Decision trees with Weka

Lab 1.1

Using Command Line Interface

CLI. ID3 decision tree

Build decision tree with Id3 algorithm (no numeric values, splits are based on information gain):

java weka.classifiers.trees.Id3 -t data/weather.nominal.arff

Build decision tree with J48 algorithm (allows numeric values, uses gain ratio, pruning):

java weka.classifiers.trees.J48 -t data/weather.arff

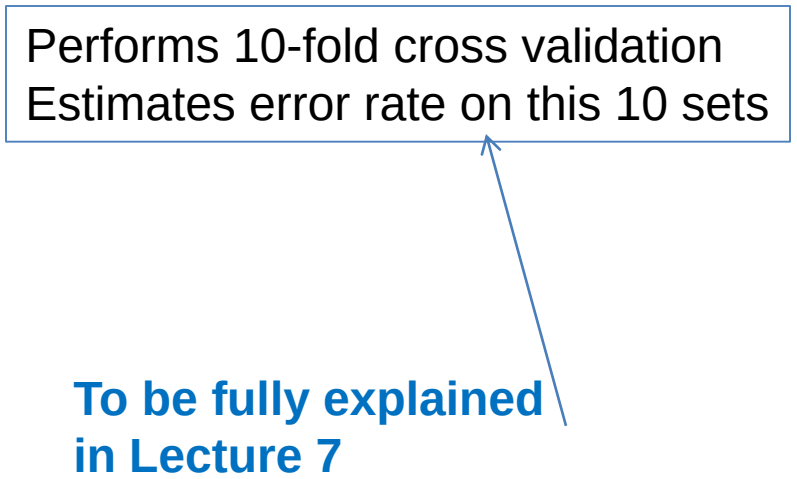
Save model to disk:

java weka.classifiers.trees.Id3 -t data/weather.nominal.arff -d data/Id3.model

CLI: J48 algorithm

Run decision tree induction algorithm J48 (allows numeric values, uses gain ratio, pruning):

java weka.classifiers.trees.J48 -t data/weather.arff



Classify a new record

- From the original arff file create a new file test1.arff where in data section put the records to classify, and put ? for the class value:

```
@relation weather.symbolic

@attribute outlook {sunny, overcast, rainy}
@attribute temperature {hot, mild, cool}
@attribute humidity {high, normal}
@attribute windy {TRUE, FALSE}
@attribute play {yes, no}

@data
sunny,hot,high,FALSE,?
sunny,hot,high,TRUE,?
```

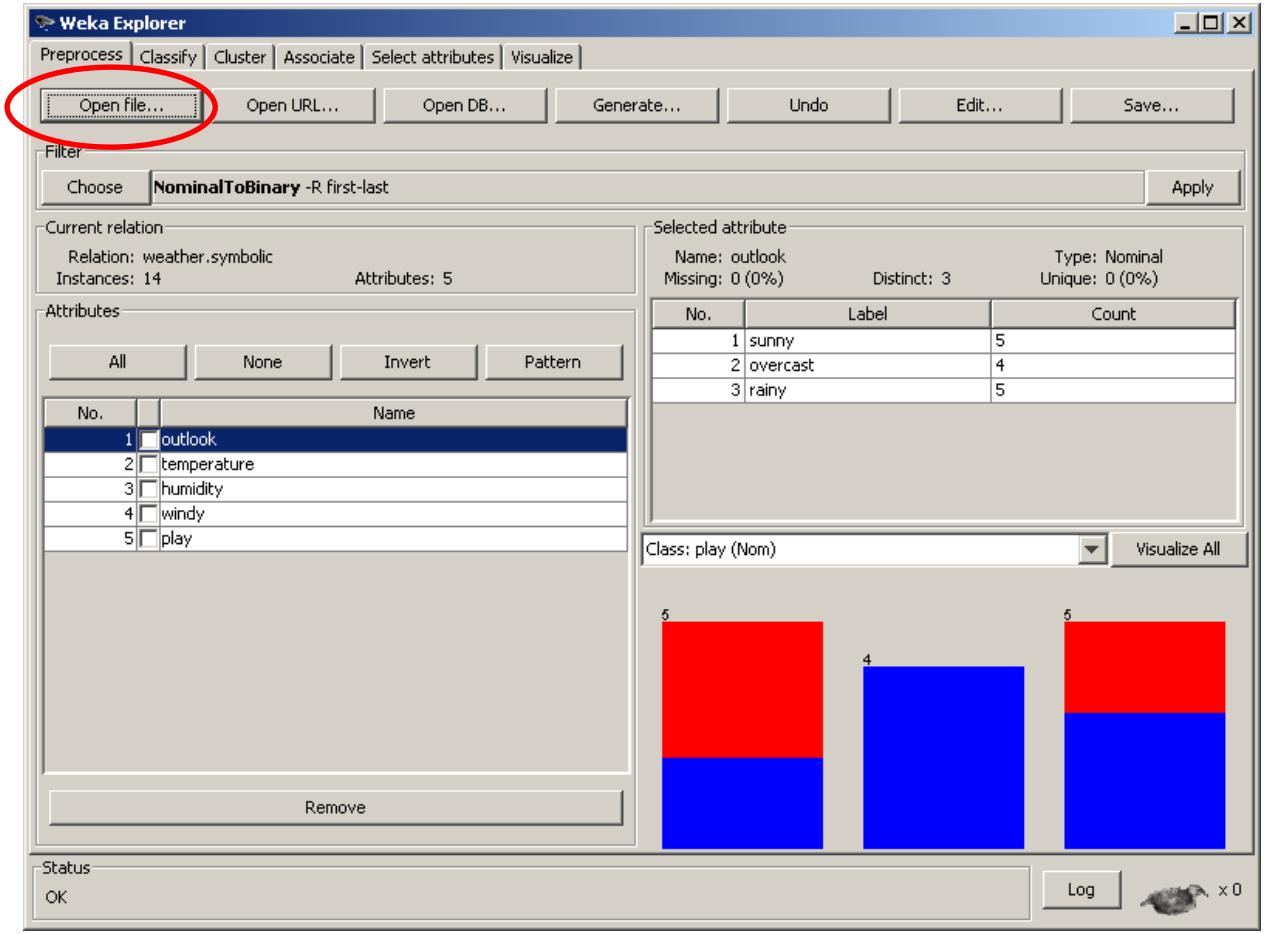
CLI: classify records

```
java weka.classifiers.trees.Id3 -T data/test1.arff -l
    models/Id3.model -p 0
```

Using GUI

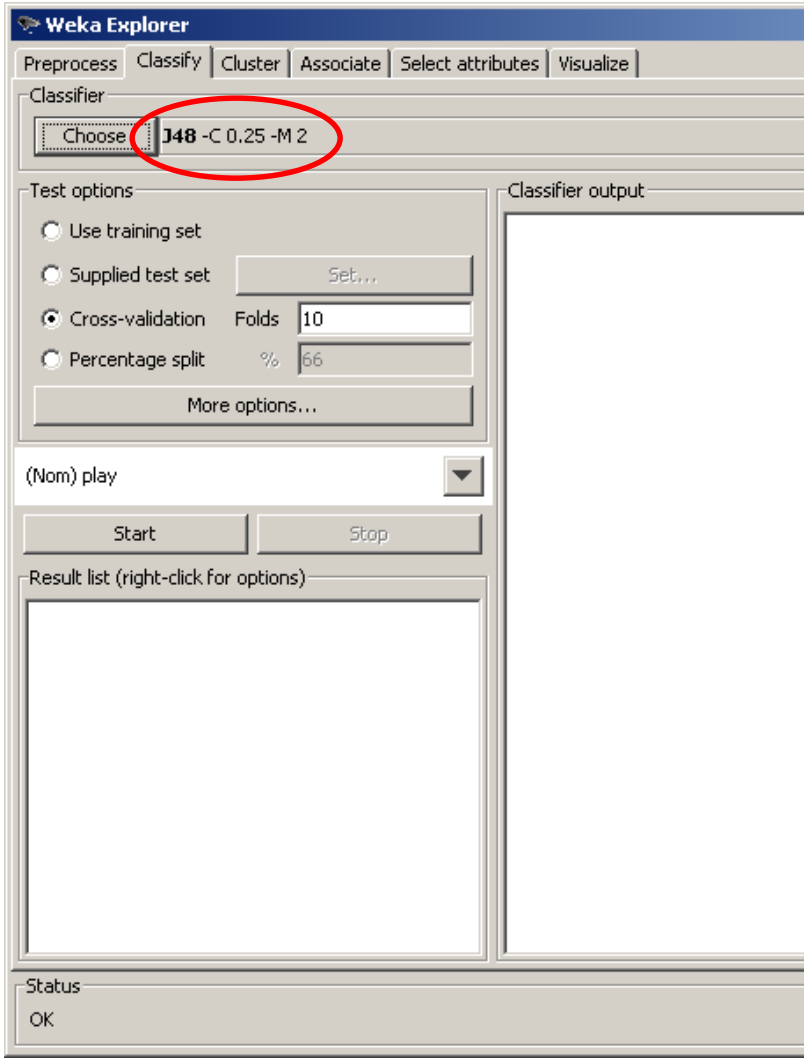
Weka explorer

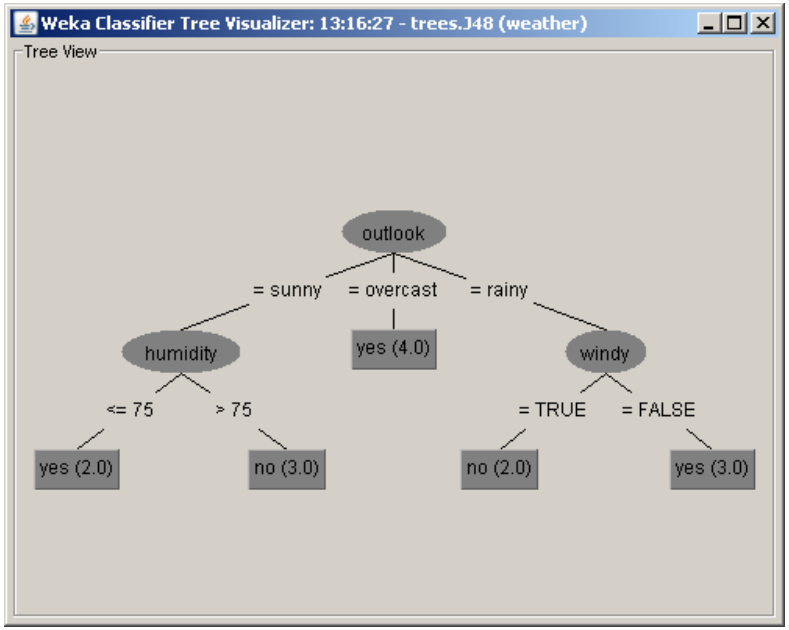
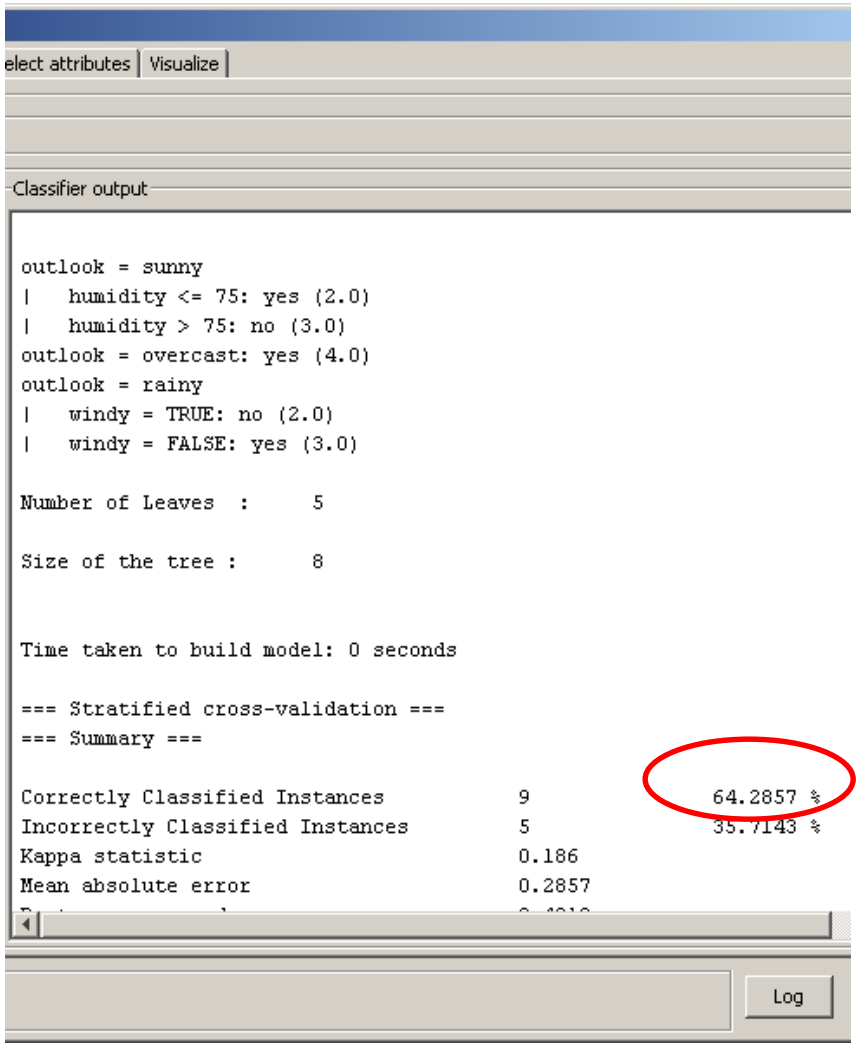
- Load file weather.arff
- Explore attributes



Weka explorer

- Run decision tree classifiers:
  - Id3 (simple, based on information gain): no numeric attributes – weather.nominal
  - J48 (C4.5 – gain ratio, missing values, tree pruning)
  - SimpleCart (numeric attributes)



Weka explorer

- Compare error rates
- Right-click on model line – visualize tree
- Can save the model and load it