WEKA Tutorial

Editor: Dongqing Xiao

Input File Format: .csv

Demo.csv

```
Where,When,Fred Starts,Joe offense,Joe defense,Opp C,OutCome
Home,7pm,Yes,Center,Forward,Tall,1
Home,7pm,Yes,Forward,Center,Short,1
Away,7pm,Yes,Forward,Forward,Tall,1
Home,5pm,No,Forward,Center,Tall,0
Away,9pm,Yes,Forward,Forward,Short,0
Away,7pm,No,Center,Forward,Tall,1
Home,7pm,No,Forward,Center,Tall,0
Home,7pm,Yes,Center,Center,Talls,1
Away,7pm,Yes,Center,Center,Short,1
```

Header

Data Instances

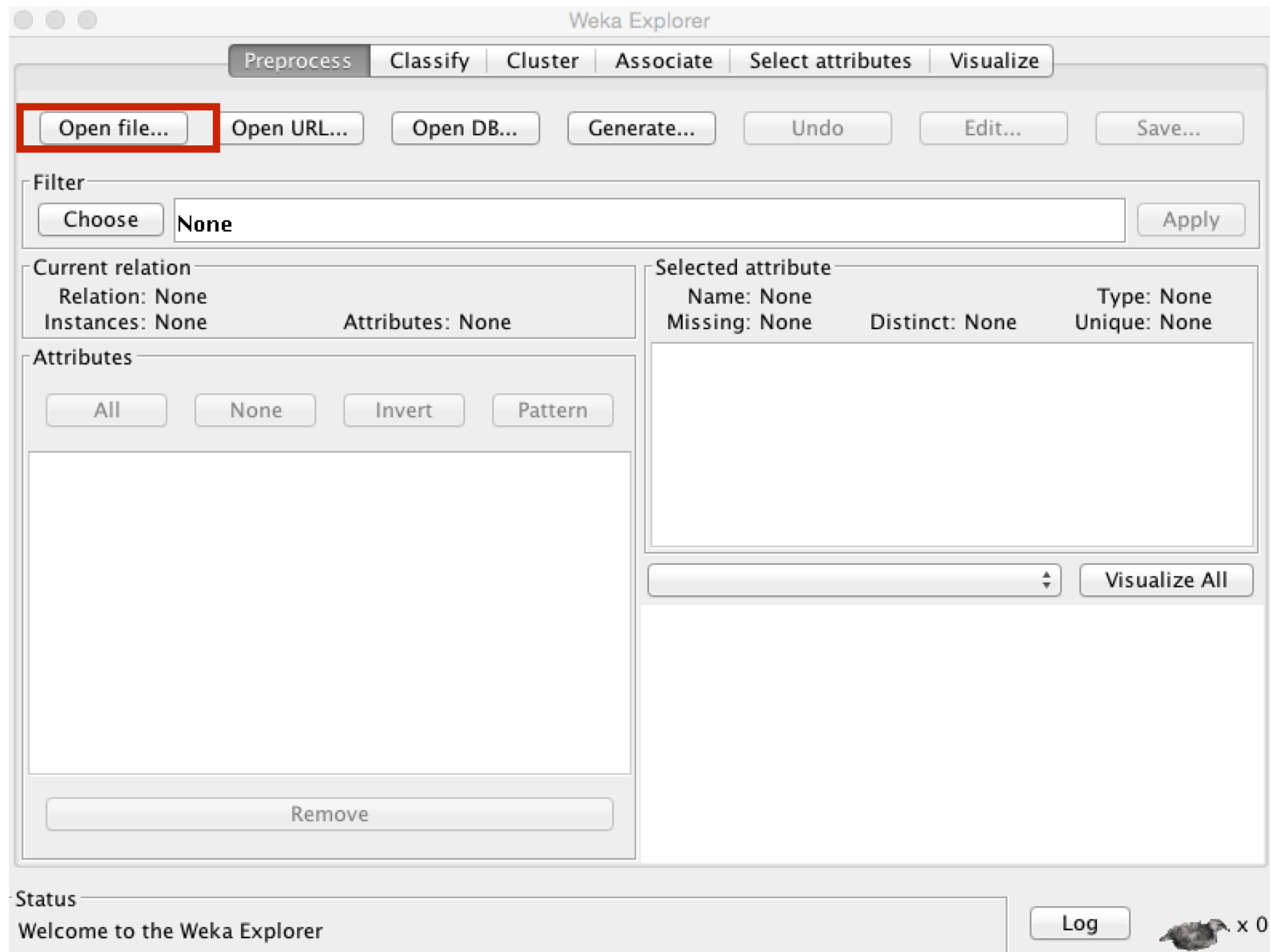Header: a list of attribute names, separated by comma","

Data Instances: each line is one data instance. A list of attribute value, separated by comma ","
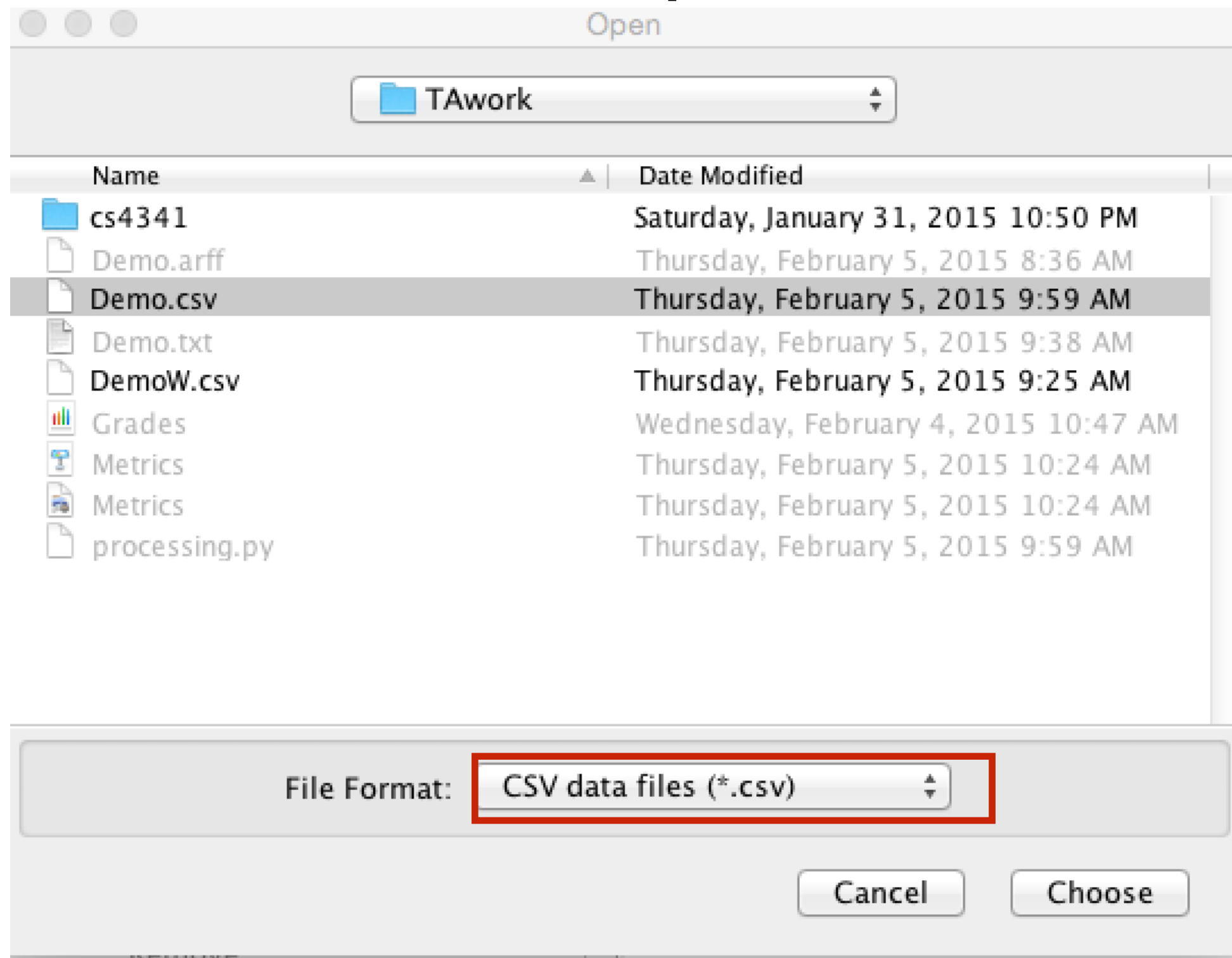
Welcome GUI

Choose the "Explorer" option. "Experimenter" is used to quickly make many trees, but as a result does not provide as much information.

Import Data

Click "Open file"

Import Data

Preprocessing Data

Check data type of each attribute

Preprocessing Data

Weka Explorer

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

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

Filter

2. Click Filter-Unsupervised-Attribute
3. Chose "Numeric to Nominal"

4. Apply Filter

Apply

weka
filters
AllFilter
MultiFilter
supervised
unsupervised
attribute
Add
AddCluster
AddExpression
AddID
AddNoise
AddValues
Center
ChangeDateFormat
ClassAssigner
ClusterMembership
Copy
Discretize
FirstOrder
InterquartileRange

Invert | Pattern
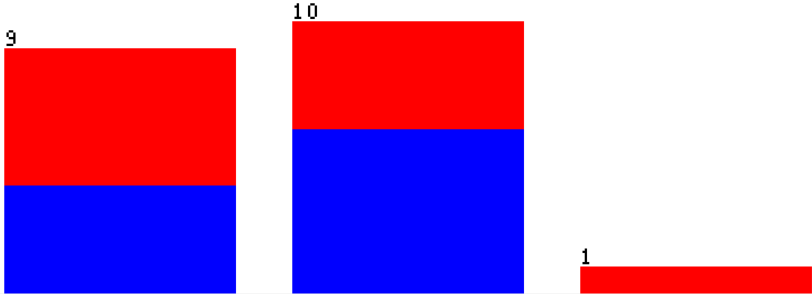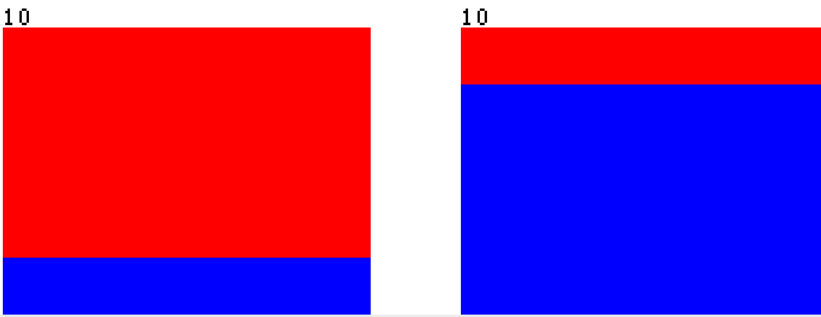
Selected attribute
Name: OutCome   Type: Numeric
Missing: 0 (0%)   Distinct: 2   Unique: 0 (0%)

Statistic	Value
Minimum	0
Maximum	1
Mean	0.5
StdDev	0.513

1. Select specific attribute/s

Filter... | Remove filter | Close

Class: OutCome (Num) | Visualize All

10 | 10

Remove

0 | 0.5 | 1

Status
OK

Log

Preprocessing Data

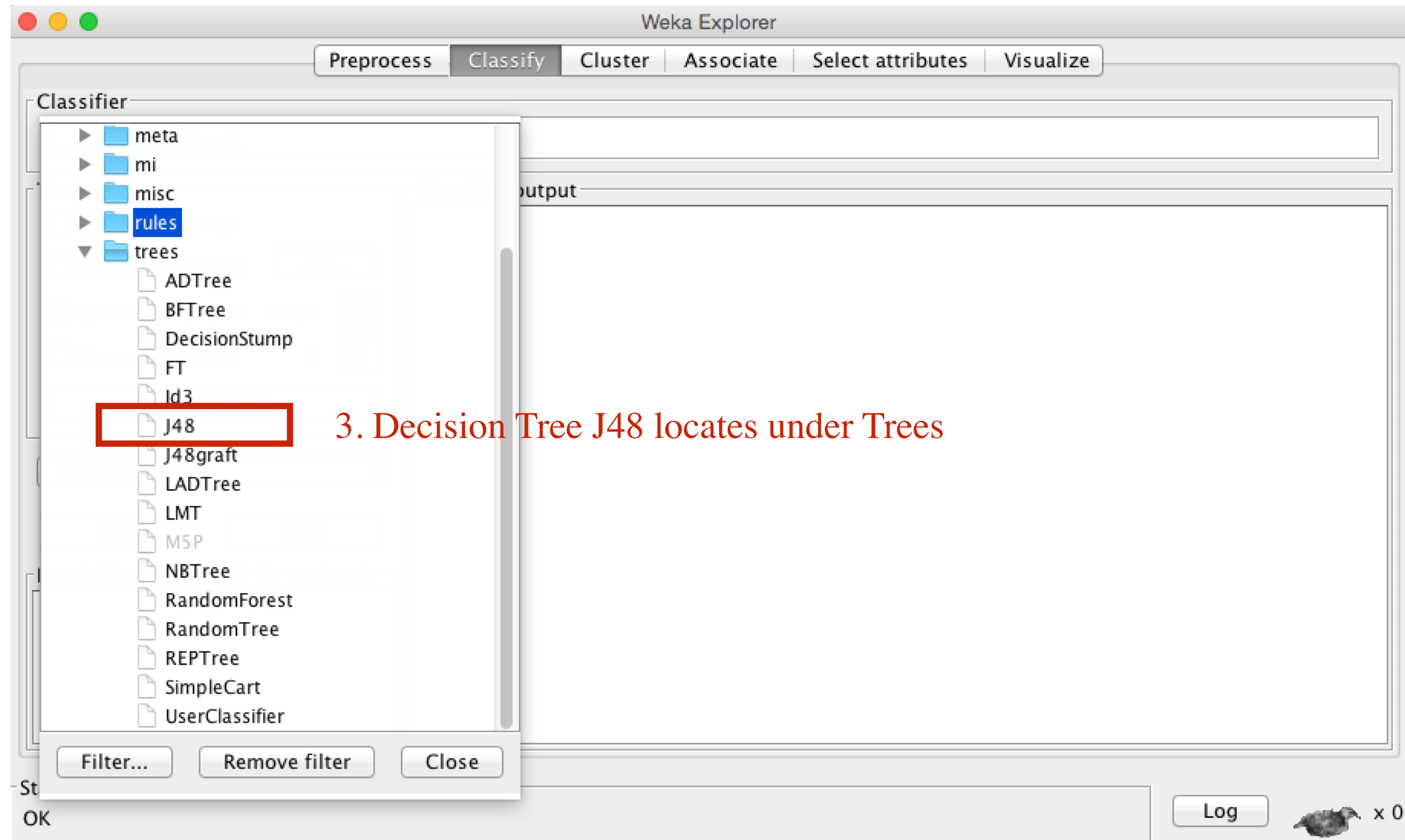Check data type of each attribute again

Classification

1. Click "Classify"

2. Click "classifier" to chose classifier

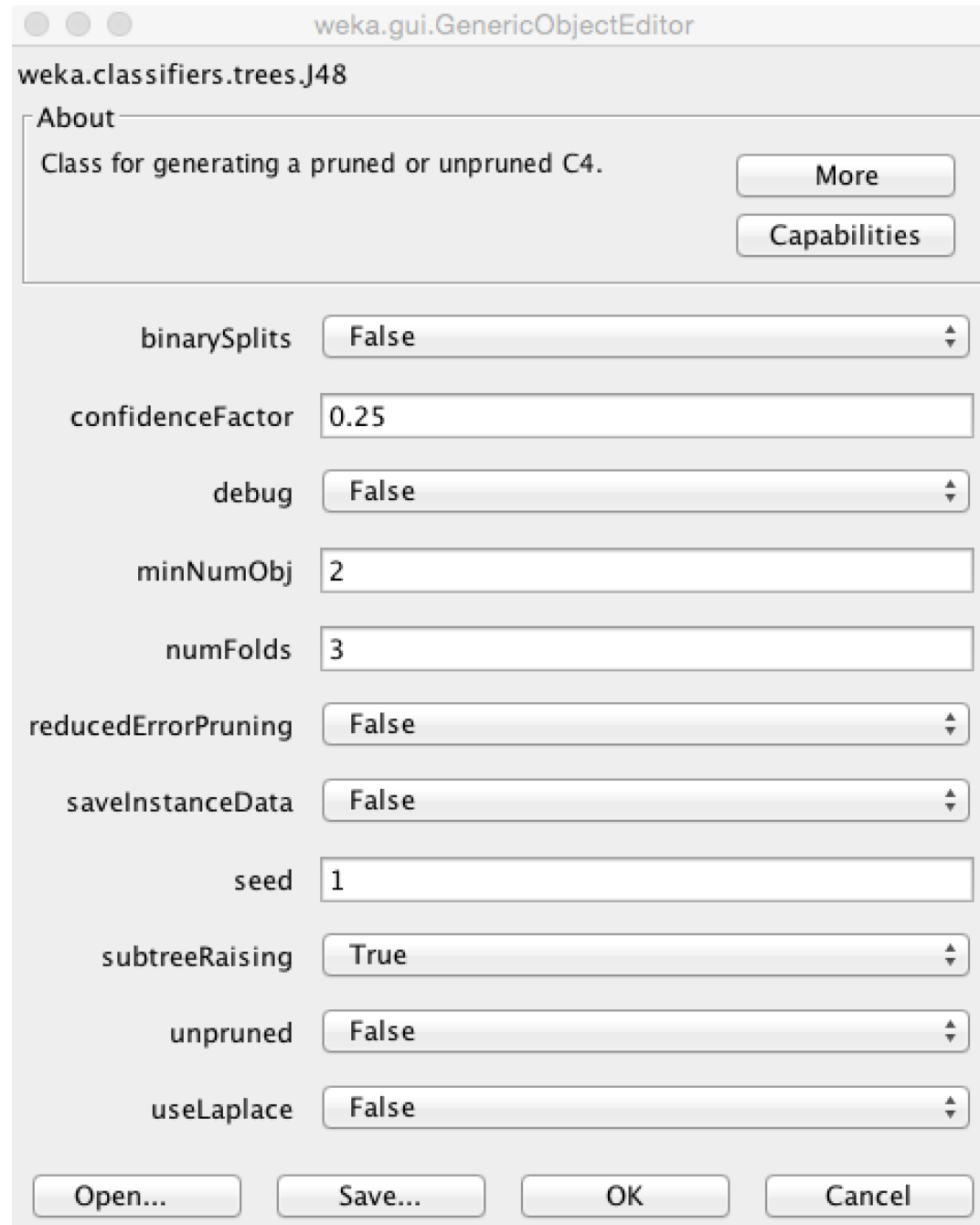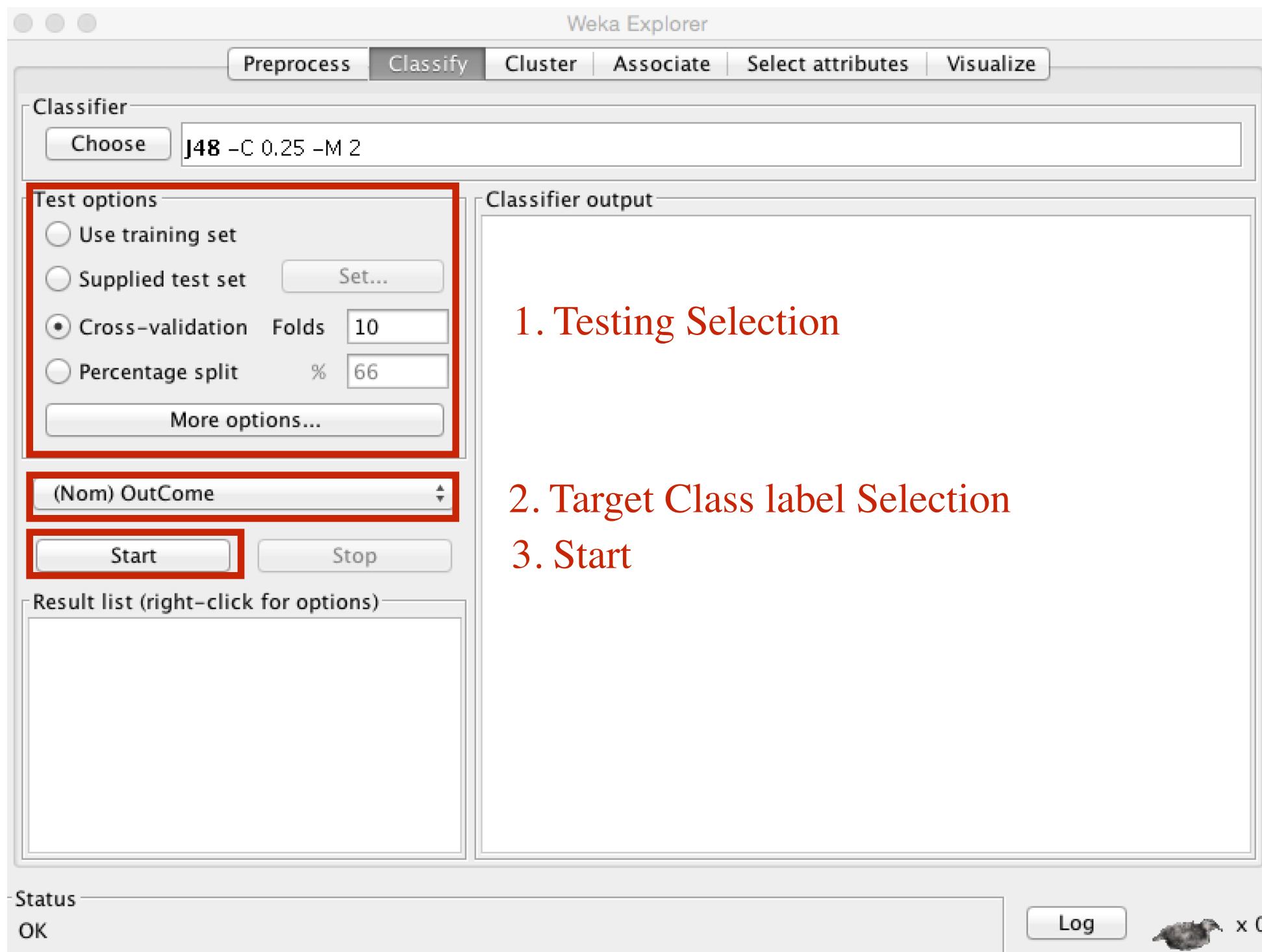3. Decision Tree J48 locates under Trees
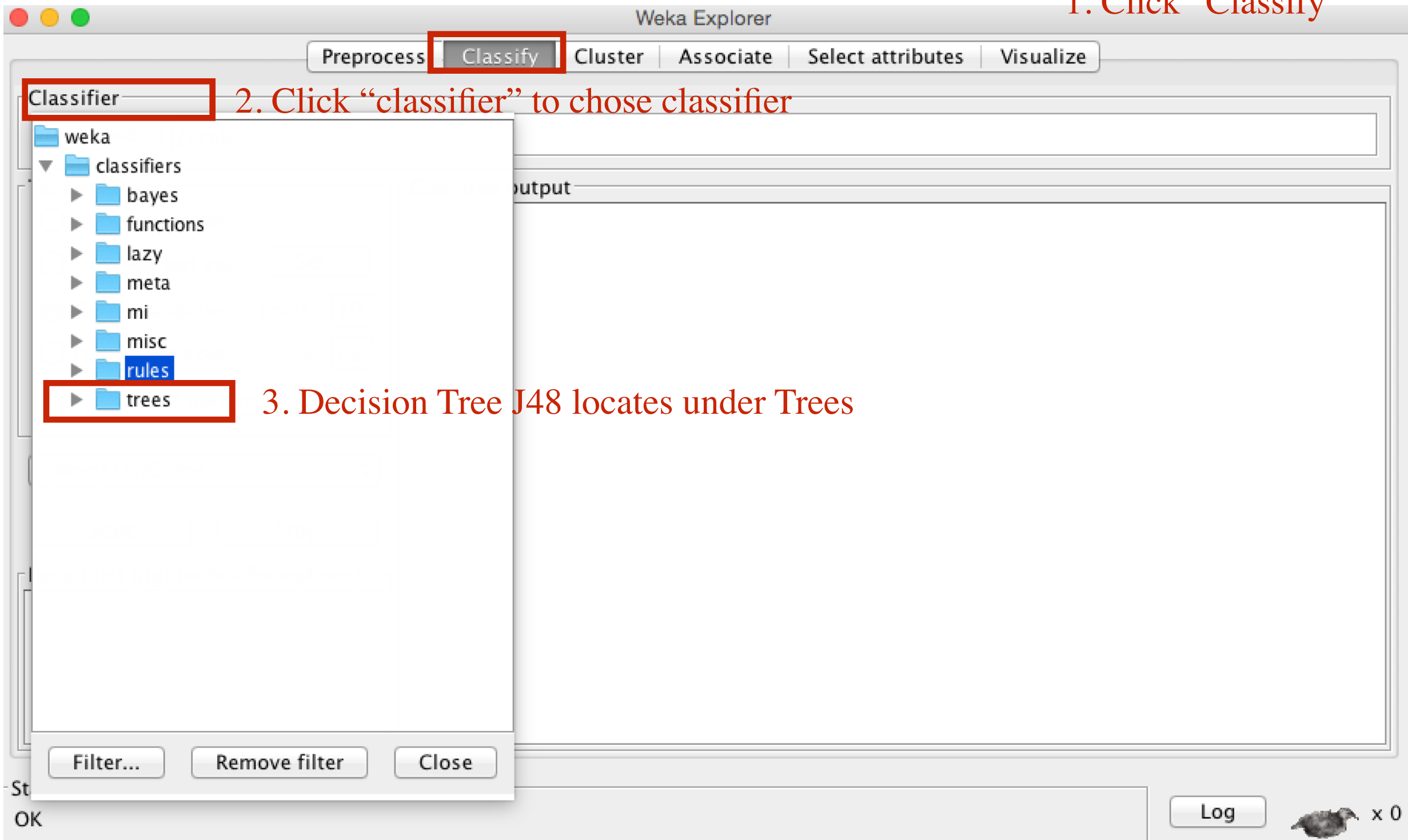
Classification-DT

Classification-DT

Before apply J48, check its configuration and change it as you want

Testing

Experiment Result

Experiment Result
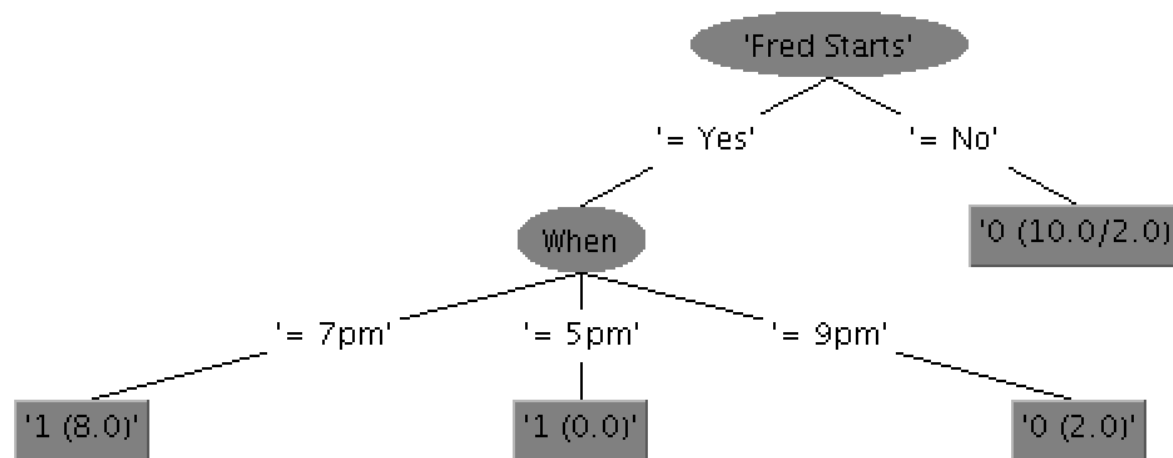
Tree Visualization

Single Predicability

To evaluate the value of "Fred Stars"

1. Select other features

2. Remove them

Single Predicability

Apply the same procedure as previous one

Undo

Use "undo" to withdraw previous remove