An Introduction to WEKA

a popular suite of machine learning software

Waikato Environment for Knowledge Analysis

Download and Install WEKA

- **Website: http://www.cs.waikato.ac.nz/~ml/weka/index.html**

- **SourceForge:**

http://sourceforge.net/projects/weka/

- **3.6 is the latest stable version**

Content

- **Intro and background**
- **Exploring WEKA**
  - Data Preparation
  - Creating Models/ Applying Algorithms
  - Evaluating Results

Available Data Mining Tools

COTs:

- IBM Intelligent Miner
- SAS Enterprise Miner
- Oracle ODM
- Microstrategy
- Microsoft DBMiner
- Pentaho
- Matlab
- Teradata

Open Source:

- WEKA
- KNIME
- Orange
- RapidMiner
- NLTK
- R
- Rattle

What is WEKA?

- **Waikato Environment for Knowledge Analysis**
  - WEKA is a data mining/machine learning application developed by Department of Computer Science, University of Waikato, New Zealand
  - WEKA is open source software in JAVA issued under the GNU General Public License
  - WEKA is a collection tools for data pre-processing, classification, regression, clustering, association, and visualization.
    - WEKA is a collection of machine learning algorithms for data mining tasks
  - WEKA is well-suited for developing new machine learning schemes
- **WEKA is a bird found only in New Zealand**

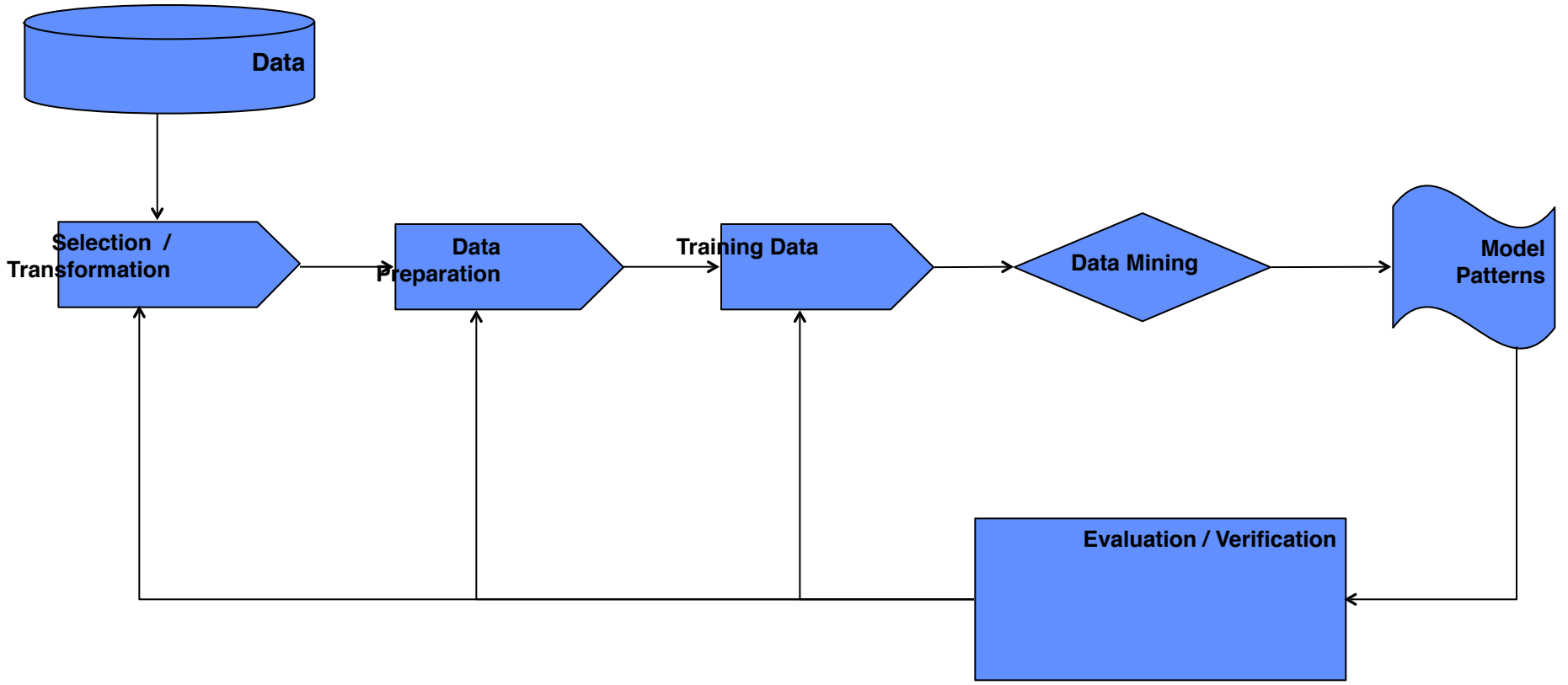
Advantages of Weka

- **Free availability**
  - Under the GNU General Public License
- **Portability**
  - Fully implemented in the Java programming language and thus runs on almost any modern computing platforms
    - Windows, Mac OS X and Linux
- **Comprehensive collection of data preprocessing and modeling techniques**
  - Supports standard data mining tasks: data preprocessing, clustering, classification, regression, visualization, and feature selection
- **Easy to use GUI**
- **Provides access to SQL databases**
  - Using Java Database Connectivity and can process the result returned by a database query

Disadvantages of Weka

- **Sequence modeling is not covered by the algorithms included in the Weka distribution**
- **Not capable of multi-relational data mining**
- **Memory bound**

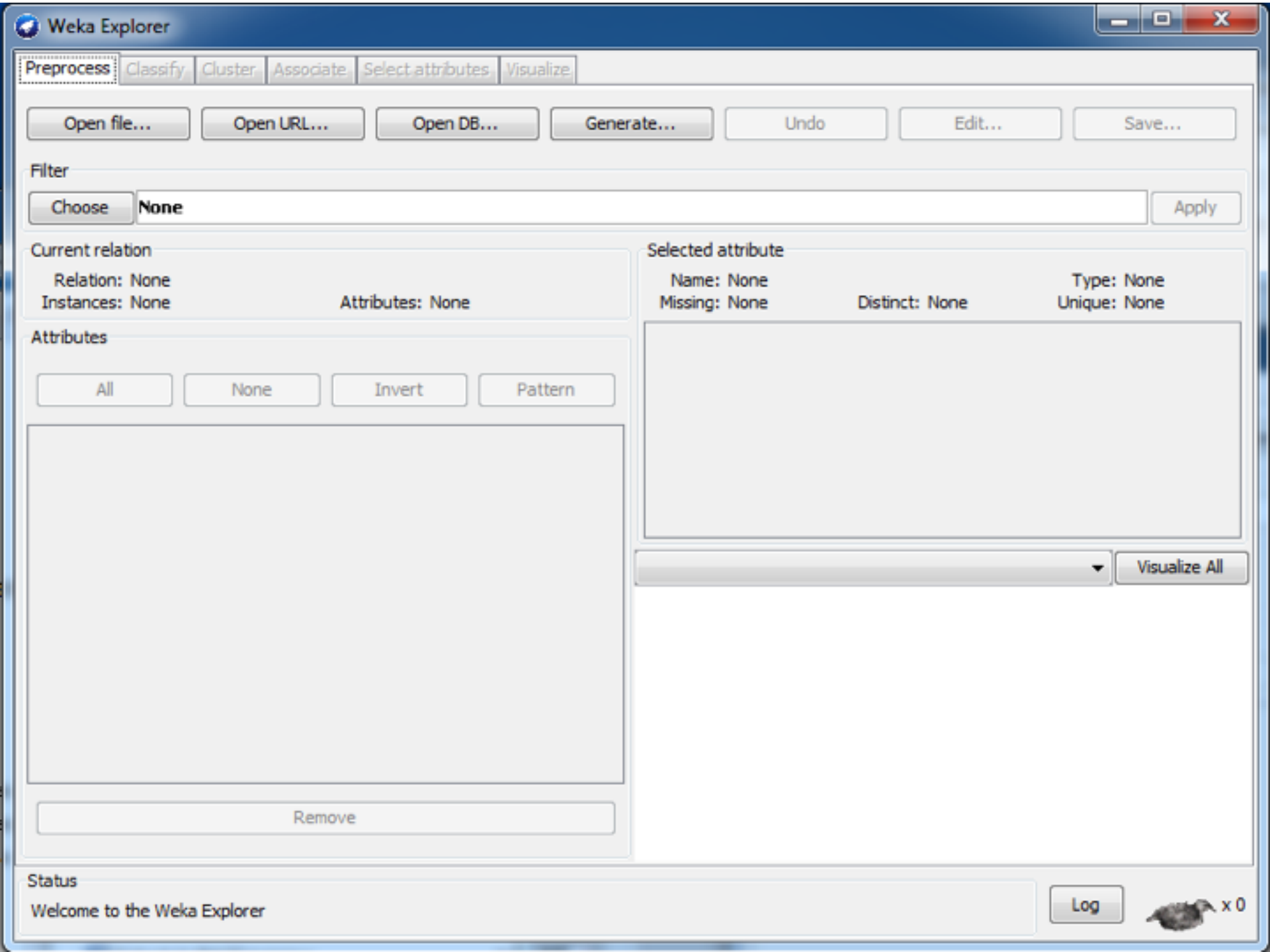
KDD Process: How does WEKA fit in?



WEKA Walk Through: Main GUI

- **Three graphical user interfaces**
  - "The Explorer" (exploratory data analysis)
    - pre-process data
    - build "classifiers"
    - cluster data
    - find associations
    - attribute selection
    - data visualization
  - "The Experimenter" (experimental environment)
    - used to compare performance of different learning schemes
  - "The KnowledgeFlow" (new process model inspired interface)
    - Java-Beans-based interface for setting up and running machine learning experiments.
- **Command line Interface ("Simple CLI")**

More at: http://www.cs.waikato.ac.nz/ml/weka/index_documentation.html


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Generate...
Undo
Edit...
Save...
Filter
Choose
None
Apply
Current relation
Relation: None
Instances: None
Attributes: None
Selected attribute
Name: None
Type: None
Missing: None
Distinct: None
Unique: None
Attributes
All
None
Invert
Pattern
Remove
Visualize All
Status
Welcome to the Weka Explorer
Log
x 0

WEKA:: Explorer: Preprocess

- **Importing data**
  - Data format
  - Uses flat text files to describe the data
  - Data can be imported from a file in various formats:
    - ARFF, CSV, C4.5, binary
  - Data can also be read from a URL or from an SQL database (using JDBC)

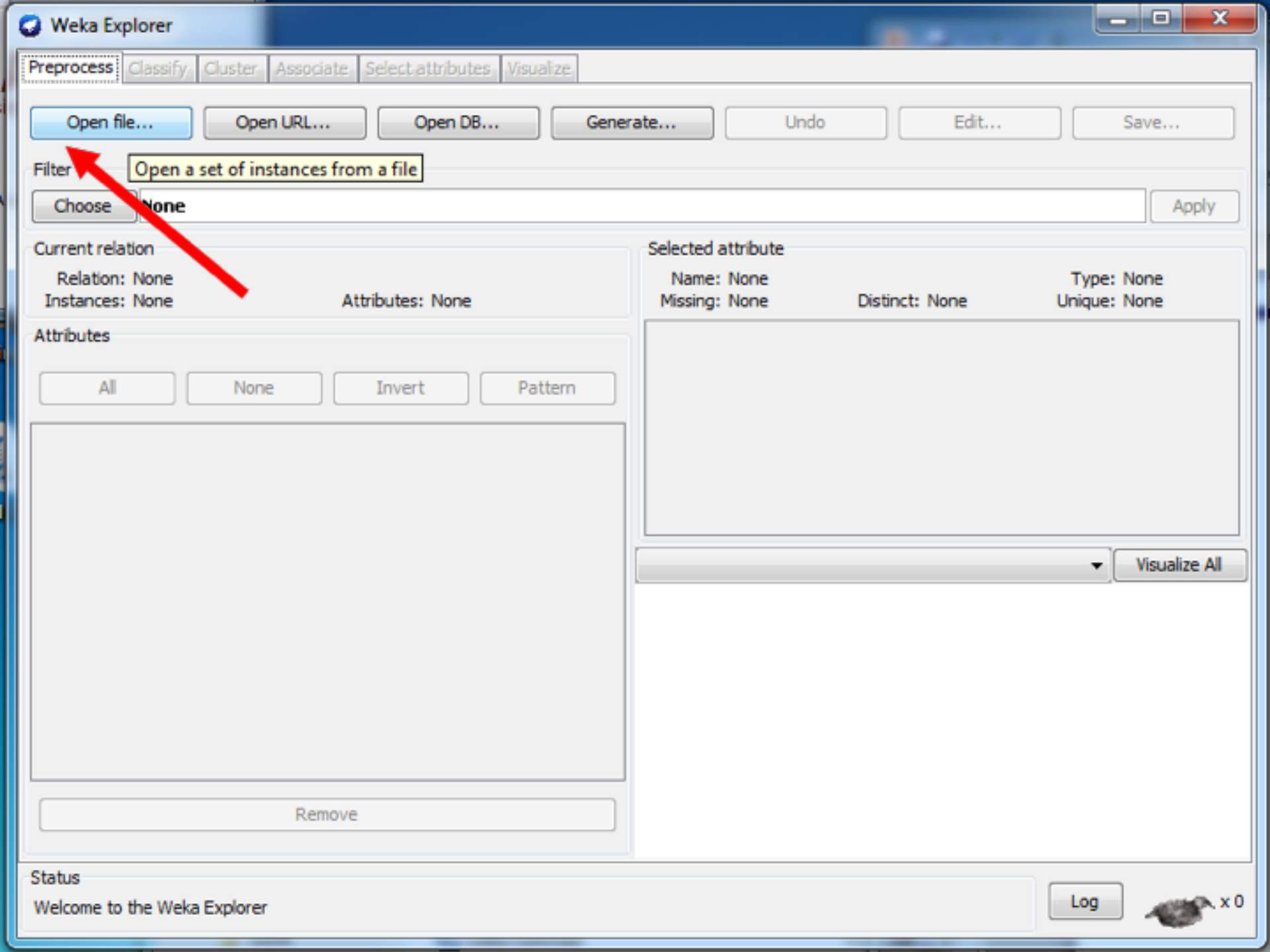
WEKA:: ARFF file format

```
@relation heart-disease-simplified

@attribute age numeric
@attribute sex { female, male}
@attribute chest_pain_type { typ_angina, asympt, non_anginal, atyp_angina}
@attribute cholesterol numeric
@attribute exercise_induced_angina { no, yes}
@attribute class { present, not_present}

@data
63,male,typ_angina,233,no,not_present
67,male,asympt,286,yes,present
67,male,asympt,229,yes,present
38,female,non_anginal,?,no,not_present
...
```

A more thorough description is available here
http://www.cs.waikato.ac.nz/~ml/weka/arff.html


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Generate...
Undo
Edit...
Save...
Filter
Open a set of instances from a file
Choose
None
Apply
Current relation
Relation: None
Instances: None
Attributes: None
Selected attribute
Name: None
Type: None
Missing: None
Distinct: None
Unique: None
Attributes
All
None
Invert
Pattern
Remove
Visualize All
Status
Welcome to the Weka Explorer
Log
x 0

Weka Explorer

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

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

Filter

Choose | None | Apply

Current relation

Relation: weather
Instances: 14 Attributes: 5

Attributes

All | None | Invert | Pattern

No.		Name
1	☐	outlook
2	☐	temperature
3	☐	humidity
4	☐	windy
5	☐	play

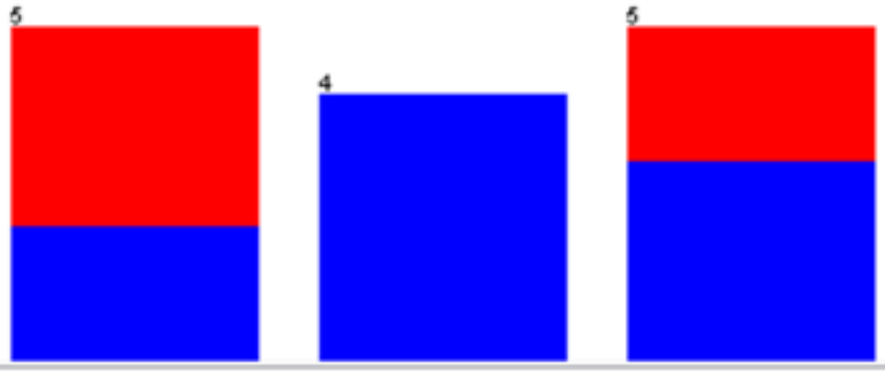
Remove

Selected attribute

Name: outlook Type: Nominal
Missing: 0 (0%) Distinct: 3 Unique: 0 (0%)

No.	Label	Count
1	sunny	5
2	overcast	4
3	rainy	5

Class: play (Nom) | Visualize All



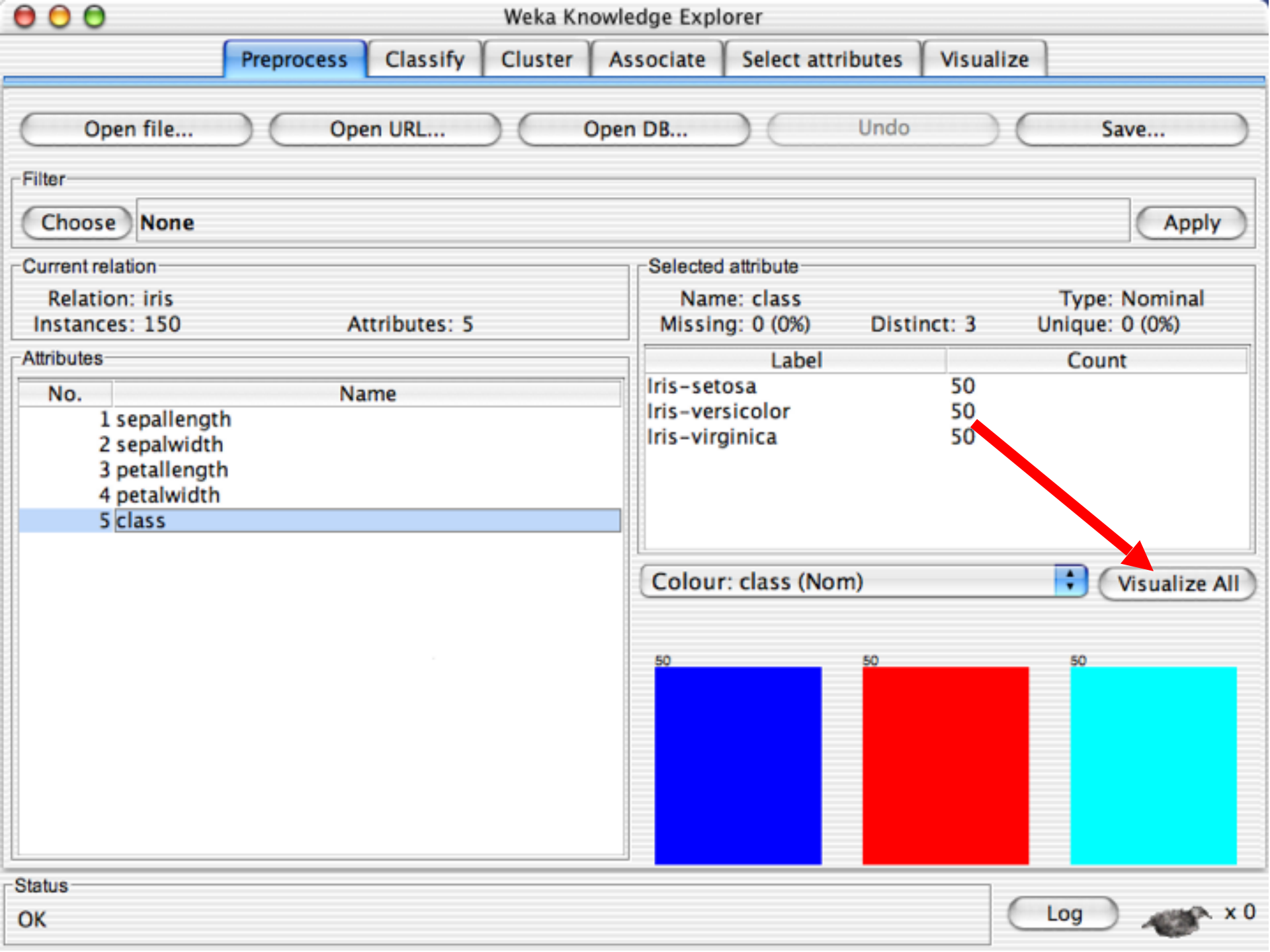
Status

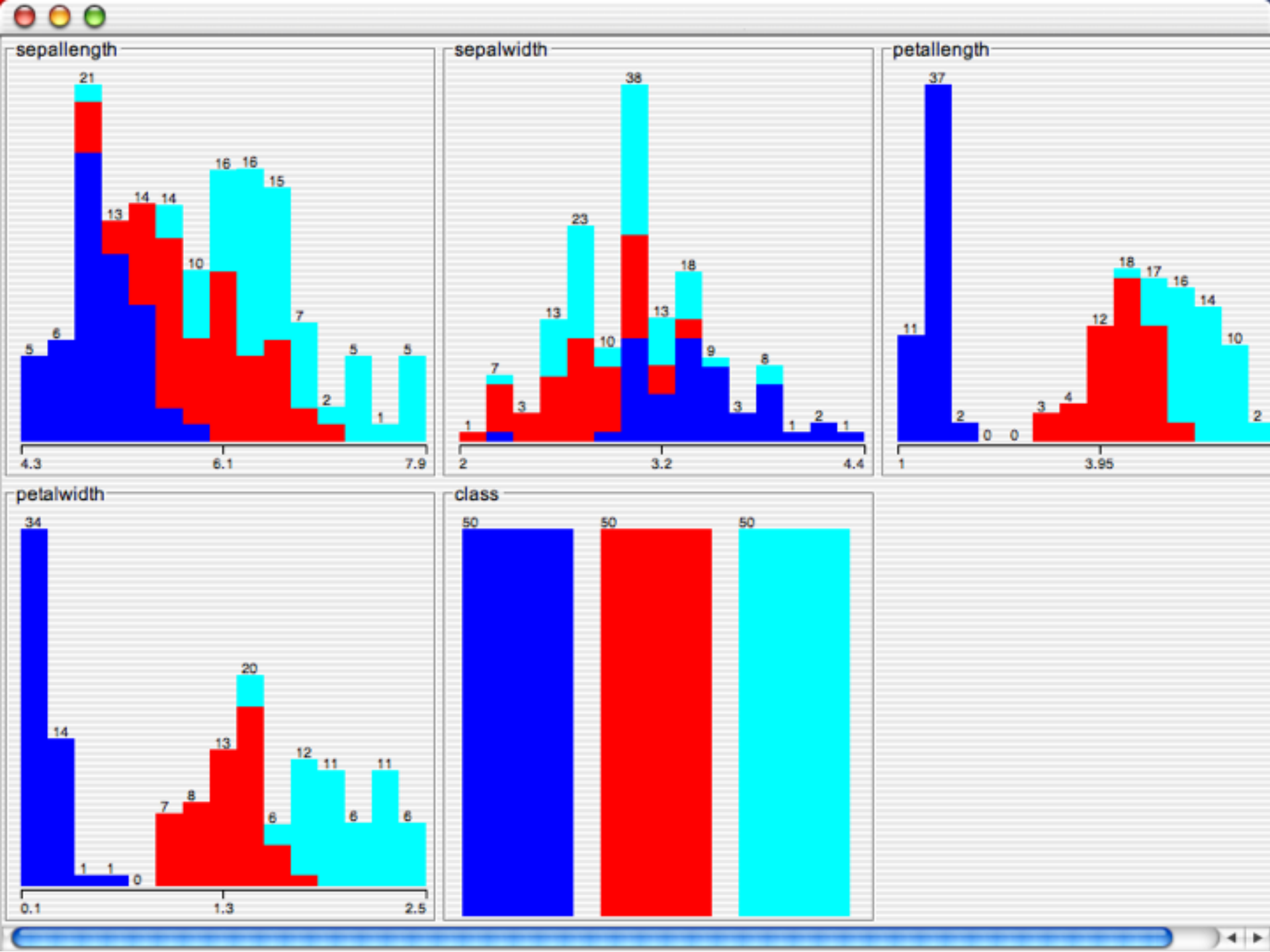
OK

Log x 0

Weka: Explorer:Preprocess

- **Preprocessing data**
  - Visualization
  - Filtering algorithms
    - filters can be used to transform the data (e.g., turning numeric attributes into discrete ones) and make it possible to delete instances and attributes according to specific criteria.
  - Removing Noisy Data
  - Adding Additional Attributes
  - Remove Attributes


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Undo
Save...
Filter
Choose
None
Apply
Current relation
Relation: iris
Instances: 150
Attributes: 5
Attributes
No.
Name
1 sepallength
2 sepalwidth
3 petallength
4 petalwidth
5 class
Selected attribute
Name: class
Type: Nominal
Missing: 0 (0%)
Distinct: 3
Unique: 0 (0%)
Label
Count
Iris-setosa
50
Iris-versicolor
50
Iris-virginica
50
Colour: class (Nom)
Visualize All
50
50
50
Status
OK
Log
x 0


sepallength
21
16 16
14 14
13
15
10
7
6
5
5
5
2
1
4.3
6.1
7.9
sepalwidth
38
23
18
13
13
10
9
8
7
3
3
2
1
1
1
2
3.2
4.4
petallength
37
18
17
16
14
12
11
10
4
3
2
2
0
0
1
3.95
petalwidth
34
20
14
13
12
11
11
8
7
6
6
6
1
1
0
0.1
1.3
2.5
class
50
50
50

Weka Explorer

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

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

Filter

Choose | None | Apply

Current relation

Relation: weather
Instances: 14 Attributes: 5

Attributes

All | None | Invert | Pattern

No.		Name
1	☐	outlook
2	☐	temperature
3	☐	humidity
4	☐	windy
5	☐	play

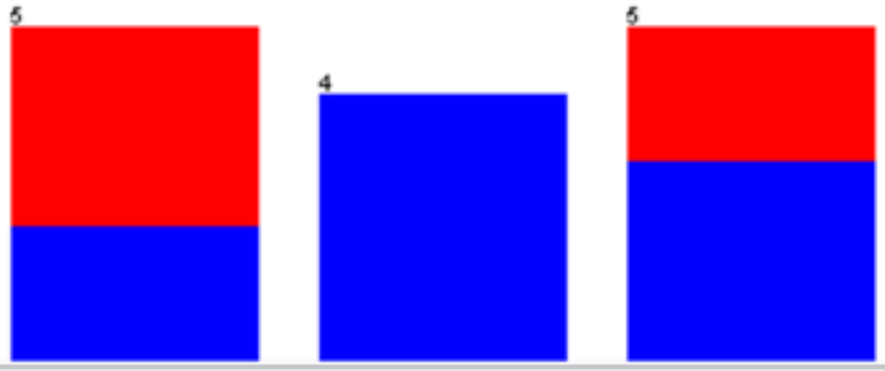
Remove

Selected attribute

Name: outlook Type: Nominal
Missing: 0 (0%) Distinct: 3 Unique: 0 (0%)

No.	Label	Count
1	sunny	5
2	overcast	4
3	rainy	5

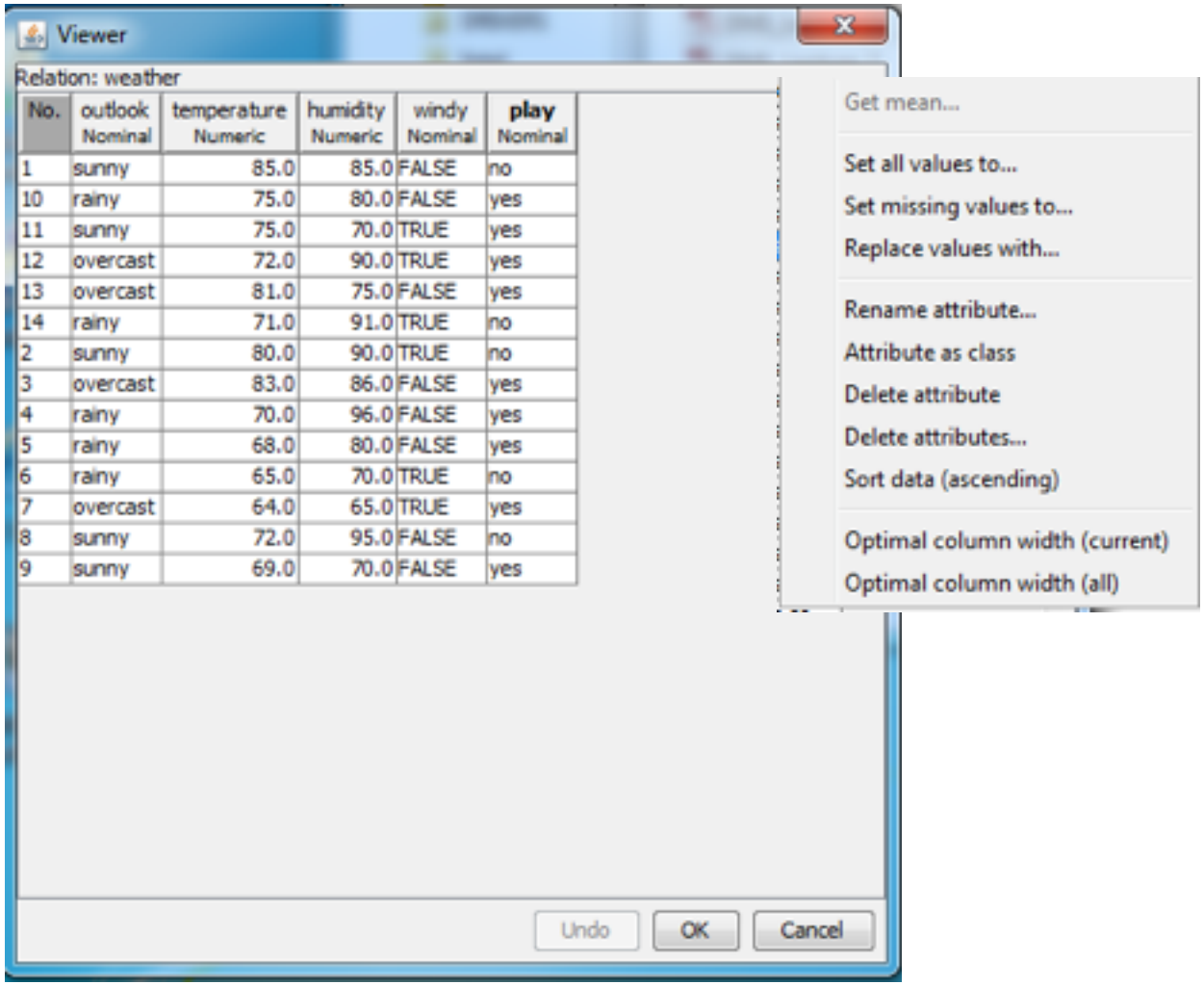
Class: play (Nom) | Visualize All



Status

OK

Log x 0


Viewer
Relation: weather
No. outlook Nominal temperature Numeric humidity Numeric windy Nominal play Nominal
1 sunny 85.0 85.0 FALSE no
10 rainy 75.0 80.0 FALSE yes
11 sunny 75.0 70.0 TRUE yes
12 overcast 72.0 90.0 TRUE yes
13 overcast 81.0 75.0 FALSE yes
14 rainy 71.0 91.0 TRUE no
2 sunny 80.0 90.0 TRUE no
3 overcast 83.0 86.0 FALSE yes
4 rainy 70.0 96.0 FALSE yes
5 rainy 68.0 80.0 FALSE yes
6 rainy 65.0 70.0 TRUE no
7 overcast 64.0 65.0 TRUE yes
8 sunny 72.0 95.0 FALSE no
9 sunny 69.0 70.0 FALSE yes
Get mean...
Set all values to...
Set missing values to...
Replace values with...
Rename attribute...
Attribute as class
Delete attribute
Delete attributes...
Sort data (ascending)
Optimal column width (current)
Optimal column width (all)
Undo
OK
Cancel

WEKA:: Explorer: Preprocess

- **Used to define filters to transform Data.**
- **WEKA contains filters for:**
  - Discretization, normalization, resampling, attribute selection, transforming, combining attributes, etc

Weka Knowledge Explorer

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

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

Filter

- weka
  - filters
    - unsupervised
      - attribute
        - Add
        - AddCluster
        - AddExpression
        - AddNoise
        - Copy
        - Discretize
        - FirstOrder
        - MakeIndicator
        - MergeTwoValues
        - NominalToBinary
        - Normalize
        - NumericToBinary
        - NumericTransform
        - Obfuscate
        - PKIDiscretize
        - Remove
        - RemoveType

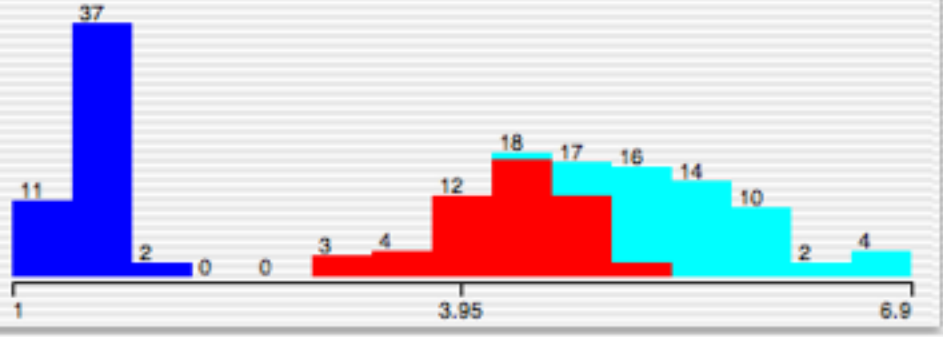
Apply

Selected attribute

Name: petallength Type: Numeric
Missing: 0 (0%) Distinct: 43 Unique: 10 (7%)

Statistic	Value
Minimum	1
Maximum	6.9
Mean	3.759
StdDev	1.764

Colour: class (Nom) | Visualize All



Status

OK

Log | x 0

Weka Knowledge Explorer

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

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

Filter

Choose | **Discretize** -B 10 -R first-last | Apply

Current relation

Relation: iris
Instances: 150 Attributes: 5

Attributes

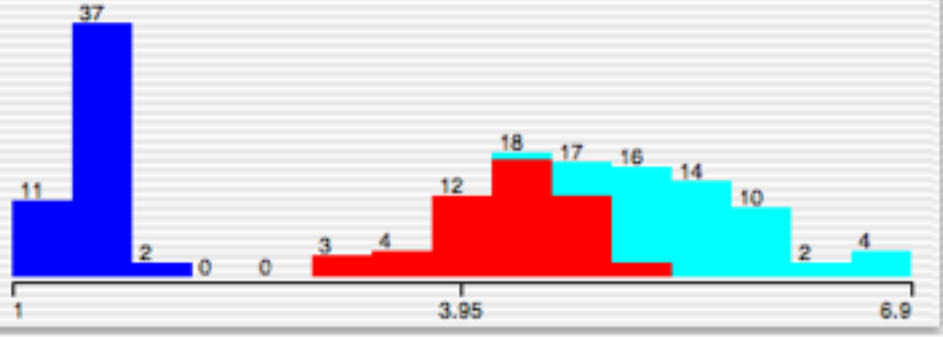
No.	Name
1	sepallength
2	sepalwidth
3	petallength
4	petalwidth
5	class

Selected attribute

Name: petallength Type: Numeric
Missing: 0 (0%) Distinct: 43 Unique: 10 (7%)

Statistic	Value
Minimum	1
Maximum	6.9
Mean	3.759
StdDev	1.764

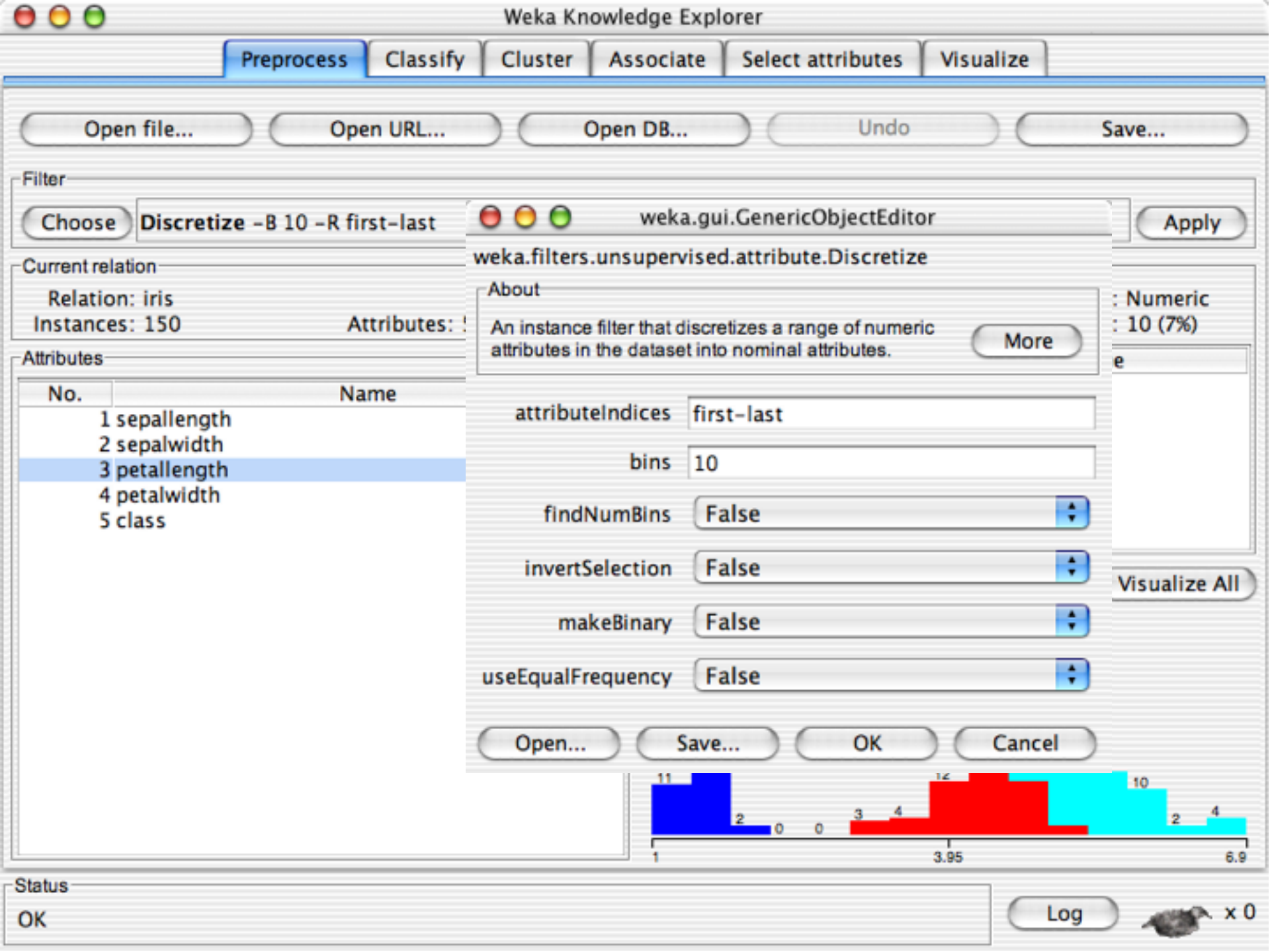
Colour: class (Nom) | Visualize All



Status

OK

Log | x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Open file...
Open URL...
Open DB...
Undo
Save...
Filter
Choose
Discretize -B 10 -R first-last
Apply
Current relation
Relation: iris
Instances: 150
Attributes:
Attributes
No.
Name
1 sepallength
2 sepalwidth
3 petallength
4 petalwidth
5 class
weka.gui.GenericObjectEditor
weka.filters.unsupervised.attribute.Discretize
About
An instance filter that discretizes a range of numeric attributes in the dataset into nominal attributes.
More
attributeIndices
first-last
bins
10
findNumBins
False
invertSelection
False
makeBinary
False
useEqualFrequency
False
Open...
Save...
OK
Cancel
: Numeric
: 10 (7%)
Visualize All
11
2
0
0
3
4
12
10
2
4
1
3.95
6.9
Status
OK
Log
x 0

Weka Knowledge Explorer

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

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

Filter

Choose | **Discretize** -F -B 10 -R first-last | Apply

Current relation

Relation: iris
Instances: 150 Attributes: 5

Attributes

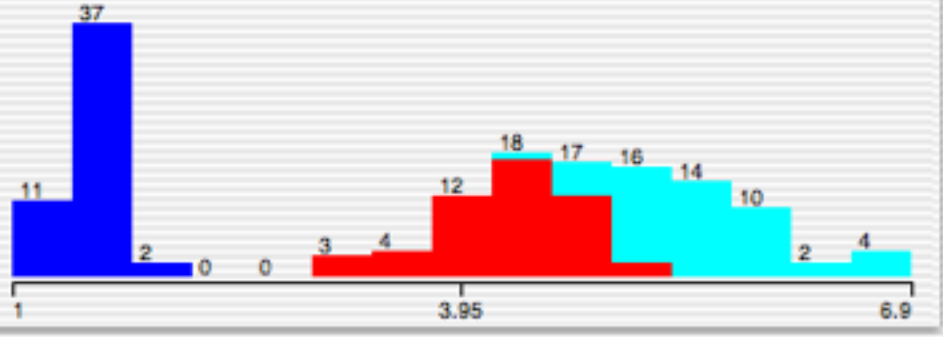
No.	Name
1	sepallength
2	sepalwidth
3	petallength
4	petalwidth
5	class

Selected attribute

Name: petallength Type: Numeric
Missing: 0 (0%) Distinct: 43 Unique: 10 (7%)

Statistic	Value
Minimum	1
Maximum	6.9
Mean	3.759
StdDev	1.764

Colour: class (Nom) | Visualize All



Status

OK

Log | x 0

Weka Knowledge Explorer

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

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

Filter

Choose **Discretize** -F -B 10 -R first-last | Apply

Current relation

Relation: iris-weka.filters.unsupervised.attribute.Disc...
Instances: 150 Attributes: 5

Attributes

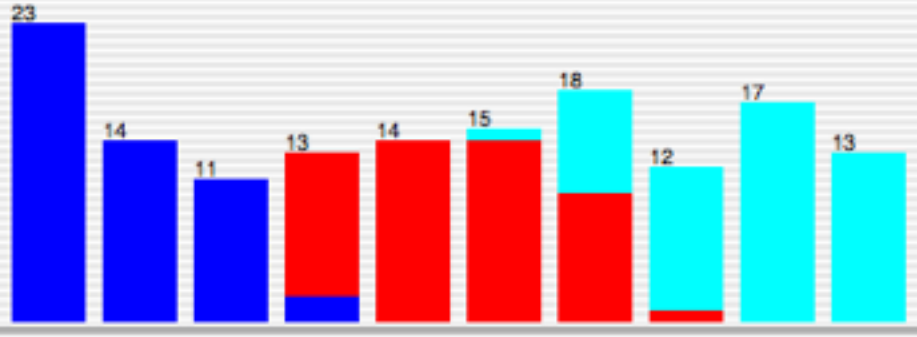
No.	Name
1	sepallength
2	sepalwidth
3	petallength
4	petalwidth
5	class

Selected attribute

Name: petallength Type: Nominal
Missing: 0 (0%) Distinct: 10 Unique: 0 (0%)

Label	Count
'(-inf-1.45]'	23
'(1.45-1.55]'	14
'(1.55-1.8]'	11
'(1.8-3.95]'	13
'(3.95-4.35]'	14
'(4.35-4.65]'	15
'(4.65-5.05]'	18

Colour: class (Nom) | Visualize All

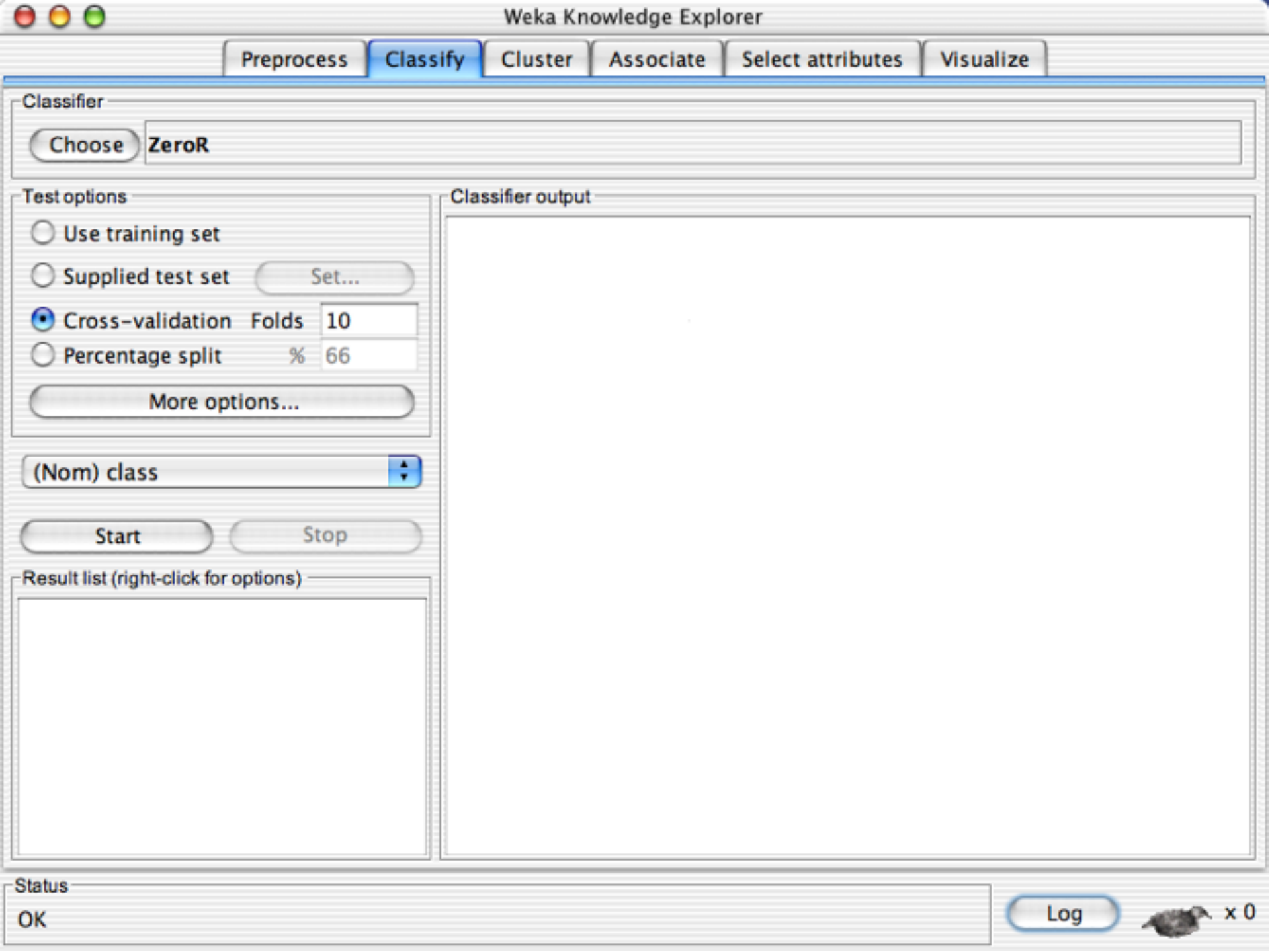


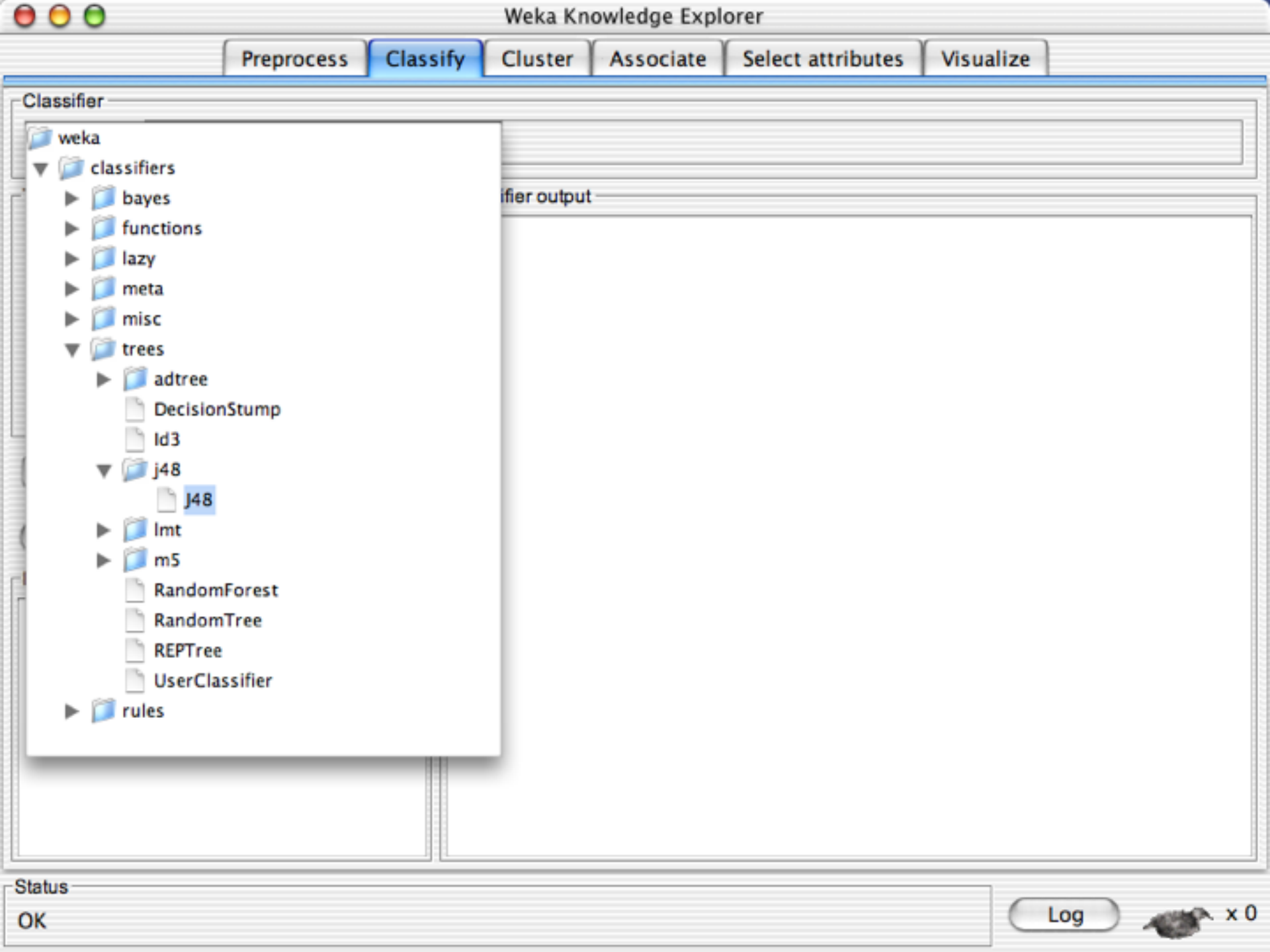
Status

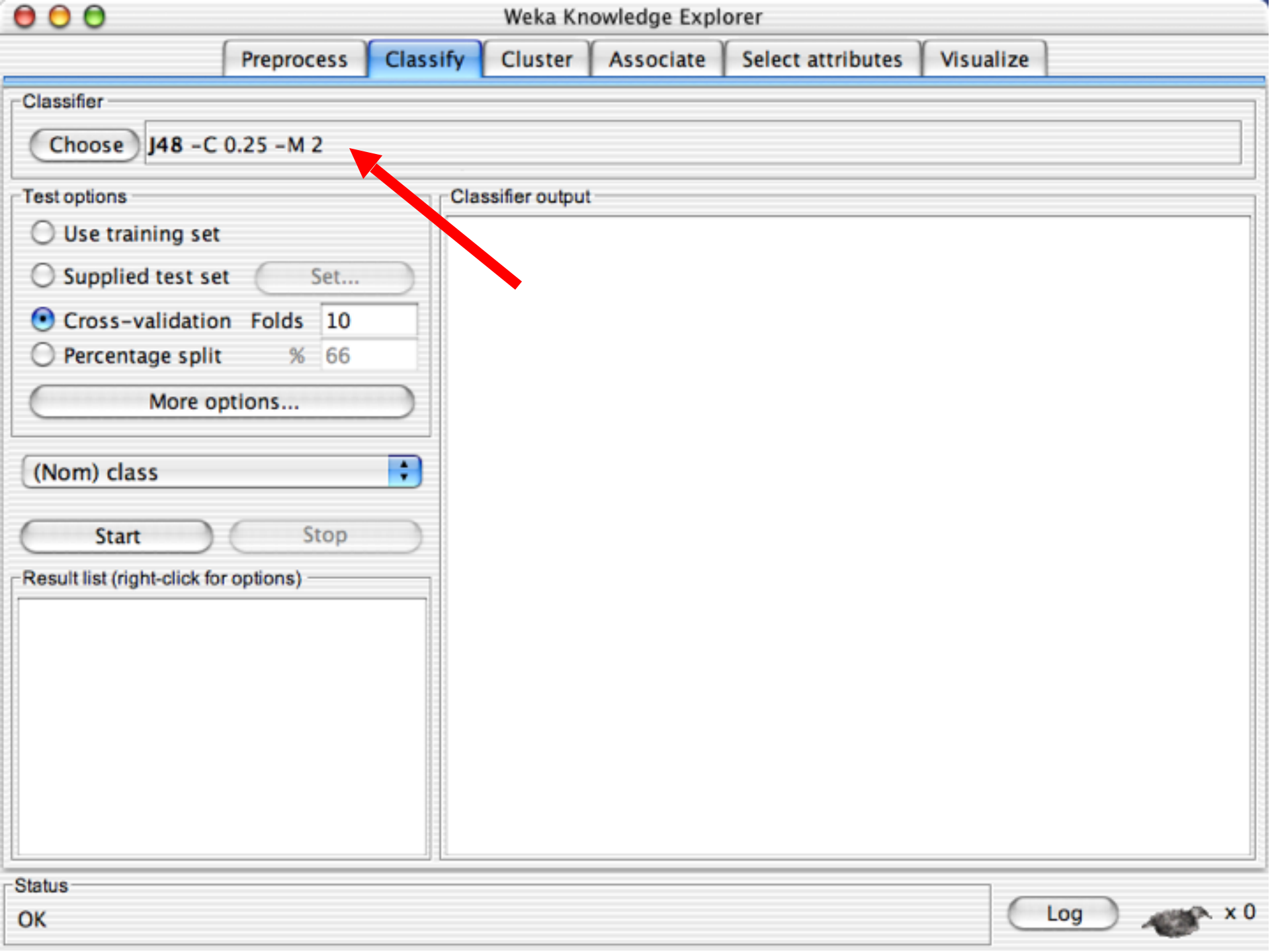
OK | Log | x 0

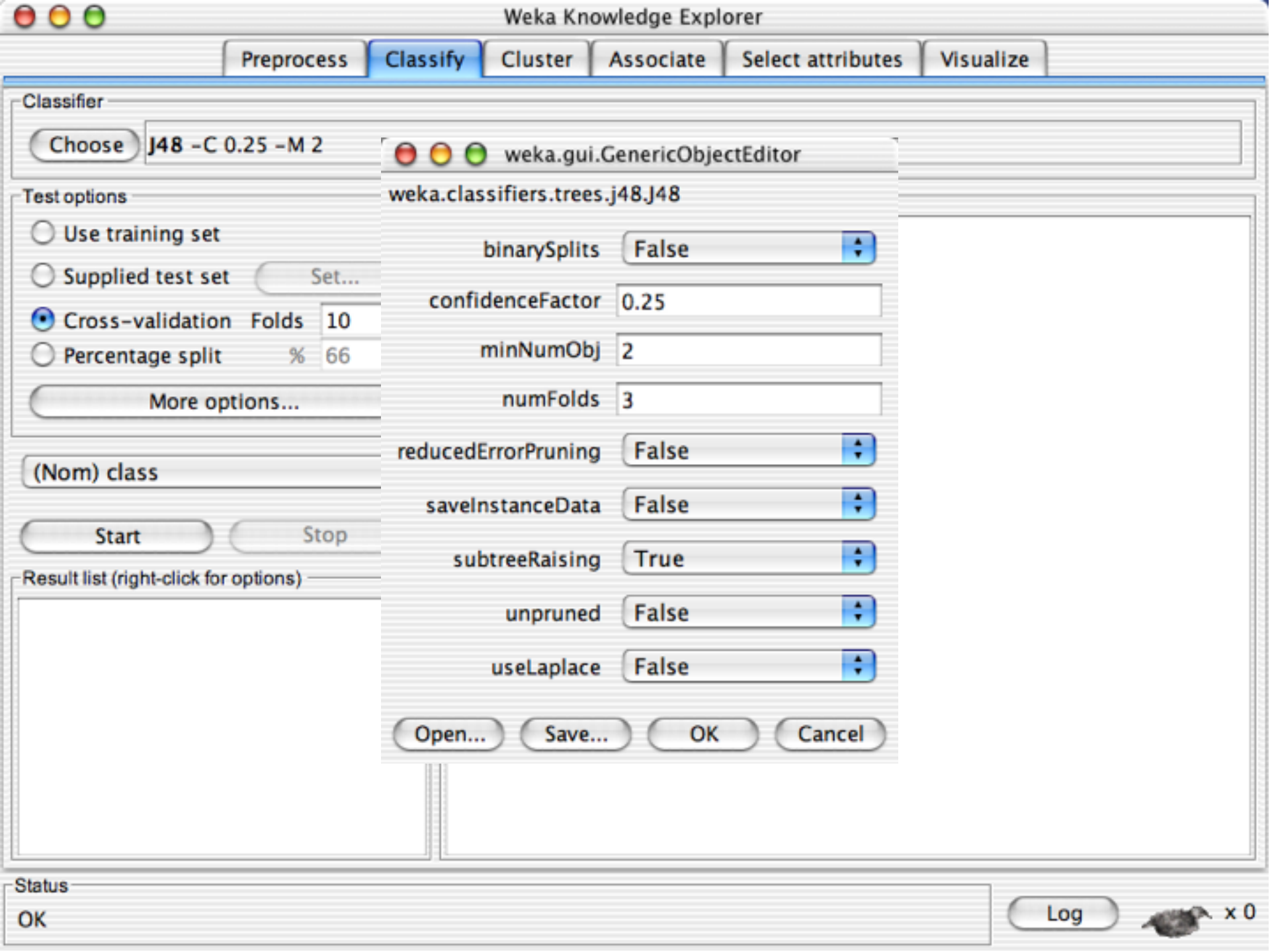
WEKA:: Explorer: building "classifiers"

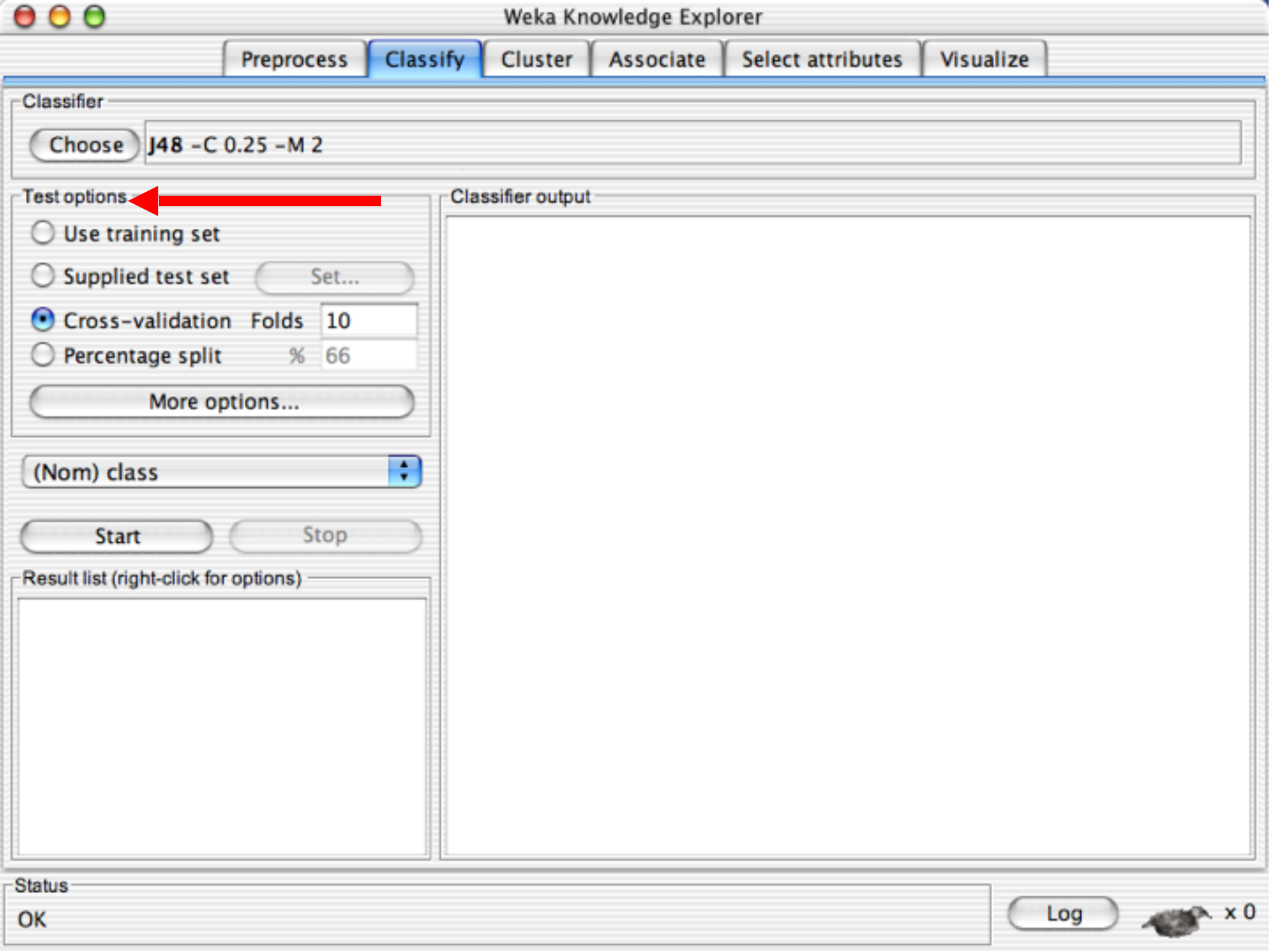
- **Classifiers in WEKA are models for predicting nominal or numeric quantities**
- **Implemented learning schemes include:**
  - Decision trees and lists, instance-based classifiers, support vector machines, multi-layer perceptrons, logistic regression, Bayes' nets, …
- **"Meta"-classifiers include:**
  - Bagging, boosting, stacking, error-correcting output codes, locally weighted learning, …

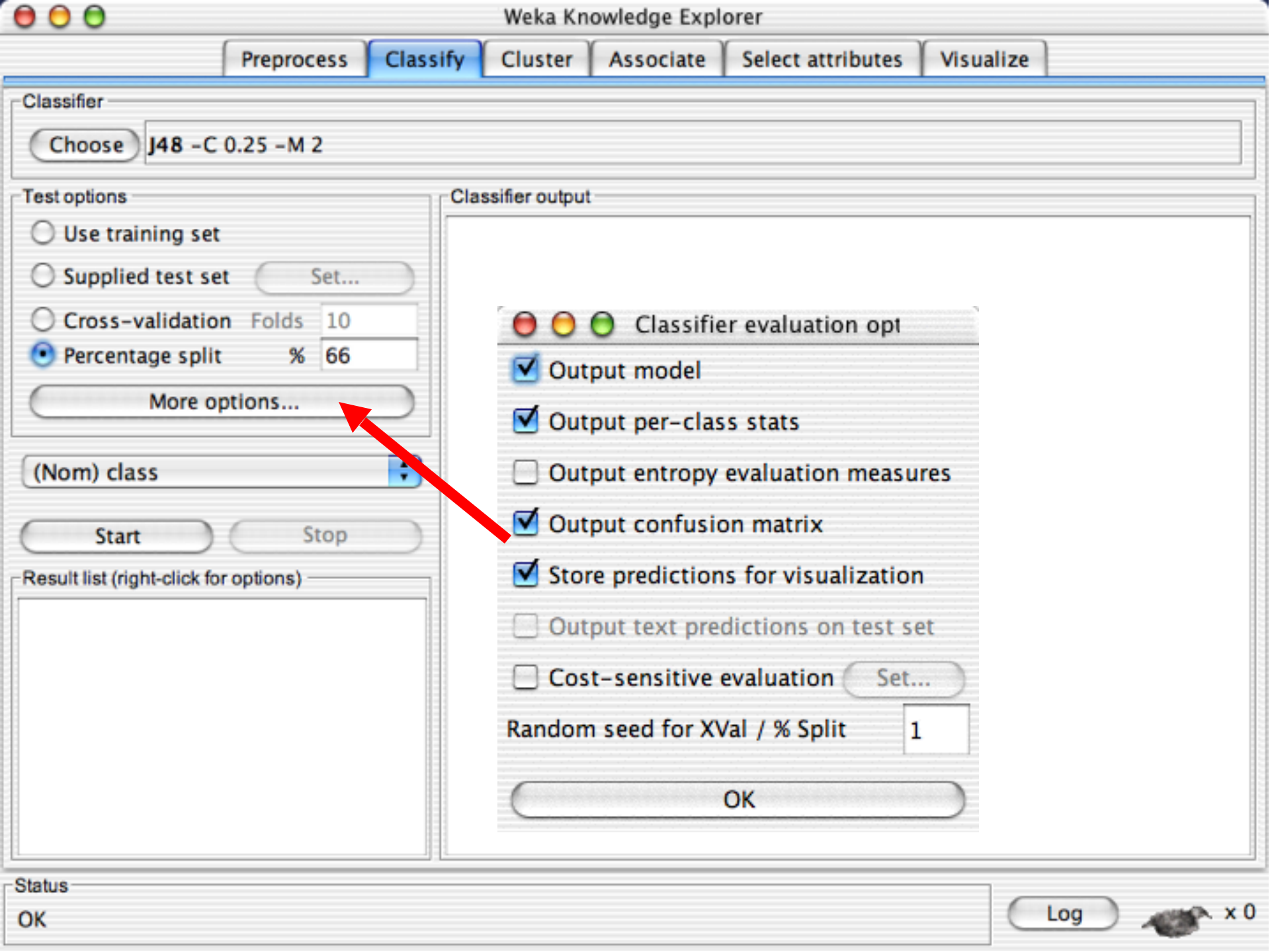

Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
ZeroR
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
Classifier output
(Nom) class
Start
Stop
Result list (right-click for options)
Status
OK
Log
x 0

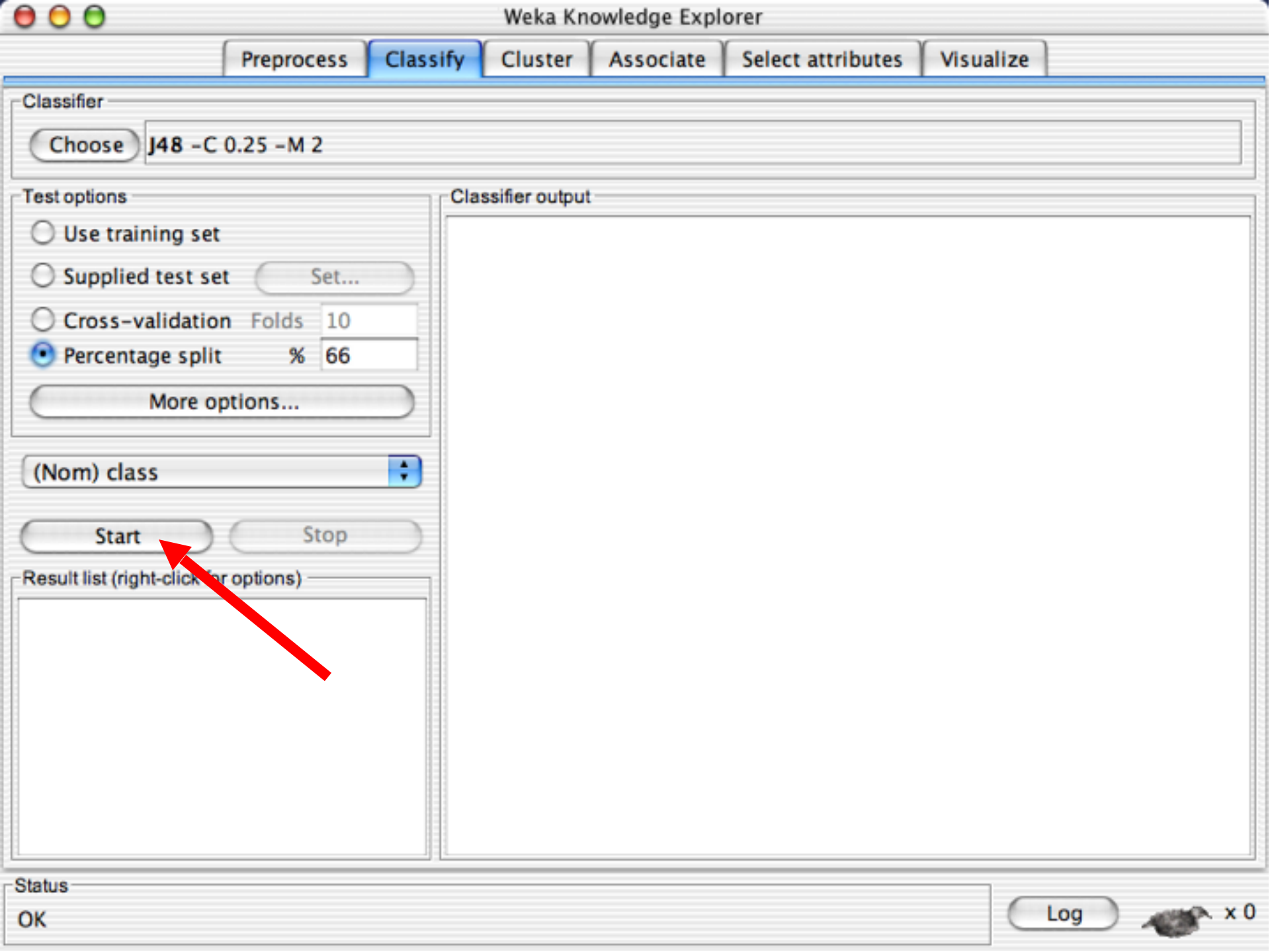

Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
weka
classifiers
bayes
functions
lazy
meta
misc
trees
adtree
DecisionStump
Id3
j48
J48
lmt
m5
RandomForest
RandomTree
REPTree
UserClassifier
rules
ifier output
Status
OK
Log
x 0

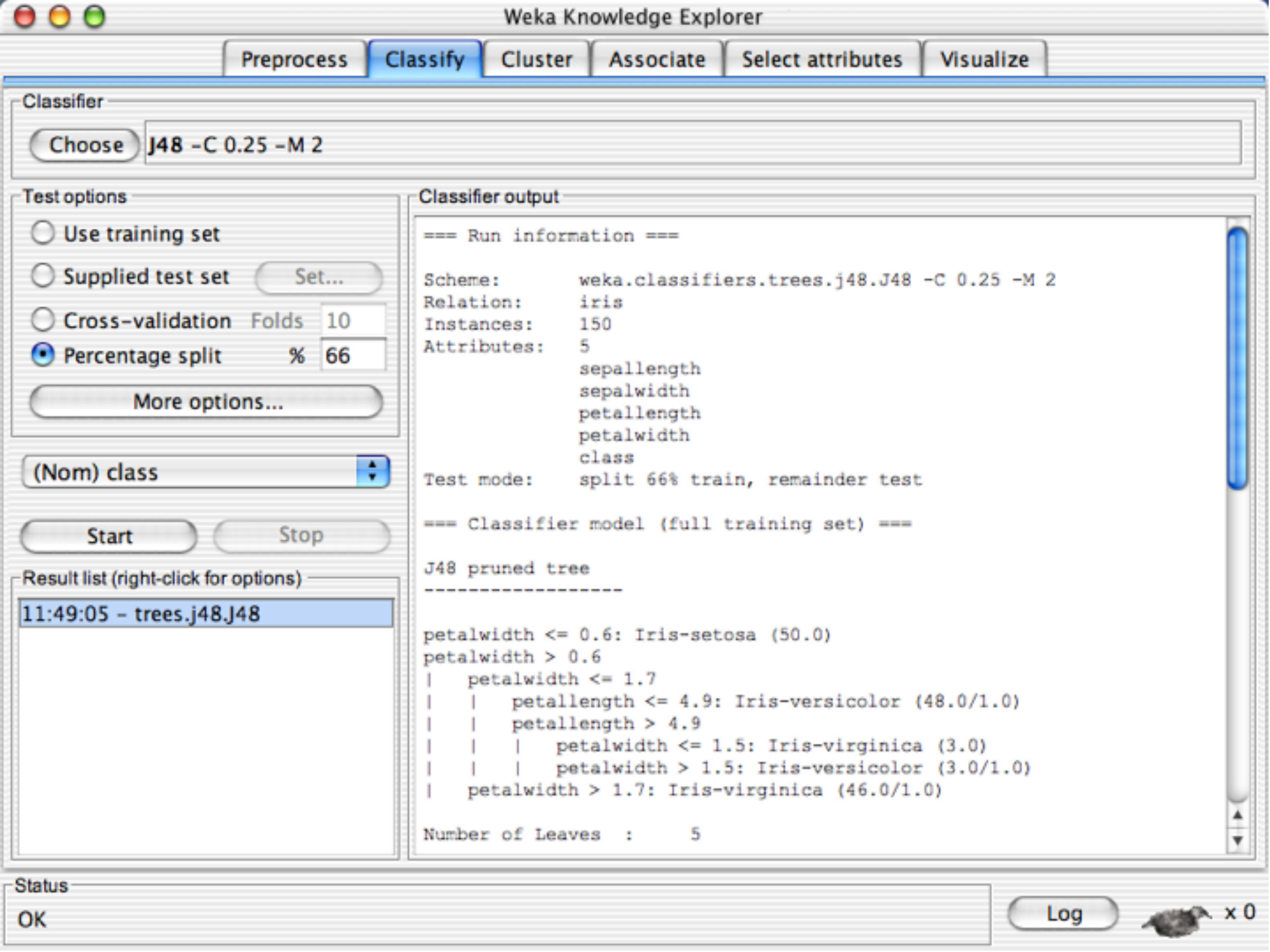

Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
J48 -C 0.25 -M 2
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
Classifier output
(Nom) class
Start
Stop
Result list (right-click for options)
Status
OK
Log
x 0

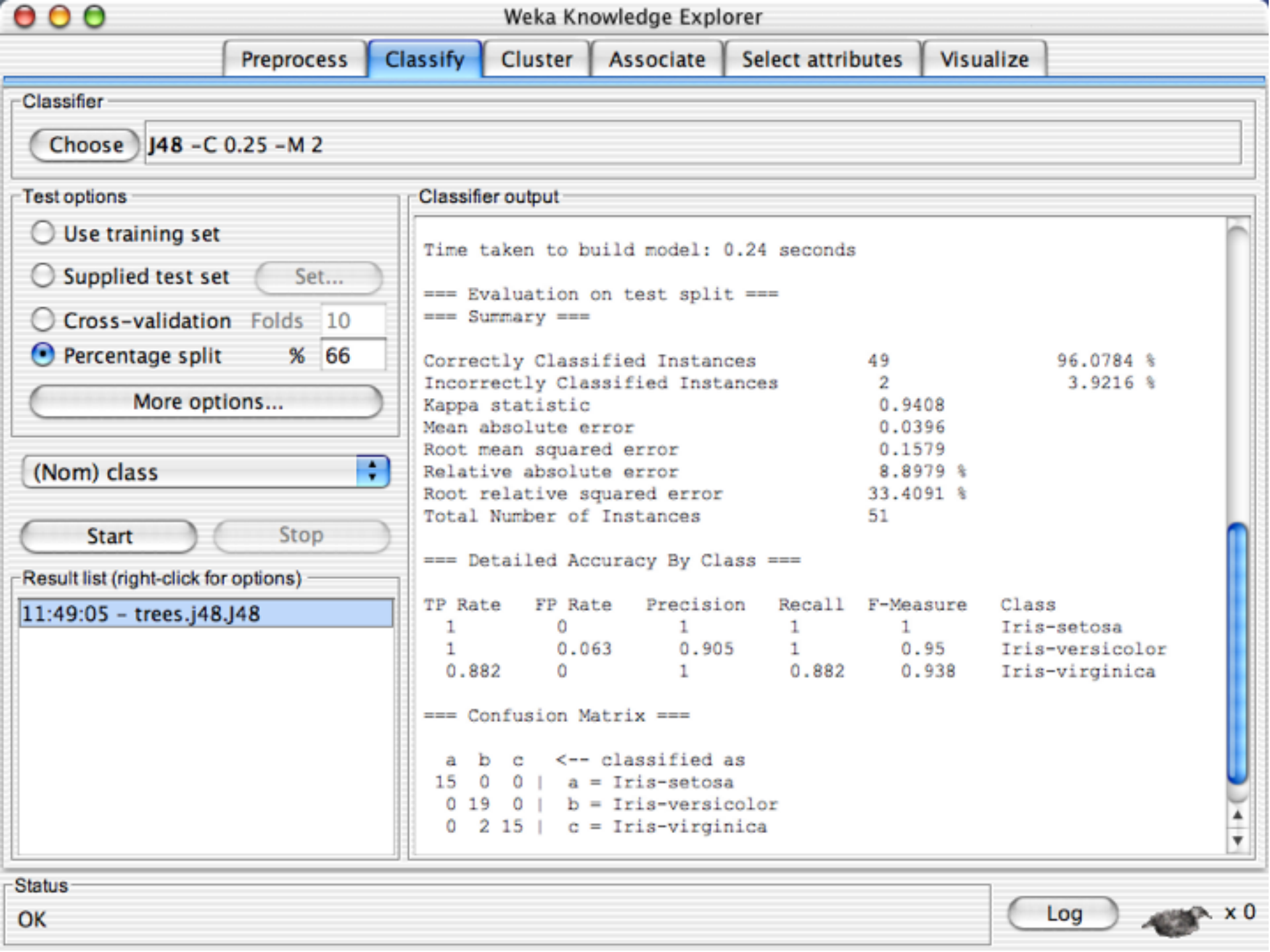

Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
J48 -C 0.25 -M 2
weka.gui.GenericObjectEditor
weka.classifiers.trees.j48.J48
binarySplits
False
confidenceFactor
0.25
minNumObj
2
numFolds
3
reducedErrorPruning
False
saveInstanceData
False
subtreeRaising
True
unpruned
False
useLaplace
False
Open...
Save...
OK
Cancel
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) class
Start
Stop
Result list (right-click for options)
Status
OK
Log
x 0

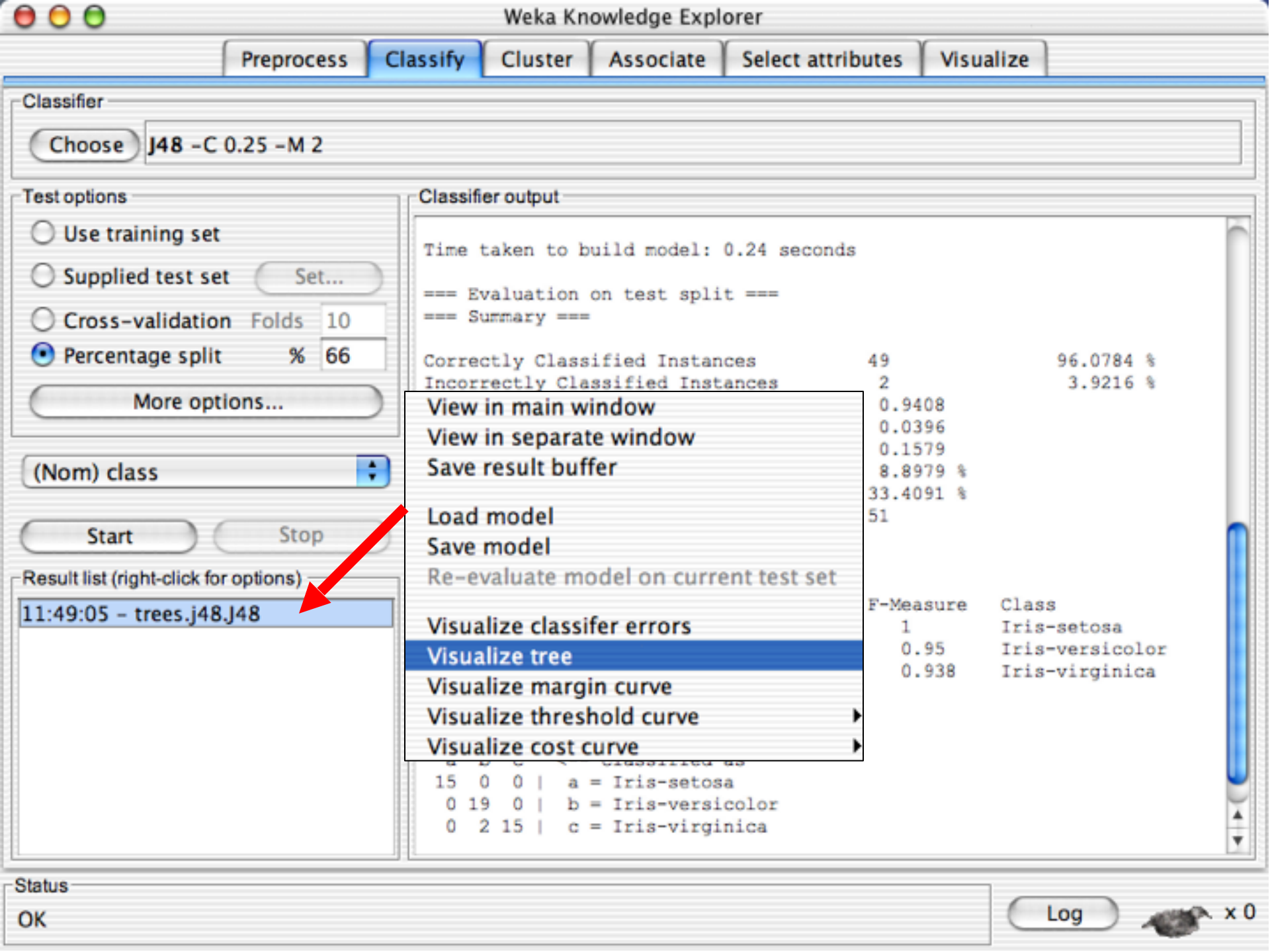

Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
J48 -C 0.25 -M 2
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
Classifier output
(Nom) class
Start
Stop
Result list (right-click for options)
Status
OK
Log
x 0

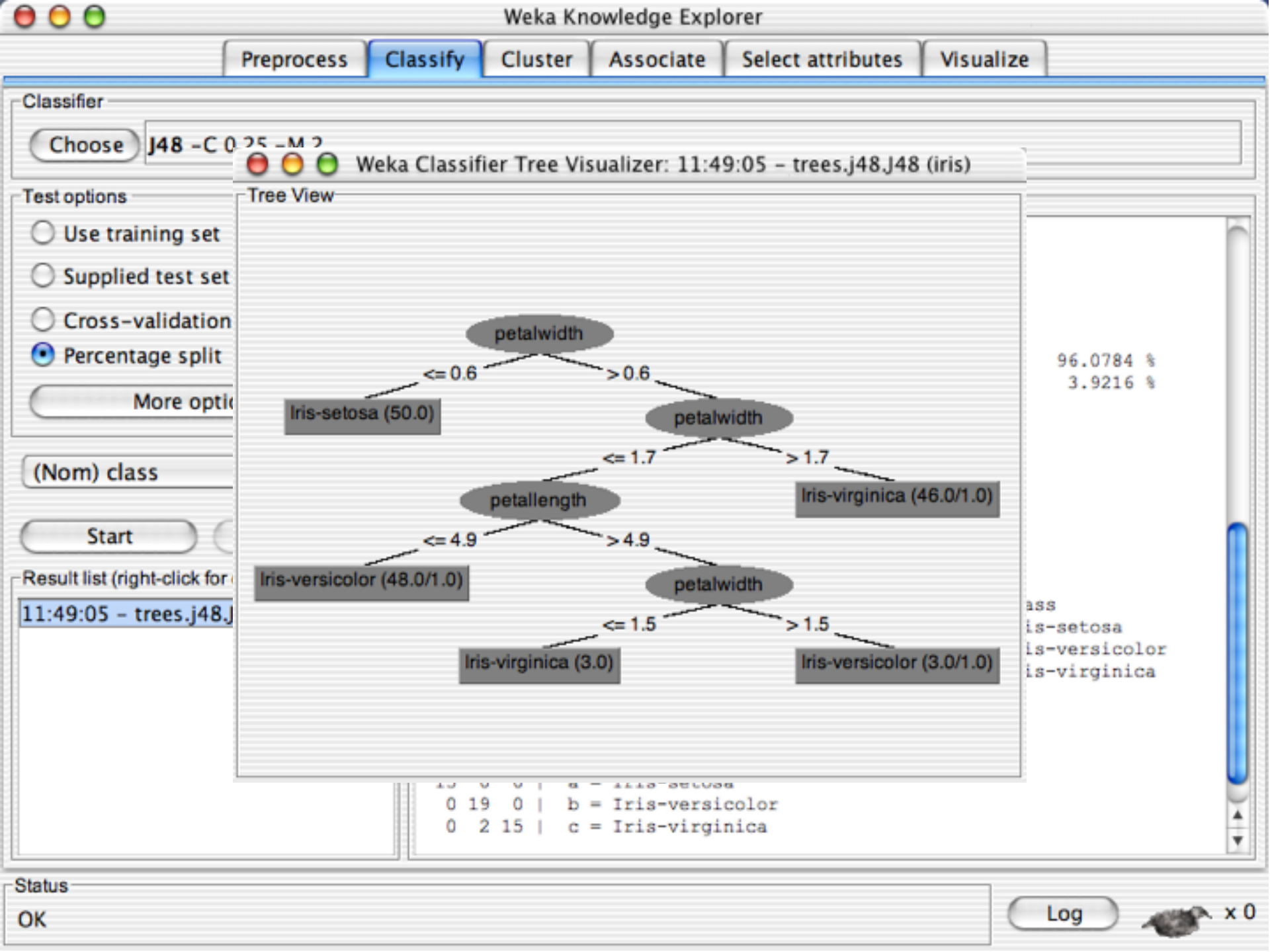

Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
J48 -C 0.25 -M 2
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
(Nom) class
Start
Stop
Result list (right-click for options)
Classifier output
Classifier evaluation opt
Output model
Output per-class stats
Output entropy evaluation measures
Output confusion matrix
Store predictions for visualization
Output text predictions on test set
Cost-sensitive evaluation
Set...
Random seed for XVal / % Split
1
OK
Status
OK
Log
x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
J48 -C 0.25 -M 2
Test options
Use training set
Supplied test set
Set...
Cross-validation
Folds
10
Percentage split
%
66
More options...
Classifier output
(Nom) class
Start
Stop
Result list (right-click for options)
Status
OK
Log
x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
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) class
Start
Stop
Result list (right-click for options)
11:49:05 - trees.j48.J48
Classifier output
=== Run information ===
Scheme: weka.classifiers.trees.j48.J48 -C 0.25 -M 2
Relation: iris
Instances: 150
Attributes: 5
sepallength
sepalwidth
petallength
petalwidth
class
Test mode: split 66% train, remainder test
=== Classifier model (full training set) ===
J48 pruned tree
petalwidth <= 0.6: Iris-setosa (50.0)
petalwidth > 0.6
| petalwidth <= 1.7
| | petallength <= 4.9: Iris-versicolor (48.0/1.0)
| | petallength > 4.9
| | | petalwidth <= 1.5: Iris-virginica (3.0)
| | | petalwidth > 1.5: Iris-versicolor (3.0/1.0)
| petalwidth > 1.7: Iris-virginica (46.0/1.0)
Number of Leaves : 5
Status
OK
Log
x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
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) class
Start
Stop
Result list (right-click for options)
11:49:05 - trees.j48.J48
Classifier output
Time taken to build model: 0.24 seconds
=== Evaluation on test split ===
=== Summary ===
Correctly Classified Instances 49 96.0784 %
Incorrectly Classified Instances 2 3.9216 %
Kappa statistic 0.9408
Mean absolute error 0.0396
Root mean squared error 0.1579
Relative absolute error 8.8979 %
Root relative squared error 33.4091 %
Total Number of Instances 51
=== Detailed Accuracy By Class ===
TP Rate FP Rate Precision Recall F-Measure Class
1 0 1 1 1 Iris-setosa
1 0.063 0.905 1 0.95 Iris-versicolor
0.882 0 1 0.882 0.938 Iris-virginica
=== Confusion Matrix ===
a b c <-- classified as
15 0 0 | a = Iris-setosa
0 19 0 | b = Iris-versicolor
0 2 15 | c = Iris-virginica
Status
OK
Log
x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
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) class
Start
Stop
Result list (right-click for options)
11:49:05 - trees.j48.J48
Classifier output
Time taken to build model: 0.24 seconds
=== Evaluation on test split ===
=== Summary ===
Correctly Classified Instances 49 96.0784 %
Incorrectly Classified Instances 2 3.9216 %
0.9408
0.0396
0.1579
8.8979 %
33.4091 %
51
F-Measure Class
1 Iris-setosa
0.95 Iris-versicolor
0.938 Iris-virginica
15 0 0 | a = Iris-setosa
0 19 0 | b = Iris-versicolor
0 2 15 | c = Iris-virginica
View in main window
View in separate window
Save result buffer
Load model
Save model
Re-evaluate model on current test set
Visualize classifer errors
Visualize tree
Visualize margin curve
Visualize threshold curve
Visualize cost curve
Status
OK
Log
x 0


Weka Knowledge Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Classifier
Choose
J48 -C 0.25 -M 2
Weka Classifier Tree Visualizer: 11:49:05 - trees.j48.J48 (iris)
Tree View
petalwidth
<= 0.6
> 0.6
Iris-setosa (50.0)
petalwidth
<= 1.7
> 1.7
petallength
Iris-virginica (46.0/1.0)
<= 4.9
> 4.9
Iris-versicolor (48.0/1.0)
petalwidth
<= 1.5
> 1.5
Iris-virginica (3.0)
Iris-versicolor (3.0/1.0)
Test options
Use training set
Supplied test set
Cross-validation
Percentage split
More optic
(Nom) class
Start
Result list (right-click for
11:49:05 - trees.j48.J
96.0784 %
3.9216 %
ass
is-setosa
is-versicolor
is-virginica
0 19 0 | b = Iris-versicolor
0 2 15 | c = Iris-virginica
Status
OK
Log
x 0

WEKA:: Explorer: building "Cluster"

- **WEKA contains "clusters" for finding groups of similar instances in a dataset**
- **Implemented schemes are:**
  - k-Means, EM, Cobweb, X-means, FarthestFirst
- **Clusters can be visualized and compared to "true" clusters (if given)**
- **Evaluation based on loglikelihood if clustering scheme produces a probability distribution**

Explorer: Finding associations

- **WEKA contains an implementation of the Apriori algorithm for learning association rules**
  - Works only with discrete data
- **Can identify statistical dependencies between groups of attributes:**
  - milk, butter ⇒ bread, eggs (with confidence 0.9 and support 2000)
- **Apriori can compute all rules that have a given minimum support and exceed a given confidence**

Explorer: Attribute Selection

- **Panel that can be used to investigate which (subsets of) attributes are the most predictive ones**
- **Attribute selection methods contain two parts:**
  - A search method: best-first, forward selection, random, exhaustive, genetic algorithm, ranking
  - An evaluation method: correlation-based, wrapper, information gain, chi-squared, …
- **Very flexible: WEKA allows (almost) arbitrary combinations of these two**

Explorer: Visualize

- **Visualization very useful in practice: e.g. helps to determine difficulty of the learning problem**
- **WEKA can visualize single attributes (1-d) and pairs of attributes (2-d)**
  - To do: rotating 3-d visualizations (Xgobi-style)
- **Color-coded class values**
- **"Jitter" option to deal with nominal attributes (and to detect "hidden" data points)**
- **"Zoom-in" function**

Weka Knowledge Explorer

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

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

Filter

Choose **None** | Apply

Current relation

Relation: Glass
Instances: 214 Attributes: 10

Attributes

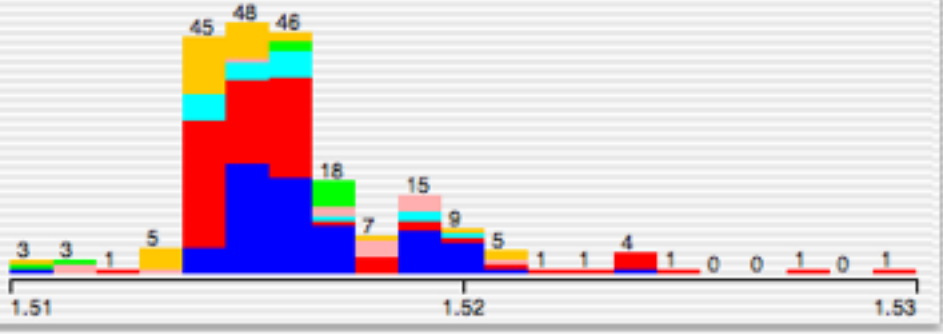
No.	Name
1	RI
2	Na
3	Mg
4	Al
5	Si
6	K
7	Ca
8	Ba
9	Fe
10	Type

Selected attribute

Name: RI Type: Numeric
Missing: 0 (0%) Distinct: 178 Unique: 145 (68%)

Statistic	Value
Minimum	1.511
Maximum	1.534
Mean	1.518
StdDev	0.003

Colour: Type (Nom) | Visualize All



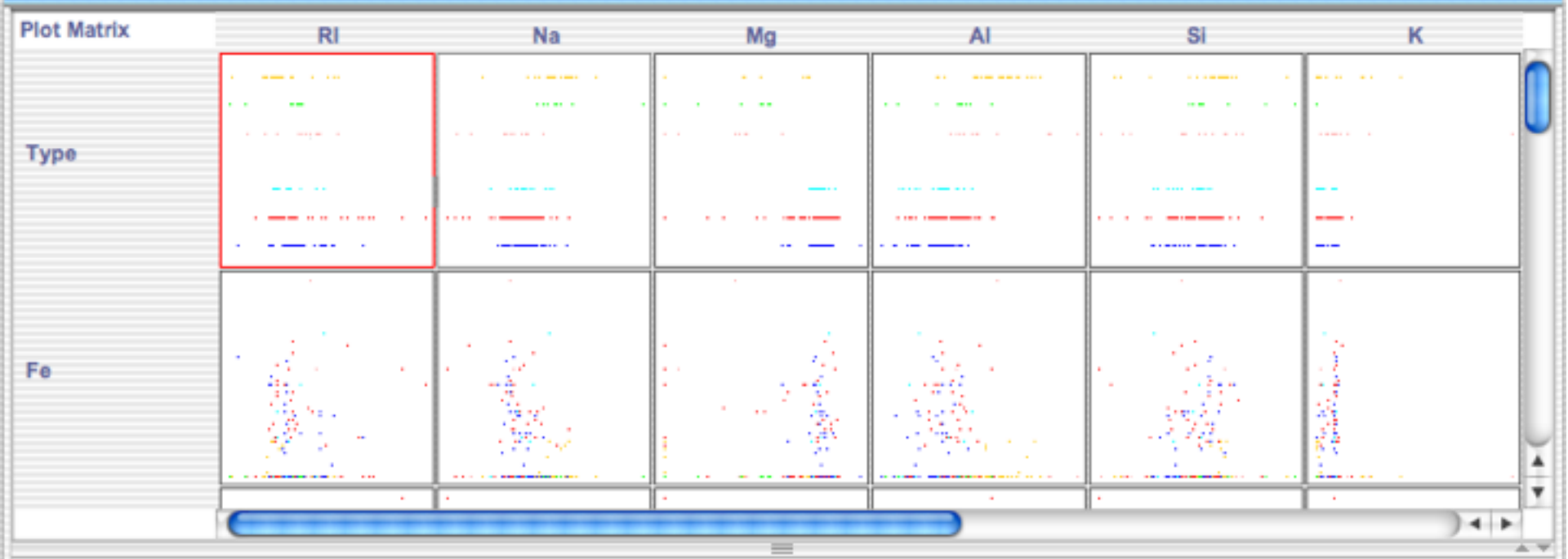
Status

OK

Log | x 0

Weka Knowledge Explorer

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



PlotSize: [100]

PointSize: [1]

Jitter:

Update

Select Attributes

Colour: Type (Nom)

SubSample % : 100

Class Colour

build wind float build wind non-float vehic wind float vehic wind non-float containers tableware headlamps

Status

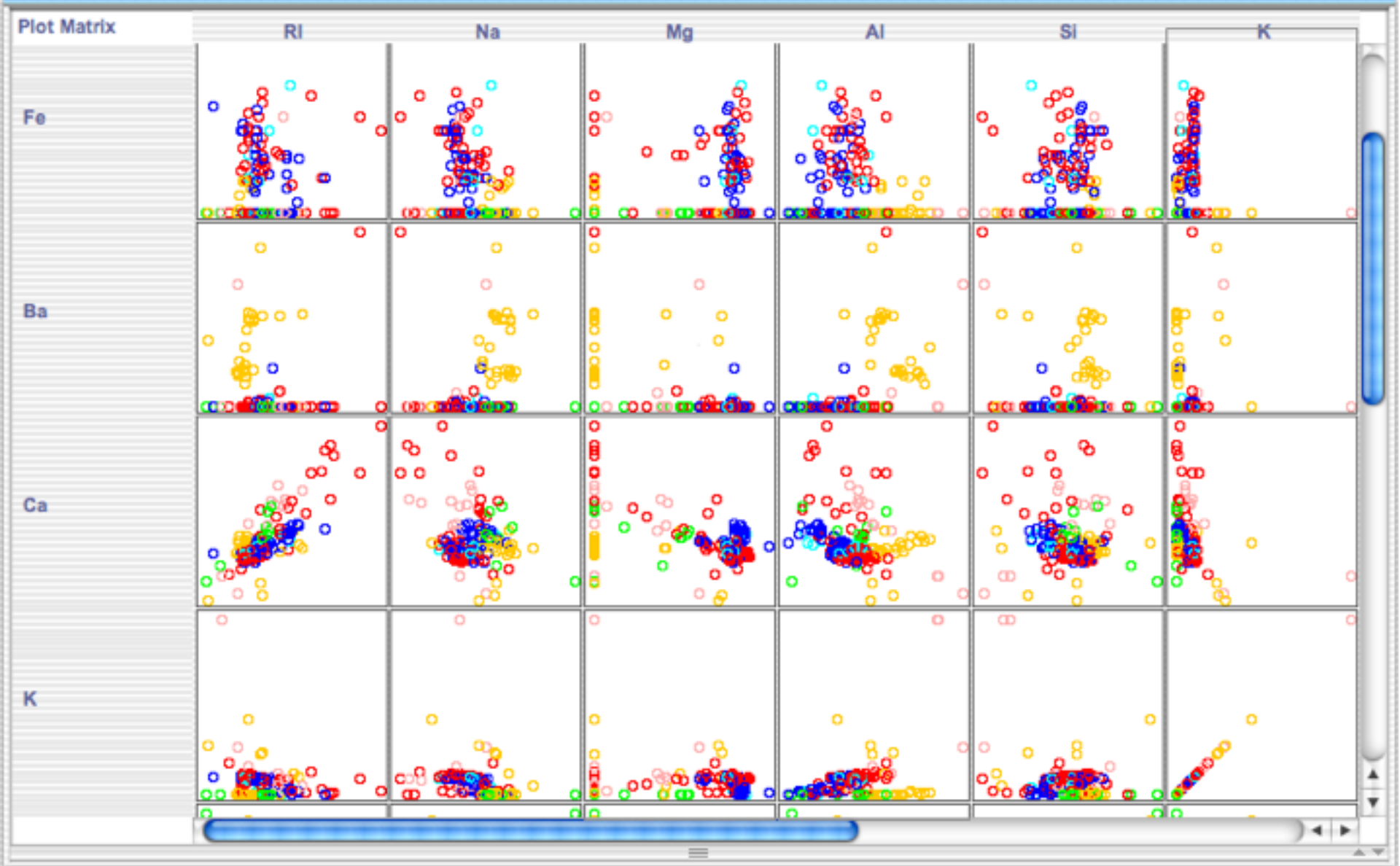
OK

Log

x 0

Weka Knowledge Explorer

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



Status

OK

Log

x 0

Weka Knowledge Explorer: Visualizing Glass

X: Al (Num)

Y: Ca (Num)

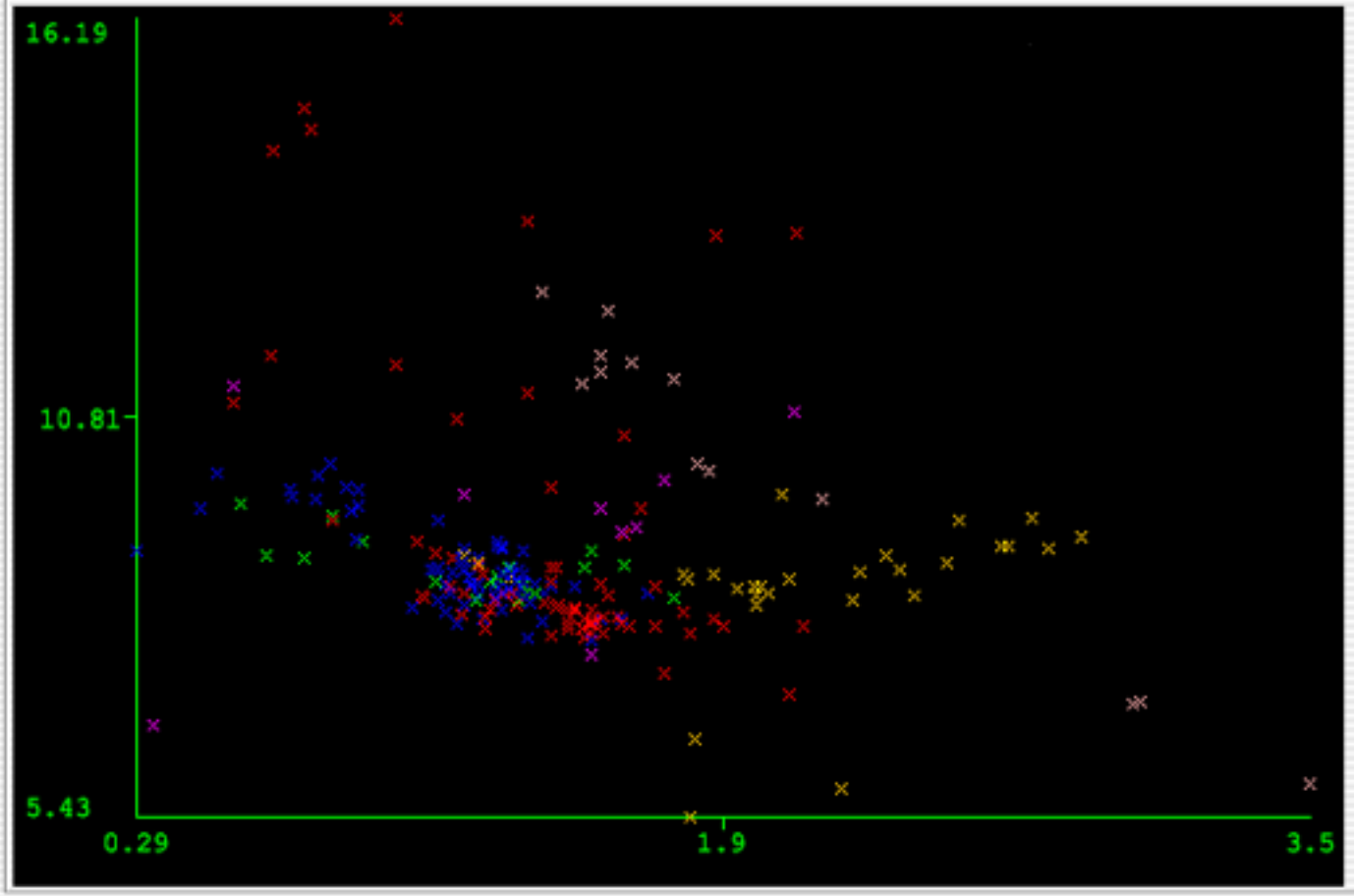
Colour: Type (Nom)

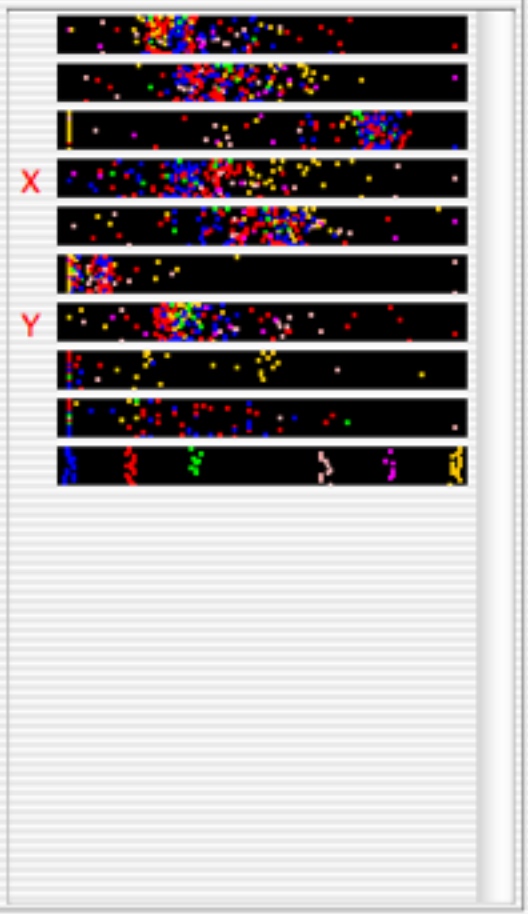
Select Instance

Reset | Clear | Save

Jitter

Plot: Glass

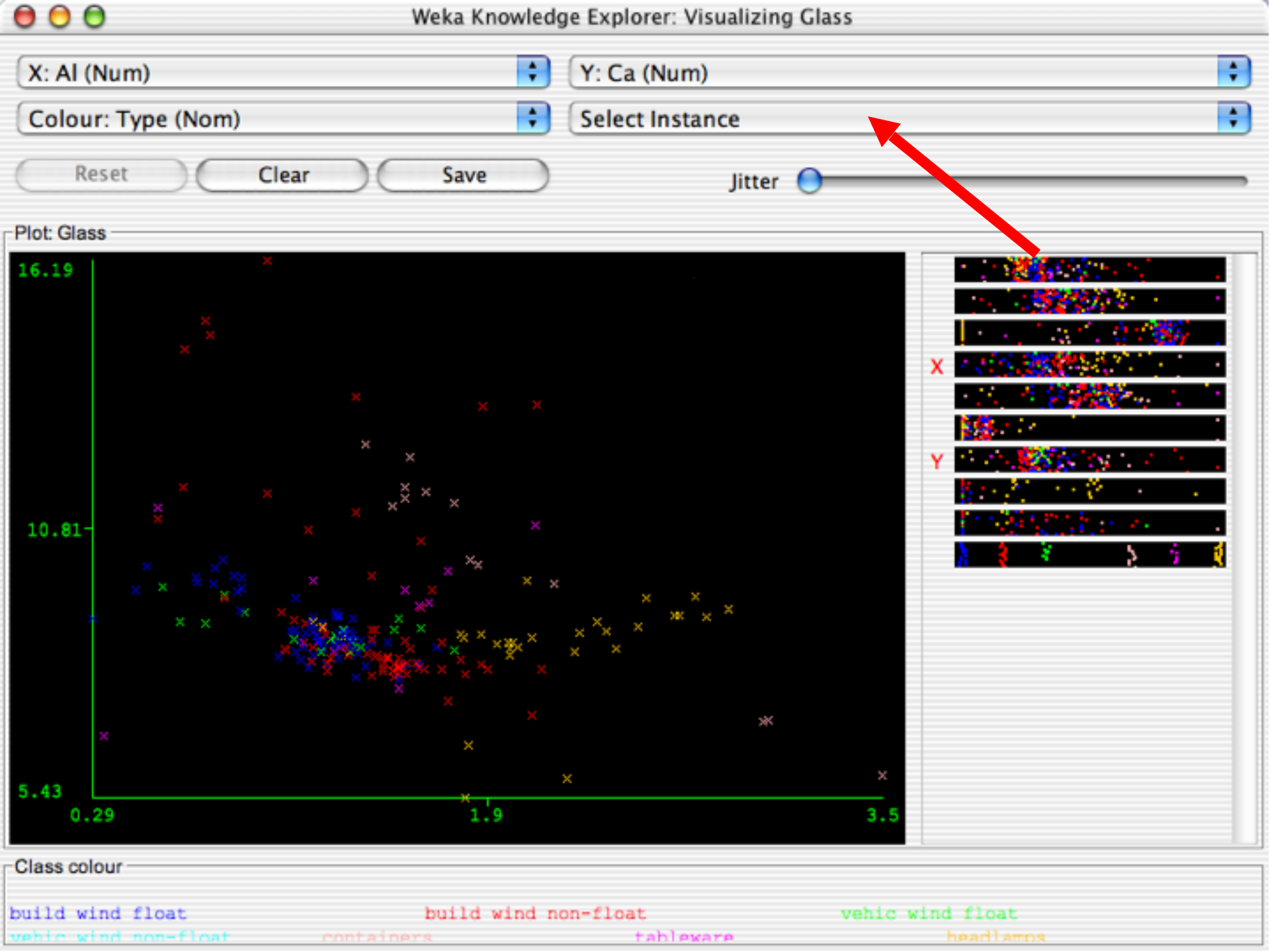




Class colour

build wind float build wind non-float vehic wind float

vehic wind non-float containers tableware headlamps


Weka Knowledge Explorer: Visualizing Glass
X: Al (Num)
Y: Ca (Num)
Colour: Type (Nom)
Select Instance
Reset
Clear
Save
Jitter
Plot: Glass
16.19
10.81
5.43
0.29
1.9
3.5
X
Y
Class colour
build wind float
build wind non-float
vehic wind float
vehic wind non-float
containers
tableware
headlamps

Weka Knowledge Explorer: Visualizing Glass

X: Al (Num)

Y: Ca (Num)

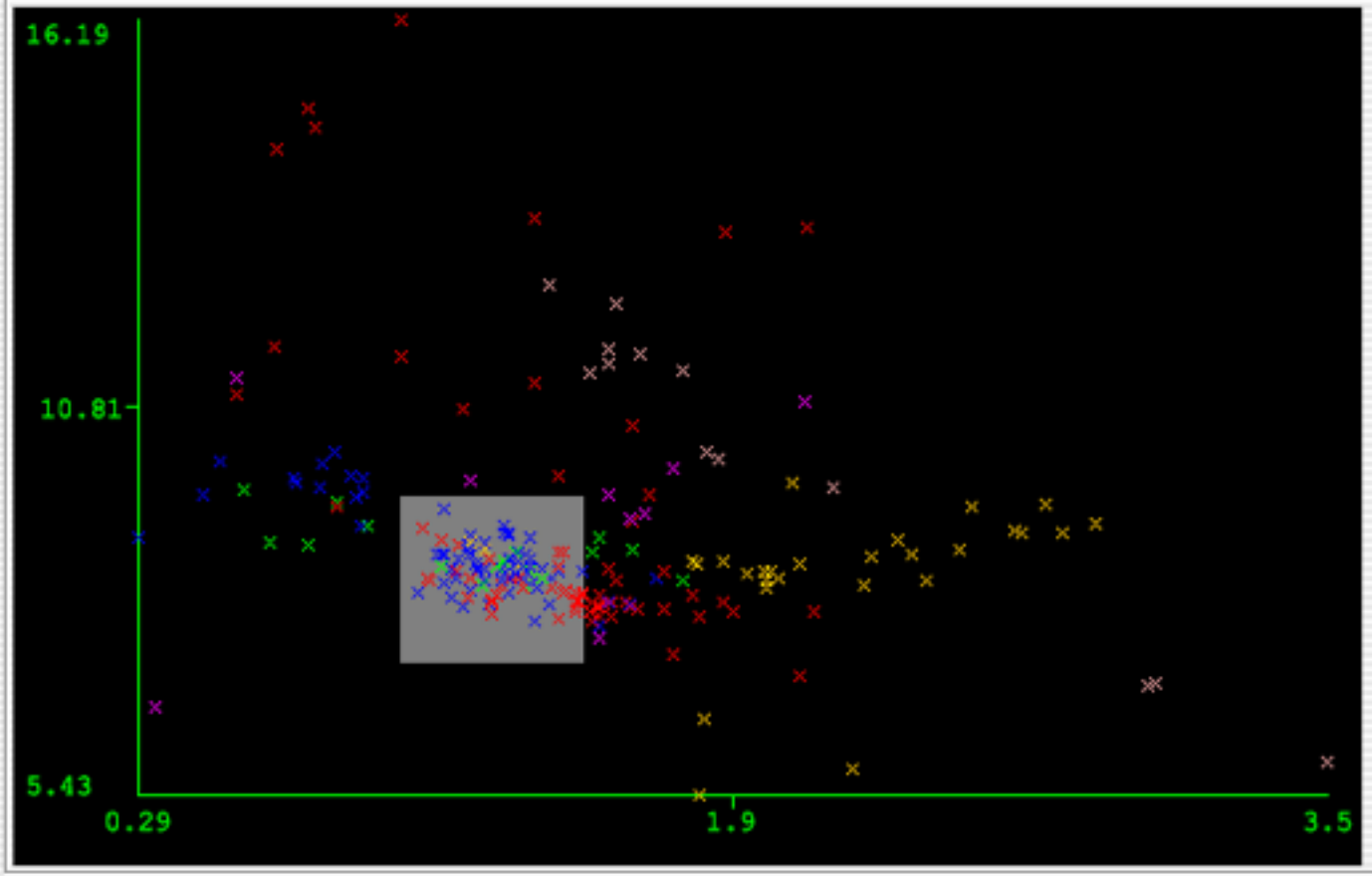
Colour: Type (Nom)

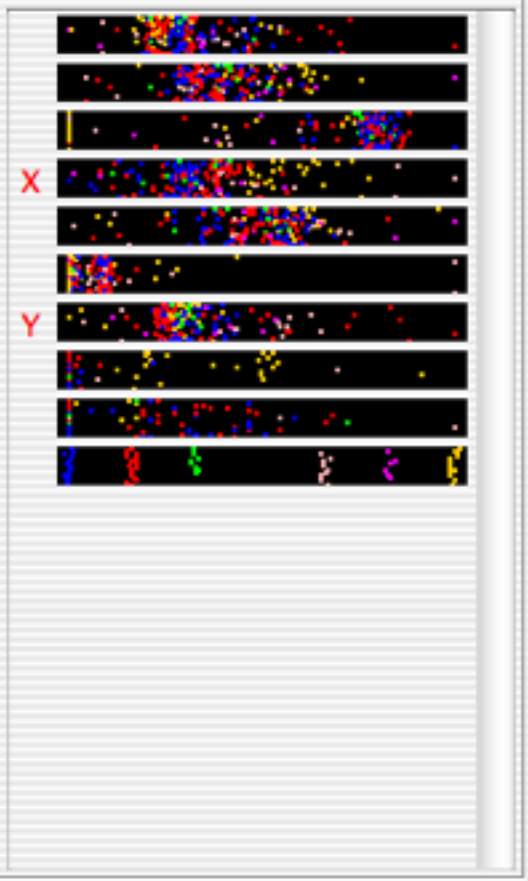
Rectangle

Submit Clear Save

Jitter

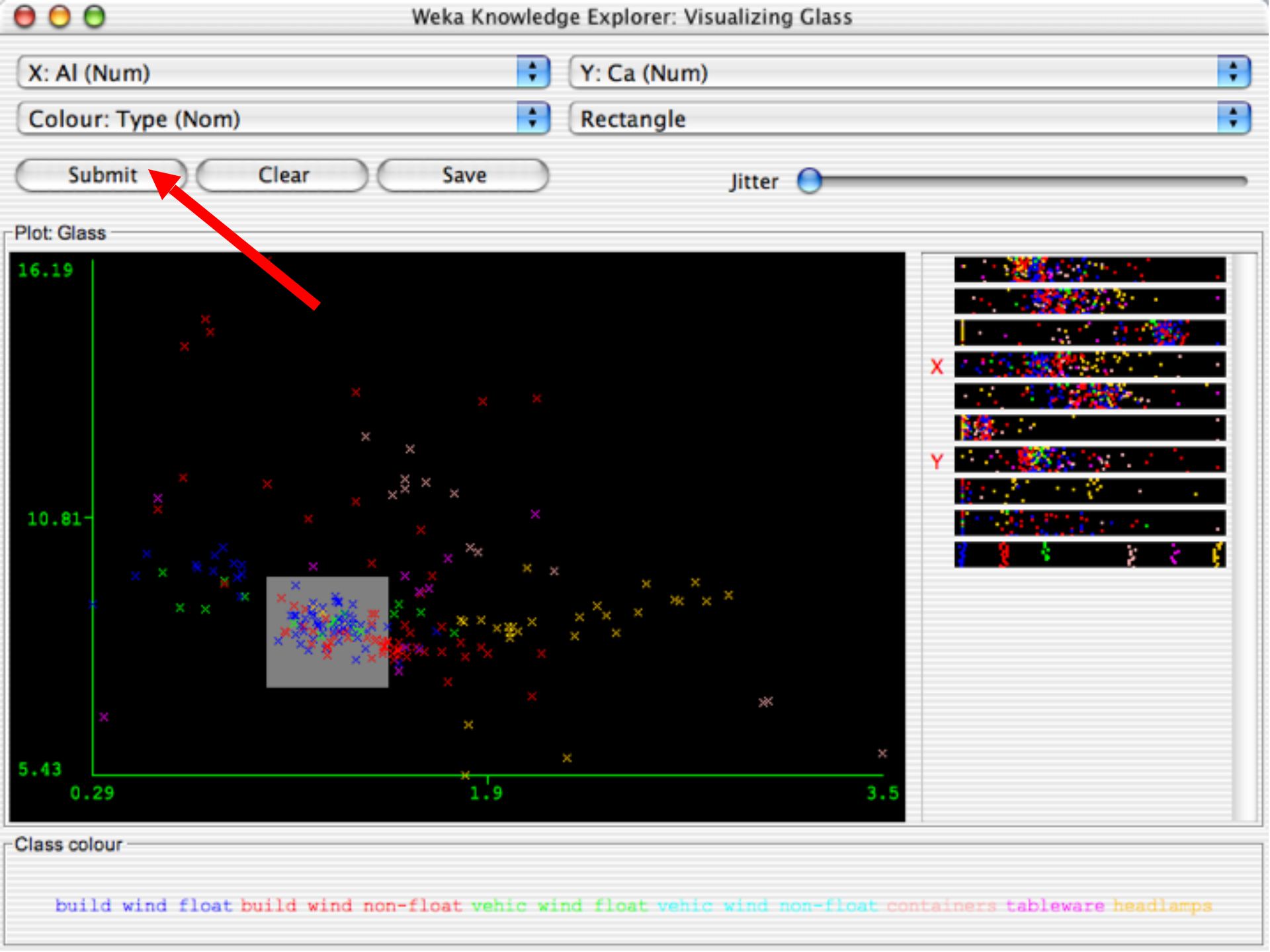
Plot: Glass





Class colour

build wind float build wind non-float vehic wind float vehic wind non-float containers tableware headlamps


Weka Knowledge Explorer: Visualizing Glass
X: Al (Num)
Y: Ca (Num)
Colour: Type (Nom)
Rectangle
Submit
Clear
Save
Jitter
Plot: Glass
16.19
10.81
5.43
0.29
1.9
3.5
X
Y
Class colour
build wind float
build wind non-float
vehic wind float
vehic wind non-float
containers
tableware
headlamps

Weka Knowledge Explorer: Visualizing Glass

X: Al (Num)

Y: Ca (Num)

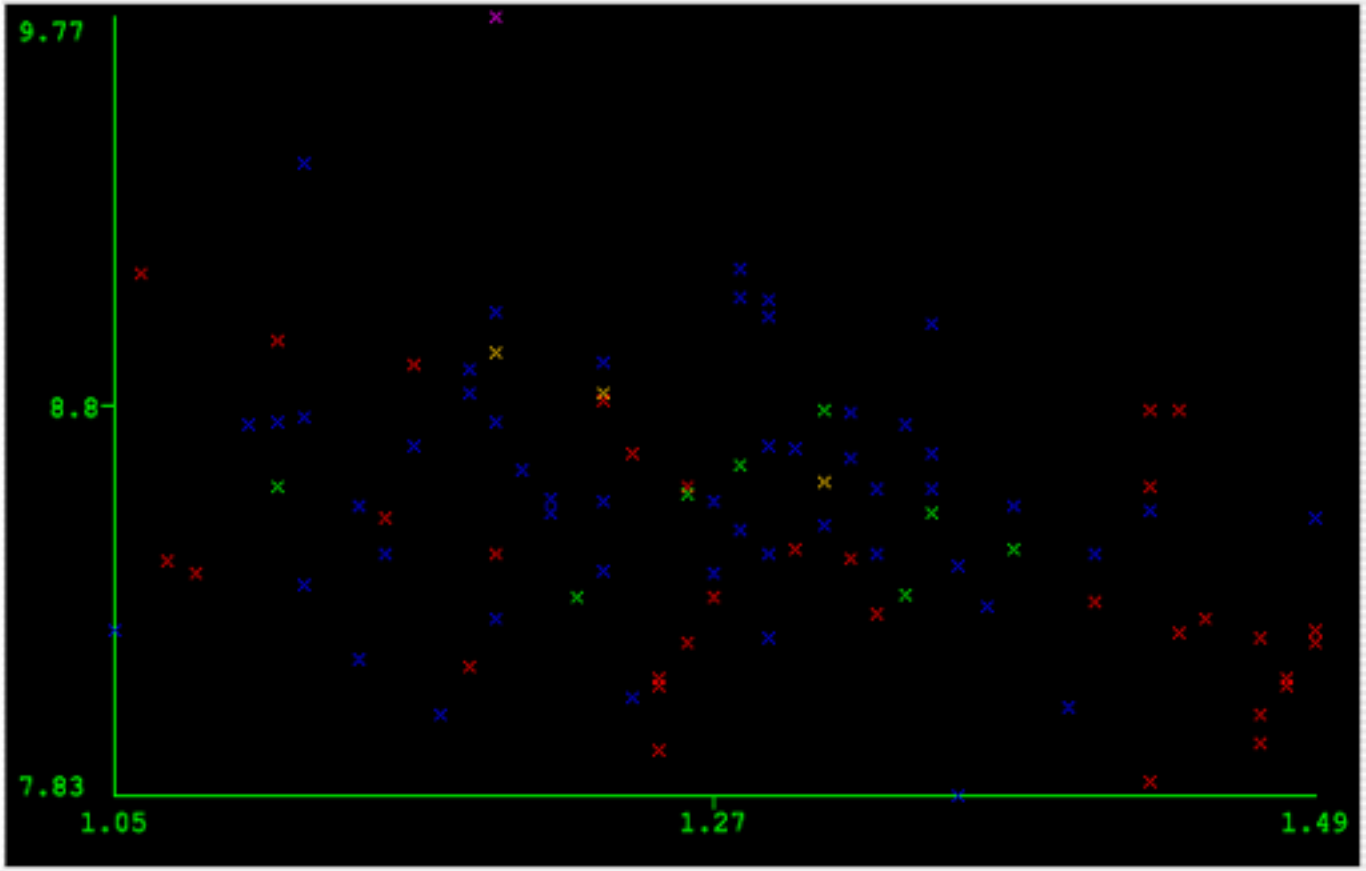
Colour: Type (Nom)

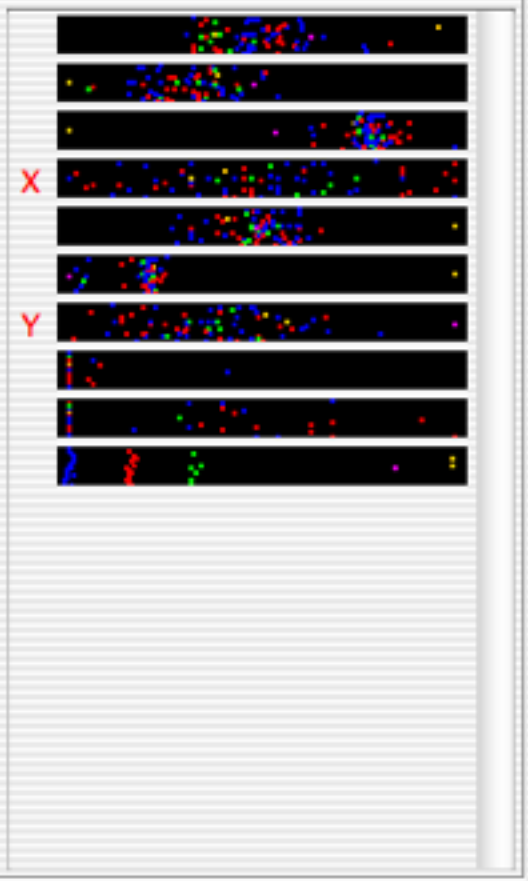
Rectangle

Reset | Clear | Save

Jitter

Plot: Glass





Class colour

build wind float
build wind non-float
vehic wind float
vehic wind non-float
containers
tableware
headlamps

References and Resources

- **References:**
  - WEKA website: http://www.cs.waikato.ac.nz/~ml/weka/index.html
  - WEKA Tutorial:
    - Machine Learning with WEKA: A presentation demonstrating all graphical user interfaces (GUI) in Weka.
    - A presentation which explains how to use Weka for exploratory data mining.
  - WEKA Data Mining Book:
    - Ian H. Witten and Eibe Frank, Data Mining: Practical Machine Learning Tools and Techniques (Second Edition)
  - WEKA Wiki: http://weka.sourceforge.net/wiki/index.php/Main_Page