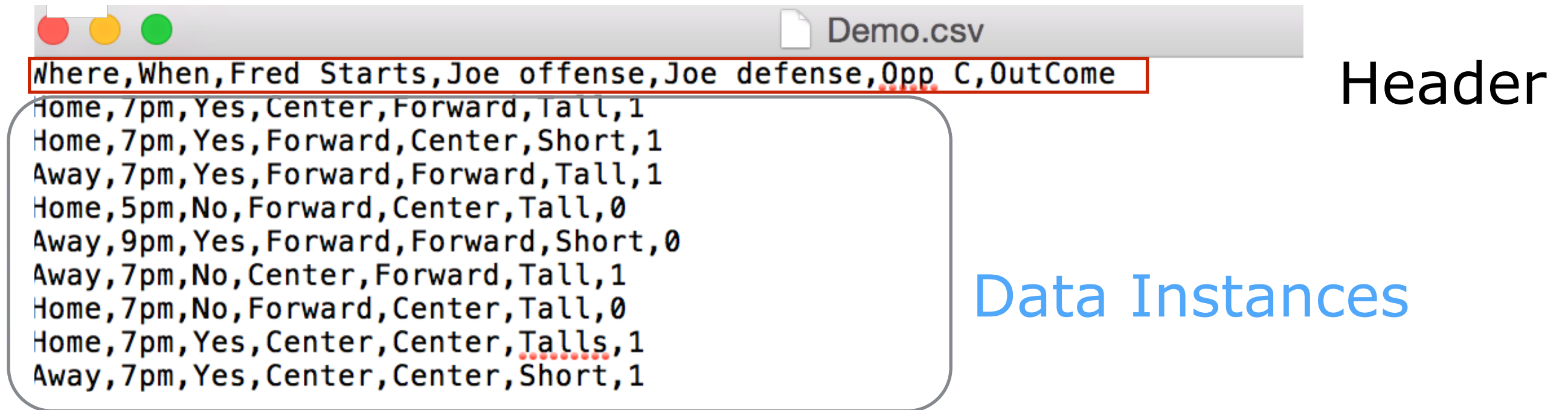


WEKA Tutorial

Editor: Dongqing Xiao

Input File Format: .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: a list of attribute names, separated by comma","

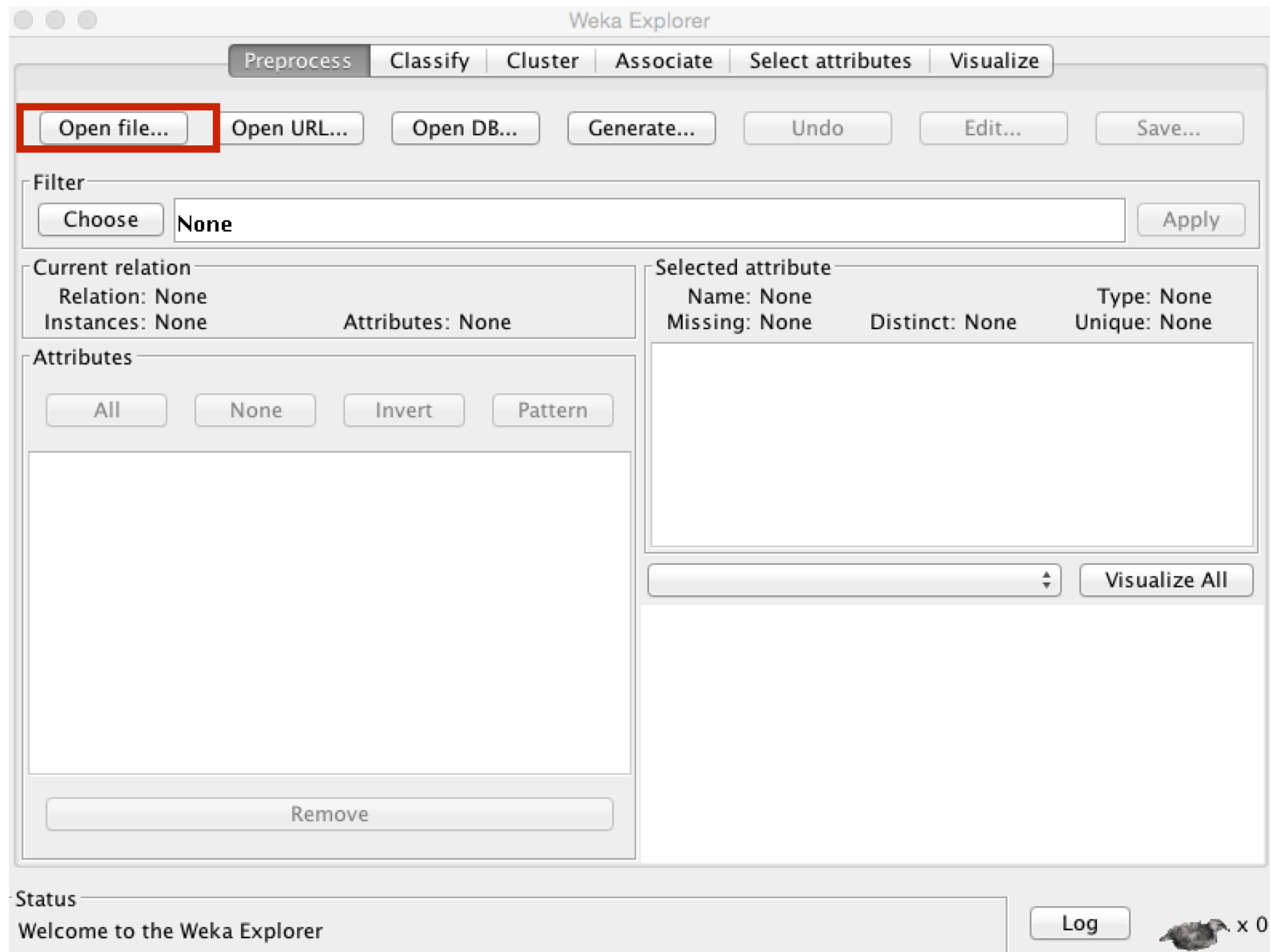
Data Instances: each line is one data instance. A list of attribute value, separated by comma ",,"

Welcome GUI



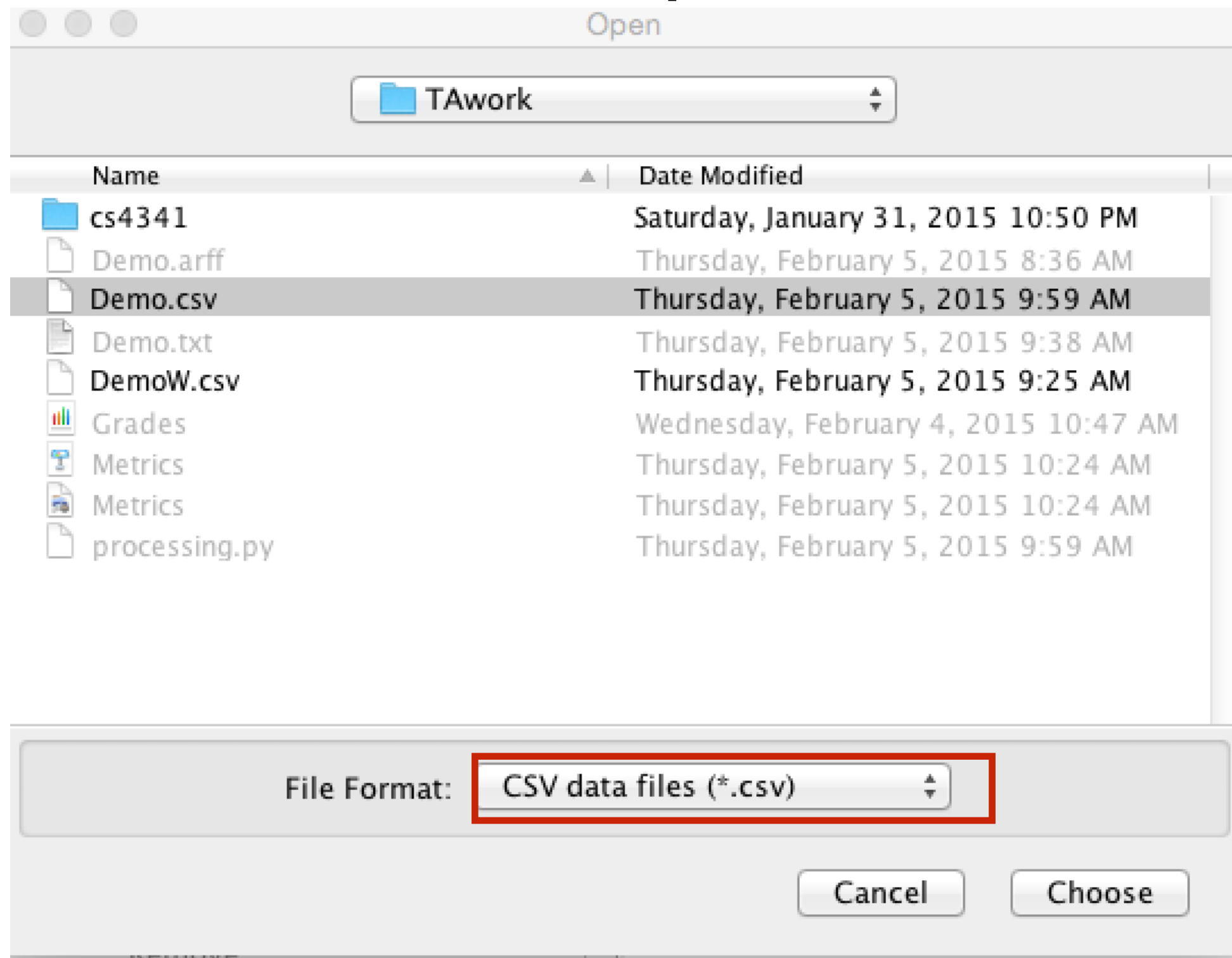
Choose the “Explorer” option. “Experimenter” is used to quickly make many trees, but as a result does not provide as much information.

Import Data



Click “Open file”

Import Data



Choose your file

Choose file format

Preprocessing Data

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: Demo Instances: 20 Attributes: 7

Attributes: All None Invert Pattern

No.	Name
1	<input type="checkbox"/> Where
2	<input type="checkbox"/> When
3	<input type="checkbox"/> Fred Starts
4	<input type="checkbox"/> Joe offense
5	<input type="checkbox"/> Joe defense
6	<input type="checkbox"/> Opp C
7	<input checked="" type="checkbox"/> OutCome

Remove

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

Statistic	Value
Minimum	0
Maximum	1
Mean	0.5
StdDev	0.513

Class: OutCome (Num) Visualize All

Status: OK Log x 0

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

weka

- filters
 - AllFilter
 - MultiFilter
- supervised
- unsupervised
 - attribute

2. Click Filter-Unsupervised-Attribute

3. Chose "Numeric to Nominal"

1. Select specific attribute/s

4. Apply Filter

Apply

Selected attribute

Statistic	Value
Minimum	0
Maximum	1
Mean	0.5
StdDev	0.513

Name: OutCome
Missing: 0 (0%)
Distinct: 2
Type: Numeric
Unique: 0 (0%)

Invert | Pattern

Class: OutCome (Num) | Visualize All

10 10

0 0.5 1

Status OK

Log x 0

Preprocessing Data

Weka Explorer

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

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

Filter: Choose **NumericToNominal -R first-last** Apply

Current relation
Relation: Demo-weka.filters.unsupervised.attribute.NumericToNominal
Instances: 20 Attributes: 7

Attributes

All None Invert Pattern

No.	Name
1	<input type="checkbox"/> Where
2	<input type="checkbox"/> When
3	<input type="checkbox"/> Fred Starts
4	<input type="checkbox"/> Joe offense
5	<input type="checkbox"/> Joe defense
6	<input type="checkbox"/> Opp C
7	<input type="checkbox"/> OutCome

Remove

Selected attribute
Name: OutCome
Missing: 0 (0%) Distinct: 2 Type: Nominal Unique: 0 (0%)

No.	Label	Count
1	0	10
2	1	10

Class: OutCome (Nom) Visualize All



Status: OK Log x 0

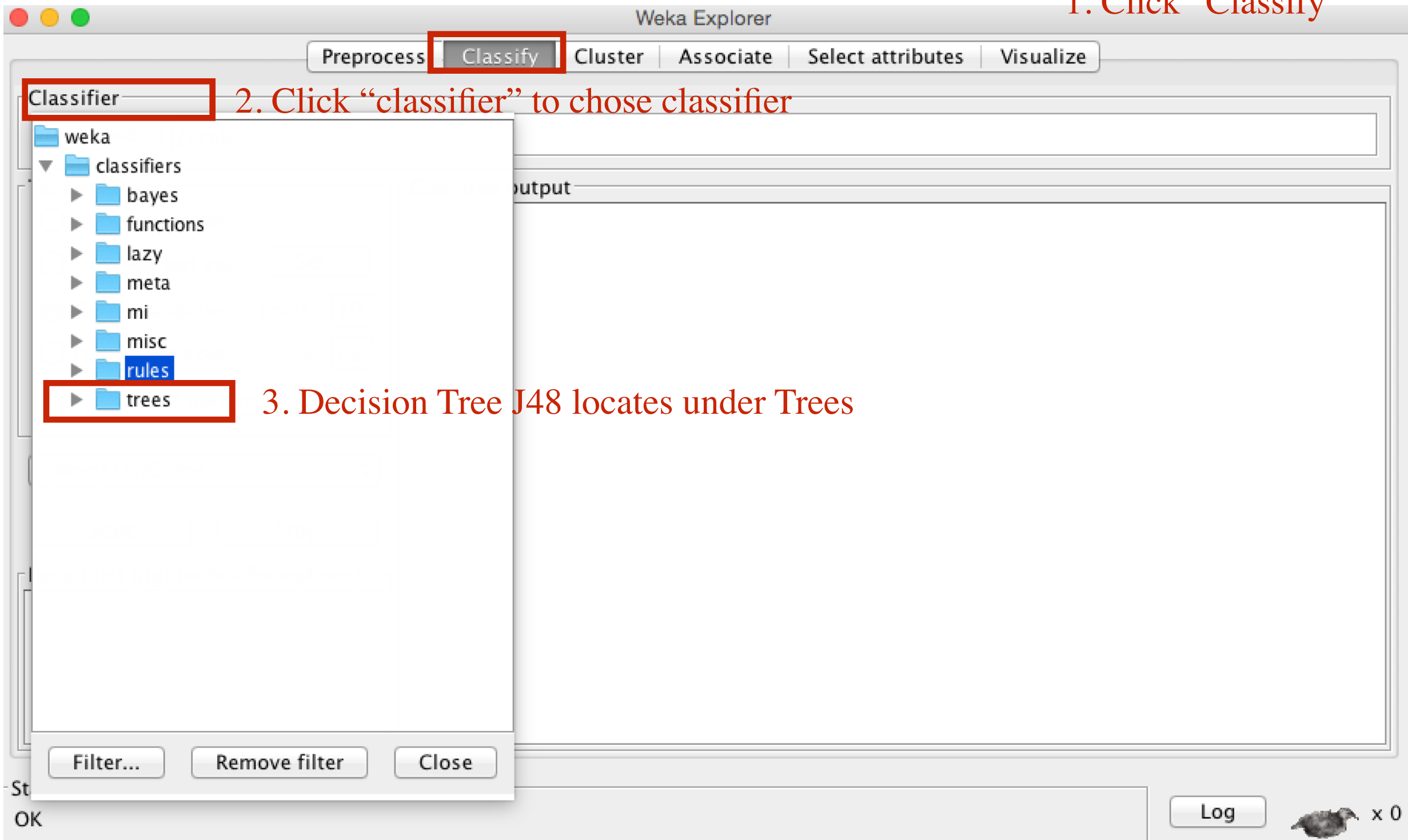
Check data
type of each
attribute again

Classification

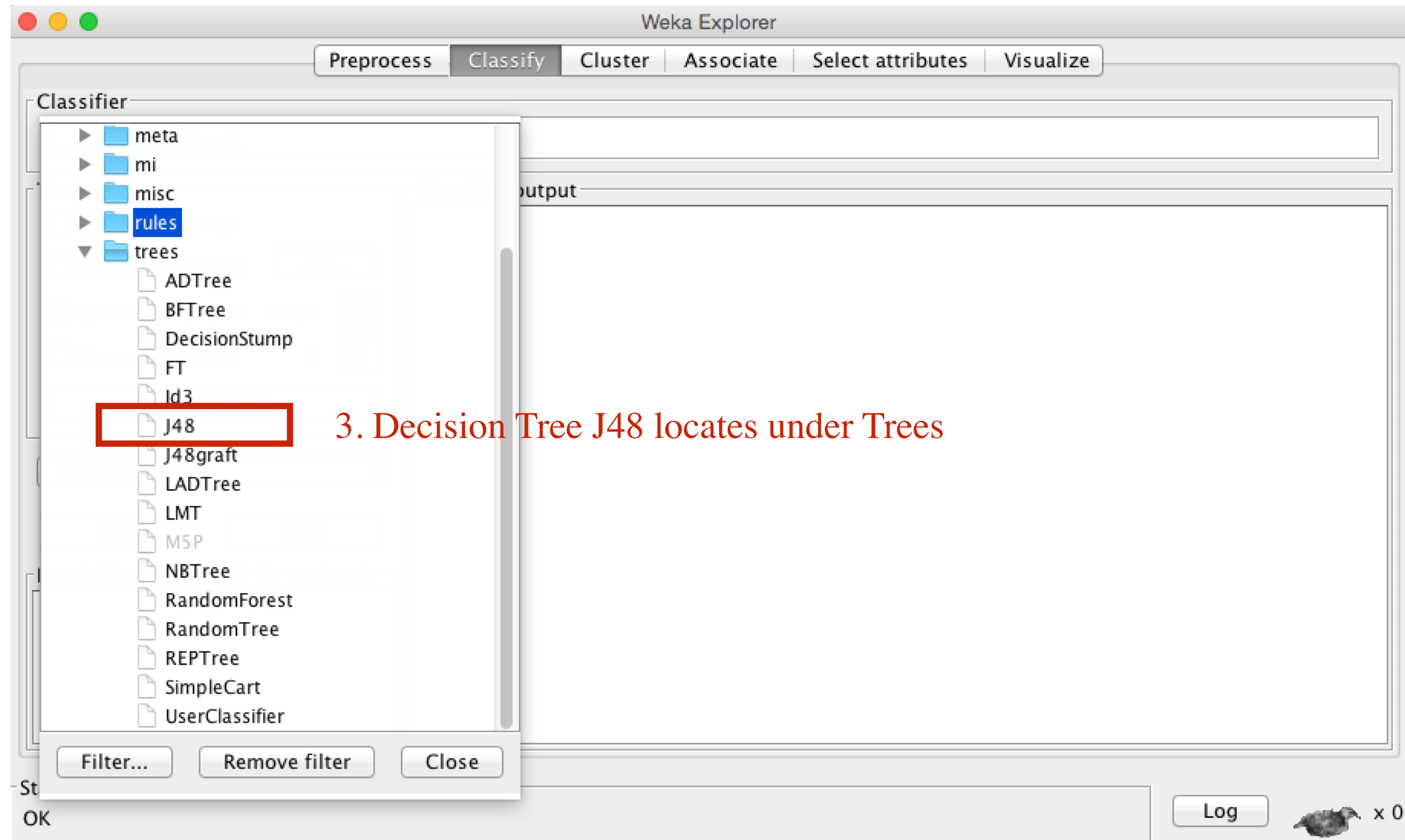
1. Click “Classify”

2. Click “classifier” to chose classifier

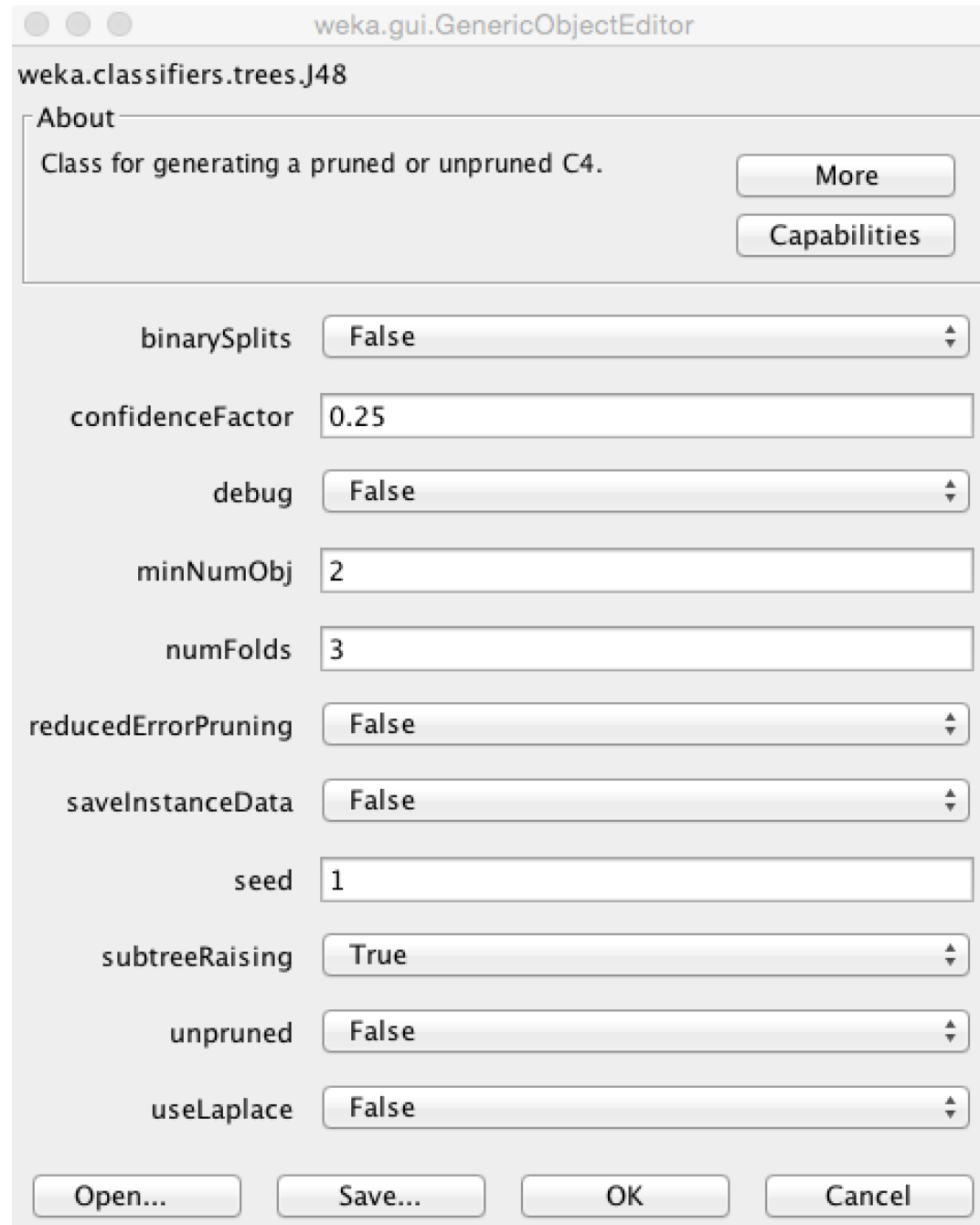
3. Decision Tree J48 locates under Trees



Classification-DT



Classification-DT



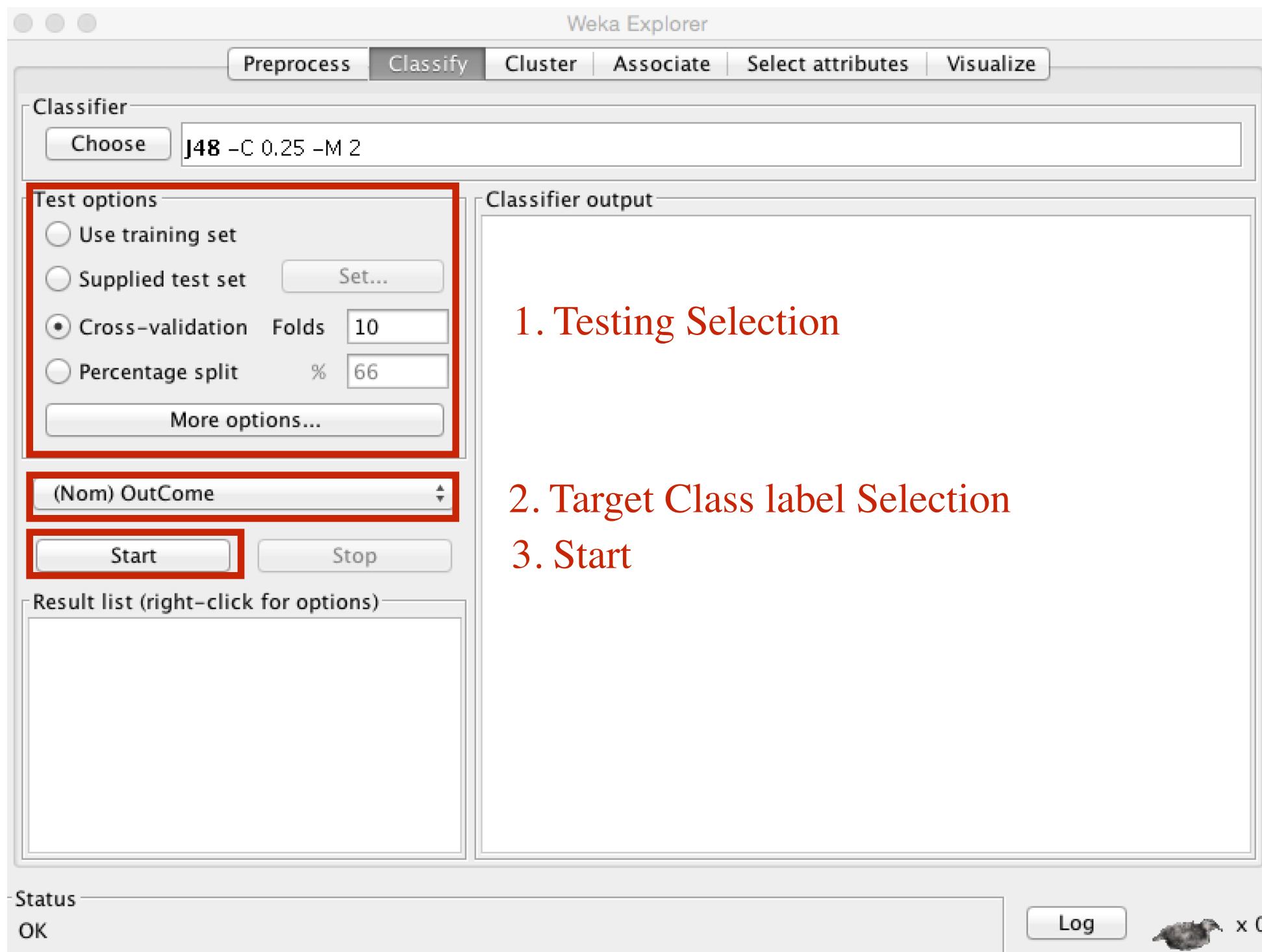
The screenshot shows the 'weka.gui.GenericObjectEditor' window for the 'weka.classifiers.trees.J48' classifier. The 'About' tab is active, displaying the description 'Class for generating a pruned or unpruned C4.' and buttons for 'More' and 'Capabilities'. Below this, various configuration parameters are listed with their current values in a table-like structure.

Parameter	Value
binarySplits	False
confidenceFactor	0.25
debug	False
minNumObj	2
numFolds	3
reducedErrorPruning	False
saveInstanceData	False
seed	1
subtreeRaising	True
unpruned	False
useLaplace	False

At the bottom of the window are four buttons: 'Open...', 'Save...', 'OK', and 'Cancel'.

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

Testing



Experiment Result

Classifier
Choose J48 -C 0.25 -M 2

Test options
☐ Use training set
☐ Supplied test set Set...
☒ Cross-validation Folds 2
☐ Percentage split % 66
More options...

Classifier output

=== Summary ===

Correctly Classified Instances	15	75	%
Incorrectly Classified Instances	5	25	%
Kappa statistic	0.5		
Mean absolute error	0.3467		
Root mean squared error	0.4877		
Relative absolute error	69.3333 %		
Root relative squared error	97.5477 %		
Total Number of Instances	20		

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.6	0.1	0.857	0.6	0.706	0.735	0
	0.9	0.4	0.692	0.9	0.783	0.735	1
Weighted Avg.	0.75	0.25	0.775	0.75	0.744	0.735	

Result list (right-click for options)
14-17-38 - trees.J48

- View in main window
- View in separate window
- Save result buffer
- Delete result buffer
- Load model
- Save model
- Re-evaluate model on current test set
- Visualize classifier errors
- Visualize tree
- Visualize margin curve
- Visualize threshold curve
- Cost/Benefit analysis
- Visualize cost curve

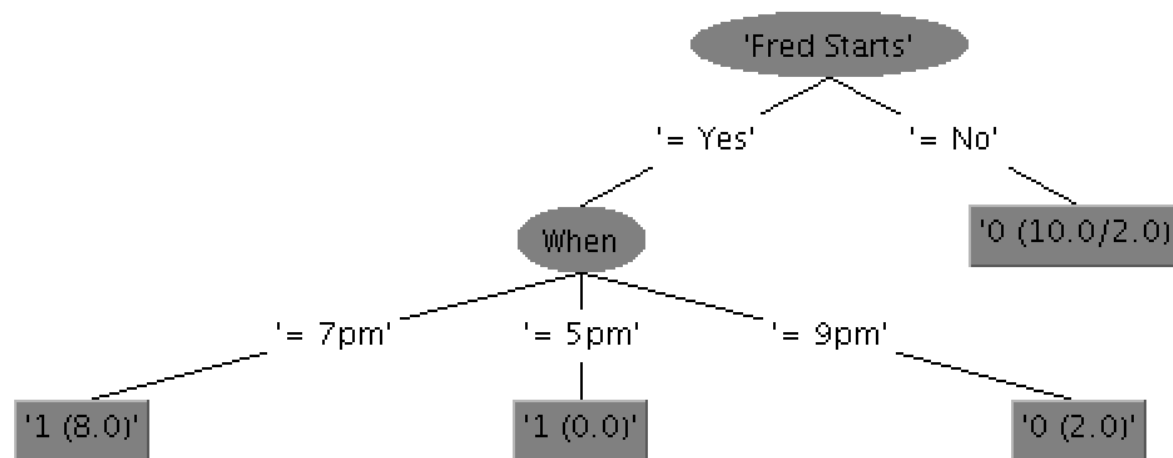
Status OK

Log x 0

Experiment Result

Experiment Result

Weka Classifier Tree Visualizer: 14:17:38 - trees.J48 (Demo-weka....)
- Tree View -



Tree Visualization

Single Predicability

Weka Explorer

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

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

Filter: Choose NumericToNominal -R first-last Apply

Current relation: Relation: Demo-weka.filters.unsupervised.attribute.... Instances: 20 Attributes: 7

Attributes: All None Invert Pattern

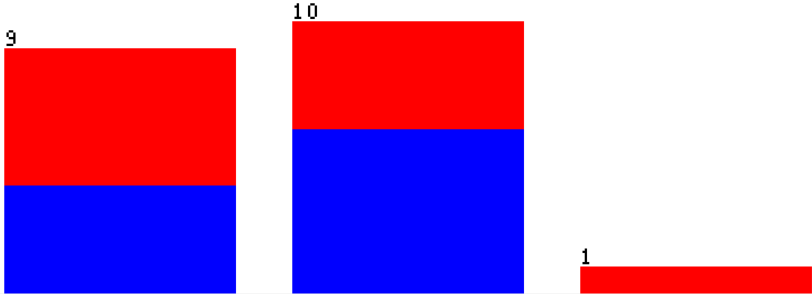
No.	Name
1	<input checked="" type="checkbox"/> Where
2	<input checked="" type="checkbox"/> When
3	<input type="checkbox"/> Fred Starts
4	<input checked="" type="checkbox"/> Joe offense
5	<input checked="" type="checkbox"/> Joe defense
6	<input checked="" type="checkbox"/> Opp C
7	<input type="checkbox"/> OutCome

Remove

Selected attribute: Name: Opp C Missing: 0 (0%) Distinct: 3 Type: Nominal Unique: 1 (5%)

No.	Label	Count
1	Tall	9
2	Short	10
3	Talls	1

Class: OutCome (Nom) Visualize All



Status: OK Log x 0

To evaluate the value of
“Fred Stars”

1. Select other features

2. Remove them

Single Predicability

Weka Explorer

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

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

Filter
Choose **NumericToNominal -R first-last** Apply

Current relation
Relation: Demo-weka.filters.unsupervised.attribute....
Instances: 20 Attributes: 2

Attributes
All None Invert Pattern

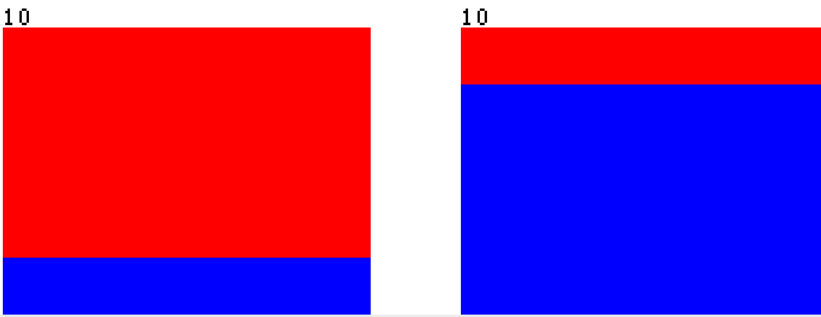
No.	Name
1	<input checked="" type="checkbox"/> Fred Starts
2	<input type="checkbox"/> OutCome

Remove

Selected attribute
Name: Fred Starts Type: Nominal
Missing: 0 (0%) Distinct: 2 Unique: 0 (0%)

No.	Label	Count
1	Yes	10
2	No	10

Class: OutCome (Nom) Visualize All



Status
OK

Log x 0

**Apply the same
procedure as
previous one**

Undo

Use “undo” to withdraw previous remove

The screenshot shows the Weka Explorer interface with the 'Preprocess' tab selected. The 'Undo' button in the top toolbar is highlighted with a red box. The 'Filter' section shows 'NumericToNominal -R first-last' applied. The 'Current relation' is 'Demo-weka.filters.unsupervised.attribute....' with 20 instances and 2 attributes. The 'Attributes' list shows 'Fred Starts' and 'OutCome'. The 'Selected attribute' section shows 'Fred Starts' with a count of 10 for 'Yes' and 10 for 'No'. The 'Class' is 'OutCome (Nom)'. The 'Remove' button is visible at the bottom.

No.	Label	Count
1	Yes	10
2	No	10

Class: OutCome (Nom)

Remove

The screenshot shows the Weka Explorer interface with the 'Preprocess' tab selected. The 'Undo' button in the top toolbar is highlighted with a red box. The 'Filter' section shows 'NumericToNominal -R first-last' applied. The 'Current relation' is 'Demo-weka.filters.unsupervised.attribute....' with 20 instances and 7 attributes. The 'Attributes' list shows 'Where', 'When', 'Fred Starts', 'Joe offense', 'Joe defense', 'Opp C', and 'OutCome'. The 'Selected attribute' section shows 'Where' with a count of 12 for 'Home' and 8 for 'Away'. The 'Class' is 'OutCome (Nom)'. The 'Remove' button is visible at the bottom.

No.	Label	Count
1	Home	12
2	Away	8

Class: OutCome (Nom)

Remove