

WEKA Instructions

Where to download?

Download the self-extracting executable from <http://www.cs.waikato.ac.nz/ml/weka/>

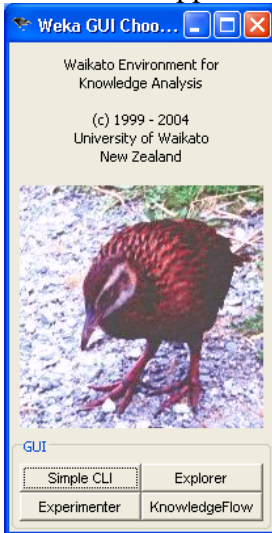
How to install?

Double click on the exe file and just follow the instructions.

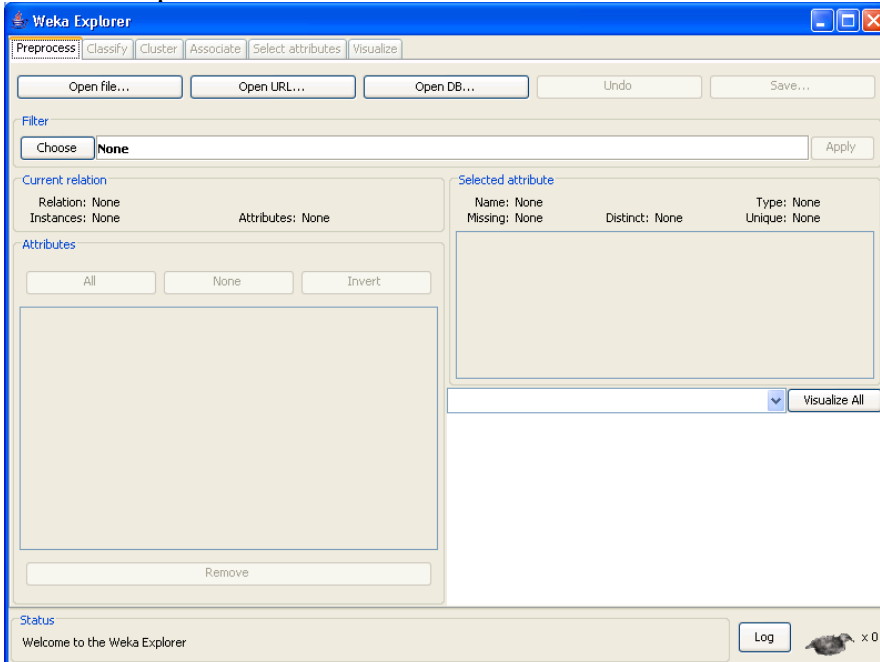
Start up the application

Go to Start->All Programs -> WEKA -> weka3-4

This should appear:

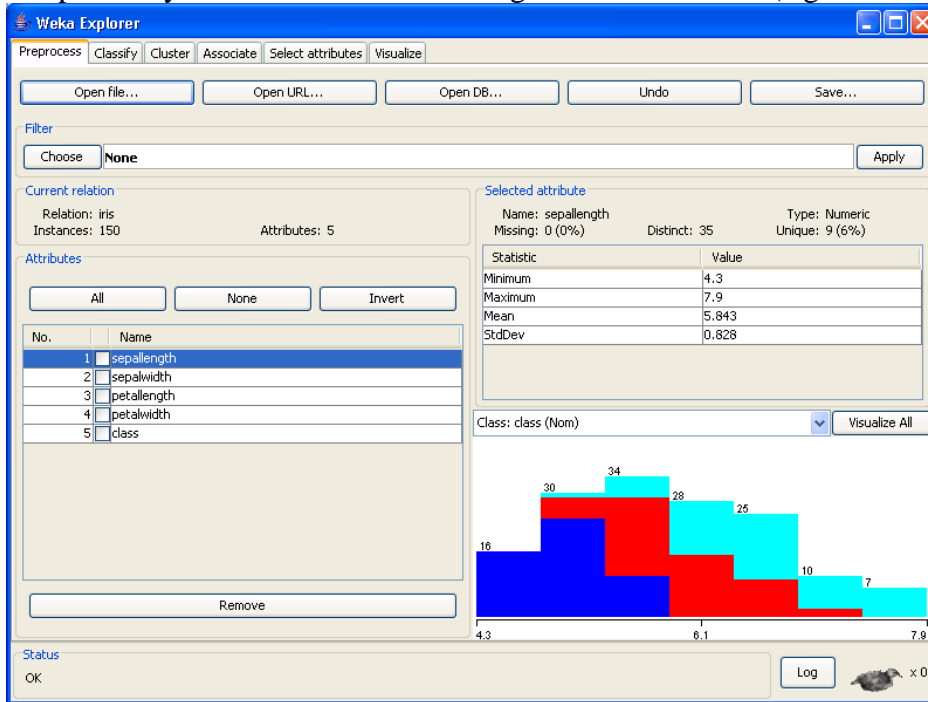


Click on Explorer

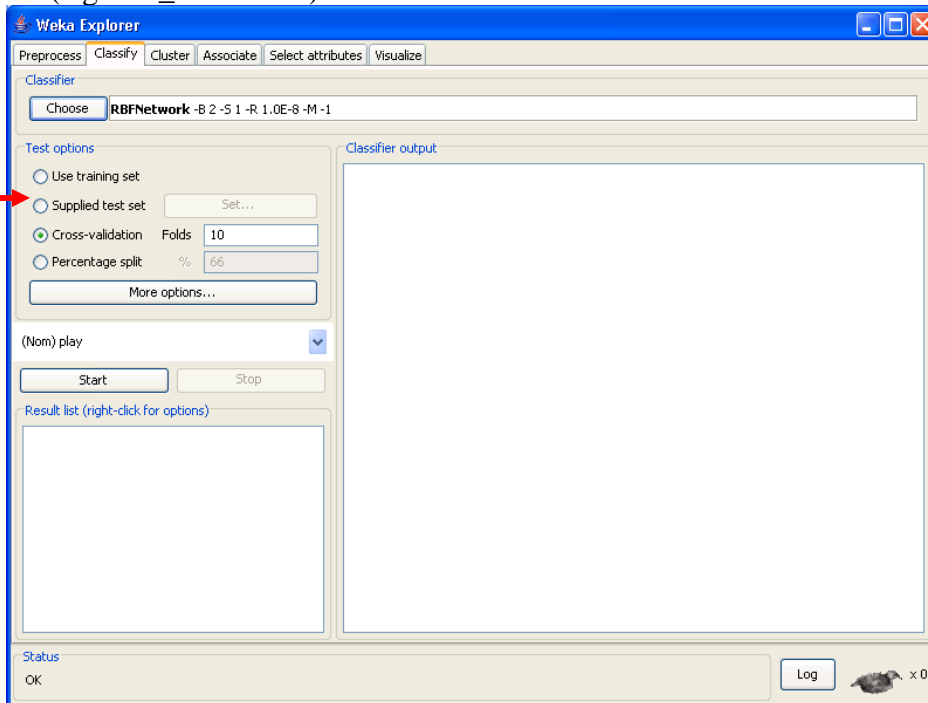


Open a training dataset by clicking on the Open File button in the 'Preprocess' tab.

We provide you with 3 different training sets in arff format (e.g. train_caldor.arff).



To classify the data, click on the Classify tab. It comes up with Cross-validation selected. Change that to Supplied test set. Click on Set, and it will ask you to choose the test data set (e.g. test_caldor.arff).

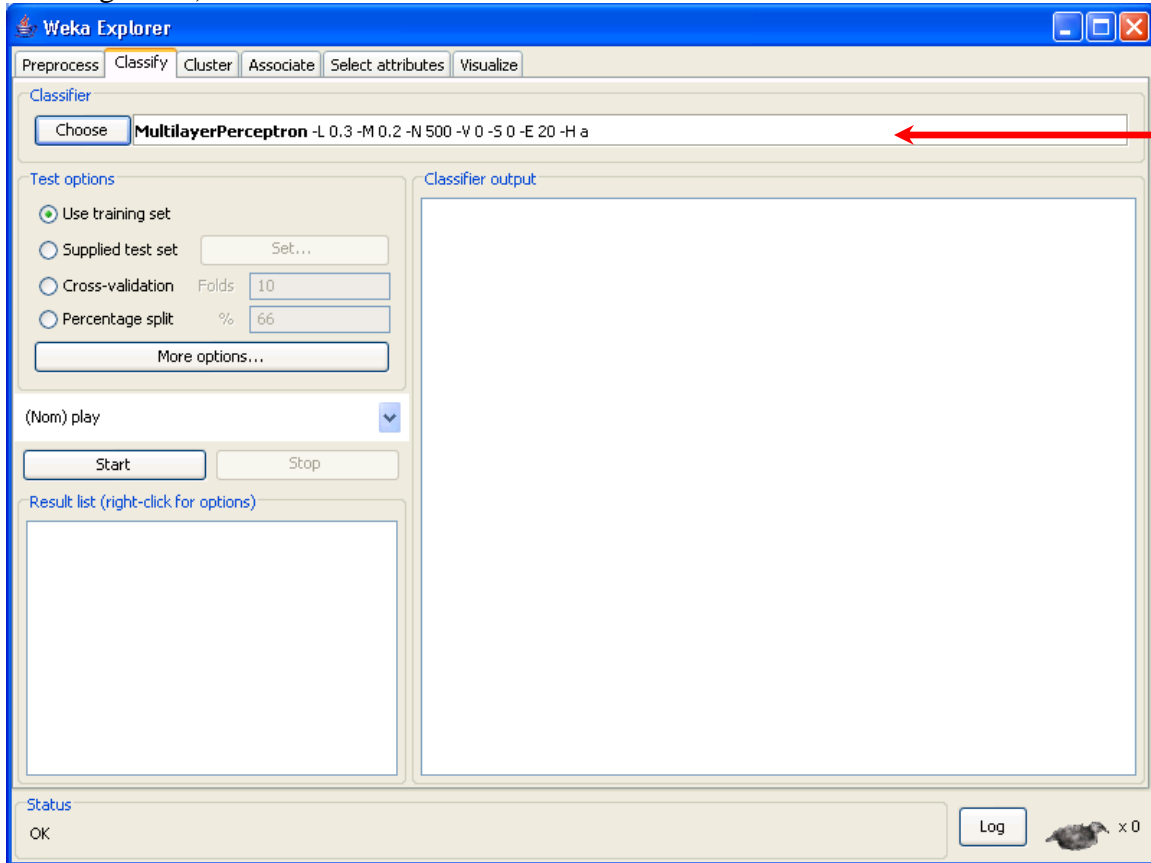


Choose a classifier.

Click on choose:

- for decision tree click on trees-> REPtrees
- for neural net click on functions-> multilayer perceptron

To change the parameters of the classifier right click on the classifier name (not required for assignment).



To start the testing, click Start. Note that the multi-layered perceptron trains very slowly. The bird in the lower right shows you that it is still executing. Eventually the results appear. The decision tree I tried (RepTree) was much faster.

The screenshot shows the Weka Explorer interface with the MultilayerPerceptron classifier selected. The classifier output window displays the following statistics:

```

Kappa statistic           0.96
Mean absolute error      0.0327
Root mean squared error  0.1291
Relative absolute error   7.3555 %
Root relative squared error 27.3796 %
Total Number of Instances 150
  
```

Below these statistics is a table titled "Detailed Accuracy By Class":

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
1	0	1	1	1	Iris-setosa
0.96	0.02	0.96	0.96	0.96	Iris-versicolor
0.96	0.02	0.96	0.96	0.96	Iris-virginica

At the bottom, the Confusion Matrix is shown:

```

=== Confusion Matrix ===
 a b c <-- classified as
50 0 0 | a = Iris-setosa
 0 48 2 | b = Iris-versicolor
 0 2 48 | c = Iris-virginica
  
```

The output gives you a lot of statistics. Report on the correct and incorrect classification rates and the confusion matrix.