.000 1.000 virginica
Weighted Avg. 0.967 0.017 0.970 0.967 0.966 0.953 1.000 1.000
=== Confusion Matrix ===
a b c <-- classified as
11 0 0 | a = setosa
0 10 0 | b = versicolor
0 1 8 | c = virginica
Status
OK
Log
x 0

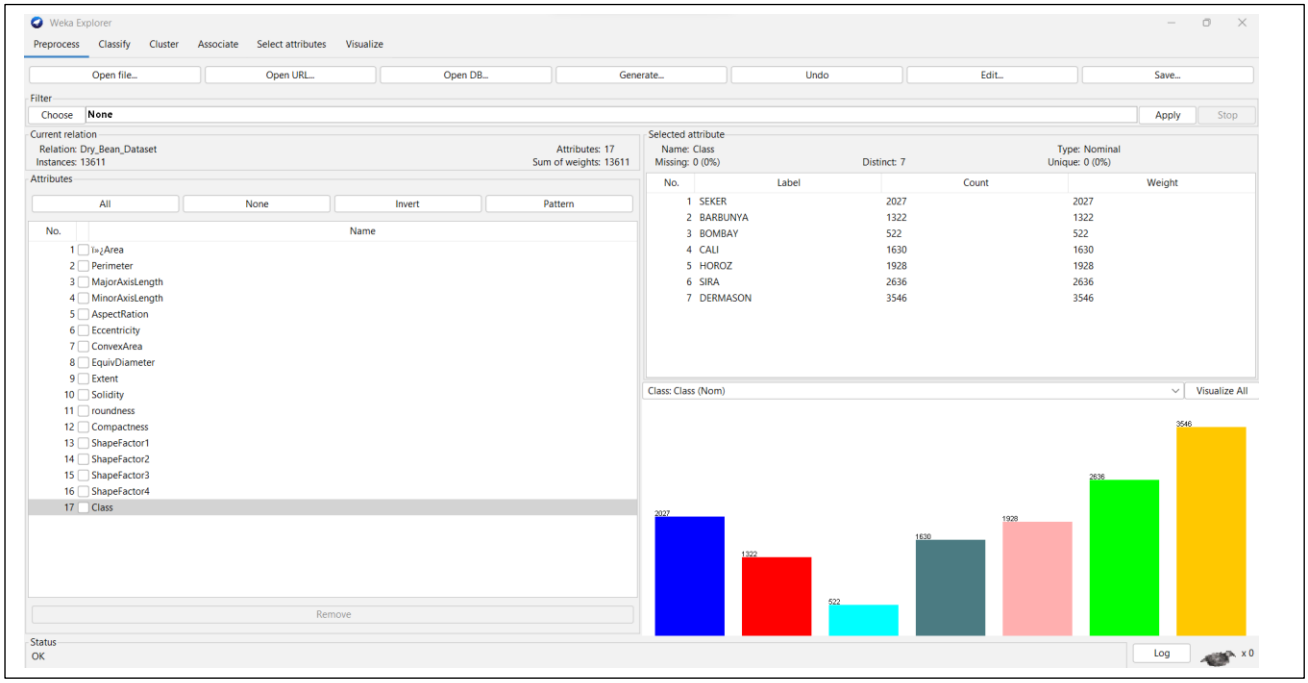
Feature Ranking

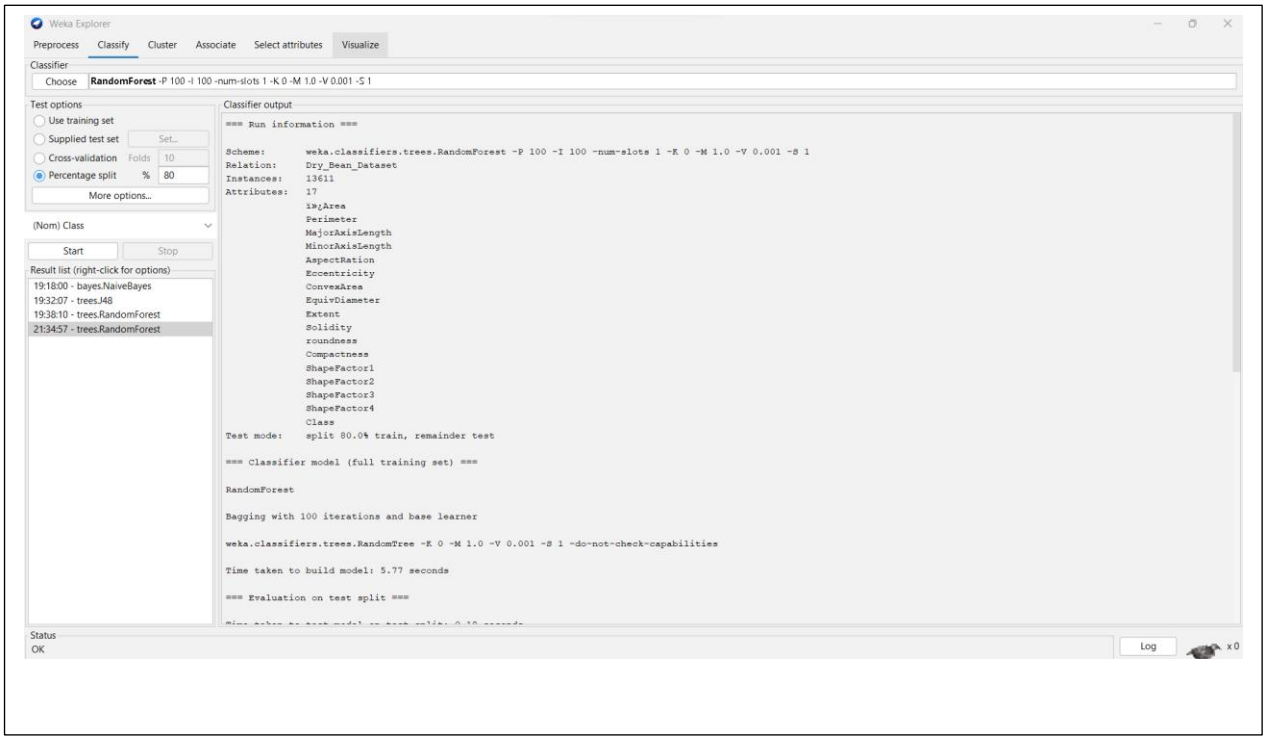


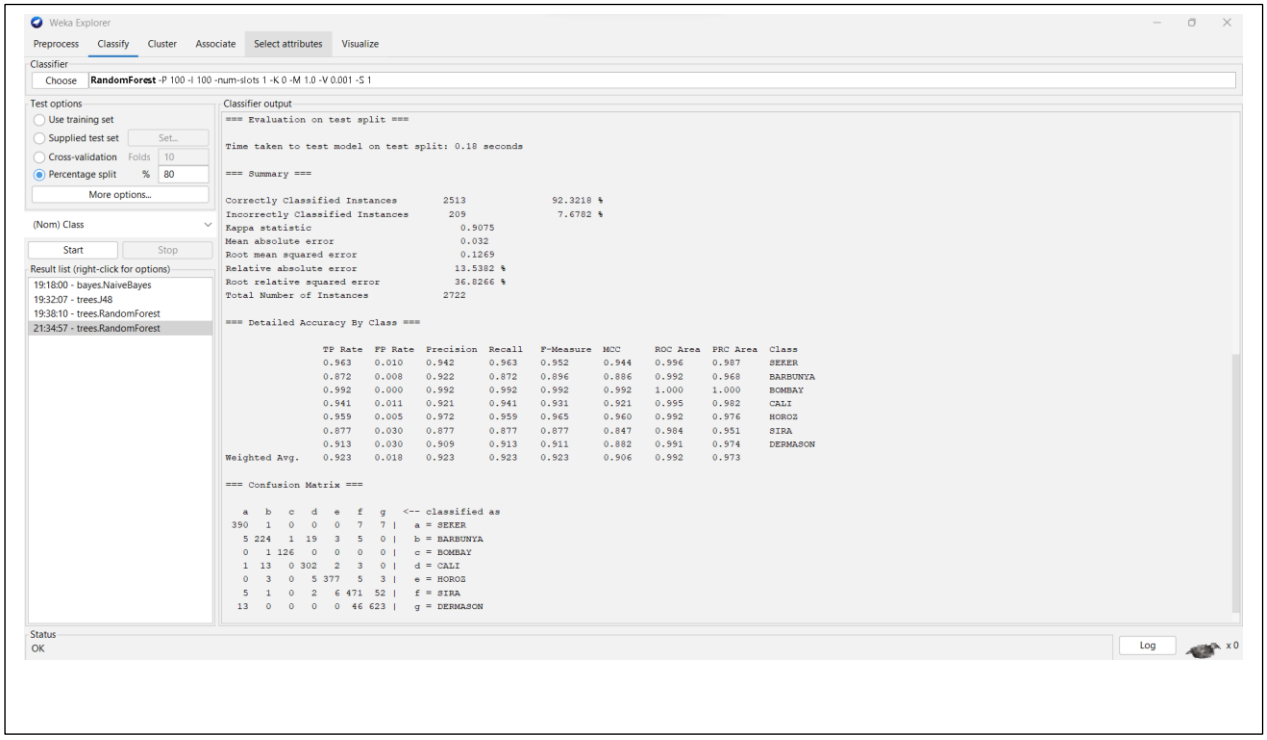
Dry Bean Dataset

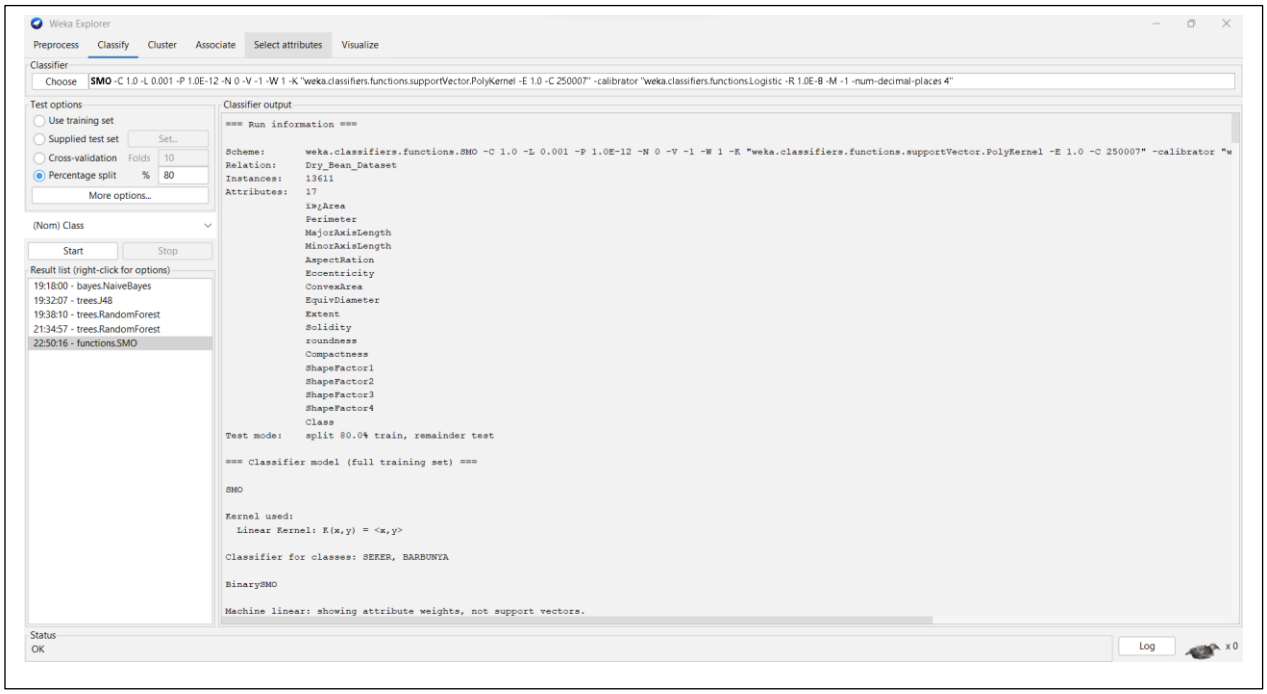
The performance of three machine learning models—Naive Bayes, Support Vector Machine (SMO), and RandomForest—has been evaluated on the Dry Bean Dataset using WEKA. Here's a comparative analysis:

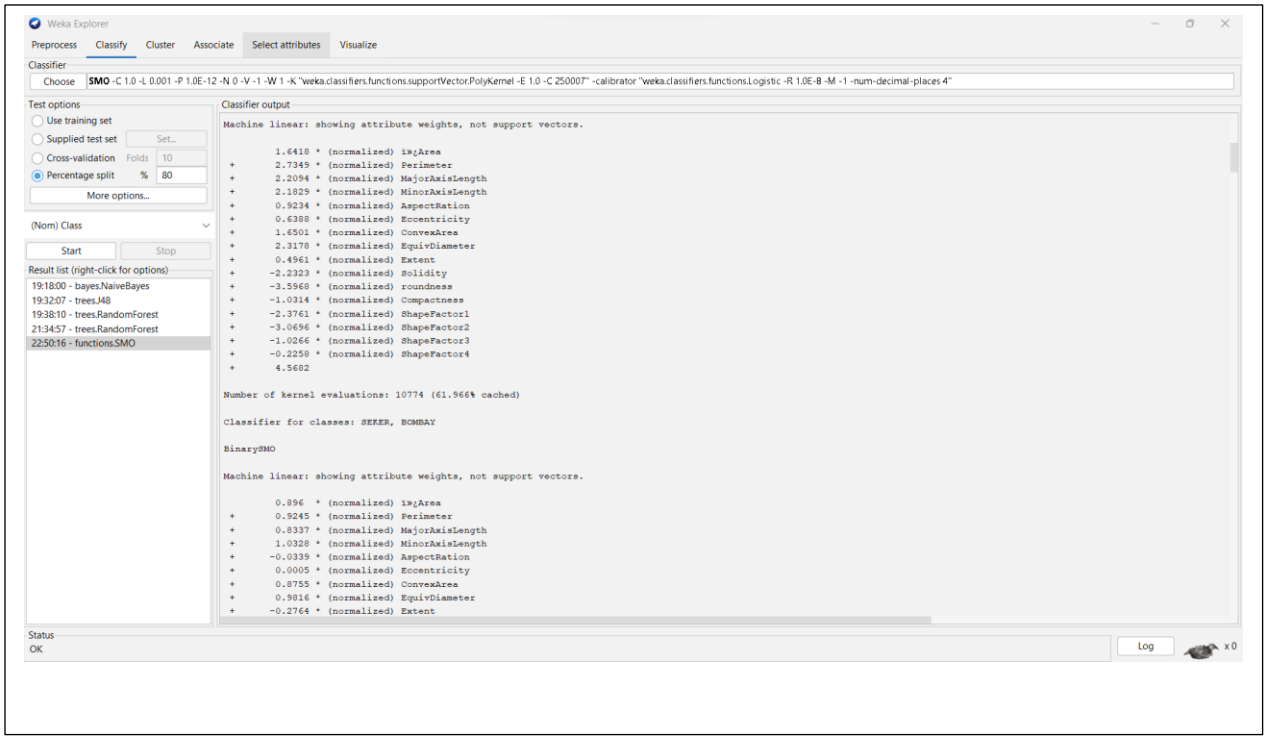
1. **Performance Metrics**:
    - **Naive Bayes**: Achieved 89.53% accuracy with a Kappa statistic of 0.8741. The detailed accuracy by class showed reasonably good performance across different classes, but there were notable errors in classification.
    - **Support Vector Machine (SMO)**: Demonstrated a higher accuracy of 91.95% and a Kappa statistic of 0.9031. This model showed an improvement in class-specific accuracies and precision, indicating a better handling of the dataset's variance.
    - **RandomForest**: Exhibited the highest accuracy among the three at 92.32% and a Kappa statistic of 0.9075. It also showed the lowest error rates and the most balanced class-wise performance, indicating its effectiveness in managing the dataset's complexities.
2. **Model Complexity and Interpretability**:
    - **Naive Bayes** is known for its simplicity and ease of interpretation but often makes strong independence assumptions that may not hold in all datasets.
    - **SMO** is more complex and less interpretable due to its reliance on support vectors and kernel transformations, but it's effective in capturing complex relationships in the data.
    - **RandomForest** is an ensemble method that offers robustness against overfitting. However, it loses some interpretability due to its ensemble nature.
3. **Generalization Capability**:
    - The higher accuracy of **SMO** and **RandomForest** suggests they are better at capturing the nuances in the dataset, potentially offering better generalization.
    - **RandomForest**, in particular, showed excellent performance across different classes, indicating its robustness and effectiveness in handling diverse data characteristics.
4. **Computational Efficiency**:
    - **Naive Bayes** is the most computationally efficient, followed by **SMO**.
    - **RandomForest** took the longest time to build the model, which is expected due to its complex ensemble nature.

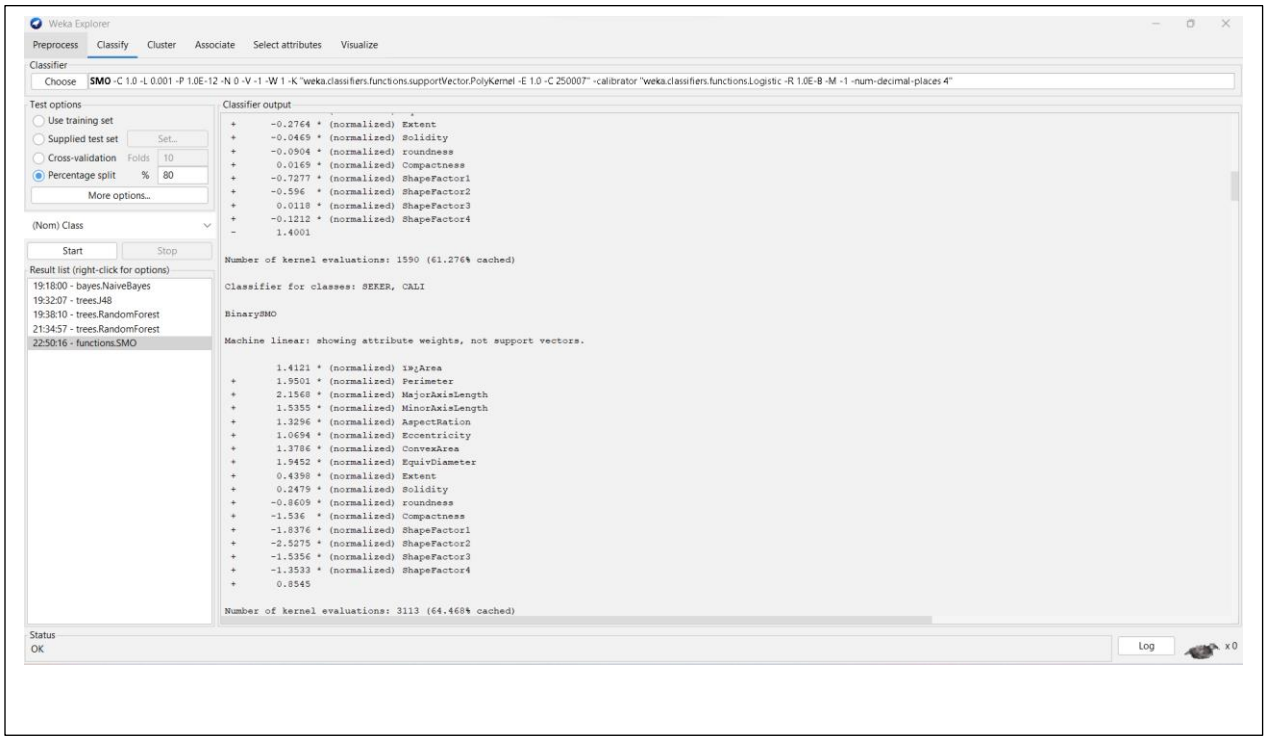

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Generate...
Undo
Edit...
Save...
Filter
Choose
None
Apply
Stop
Current relation
Relation: Dry_Bean_Dataset
Instances: 13611
Attributes: 17
Sum of weights: 13611
Selected attribute
Name: Class
Missing: 0 (0%)
Distinct: 7
Type: Nominal
Unique: 0 (0%)
Attributes
All
None
Invert
Pattern
No.
Name
1 i»¿Area
2 Perimeter
3 MajorAxisLength
4 MinorAxisLength
5 AspectRation
6 Eccentricity
7 ConvexArea
8 EquivDiameter
9 Extent
10 Solidity
11 roundness
12 Compactness
13 ShapeFactor1
14 ShapeFactor2
15 ShapeFactor3
16 ShapeFactor4
17 Class
Remove
No.
Label
Count
Weight
1 SEKER 2027 2027
2 BARBUNYA 1322 1322
3 BOMBAY 522 522
4 CALI 1630 1630
5 HOROZ 1928 1928
6 SIRA 2636 2636
7 DERMASON 3546 3546
Class: Class (Nom)
Visualize All
2027
1322
522
1630
1928
2636
3546
Status
OK
Log
x 0

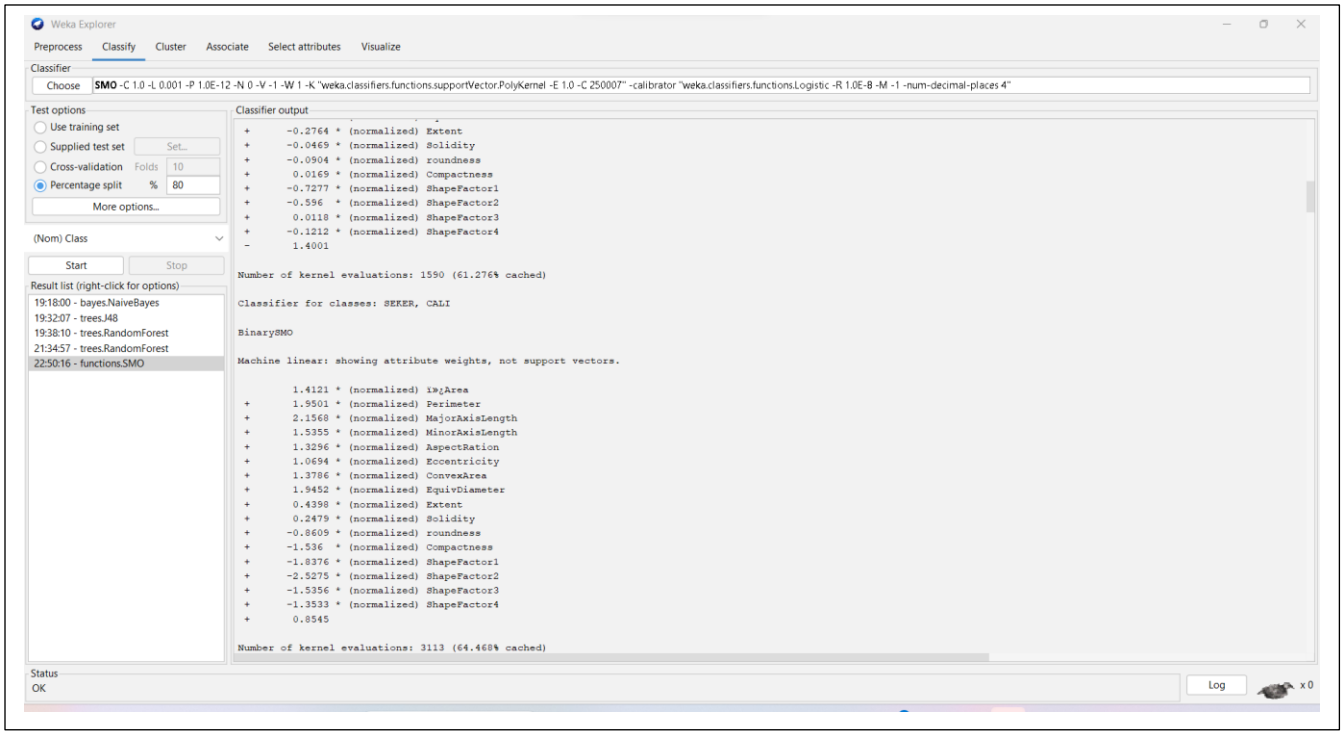

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
80
More options...
(Nom) Class
Start
Stop
Result list (right-click for options)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
Classifier output
=== Run information ===
Scheme: weka.classifiers.trees.RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Relation: Dry_Bean_Dataset
Instances: 13611
Attributes: 17
i»¿Area
Perimeter
MajorAxisLength
MinorAxisLength
AspectRation
Eccentricity
ConvexArea
EquivDiameter
Extent
Solidity
roundness
Compactness
ShapeFactor1
ShapeFactor2
ShapeFactor3
ShapeFactor4
Class
Test mode: split 80.0% train, remainder test
=== Classifier model (full training set) ===
RandomForest
Bagging with 100 iterations and base learner
weka.classifiers.trees.RandomTree -K 0 -M 1.0 -V 0.001 -S 1 -do-not-check-capabilities
Time taken to build model: 5.77 seconds
=== Evaluation on test split ===
Status
OK
Log
x 0

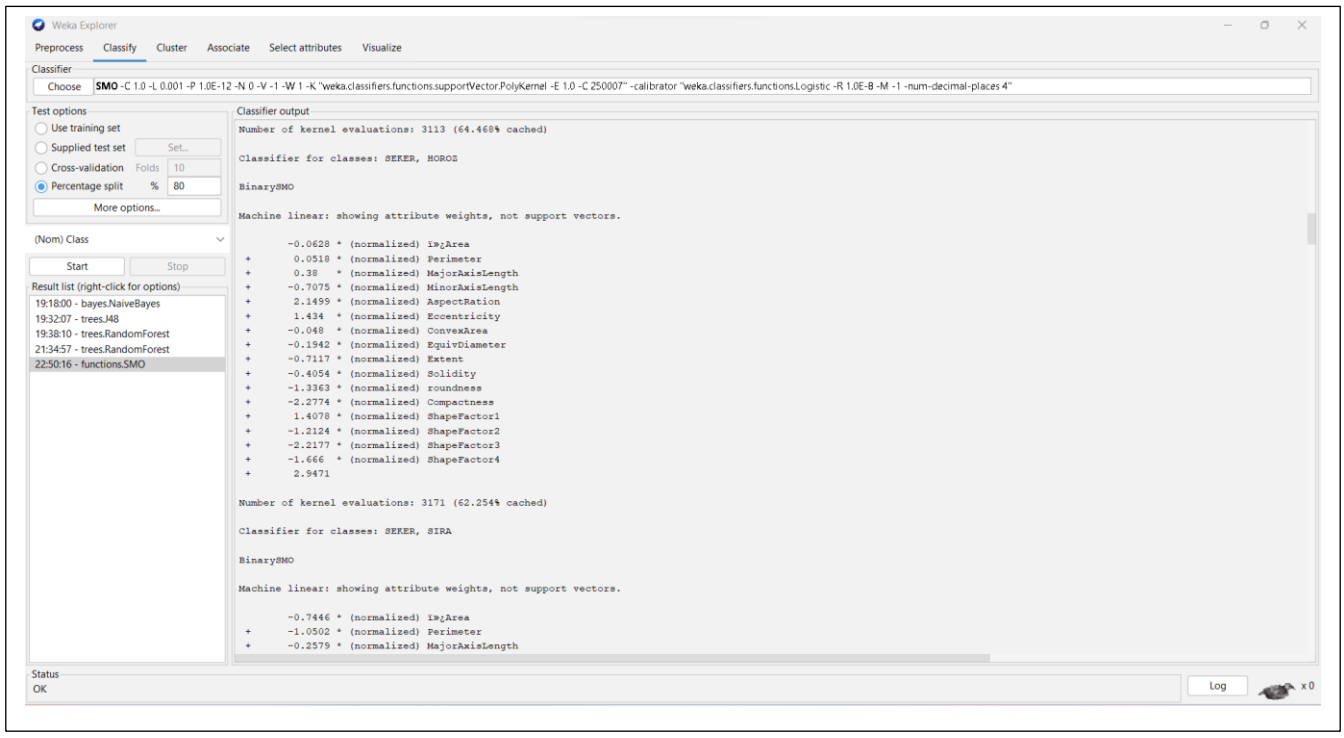

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
RandomForest -P 100 -I 100 -num-slots 1 -K 0 -M 1.0 -V 0.001 -S 1
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
80
More options...
(Nom) Class
Start
Stop
Result list (right-click for options)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
Classifier output
=== Evaluation on test split ===
Time taken to test model on test split: 0.18 seconds
=== Summary ===
Correctly Classified Instances 2513 92.3218 %
Incorrectly Classified Instances 209 7.6782 %
Kappa statistic 0.9075
Mean absolute error 0.032
Root mean squared error 0.1269
Relative absolute error 13.5382 %
Root relative squared error 36.8266 %
Total Number of Instances 2722
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.963 0.010 0.942 0.963 0.952 0.944 0.996 0.987 SEKER
0.872 0.008 0.922 0.872 0.896 0.886 0.992 0.968 BARBUNYA
0.992 0.000 0.992 0.992 0.992 0.992 1.000 1.000 BOMBAY
0.941 0.011 0.921 0.941 0.931 0.921 0.995 0.982 CALI
0.959 0.005 0.972 0.959 0.965 0.960 0.992 0.976 HOROZ
0.877 0.030 0.877 0.877 0.877 0.847 0.984 0.951 SIRA
0.913 0.030 0.909 0.913 0.911 0.882 0.991 0.974 DERMASON
Weighted Avg. 0.923 0.018 0.923 0.923 0.923 0.906 0.992 0.973
=== Confusion Matrix ===
a b c d e f g <-- classified as
390 1 0 0 0 7 7 | a = SEKER
5 224 1 19 3 5 0 | b = BARBUNYA
0 1 126 0 0 0 0 | c = BOMBAY
1 13 0 302 2 3 0 | d = CALI
0 3 0 5 377 5 3 | e = HOROZ
5 1 0 2 6 471 52 | f = SIRA
13 0 0 0 0 46 623 | g = DERMASON
Status
OK
Log
x 0

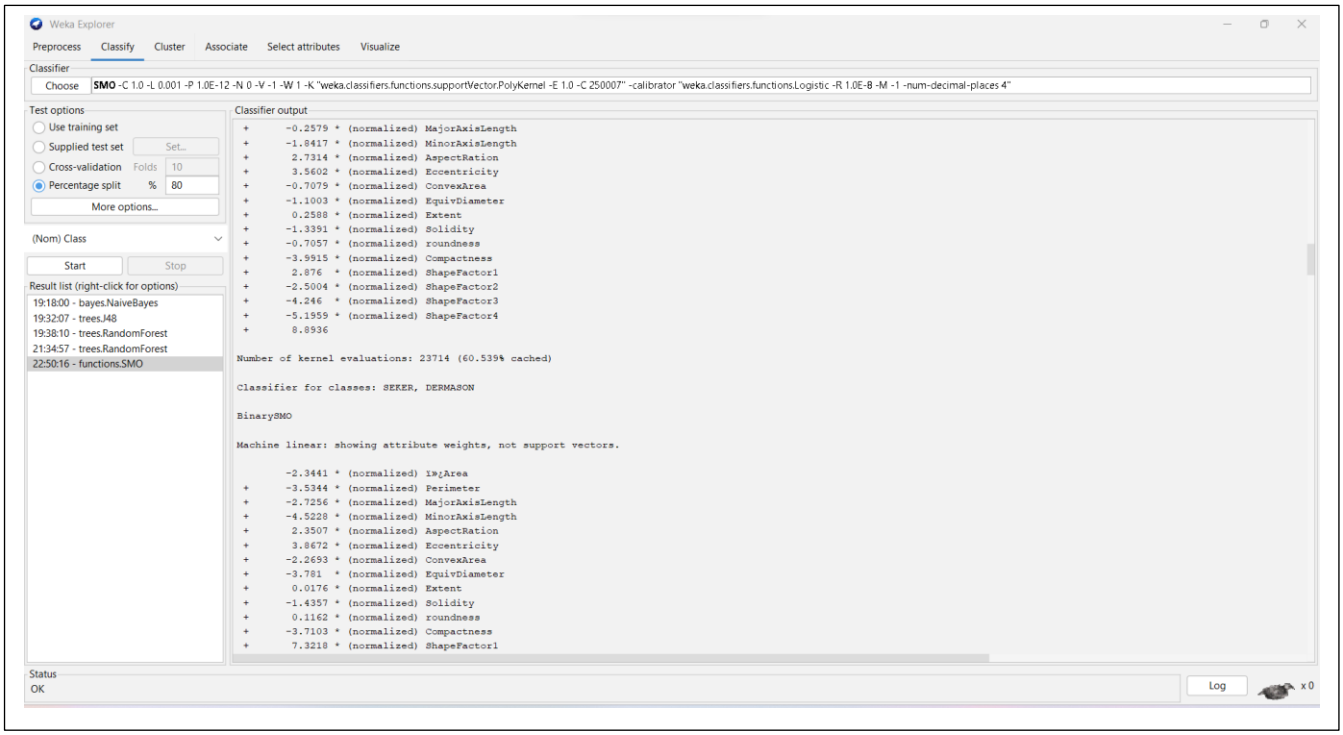

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
=== Run information ===
Scheme: weka.classifiers.functions.SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "w
Relation: Dry_Bean_Dataset
Instances: 13611
Attributes: 17
i»¿Area
Perimeter
MajorAxisLength
MinorAxisLength
AspectRation
Eccentricity
ConvexArea
EquivDiameter
Extent
Solidity
roundness
Compactness
ShapeFactor1
ShapeFactor2
ShapeFactor3
ShapeFactor4
Class
Test mode: split 80.0% train, remainder test
=== Classifier model (full training set) ===
SMO
Kernel used:
Linear Kernel: K(x,y) = <x,y>
Classifier for classes: SEKER, BARBUNYA
BinarySMO
Machine linear: showing attribute weights, not support vectors.
Status
OK
Log
x 0

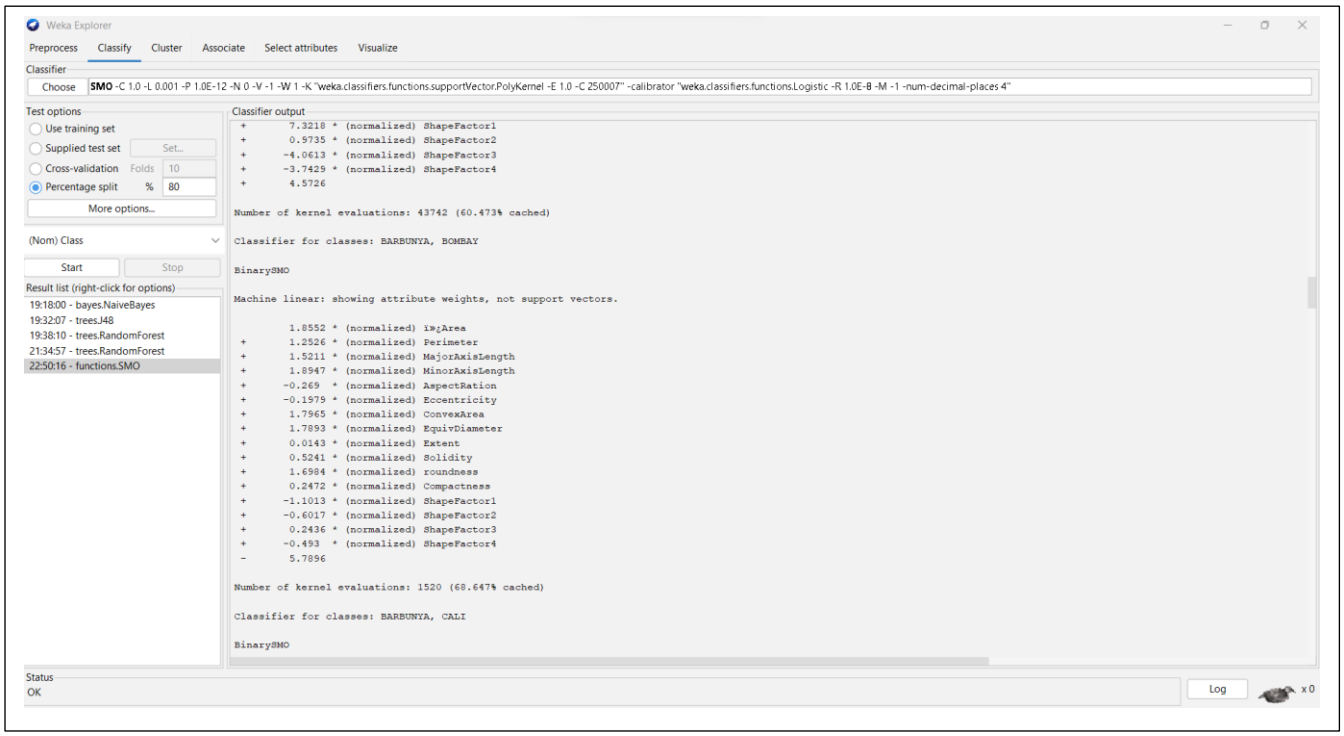

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
Machine linear: showing attribute weights, not support vectors.
1.6418 * (normalized) I▶¿Area
+ 2.7349 * (normalized) Perimeter
+ 2.2094 * (normalized) MajorAxisLength
+ 2.1829 * (normalized) MinorAxisLength
+ 0.9234 * (normalized) AspectRation
+ 0.6388 * (normalized) Eccentricity
+ 1.6501 * (normalized) ConvexArea
+ 2.3178 * (normalized) EquivDiameter
+ 0.4561 * (normalized) Extent
+ -2.2323 * (normalized) Solidity
+ -3.5968 * (normalized) roundness
+ -1.0314 * (normalized) Compactness
+ -2.3761 * (normalized) ShapeFactor1
+ -3.0696 * (normalized) ShapeFactor2
+ -1.0266 * (normalized) ShapeFactor3
+ -0.2258 * (normalized) ShapeFactor4
+ 4.5682
Number of kernel evaluations: 10774 (61.966% cached)
Classifier for classes: SEKER, BOMBAY
BinarySMO
Machine linear: showing attribute weights, not support vectors.
0.896 * (normalized) I▶¿Area
+ 0.9245 * (normalized) Perimeter
+ 0.8337 * (normalized) MajorAxisLength
+ 1.0328 * (normalized) MinorAxisLength
+ -0.0339 * (normalized) AspectRation
+ 0.0005 * (normalized) Eccentricity
+ 0.8755 * (normalized) ConvexArea
+ 0.9816 * (normalized) EquivDiameter
+ -0.2764 * (normalized) Extent
Status
OK
Log
x 0

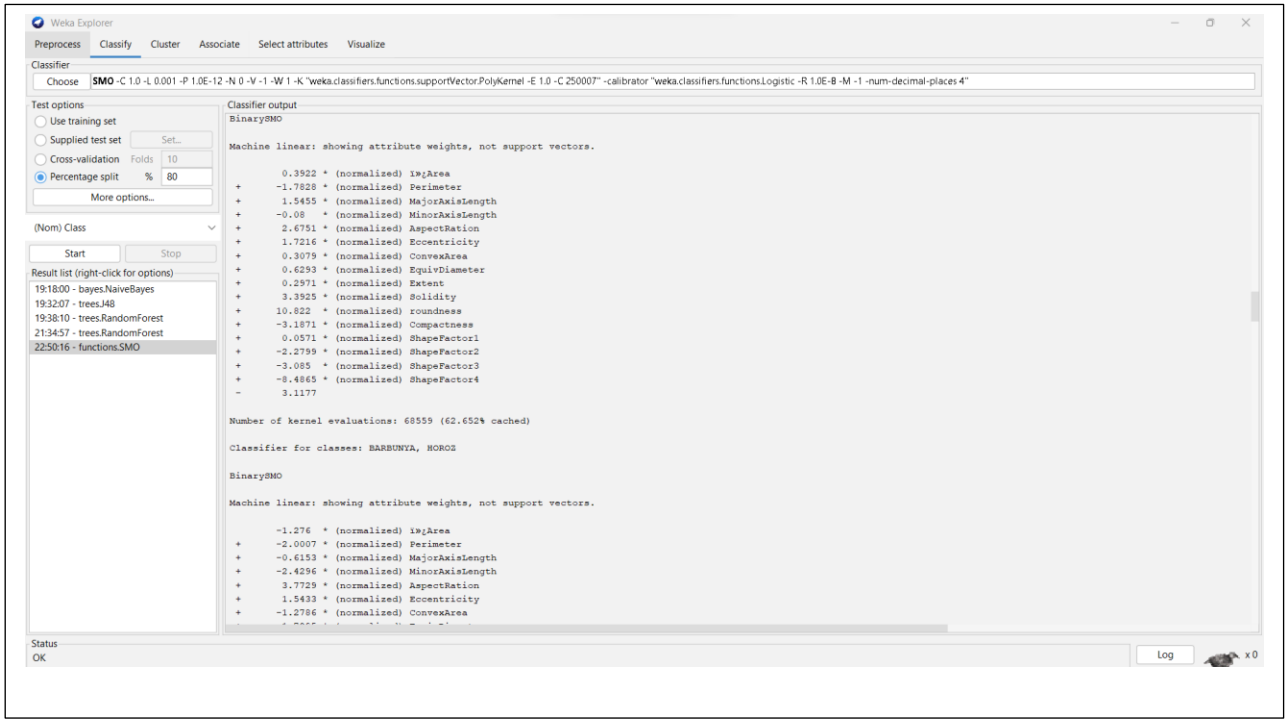

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
+ -0.2764 * (normalized) Extent
+ -0.0469 * (normalized) Solidity
+ -0.0904 * (normalized) roundness
+ 0.0169 * (normalized) Compactness
+ -0.7277 * (normalized) ShapeFactor1
+ -0.596 * (normalized) ShapeFactor2
+ 0.0118 * (normalized) ShapeFactor3
+ -0.1212 * (normalized) ShapeFactor4
- 1.4001
Number of kernel evaluations: 1590 (61.276% cached)
Classifier for classes: SEKER, CALI
BinarySMO
Machine linear: showing attribute weights, not support vectors.
1.4121 * (normalized) I▶¿Area
+ 1.9501 * (normalized) Perimeter
+ 2.1568 * (normalized) MajorAxisLength
+ 1.5355 * (normalized) MinorAxisLength
+ 1.3296 * (normalized) AspectRation
+ 1.0694 * (normalized) Eccentricity
+ 1.3786 * (normalized) ConvexArea
+ 1.9452 * (normalized) EquivDiameter
+ 0.4398 * (normalized) Extent
+ 0.2475 * (normalized) Solidity
+ -0.8609 * (normalized) roundness
+ -1.536 * (normalized) Compactness
+ -1.8376 * (normalized) ShapeFactor1
+ -2.5275 * (normalized) ShapeFactor2
+ -1.5356 * (normalized) ShapeFactor3
+ -1.3533 * (normalized) ShapeFactor4
+ 0.8545
Number of kernel evaluations: 3113 (64.468% cached)
Status
OK
Log
x 0

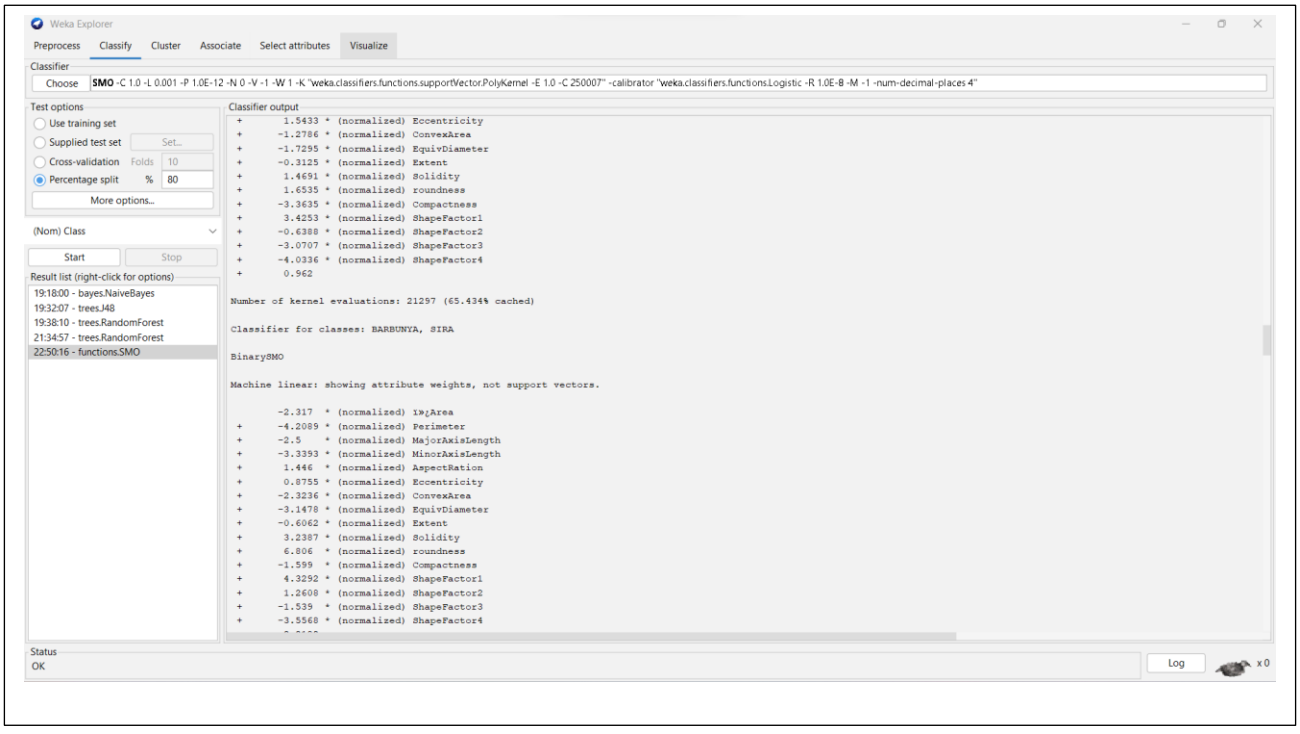

Weka Explorer
Preprocess Classify Cluster Associate Select attributes Visualize
Classifier
Choose SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
+ -0.2764 * (normalized) Extent
+ -0.0469 * (normalized) Solidity
+ -0.0904 * (normalized) roundness
+ 0.0169 * (normalized) Compactness
+ -0.7277 * (normalized) ShapeFactor1
+ -0.596 * (normalized) ShapeFactor2
+ 0.0118 * (normalized) ShapeFactor3
+ -0.1212 * (normalized) ShapeFactor4
- 1.4001
Number of kernel evaluations: 1590 (61.276% cached)
Classifier for classes: SEKER, CALI
BinarySMO
Machine linear: showing attribute weights, not support vectors.
1.4121 * (normalized) i»¿Area
+ 1.9501 * (normalized) Perimeter
+ 2.1568 * (normalized) MajorAxisLength
+ 1.5355 * (normalized) MinorAxisLength
+ 1.3296 * (normalized) AspectRation
+ 1.0694 * (normalized) Eccentricity
+ 1.3786 * (normalized) ConvexArea
+ 1.9452 * (normalized) EquivDiameter
+ 0.4398 * (normalized) Extent
+ 0.2479 * (normalized) Solidity
+ -0.8609 * (normalized) roundness
+ -1.536 * (normalized) Compactness
+ -1.8376 * (normalized) ShapeFactor1
+ -2.5275 * (normalized) ShapeFactor2
+ -1.5356 * (normalized) ShapeFactor3
+ -1.3533 * (normalized) ShapeFactor4
+ 0.8545
Number of kernel evaluations: 3113 (64.468% cached)
Status
OK
Log
x 0

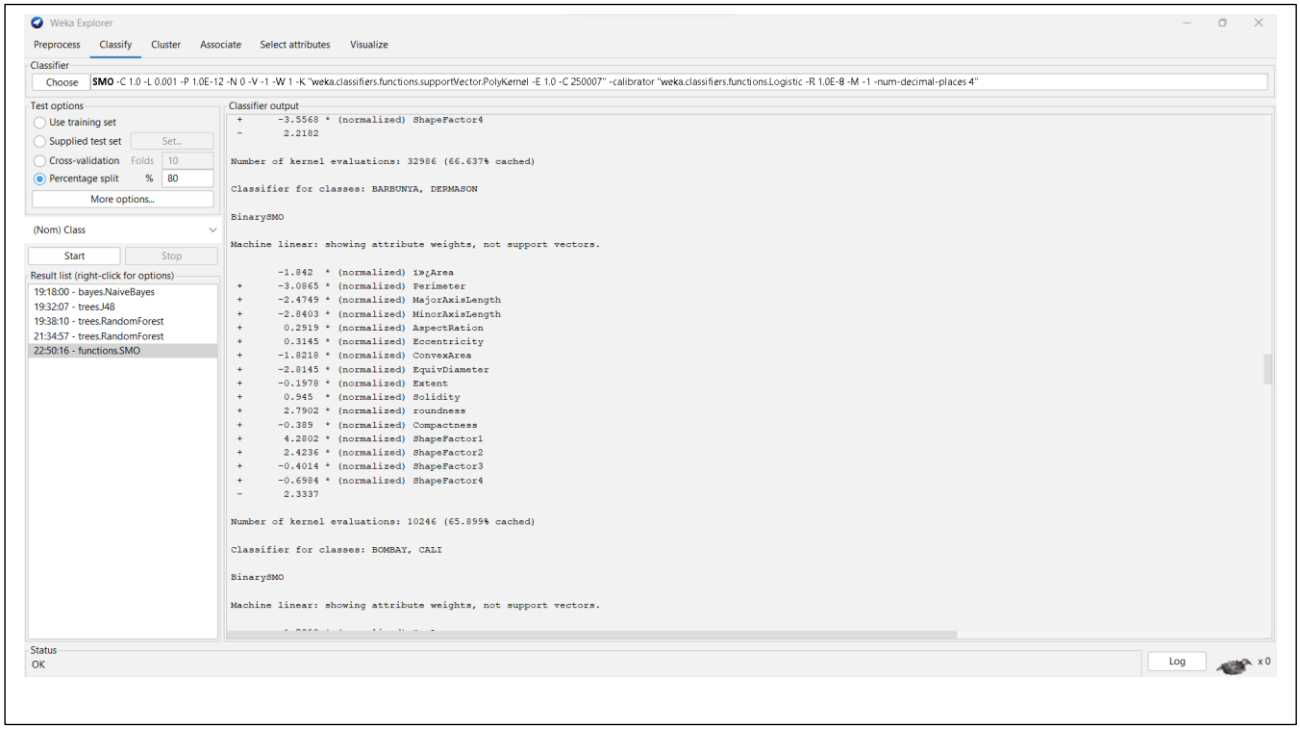

Weka Explorer
Preprocess Classify Cluster Associate Select attributes Visualize
Classifier
Choose SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
Number of kernel evaluations: 3113 (64.468% cached)
Classifier for classes: SEKER, HOROZ
BinarySMO
Machine linear: showing attribute weights, not support vectors.
-0.0628 * (normalized) i»¿Area
+ 0.0518 * (normalized) Perimeter
+ 0.38 * (normalized) MajorAxisLength
+ -0.7075 * (normalized) MinorAxisLength
+ 2.1499 * (normalized) AspectRation
+ 1.434 * (normalized) Eccentricity
+ -0.048 * (normalized) ConvexArea
+ -0.1942 * (normalized) EquivDiameter
+ -0.7117 * (normalized) Extent
+ -0.4054 * (normalized) Solidity
+ -1.3363 * (normalized) roundness
+ -2.2774 * (normalized) Compactness
+ 1.4078 * (normalized) ShapeFactor1
+ -1.2124 * (normalized) ShapeFactor2
+ -2.2177 * (normalized) ShapeFactor3
+ -1.666 * (normalized) ShapeFactor4
+ 2.9471
Number of kernel evaluations: 3171 (62.254% cached)
Classifier for classes: SEKER, SIRA
BinarySMO
Machine linear: showing attribute weights, not support vectors.
-0.7446 * (normalized) i»¿Area
+ -1.0502 * (normalized) Perimeter
+ -0.2579 * (normalized) MajorAxisLength
Status
OK
Log
x 0

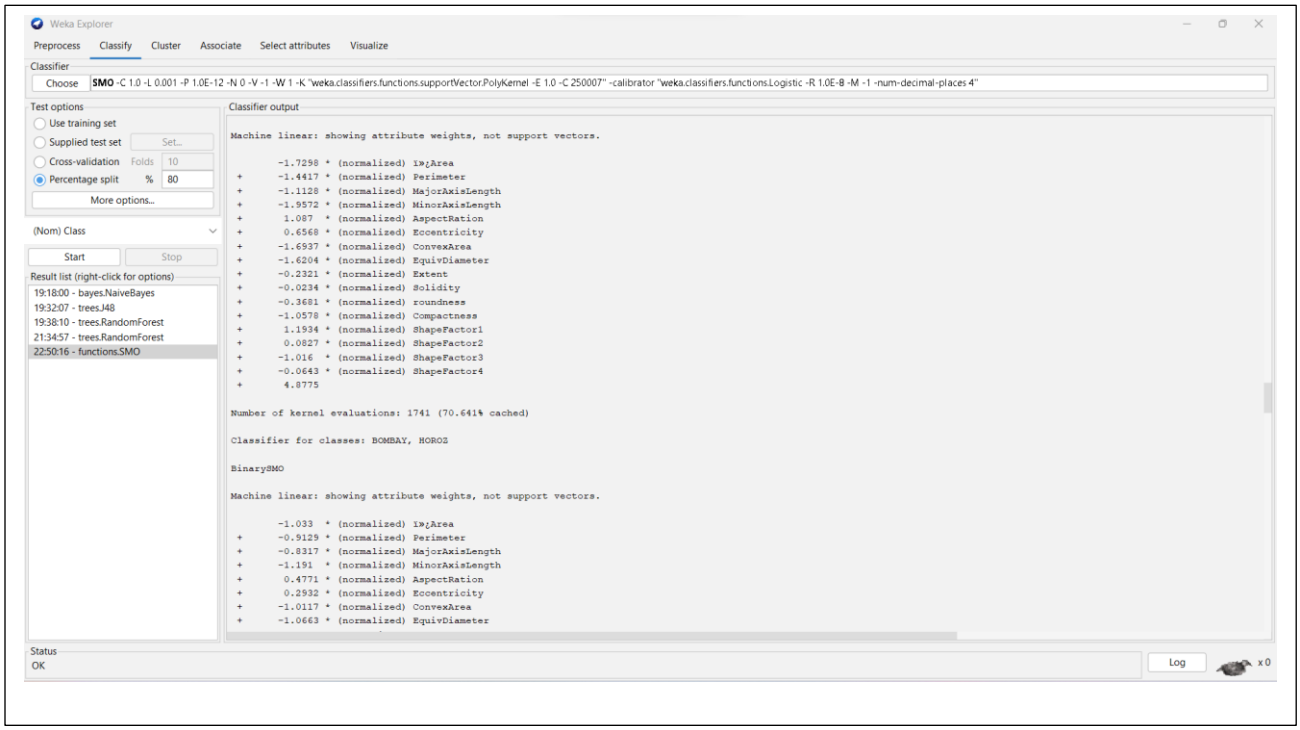

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
+ -0.2579 * (normalized) MajorAxisLength
+ -1.8417 * (normalized) MinorAxisLength
+ 2.7314 * (normalized) AspectRation
+ 3.5602 * (normalized) Eccentricity
+ -0.7079 * (normalized) ConvexArea
+ -1.1003 * (normalized) EquivDiameter
+ 0.2588 * (normalized) Extent
+ -1.3391 * (normalized) Solidity
+ -0.7057 * (normalized) roundness
+ -3.9915 * (normalized) Compactness
+ 2.876 * (normalized) ShapeFactor1
+ -2.5004 * (normalized) ShapeFactor2
+ -4.246 * (normalized) ShapeFactor3
+ -5.1959 * (normalized) ShapeFactor4
+ 8.8936
Number of kernel evaluations: 23714 (60.539% cached)
Classifier for classes: SEKER, DERMASON
BinarySMO
Machine linear: showing attribute weights, not support vectors.
-2.3441 * (normalized) I»¿Area
+ -3.5344 * (normalized) Perimeter
+ -2.7256 * (normalized) MajorAxisLength
+ -4.5228 * (normalized) MinorAxisLength
+ 2.3507 * (normalized) AspectRation
+ 3.8672 * (normalized) Eccentricity
+ -2.2693 * (normalized) ConvexArea
+ -3.781 * (normalized) EquivDiameter
+ 0.0176 * (normalized) Extent
+ -1.4357 * (normalized) Solidity
+ 0.1162 * (normalized) roundness
+ -3.7103 * (normalized) Compactness
+ 7.3218 * (normalized) ShapeFactor1
Status
OK
Log
x 0

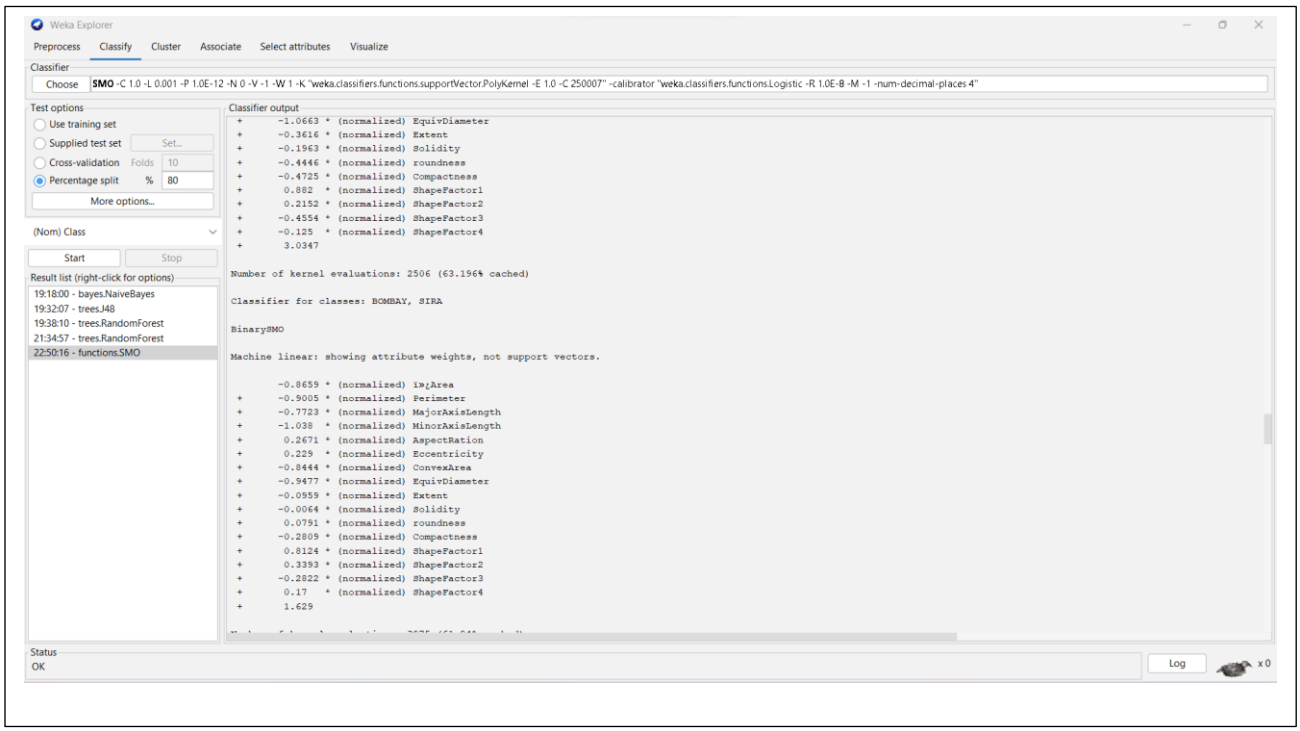

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
+ 7.3218 * (normalized) ShapeFactor1
+ 0.9735 * (normalized) ShapeFactor2
+ -4.0613 * (normalized) ShapeFactor3
+ -3.7429 * (normalized) ShapeFactor4
+ 4.5726
Number of kernel evaluations: 43742 (60.473% cached)
Classifier for classes: BARBUNYA, BOMBAY
BinarySMO
Machine linear: showing attribute weights, not support vectors.
1.8552 * (normalized) I»¿Area
+ 1.2526 * (normalized) Perimeter
+ 1.5211 * (normalized) MajorAxisLength
+ 1.8947 * (normalized) MinorAxisLength
+ -0.269 * (normalized) AspectRation
+ -0.1979 * (normalized) Eccentricity
+ 1.7965 * (normalized) ConvexArea
+ 1.7893 * (normalized) EquivDiameter
+ 0.0143 * (normalized) Extent
+ 0.5241 * (normalized) Solidity
+ 1.6984 * (normalized) roundness
+ 0.2472 * (normalized) Compactness
+ -1.1013 * (normalized) ShapeFactor1
+ -0.6017 * (normalized) ShapeFactor2
+ 0.2436 * (normalized) ShapeFactor3
+ -0.493 * (normalized) ShapeFactor4
- 5.7896
Number of kernel evaluations: 1520 (68.647% cached)
Classifier for classes: BARBUNYA, CALI
BinarySMO
Status
OK
Log
x 0

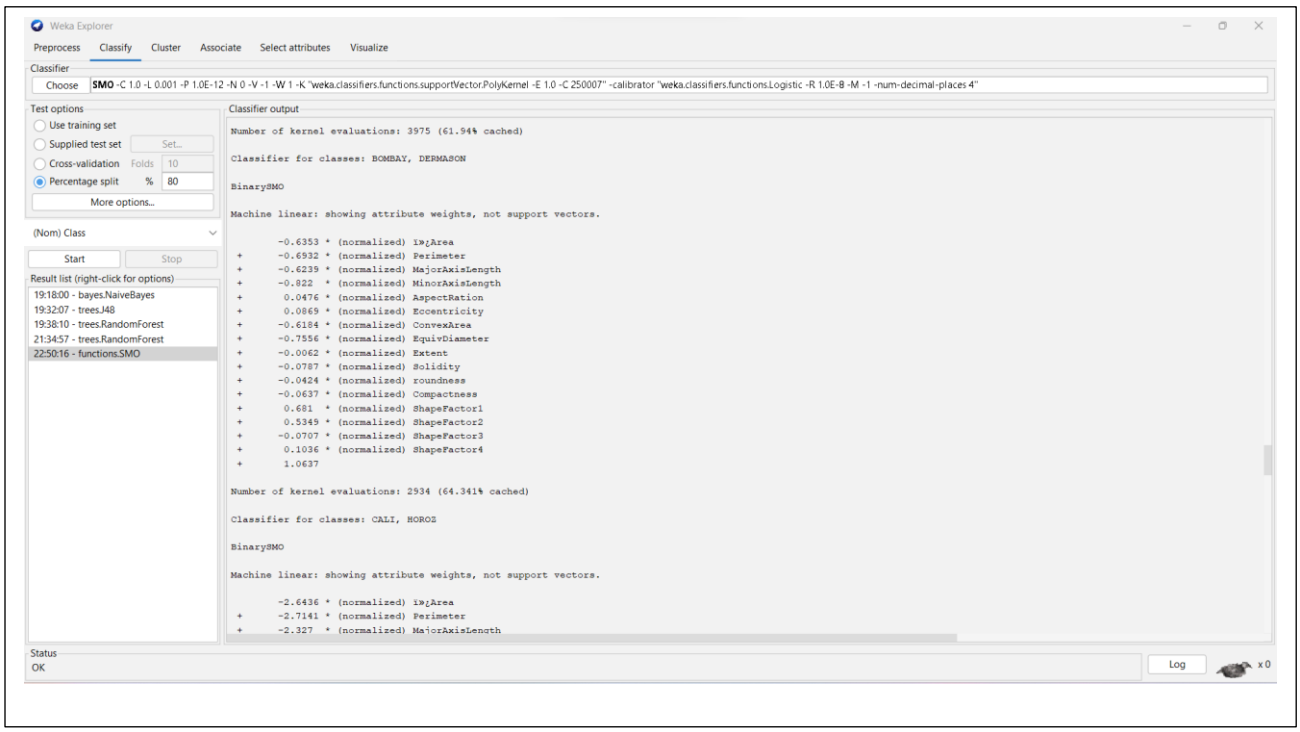

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
BinarySMO
Machine linear: showing attribute weights, not support vectors.
0.3922 * (normalized) İ»¿Area
-1.7828 * (normalized) Perimeter
1.5455 * (normalized) MajorAxisLength
-0.08 * (normalized) MinorAxisLength
2.6751 * (normalized) AspectRation
1.7216 * (normalized) Eccentricity
0.3079 * (normalized) ConvexArea
0.6293 * (normalized) EquivDiameter
0.2571 * (normalized) Extent
3.3925 * (normalized) Solidity
10.822 * (normalized) roundness
-3.1871 * (normalized) Compactness
0.0571 * (normalized) ShapeFactor1
-2.2799 * (normalized) ShapeFactor2
-3.085 * (normalized) ShapeFactor3
-8.4865 * (normalized) ShapeFactor4
3.1177
Number of kernel evaluations: 68559 (62.652% cached)
Classifier for classes: BARBUNYA, HOROZ
BinarySMO
Machine linear: showing attribute weights, not support vectors.
-1.276 * (normalized) İ»¿Area
-2.0007 * (normalized) Perimeter
-0.6153 * (normalized) MajorAxisLength
-2.4296 * (normalized) MinorAxisLength
3.7729 * (normalized) AspectRation
1.5433 * (normalized) Eccentricity
-1.2786 * (normalized) ConvexArea
Status
OK
Log
x 0

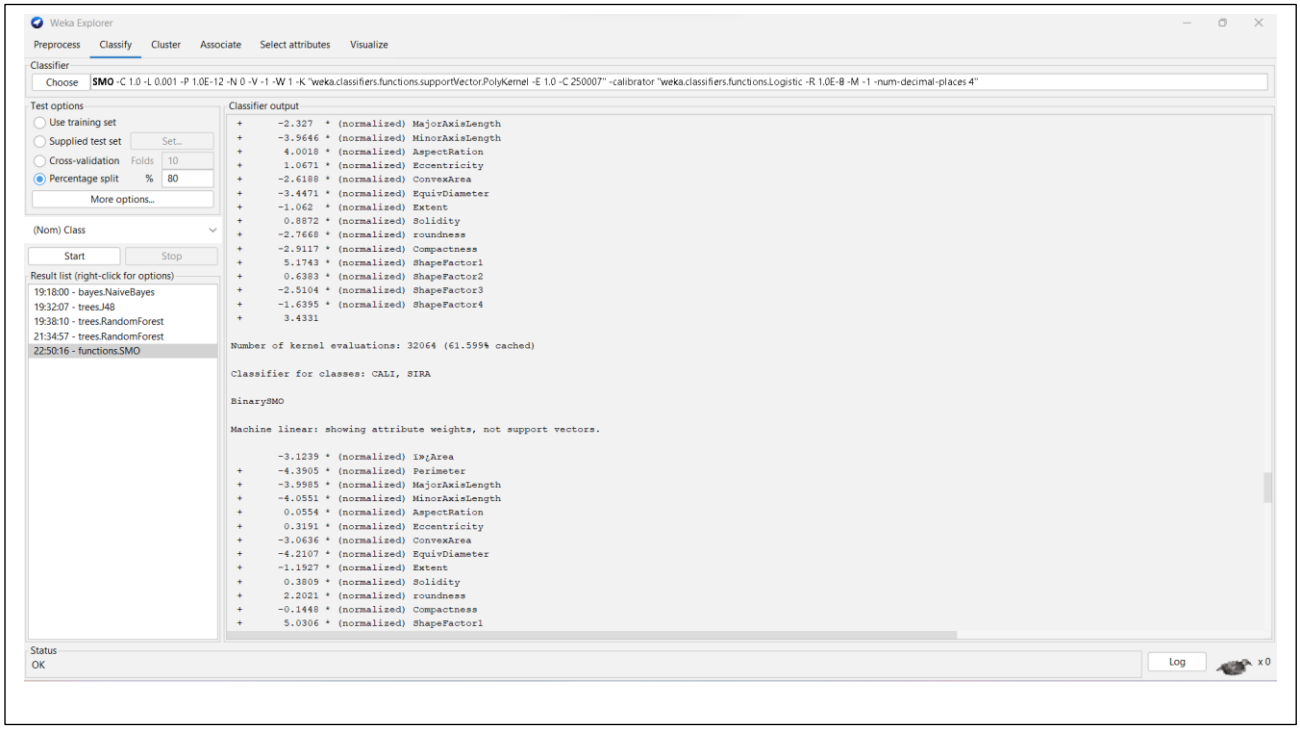

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
1.5433 * (normalized) Eccentricity
-1.2786 * (normalized) ConvexArea
-1.7295 * (normalized) EquivDiameter
-0.3125 * (normalized) Extent
1.4691 * (normalized) Solidity
1.6535 * (normalized) roundness
-3.3635 * (normalized) Compactness
3.4253 * (normalized) ShapeFactor1
-0.6388 * (normalized) ShapeFactor2
-3.0707 * (normalized) ShapeFactor3
-4.0336 * (normalized) ShapeFactor4
0.962
Number of kernel evaluations: 21297 (65.434% cached)
Classifier for classes: BARBUNYA, SIRA
BinarySMO
Machine linear: showing attribute weights, not support vectors.
-2.317 * (normalized) İ»¿Area
-4.2089 * (normalized) Perimeter
-2.5 * (normalized) MajorAxisLength
-3.3393 * (normalized) MinorAxisLength
1.446 * (normalized) AspectRation
0.8755 * (normalized) Eccentricity
-2.3236 * (normalized) ConvexArea
-3.1478 * (normalized) EquivDiameter
-0.6062 * (normalized) Extent
3.2387 * (normalized) Solidity
6.806 * (normalized) roundness
-1.599 * (normalized) Compactness
4.3292 * (normalized) ShapeFactor1
1.2608 * (normalized) ShapeFactor2
-1.539 * (normalized) ShapeFactor3
-3.5568 * (normalized) ShapeFactor4
Status
OK
Log
x 0

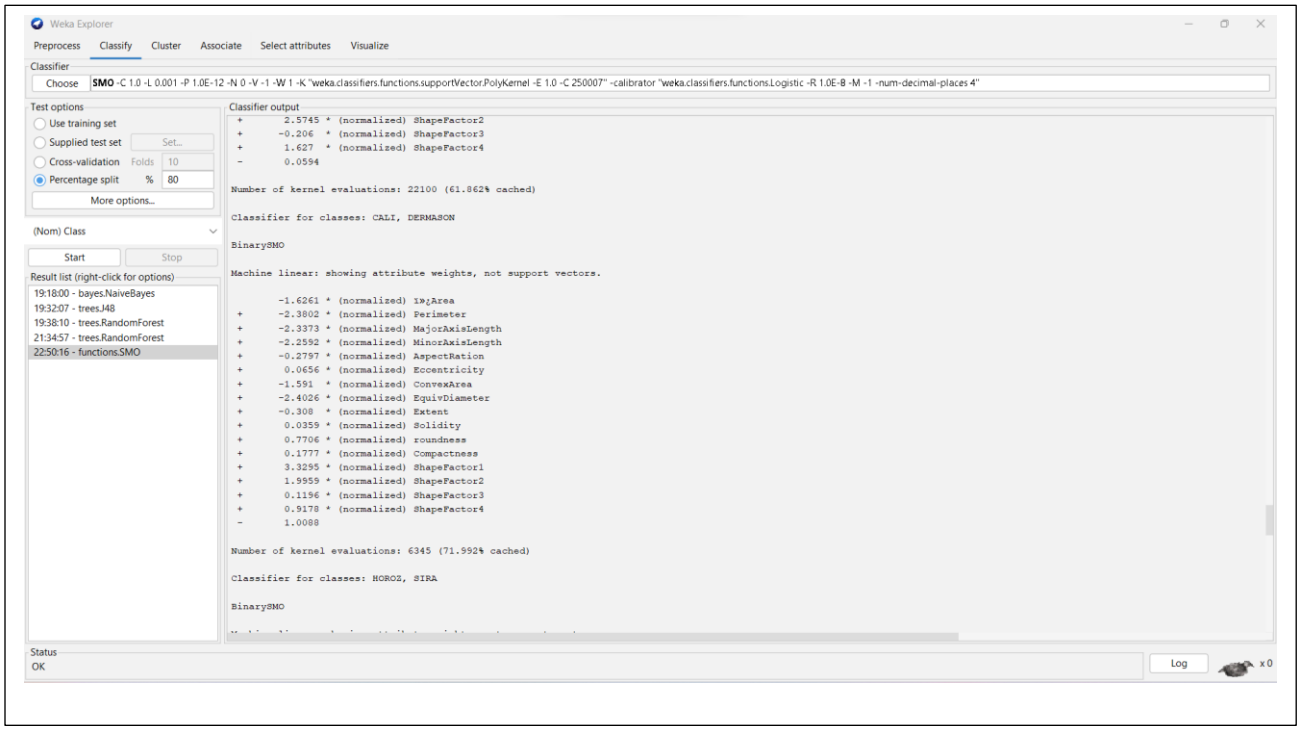

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
+ -3.5568 * (normalized) ShapeFactor4
- 2.2182
Number of kernel evaluations: 32986 (66.637% cached)
Classifier for classes: BARBUNYA, DERMASON
BinarySMO
Machine linear: showing attribute weights, not support vectors.
-1.842 * (normalized) i»¿Area
+ -3.0865 * (normalized) Perimeter
+ -2.4749 * (normalized) MajorAxisLength
+ -2.8403 * (normalized) MinorAxisLength
+ 0.2919 * (normalized) AspectRation
+ 0.3145 * (normalized) Eccentricity
+ -1.8218 * (normalized) ConvexArea
+ -2.8145 * (normalized) EquivDiameter
+ -0.1978 * (normalized) Extent
+ 0.945 * (normalized) Solidity
+ 2.7902 * (normalized) roundness
+ -0.389 * (normalized) Compactness
+ 4.2802 * (normalized) ShapeFactor1
+ 2.4236 * (normalized) ShapeFactor2
+ -0.4014 * (normalized) ShapeFactor3
+ -0.6984 * (normalized) ShapeFactor4
- 2.3337
Number of kernel evaluations: 10246 (65.899% cached)
Classifier for classes: BOMBAY, CALI
BinarySMO
Machine linear: showing attribute weights, not support vectors.
Status
OK
Log
x 0

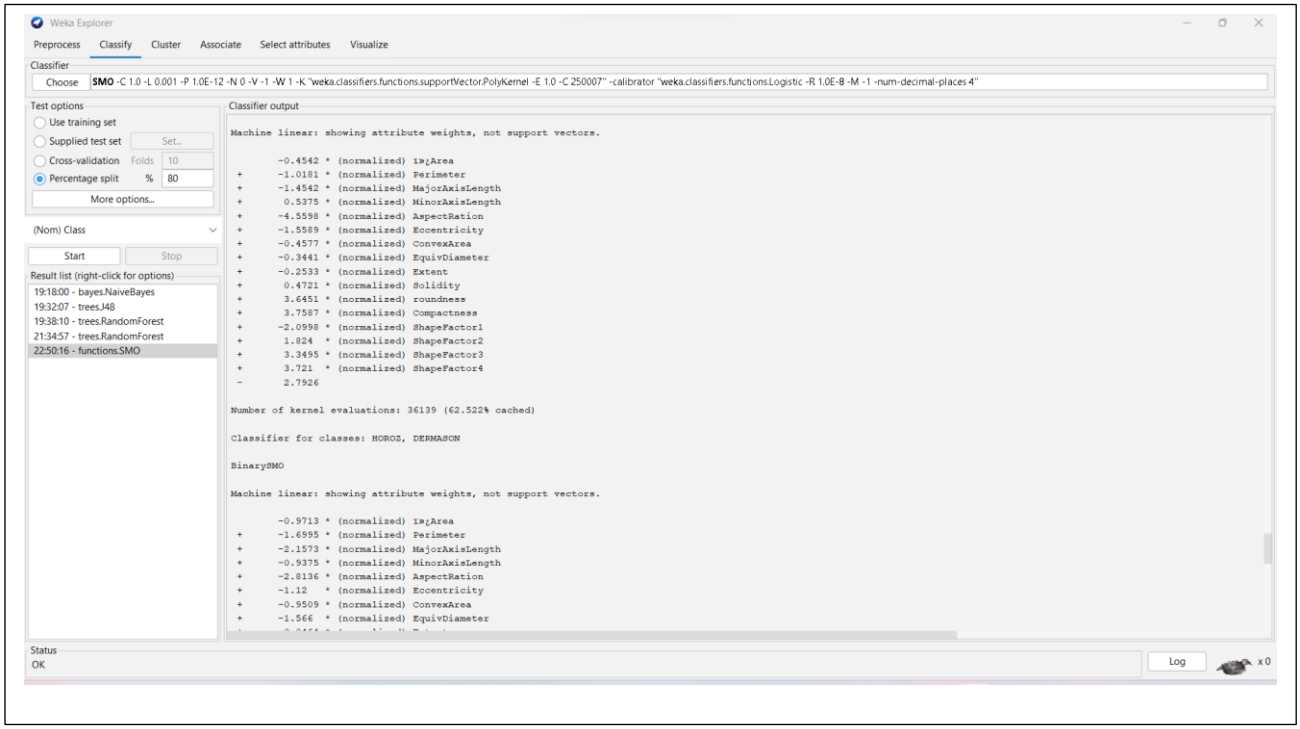

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
Machine linear: showing attribute weights, not support vectors.
-1.7298 * (normalized) i»¿Area
+ -1.4417 * (normalized) Perimeter
+ -1.1128 * (normalized) MajorAxisLength
+ -1.9572 * (normalized) MinorAxisLength
+ 1.087 * (normalized) AspectRation
+ 0.6568 * (normalized) Eccentricity
+ -1.6937 * (normalized) ConvexArea
+ -1.6204 * (normalized) EquivDiameter
+ -0.2321 * (normalized) Extent
+ -0.0234 * (normalized) Solidity
+ -0.3681 * (normalized) roundness
+ -1.0578 * (normalized) Compactness
+ 1.1934 * (normalized) ShapeFactor1
+ 0.0827 * (normalized) ShapeFactor2
+ -1.016 * (normalized) ShapeFactor3
+ -0.0643 * (normalized) ShapeFactor4
+ 4.8775
Number of kernel evaluations: 1741 (70.641% cached)
Classifier for classes: BOMBAY, HOROZ
BinarySMO
Machine linear: showing attribute weights, not support vectors.
-1.033 * (normalized) i»¿Area
+ -0.9129 * (normalized) Perimeter
+ -0.8317 * (normalized) MajorAxisLength
+ -1.191 * (normalized) MinorAxisLength
+ 0.4771 * (normalized) AspectRation
+ 0.2932 * (normalized) Eccentricity
+ -1.0117 * (normalized) ConvexArea
+ -1.0663 * (normalized) EquivDiameter
Status
OK
Log
x 0

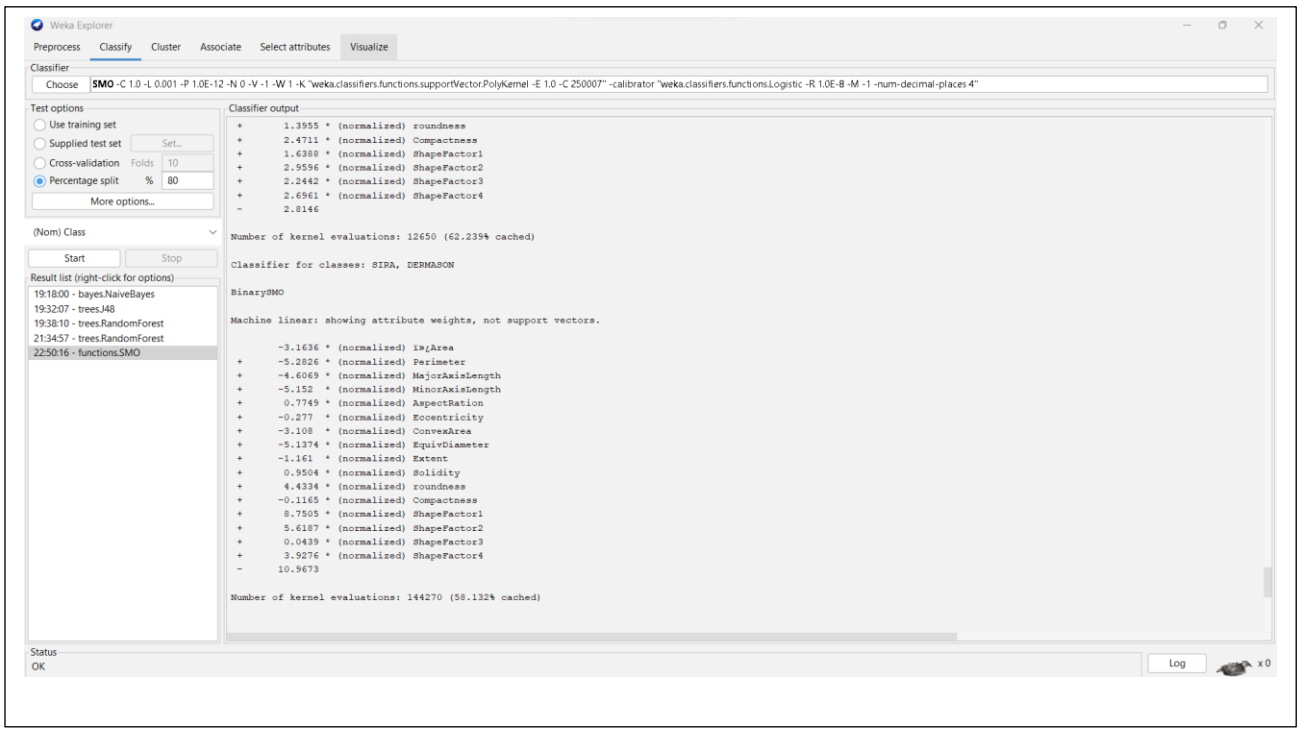

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
+ -1.0663 * (normalized) EquivDiameter
+ -0.3616 * (normalized) Extent
+ -0.1963 * (normalized) Solidity
+ -0.4446 * (normalized) roundness
+ -0.4725 * (normalized) Compactness
+ 0.882 * (normalized) ShapeFactor1
+ 0.2152 * (normalized) ShapeFactor2
+ -0.4554 * (normalized) ShapeFactor3
+ -0.125 * (normalized) ShapeFactor4
+ 3.0347
Number of kernel evaluations: 2506 (63.196% cached)
Classifier for classes: BOMBAY, SIRA
BinarySMO
Machine linear: showing attribute weights, not support vectors.
-0.8659 * (normalized) in¿Area
+ -0.9005 * (normalized) Perimeter
+ -0.7723 * (normalized) MajorAxisLength
+ -1.038 * (normalized) MinorAxisLength
+ 0.2671 * (normalized) AspectRation
+ 0.229 * (normalized) Eccentricity
+ -0.8444 * (normalized) ConvexArea
+ -0.9477 * (normalized) EquivDiameter
+ -0.0959 * (normalized) Extent
+ -0.0064 * (normalized) Solidity
+ 0.0791 * (normalized) roundness
+ -0.2809 * (normalized) Compactness
+ 0.8124 * (normalized) ShapeFactor1
+ 0.3393 * (normalized) ShapeFactor2
+ -0.2822 * (normalized) ShapeFactor3
+ 0.17 * (normalized) ShapeFactor4
+ 1.629
Status
OK
Log
x 0

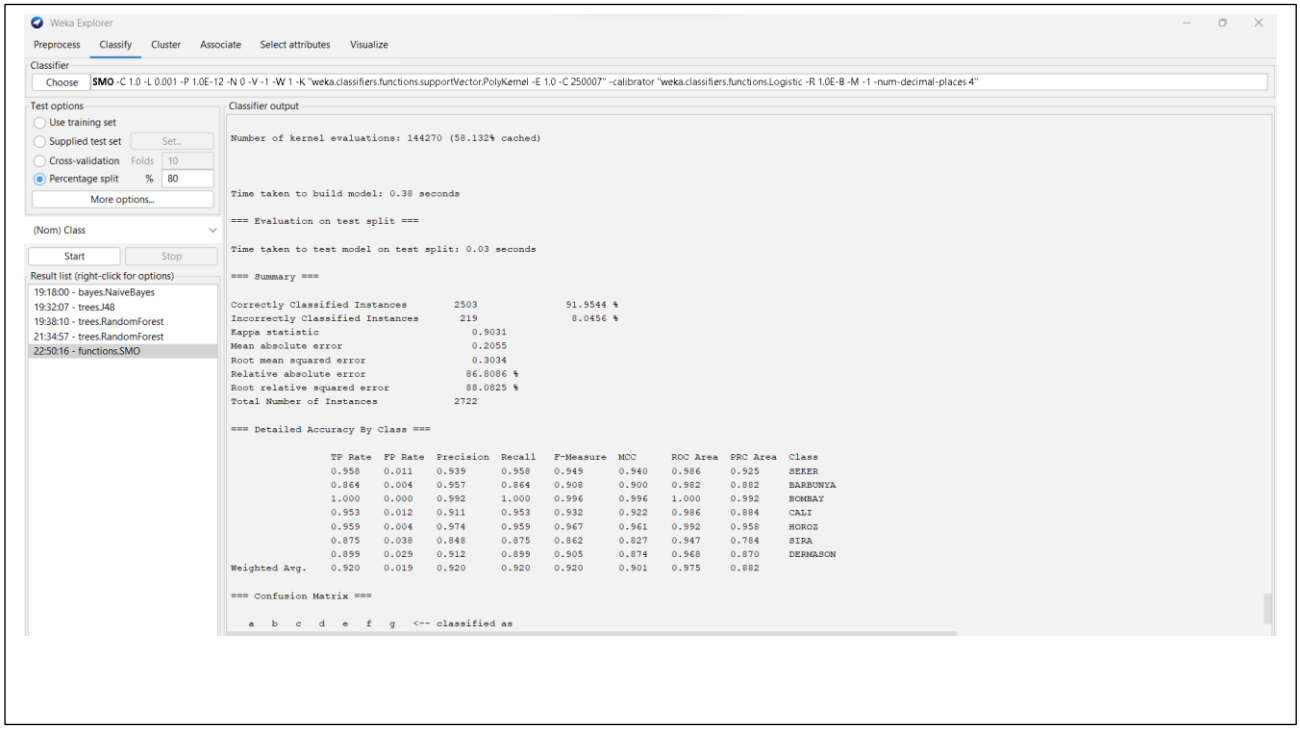

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
Number of kernel evaluations: 3975 (61.94% cached)
Classifier for classes: BOMBAY, DERMASON
BinarySMO
Machine linear: showing attribute weights, not support vectors.
-0.6353 * (normalized) in¿Area
+ -0.6932 * (normalized) Perimeter
+ -0.6239 * (normalized) MajorAxisLength
+ -0.822 * (normalized) MinorAxisLength
+ 0.0476 * (normalized) AspectRation
+ 0.0869 * (normalized) Eccentricity
+ -0.6184 * (normalized) ConvexArea
+ -0.7556 * (normalized) EquivDiameter
+ -0.0062 * (normalized) Extent
+ -0.0787 * (normalized) Solidity
+ -0.0424 * (normalized) roundness
+ -0.0637 * (normalized) Compactness
+ 0.681 * (normalized) ShapeFactor1
+ 0.5349 * (normalized) ShapeFactor2
+ -0.0707 * (normalized) ShapeFactor3
+ 0.1036 * (normalized) ShapeFactor4
+ 1.0637
Number of kernel evaluations: 2934 (64.341% cached)
Classifier for classes: CALI, HOROZ
BinarySMO
Machine linear: showing attribute weights, not support vectors.
-2.6436 * (normalized) in¿Area
+ -2.7141 * (normalized) Perimeter
+ -2.327 * (normalized) MajorAxisLength
Status
OK
Log
x 0

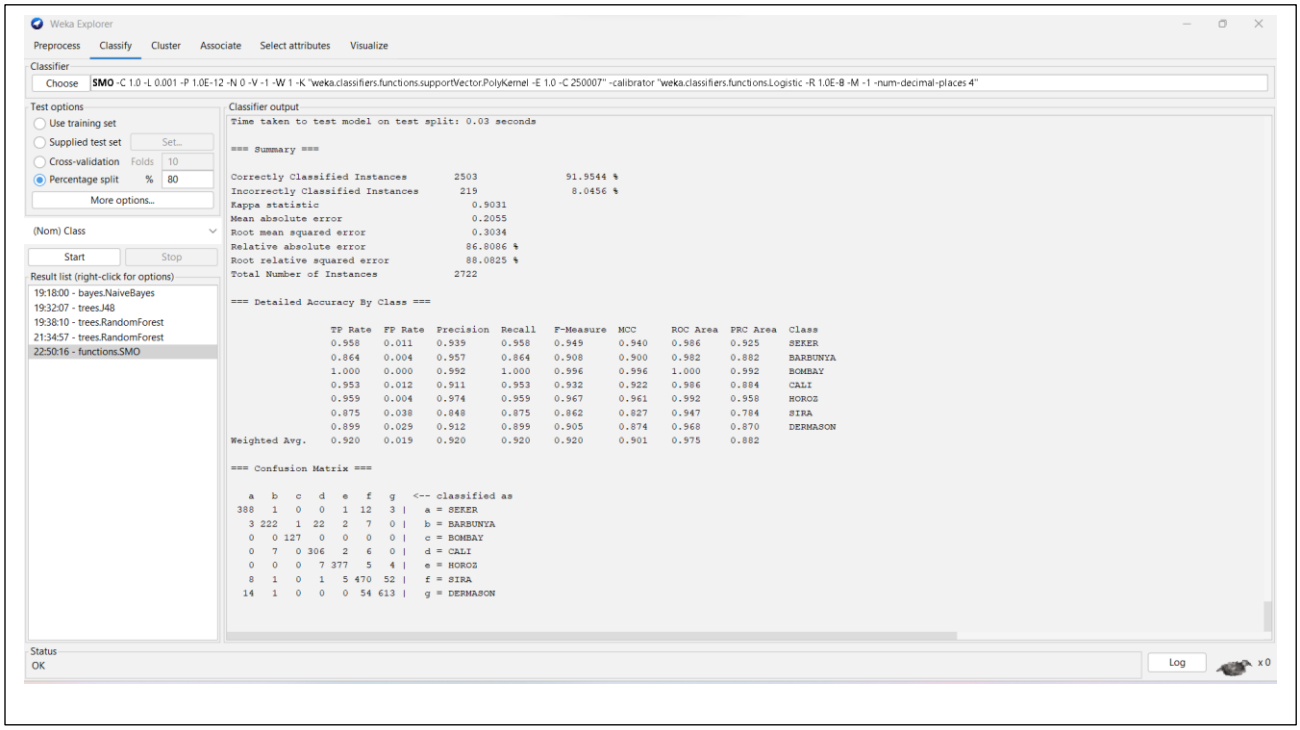

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
+ -2.327 * (normalized) MajorAxisLength
+ -3.9646 * (normalized) MinorAxisLength
+ 4.0018 * (normalized) AspectRation
+ 1.0671 * (normalized) Eccentricity
+ -2.6188 * (normalized) ConvexArea
+ -3.4471 * (normalized) EquivDiameter
+ -1.062 * (normalized) Extent
+ 0.8872 * (normalized) Solidity
+ -2.7668 * (normalized) roundness
+ -2.9117 * (normalized) Compactness
+ 5.1743 * (normalized) ShapeFactor1
+ 0.6383 * (normalized) ShapeFactor2
+ -2.5104 * (normalized) ShapeFactor3
+ -1.6395 * (normalized) ShapeFactor4
+ 3.4331
Number of kernel evaluations: 32064 (61.599% cached)
Classifier for classes: CALI, SIRA
BinarySMO
Machine linear: showing attribute weights, not support vectors.
-3.1239 * (normalized) I»¿Area
+ -4.3905 * (normalized) Perimeter
+ -3.9985 * (normalized) MajorAxisLength
+ -4.0551 * (normalized) MinorAxisLength
+ 0.0554 * (normalized) AspectRation
+ 0.3191 * (normalized) Eccentricity
+ -3.0636 * (normalized) ConvexArea
+ -4.2107 * (normalized) EquivDiameter
+ -1.1927 * (normalized) Extent
+ 0.3809 * (normalized) Solidity
+ 2.2021 * (normalized) roundness
+ -0.1448 * (normalized) Compactness
+ 5.0306 * (normalized) ShapeFactor1
Status
OK
Log
x 0

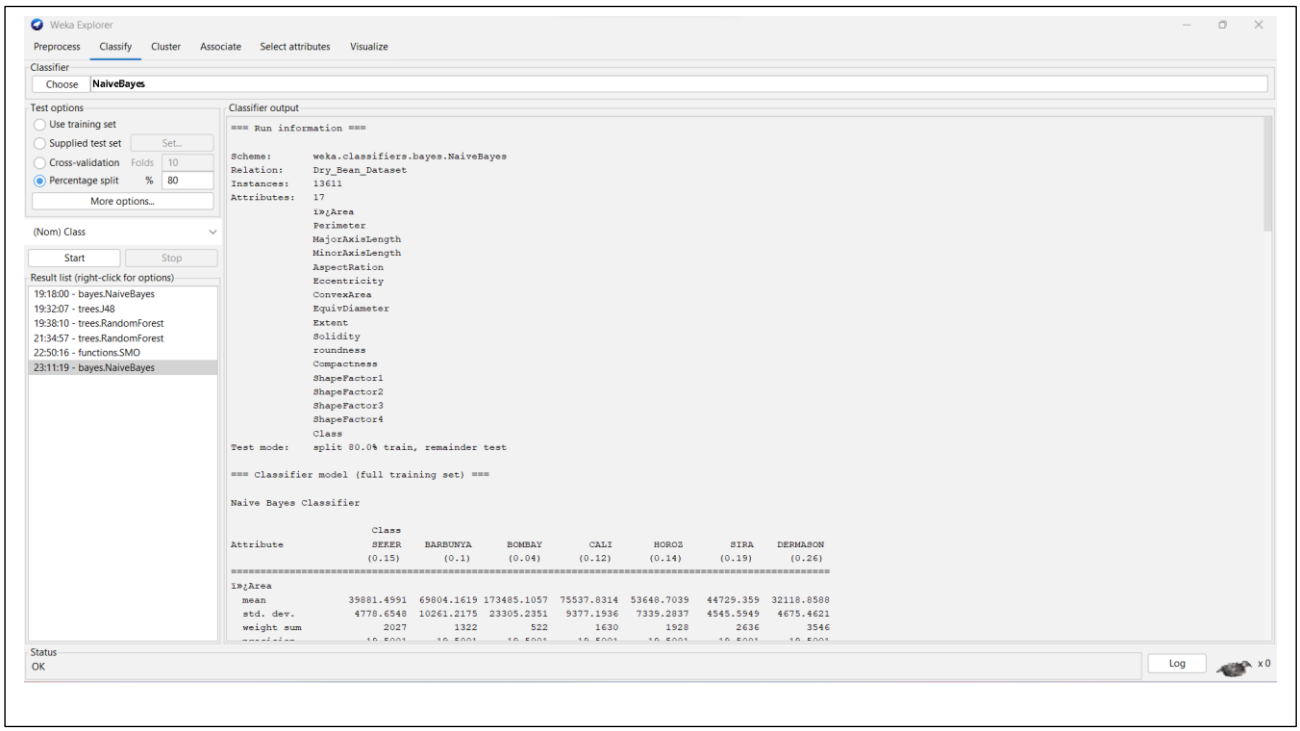

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
+ 2.5745 * (normalized) ShapeFactor2
+ -0.206 * (normalized) ShapeFactor3
+ 1.627 * (normalized) ShapeFactor4
- 0.0594
Number of kernel evaluations: 22100 (61.862% cached)
Classifier for classes: CALI, DERMASON
BinarySMO
Machine linear: showing attribute weights, not support vectors.
-1.6261 * (normalized) I»¿Area
+ -2.3802 * (normalized) Perimeter
+ -2.3373 * (normalized) MajorAxisLength
+ -2.2592 * (normalized) MinorAxisLength
+ -0.2797 * (normalized) AspectRation
+ 0.0656 * (normalized) Eccentricity
+ -1.591 * (normalized) ConvexArea
+ -2.4026 * (normalized) EquivDiameter
+ -0.308 * (normalized) Extent
+ 0.0359 * (normalized) Solidity
+ 0.7706 * (normalized) roundness
+ 0.1777 * (normalized) Compactness
+ 3.3295 * (normalized) ShapeFactor1
+ 1.9959 * (normalized) ShapeFactor2
+ 0.1196 * (normalized) ShapeFactor3
+ 0.9178 * (normalized) ShapeFactor4
- 1.0088
Number of kernel evaluations: 6345 (71.992% cached)
Classifier for classes: HOROZ, SIRA
BinarySMO
Status
OK
Log
x 0

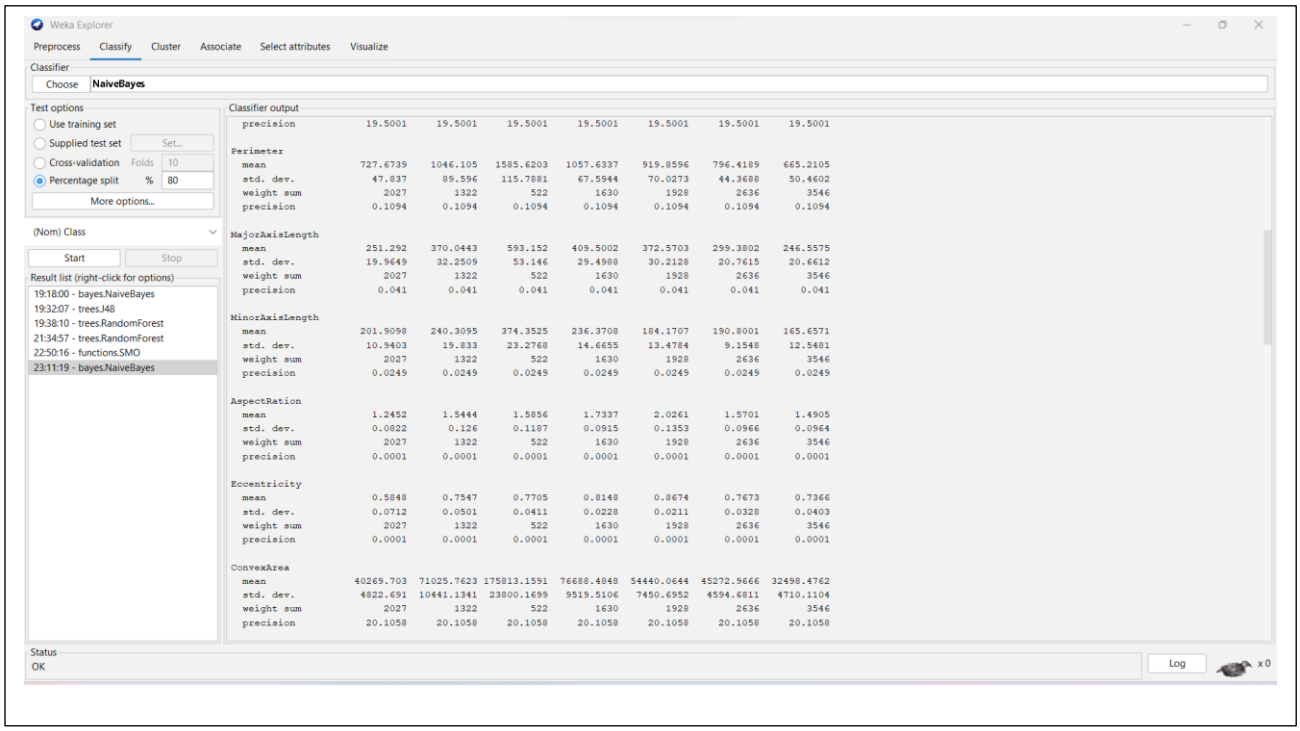

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
Machine linear: showing attribute weights, not support vectors.
−0.4542 * (normalized) 1»¿Area
+ −1.0181 * (normalized) Perimeter
+ −1.4542 * (normalized) MajorAxisLength
+ 0.5375 * (normalized) MinorAxisLength
+ −4.5598 * (normalized) AspectRation
+ −1.5589 * (normalized) Eccentricity
+ −0.4577 * (normalized) ConvexArea
+ −0.3441 * (normalized) EquivDiameter
+ −0.2533 * (normalized) Extent
+ 0.4721 * (normalized) Solidity
+ 3.6451 * (normalized) roundness
+ 3.7587 * (normalized) Compactness
+ −2.0998 * (normalized) ShapeFactor1
+ 1.824 * (normalized) ShapeFactor2
+ 3.3455 * (normalized) ShapeFactor3
+ 3.721 * (normalized) ShapeFactor4
− 2.7926
Number of kernel evaluations: 36139 (62.522% cached)
Classifier for classes: HOROZ, DERMASON
BinarySMO
Machine linear: showing attribute weights, not support vectors.
−0.9713 * (normalized) 1»¿Area
+ −1.6995 * (normalized) Perimeter
+ −2.1573 * (normalized) MajorAxisLength
+ −0.9375 * (normalized) MinorAxisLength
+ −2.8136 * (normalized) AspectRation
+ −1.12 * (normalized) Eccentricity
+ −0.9509 * (normalized) ConvexArea
+ −1.566 * (normalized) EquivDiameter
Status
OK
Log
x 0

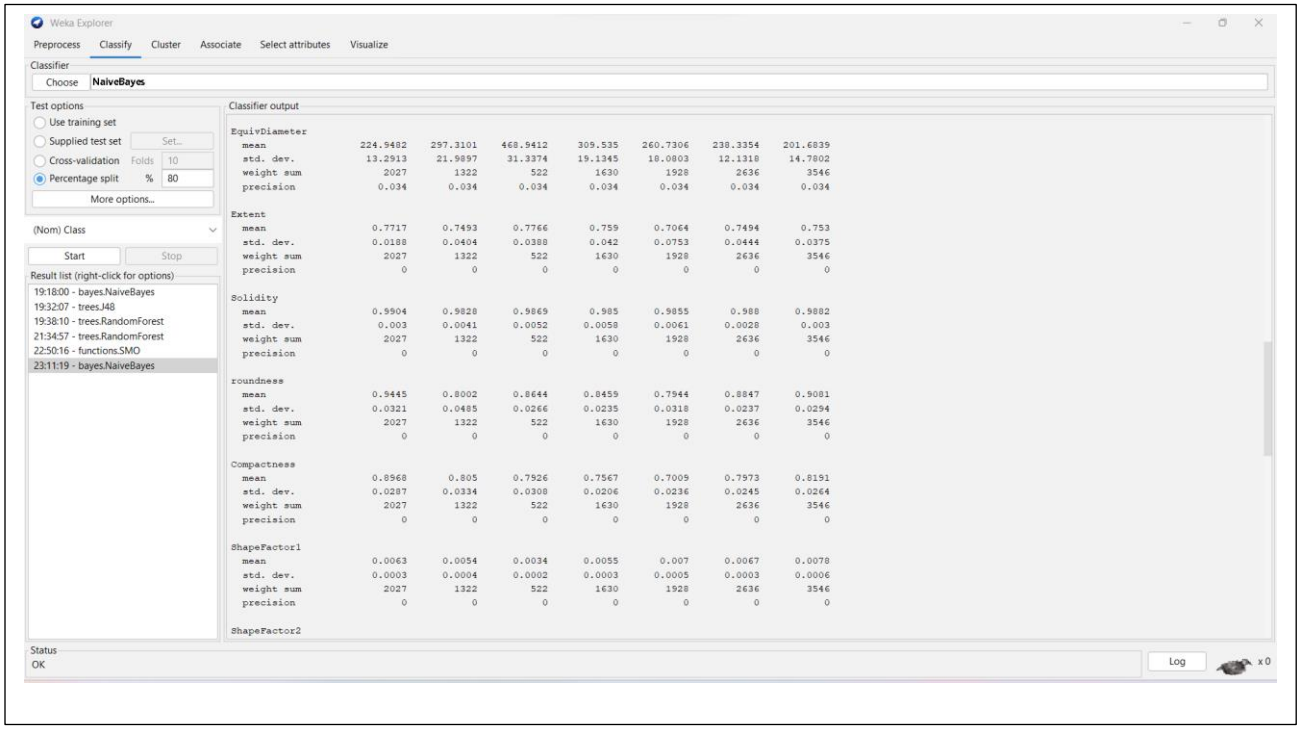

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
+ 1.3955 * (normalized) roundness
+ 2.4711 * (normalized) Compactness
+ 1.6388 * (normalized) ShapeFactor1
+ 2.9596 * (normalized) ShapeFactor2
+ 2.2442 * (normalized) ShapeFactor3
+ 2.6961 * (normalized) ShapeFactor4
− 2.8146
Number of kernel evaluations: 12650 (62.239% cached)
Classifier for classes: SIRA, DERMASON
BinarySMO
Machine linear: showing attribute weights, not support vectors.
−3.1636 * (normalized) 1»¿Area
+ −5.2826 * (normalized) Perimeter
+ −4.6069 * (normalized) MajorAxisLength
+ −5.152 * (normalized) MinorAxisLength
+ 0.7749 * (normalized) AspectRation
+ −0.277 * (normalized) Eccentricity
+ −3.108 * (normalized) ConvexArea
+ −5.1374 * (normalized) EquivDiameter
+ −1.161 * (normalized) Extent
+ 0.9504 * (normalized) Solidity
+ 4.4334 * (normalized) roundness
+ −0.1165 * (normalized) Compactness
+ 8.7505 * (normalized) ShapeFactor1
+ 5.6187 * (normalized) ShapeFactor2
+ 0.0439 * (normalized) ShapeFactor3
+ 3.9276 * (normalized) ShapeFactor4
− 10.9673
Number of kernel evaluations: 144270 (58.132% cached)
Status
OK
Log
x 0

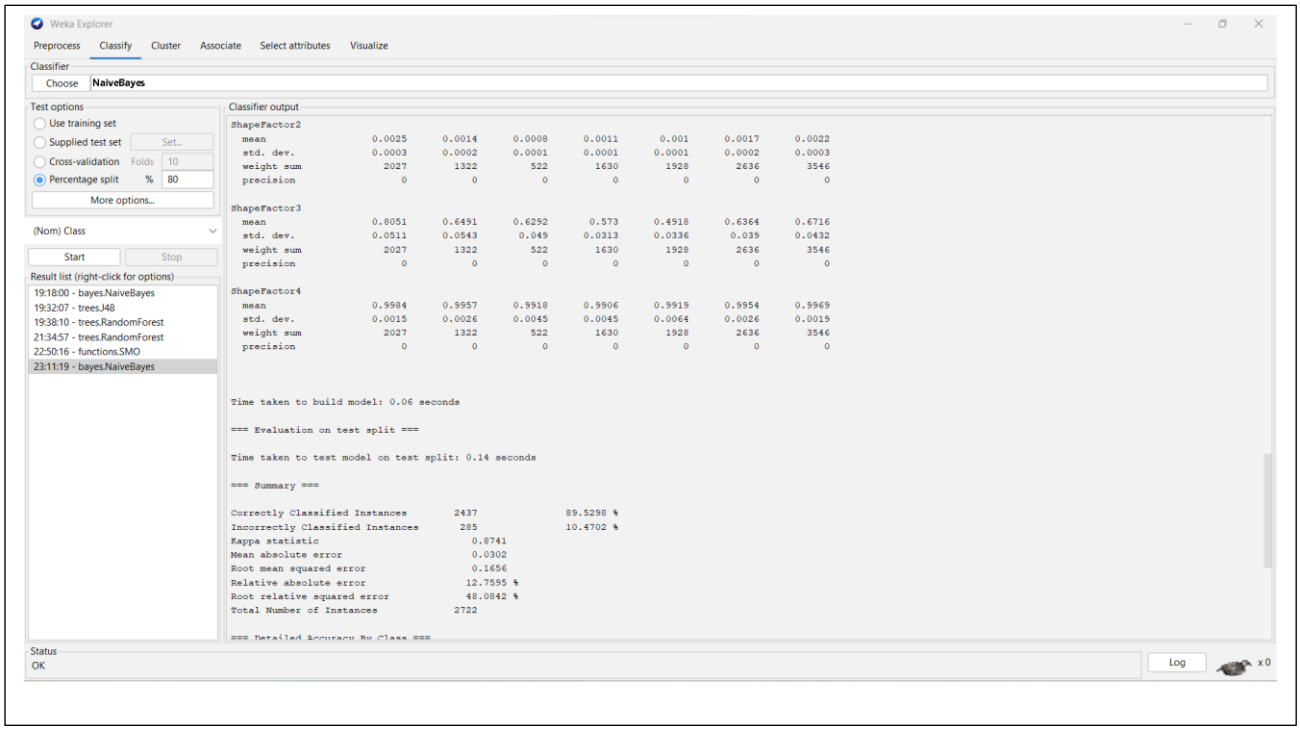

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
Number of kernel evaluations: 144270 (58.132% cached)
Time taken to build model: 0.38 seconds
=== Evaluation on test split ===
Time taken to test model on test split: 0.03 seconds
=== Summary ===
Correctly Classified Instances 2503 91.9544 %
Incorrectly Classified Instances 219 8.0456 %
Kappa statistic 0.9031
Mean absolute error 0.2055
Root mean squared error 0.3034
Relative absolute error 86.8086 %
Root relative squared error 88.0825 %
Total Number of Instances 2722
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.958 0.011 0.939 0.958 0.949 0.940 0.986 0.925 SEKER
0.864 0.004 0.957 0.864 0.908 0.900 0.982 0.882 BARBUNYA
1.000 0.000 0.992 1.000 0.996 0.996 1.000 0.992 BOMBAY
0.953 0.012 0.911 0.953 0.932 0.922 0.986 0.884 CALI
0.959 0.004 0.974 0.959 0.967 0.961 0.992 0.958 HOROZ
0.875 0.038 0.848 0.875 0.862 0.827 0.947 0.784 SIRA
0.899 0.029 0.912 0.899 0.905 0.874 0.968 0.870 DERMASON
Weighted Avg. 0.920 0.019 0.920 0.920 0.920 0.901 0.975 0.882
=== Confusion Matrix ===
a b c d e f g <-- classified as

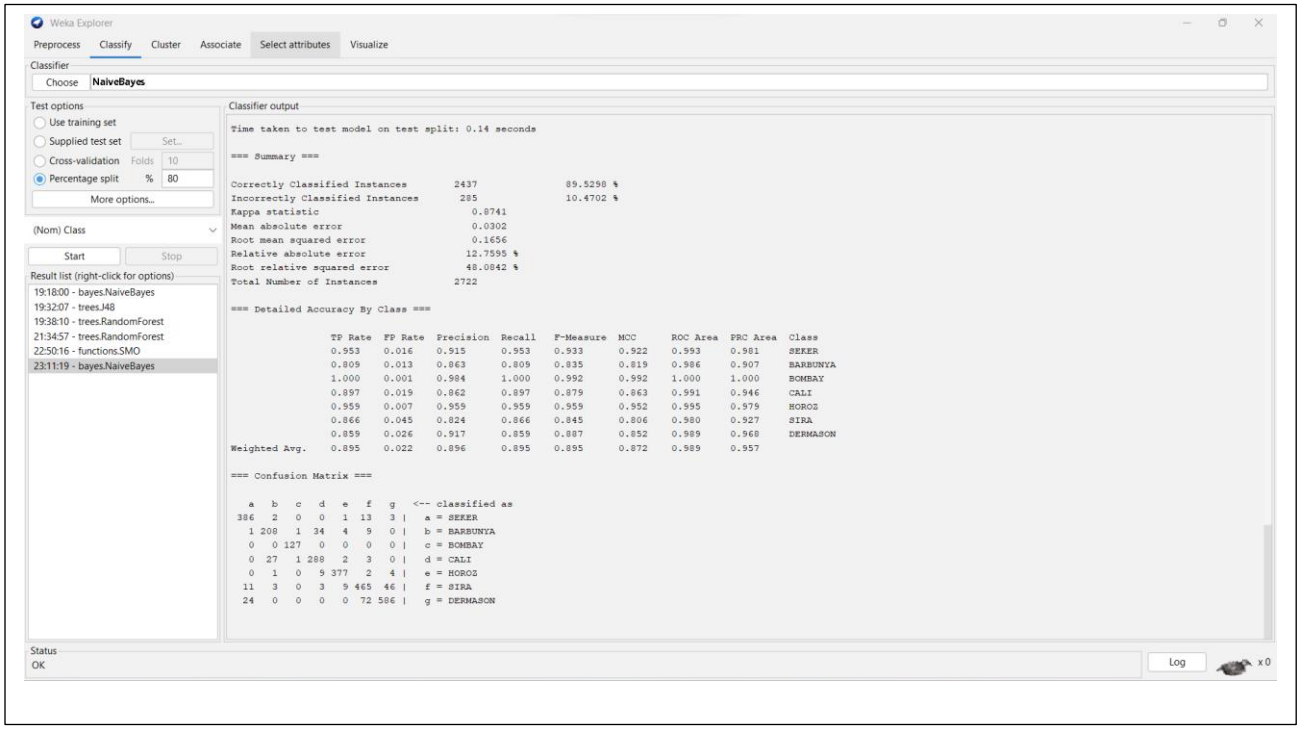

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.PolyKernel -E 1.0 -C 250007" -calibrator "weka.classifiers.functions.Logistic -R 1.0E-8 -M -1 -num-decimal-places 4"
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
Classifier output
Time taken to test model on test split: 0.03 seconds
=== Summary ===
Correctly Classified Instances 2503 91.9544 %
Incorrectly Classified Instances 219 8.0456 %
Kappa statistic 0.9031
Mean absolute error 0.2055
Root mean squared error 0.3034
Relative absolute error 86.8086 %
Root relative squared error 88.0825 %
Total Number of Instances 2722
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class
0.958 0.011 0.939 0.958 0.949 0.940 0.986 0.925 SEKER
0.864 0.004 0.957 0.864 0.908 0.900 0.982 0.882 BARBUNYA
1.000 0.000 0.992 1.000 0.996 0.996 1.000 0.992 BOMBAY
0.953 0.012 0.911 0.953 0.932 0.922 0.986 0.884 CALI
0.959 0.004 0.974 0.959 0.967 0.961 0.992 0.958 HOROZ
0.875 0.038 0.848 0.875 0.862 0.827 0.947 0.784 SIRA
0.899 0.029 0.912 0.899 0.905 0.874 0.968 0.870 DERMASON
Weighted Avg. 0.920 0.019 0.920 0.920 0.920 0.901 0.975 0.882
=== Confusion Matrix ===
a b c d e f g <-- classified as
388 1 0 0 1 12 3 | a = SEKER
3 222 1 22 2 7 0 | b = BARBUNYA
0 0 127 0 0 0 0 | c = BOMBAY
0 7 0 306 2 6 0 | d = CALI
0 0 0 7 377 5 4 | e = HOROZ
8 1 0 1 5 470 52 | f = SIRA
14 1 0 0 0 54 613 | g = DERMASON
Status
OK
Log
x 0


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
NaiveBayes
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
23:11:19 - bayes.NaiveBayes
Classifier output
=== Run information ===
Scheme: weka.classifiers.bayes.NaiveBayes
Relation: Dry_Bean_Dataset
Instances: 13611
Attributes: 17
Area
Perimeter
MajorAxisLength
MinorAxisLength
AspectRation
Eccentricity
ConvexArea
EquivDiameter
Extent
Solidity
roundness
Compactness
ShapeFactor1
ShapeFactor2
ShapeFactor3
ShapeFactor4
Class
Test mode: split 80.0% train, remainder test
=== Classifier model (full training set) ===
Naive Bayes Classifier
Class
Attribute SEKER BARBUNYA BOMBAY CALI HOROZ SIRA DERMASON
(0.15) (0.1) (0.04) (0.12) (0.14) (0.19) (0.26)
Area
mean 39881.4991 69804.1619 173485.1057 75537.8314 53648.7039 44729.359 32118.8588
std. dev. 4778.6548 10261.2175 23305.2351 9377.1936 7339.2837 4545.5949 4675.4621
weight sum 2027 1322 522 1630 1928 2636 3546
Status
OK
Log
x 0

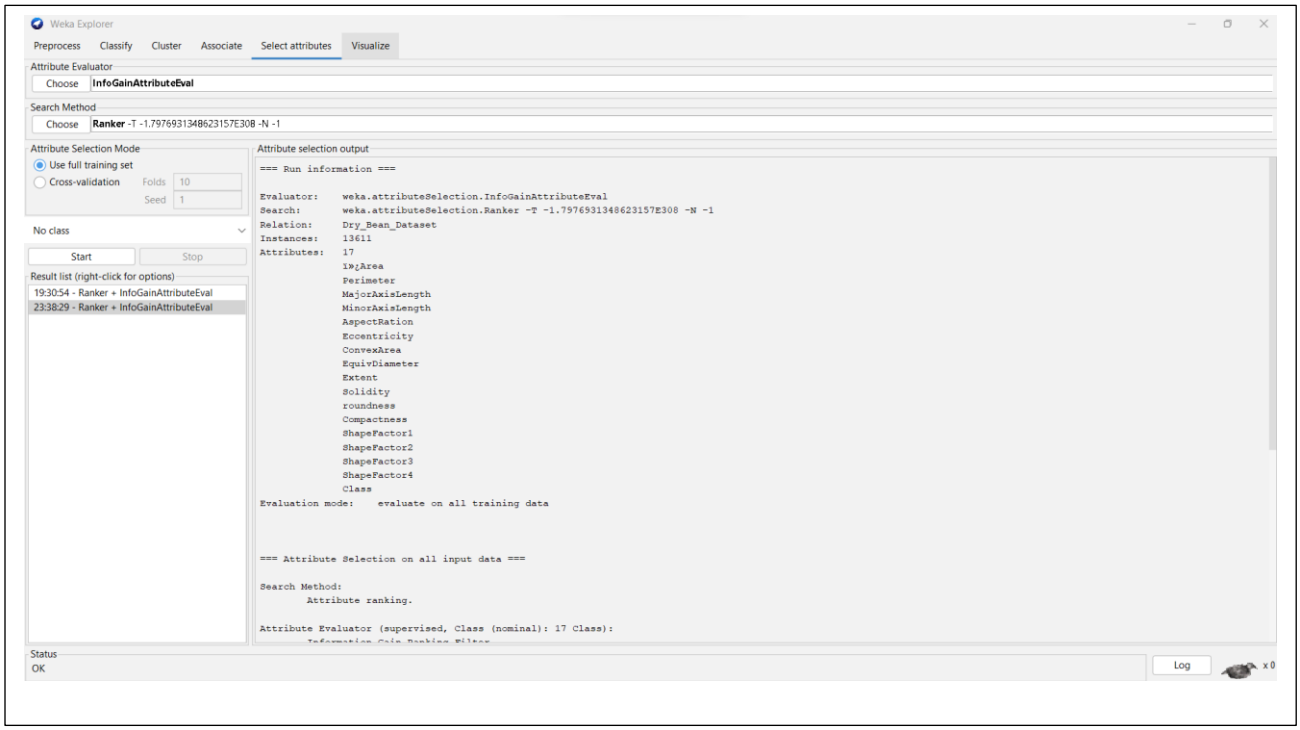

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
NaiveBayes
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
80
More options...
(Nom) Class
Start
Stop
Result list (right-click for options)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
23:11:19 - bayes.NaiveBayes
Classifier output
precision 19.5001 19.5001 19.5001 19.5001 19.5001 19.5001 19.5001
Perimeter
mean 727.6739 1046.105 1585.6203 1057.6337 919.8596 796.4189 665.2105
std. dev. 47.837 89.596 115.7881 67.5944 70.0273 44.3688 50.4602
weight sum 2027 1322 522 1630 1928 2636 3546
precision 0.1094 0.1094 0.1094 0.1094 0.1094 0.1094 0.1094
MajorAxisLength
mean 251.292 370.0443 593.152 409.5002 372.5703 299.3802 246.5575
std. dev. 19.9649 32.2509 53.146 29.4988 30.2128 20.7615 20.6612
weight sum 2027 1322 522 1630 1928 2636 3546
precision 0.041 0.041 0.041 0.041 0.041 0.041 0.041
MinorAxisLength
mean 201.9098 240.3095 374.3525 236.3708 184.1707 190.8001 165.6571
std. dev. 10.9403 19.833 23.2768 14.6655 13.4784 9.1548 12.5481
weight sum 2027 1322 522 1630 1928 2636 3546
precision 0.0249 0.0249 0.0249 0.0249 0.0249 0.0249 0.0249
AspectRation
mean 1.2452 1.5444 1.5856 1.7337 2.0261 1.5701 1.4905
std. dev. 0.0822 0.126 0.1187 0.0915 0.1353 0.0966 0.0964
weight sum 2027 1322 522 1630 1928 2636 3546
precision 0.0001 0.0001 0.0001 0.0001 0.0001 0.0001 0.0001
Eccentricity
mean 0.5848 0.7547 0.7705 0.8148 0.8674 0.7673 0.7366
std. dev. 0.0712 0.0501 0.0411 0.0228 0.0211 0.0328 0.0403
weight sum 2027 1322 522 1630 1928 2636 3546
precision 0.0001 0.0001 0.0001 0.0001 0.0001 0.0001 0.0001
ConvexArea
mean 40269.703 71025.7623 175813.1591 76688.4848 54440.0644 45272.9666 32498.4762
std. dev. 4822.691 10441.1341 23800.1699 9519.5106 7450.6952 4594.6811 4710.1104
weight sum 2027 1322 522 1630 1928 2636 3546
precision 20.1058 20.1058 20.1058 20.1058 20.1058 20.1058 20.1058
Status
OK
Log
x 0

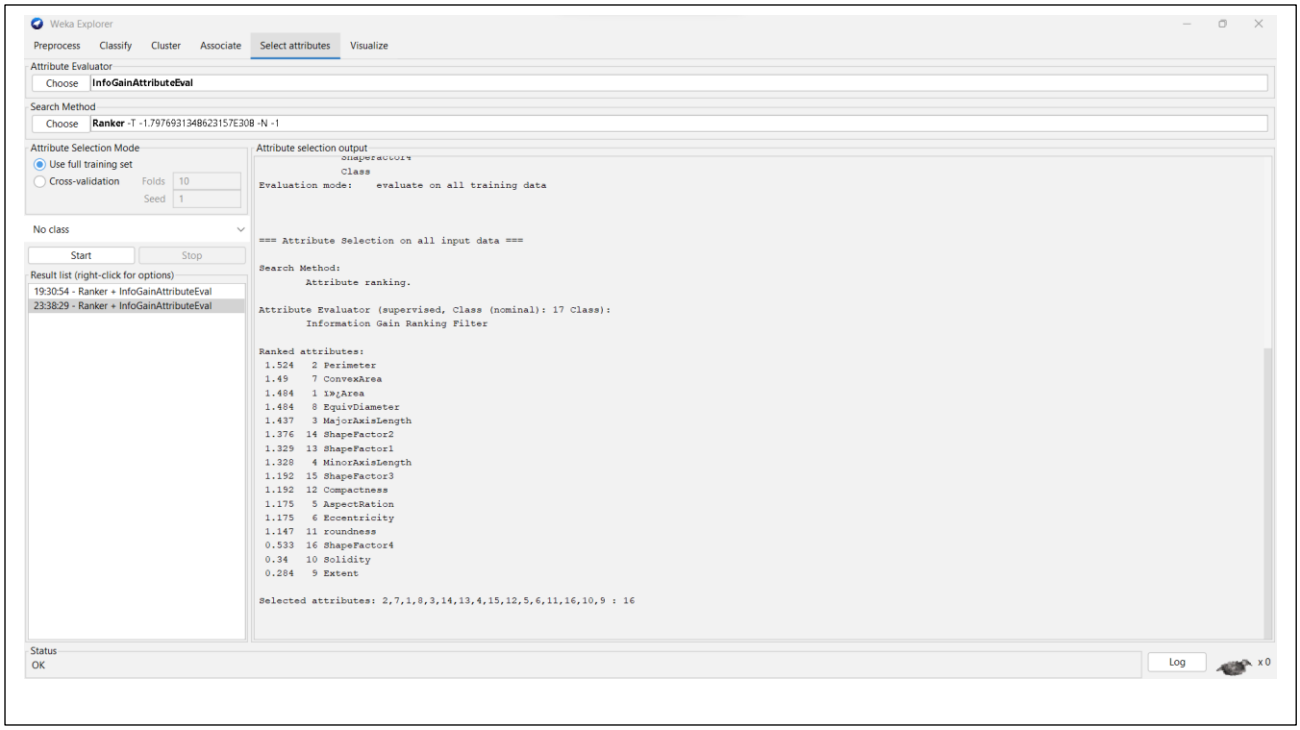

Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
NaiveBayes
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
80
More options...
(Nom) Class
Start
Stop
Result list (right-click for options)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
23:11:19 - bayes.NaiveBayes
Classifier output
EquivDiameter
mean 224.9482 297.3101 468.9412 309.535 260.7306 238.3354 201.6839
std. dev. 13.2913 21.9897 31.3374 19.1345 18.0803 12.1318 14.7802
weight sum 2027 1322 522 1630 1928 2636 3546
precision 0.034 0.034 0.034 0.034 0.034 0.034 0.034
Extent
mean 0.7717 0.7493 0.7766 0.759 0.7064 0.7494 0.753
std. dev. 0.0188 0.0404 0.0388 0.042 0.0753 0.0444 0.0375
weight sum 2027 1322 522 1630 1928 2636 3546
precision 0 0 0 0 0 0 0
Solidity
mean 0.9904 0.9828 0.9869 0.985 0.9855 0.988 0.9882
std. dev. 0.003 0.0041 0.0052 0.0058 0.0061 0.0028 0.003
weight sum 2027 1322 522 1630 1928 2636 3546
precision 0 0 0 0 0 0 0
roundness
mean 0.9445 0.8002 0.8644 0.8459 0.7944 0.8847 0.9081
std. dev. 0.0321 0.0485 0.0266 0.0235 0.0318 0.0237 0.0294
weight sum 2027 1322 522 1630 1928 2636 3546
precision 0 0 0 0 0 0 0
Compactness
mean 0.8568 0.805 0.7526 0.7567 0.7009 0.7973 0.8191
std. dev. 0.0287 0.0334 0.0308 0.0206 0.0236 0.0245 0.0264
weight sum 2027 1322 522 1630 1928 2636 3546
precision 0 0 0 0 0 0 0
ShapeFactor1
mean 0.0063 0.0054 0.0034 0.0055 0.007 0.0067 0.0078
std. dev. 0.0003 0.0004 0.0002 0.0003 0.0005 0.0003 0.0006
weight sum 2027 1322 522 1630 1928 2636 3546
precision 0 0 0 0 0 0 0
ShapeFactor2
Status
OK
Log
x 0


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
NaiveBayes
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)
19:18:00 - bayes.NaiveBayes
19:32:07 - trees.J48
19:38:10 - trees.RandomForest
21:34:57 - trees.RandomForest
22:50:16 - functions.SMO
23:11:19 - bayes.NaiveBayes
Classifier output
ShapeFactor2
mean 0.0025 0.0014 0.0008 0.0011 0.001 0.0017 0.0022
std. dev. 0.0003 0.0002 0.0001 0.0001 0.0001 0.0002 0.0003
weight sum 2027 1322 522 1630 1928 2636 3546
precision 0 0 0 0 0 0 0
ShapeFactor3
mean 0.8051 0.6491 0.6292 0.573 0.4918 0.6364 0.6716
std. dev. 0.0511 0.0543 0.049 0.0313 0.0336 0.039 0.0432
weight sum 2027 1322 522 1630 1928 2636 3546
precision 0 0 0 0 0 0 0
ShapeFactor4
mean 0.9984 0.9957 0.9918 0.9906 0.9919 0.9954 0.9969
std. dev. 0.0015 0.0026 0.0045 0.0045 0.0064 0.0026 0.0019
weight sum 2027 1322 522 1630 1928 2636 3546
precision 0 0 0 0 0 0 0
Time taken to build model: 0.06 seconds
=== Evaluation on test split ===
Time taken to test model on test split: 0.14 seconds
=== Summary ===
Correctly Classified Instances 2437 89.5298 %
Incorrectly Classified Instances 285 10.4702 %
Kappa statistic 0.8741
Mean absolute error 0.0302
Root mean squared error 0.1656
Relative absolute error 12.7595 %
Root relative squared error 48.0842 %
Total Number of Instances 2722
=== Detailed Accuracy By Class ===
Status
OK
Log
x 0



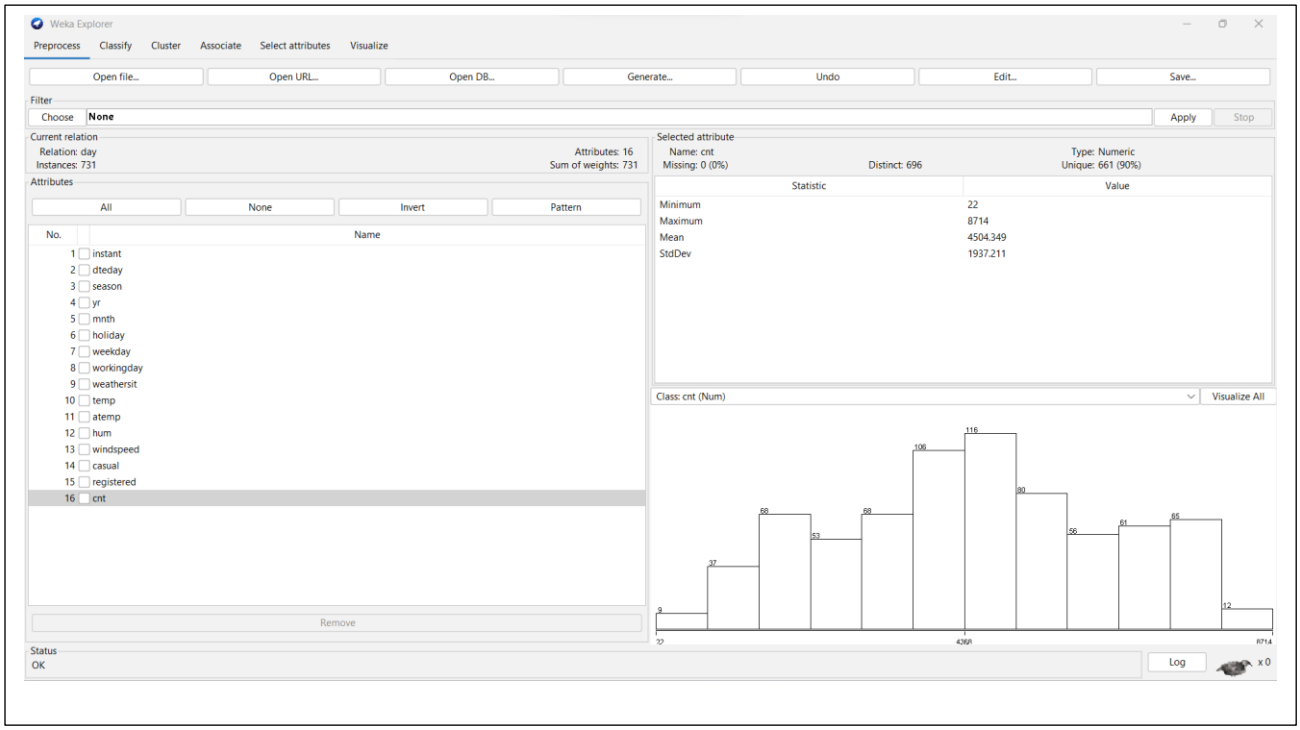
Feature Ranking

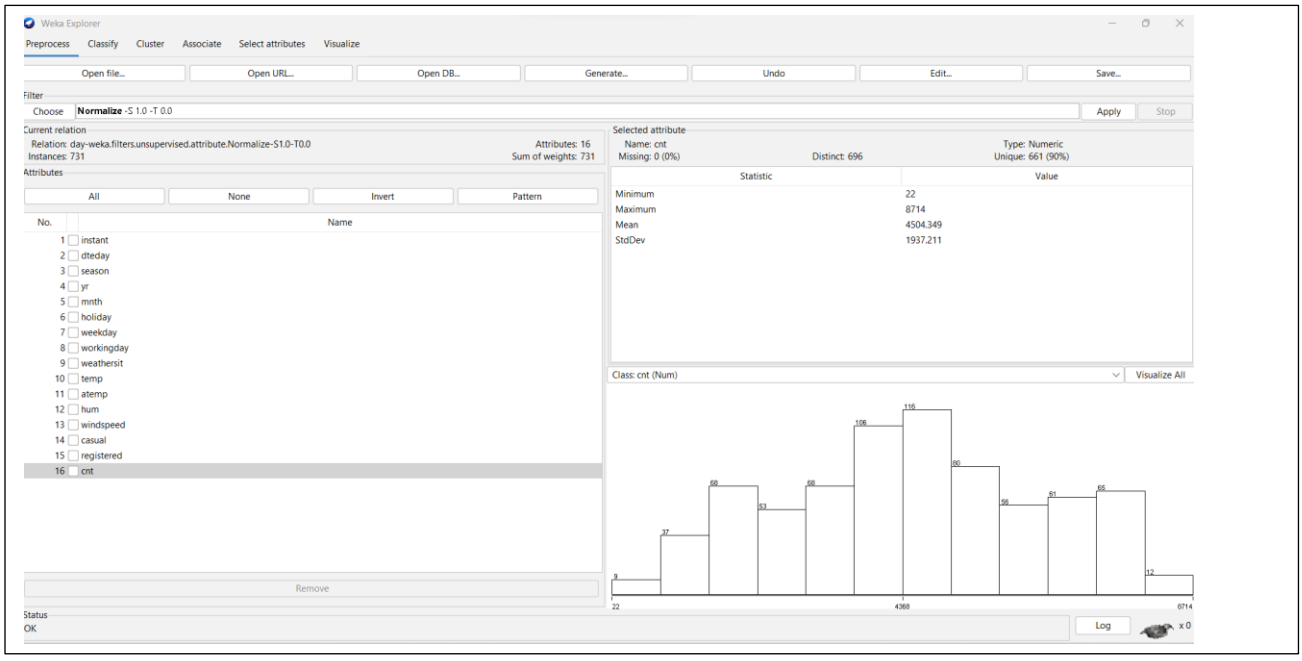


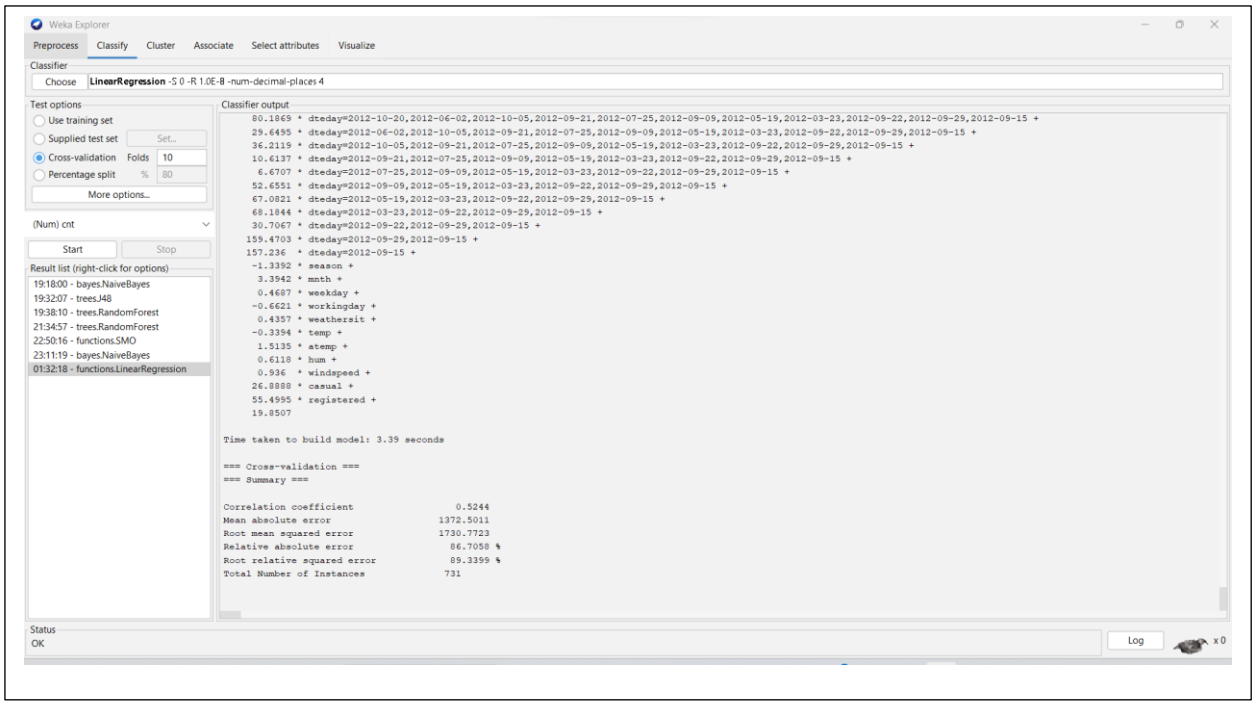


In summary, the choice of the best model depends on the balance between accuracy, interpretability, and computational efficiency. For the Dry Bean Dataset, RandomForest seems to offer the best trade-off, particularly in scenarios where accuracy is paramount and computational resources are not a major constraint.

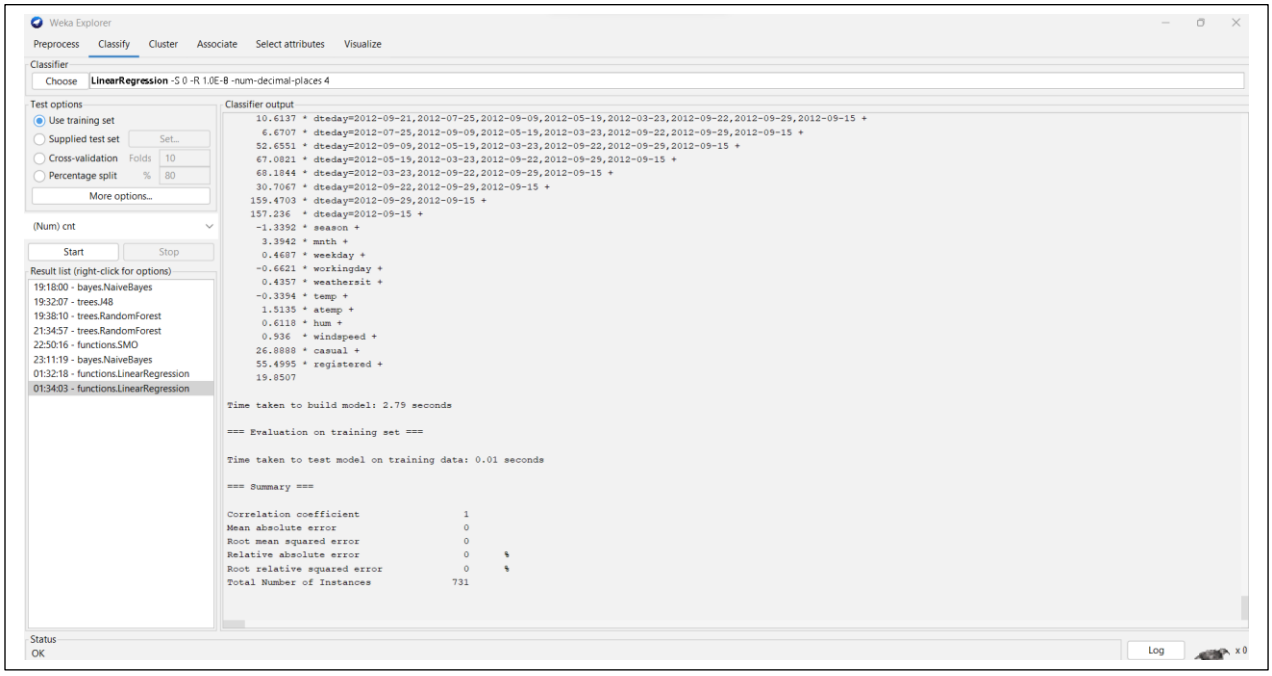
Regression Model



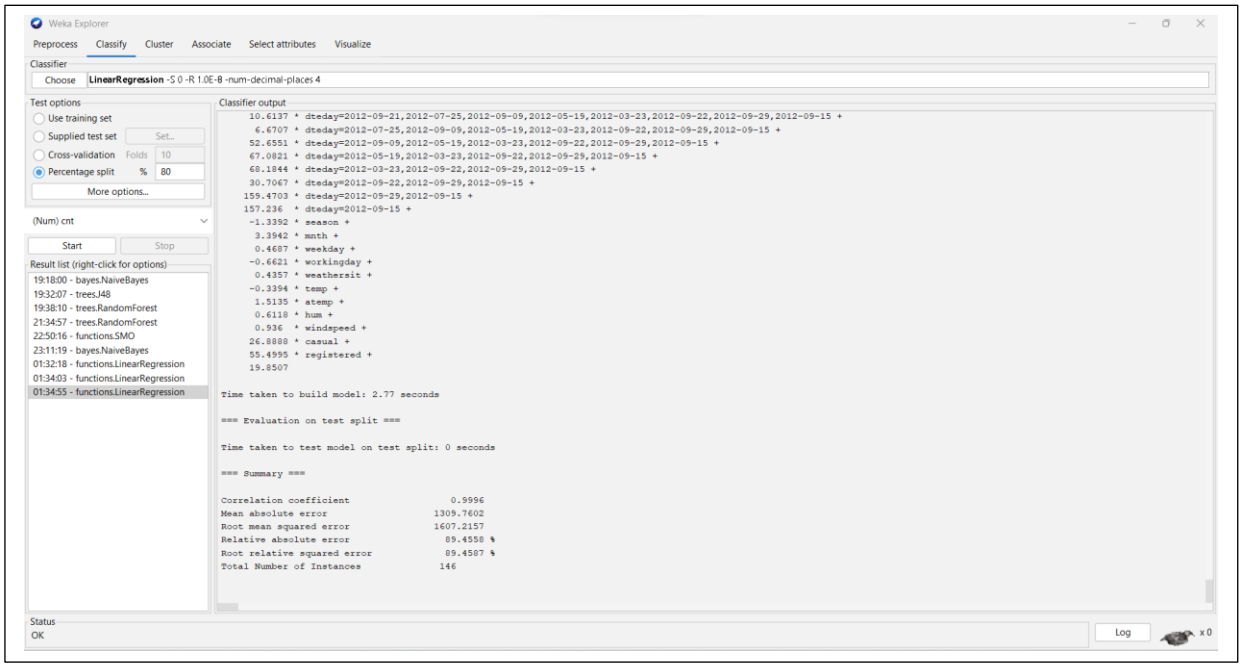




The model achieved a correlation coefficient of 0.5244, indicating a moderate linear relationship between predicted and actual values. However, the mean absolute error (MAE) and root mean squared error (RMSE) were relatively high at 1372.5011 and 1730.7723, respectively. These errors, along with the relative absolute error (RAE) and root relative squared error (RRSE) percentages nearing 87% and 89%, suggest that the model's predictions vary significantly from the actual values.



Remarkably, the training set model displayed perfect performance metrics on the training set, with a correlation coefficient of 1 and no errors. This suggests an overfitting scenario where the model has learned the training data intricacies to the extent that it perfectly predicts every instance but may struggle to generalize to unseen data.



On the test split, the model approached near-perfect performance with a correlation coefficient of 0.9996, closely mirroring the training set results. The MAE and RMSE were observed at 1309.7602 and 1607.2157, respectively, with RAE and RRSE percentages slightly higher than 89%. Although these metrics indicate a high degree of accuracy, the close resemblance to the training set performance again raises concerns about overfitting, particularly given the perfect metrics on the training data.