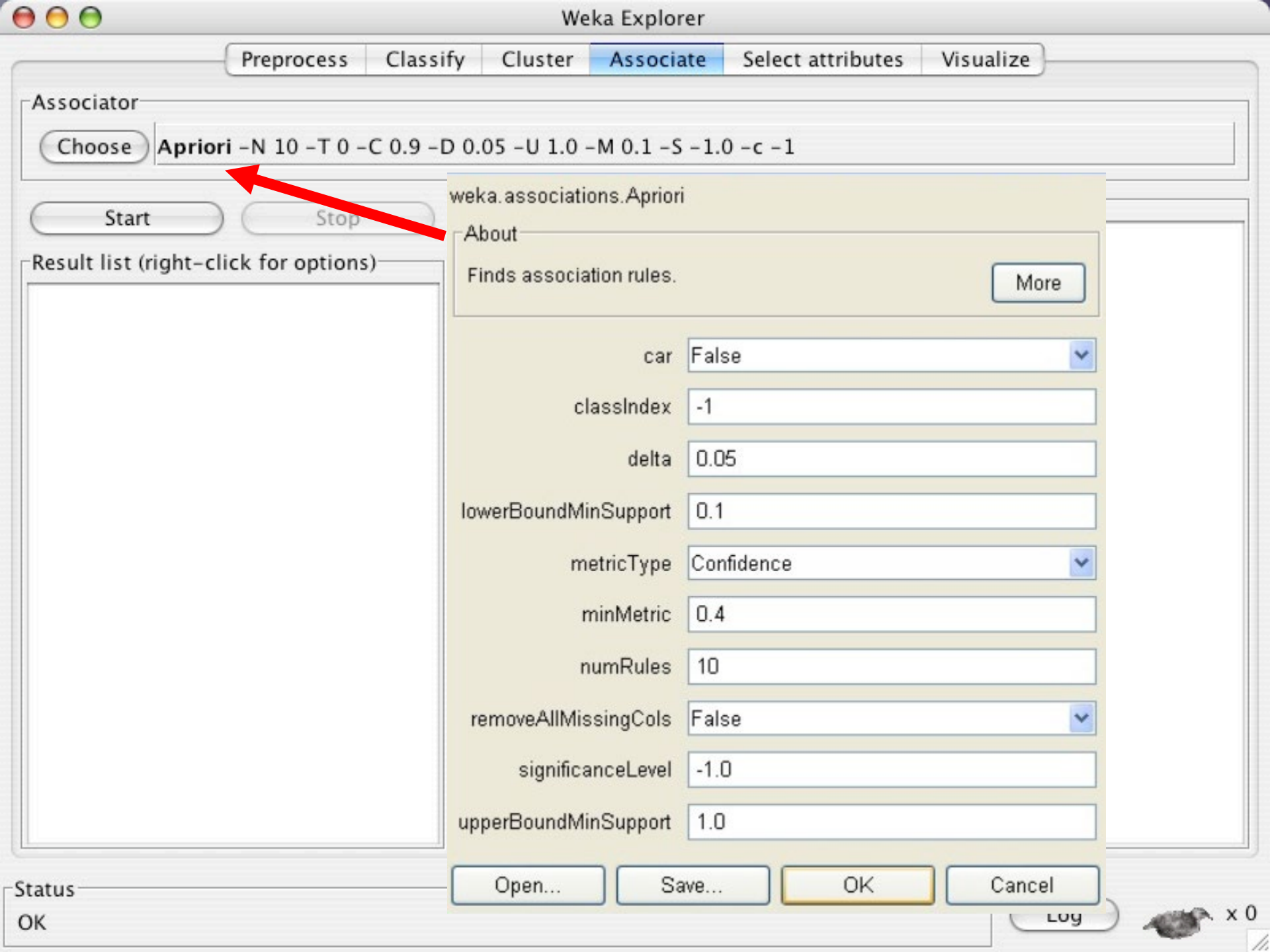
Weka: Association

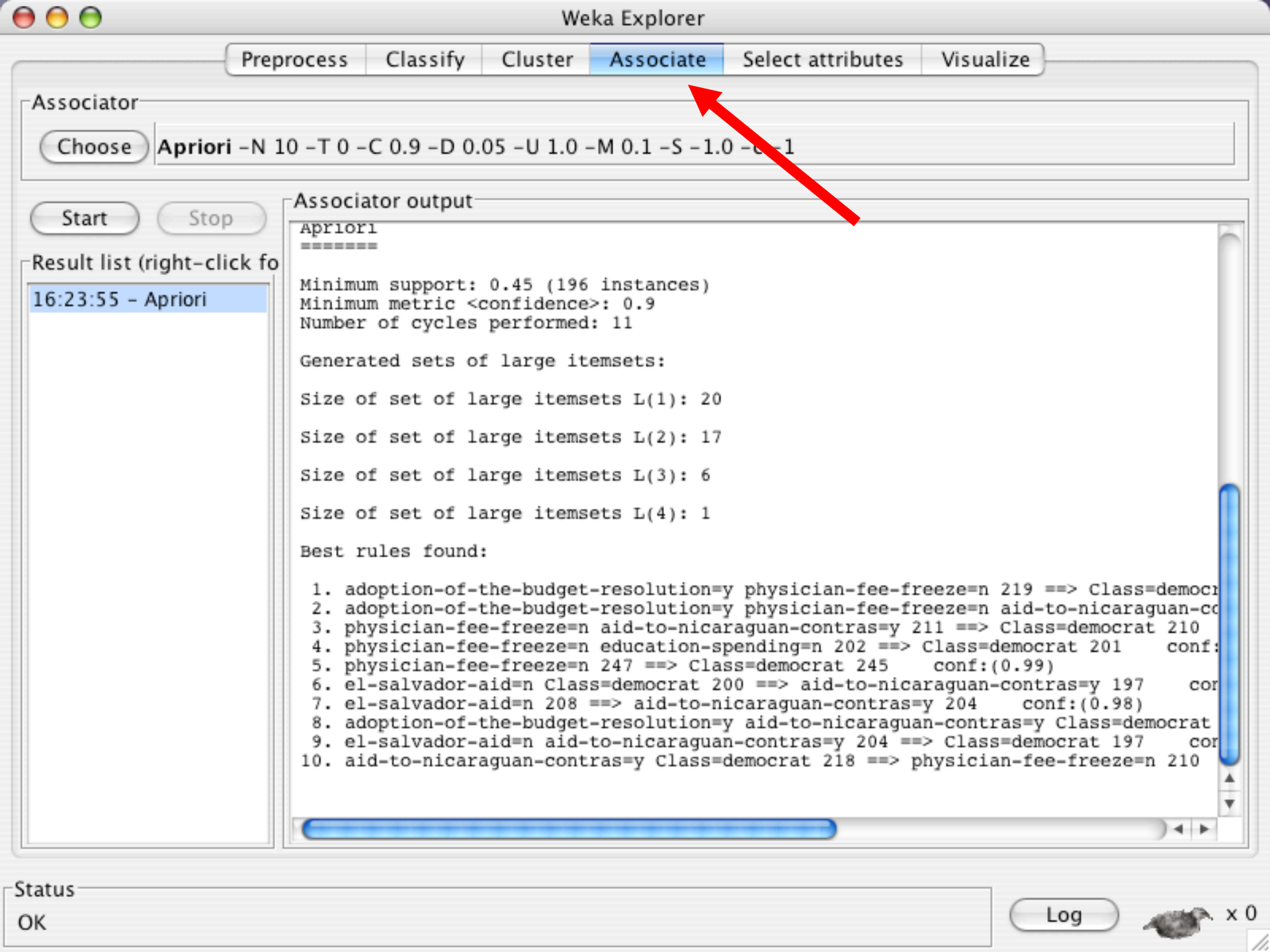
Cristóbal Romero,
Córdoba University, Campus Universitario de Rabanales, 14071, Córdoba, Spain
cromero@uco.es,

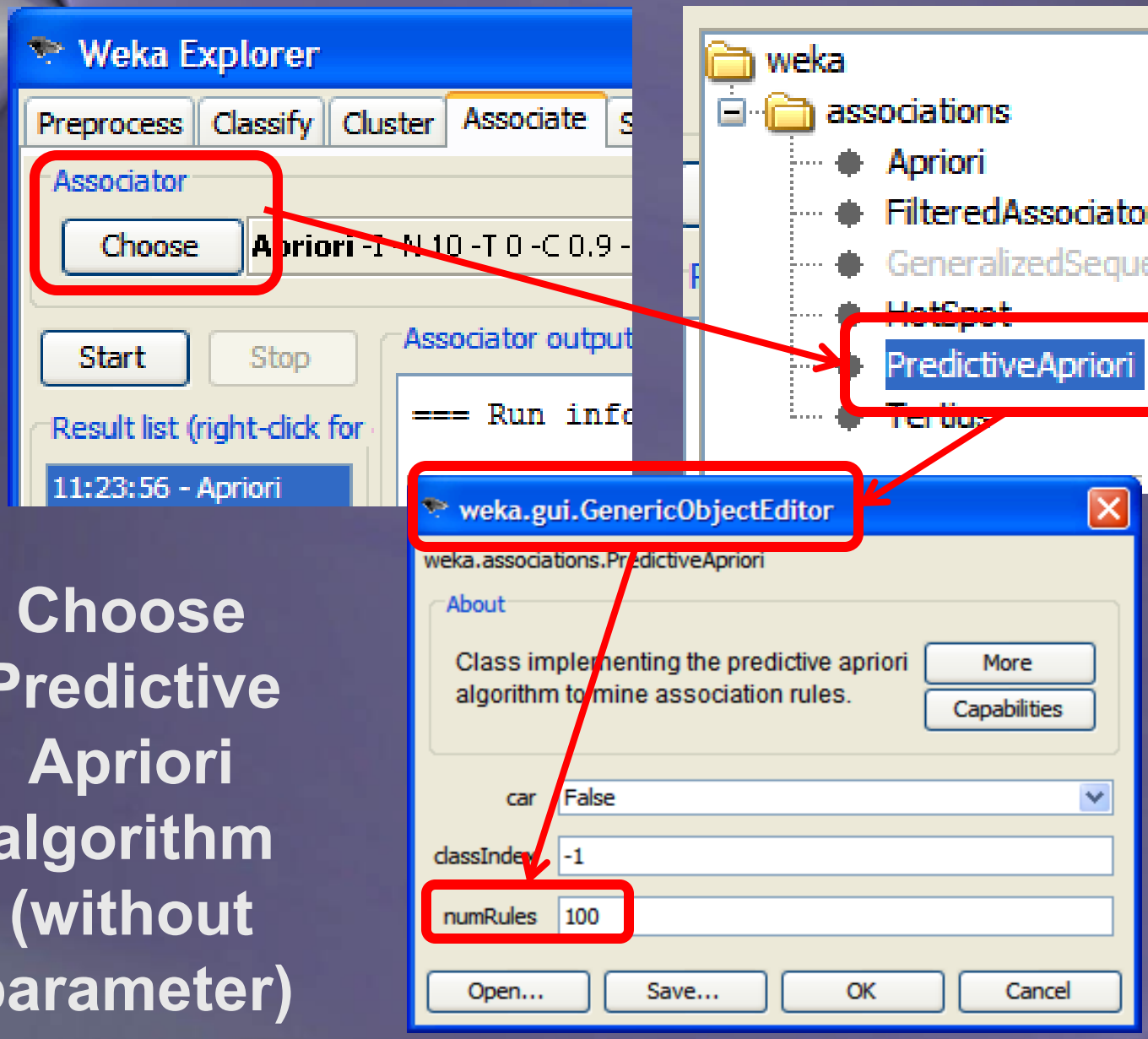
KDIS Lab
KNOWLEDGE DISCOVERY AND INTELLIGENT SYSTEMS

Weka: finding associations

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


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Associator
Choose
Apriori -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1
Start
Stop
Result list (right-click for options)
weka.associations.Apriori
About
Finds association rules.
More
car
False
classIndex
-1
delta
0.05
lowerBoundMinSupport
0.1
metricType
Confidence
minMetric
0.4
numRules
10
removeAllMissingCols
False
significanceLevel
-1.0
upperBoundMinSupport
1.0
Open...
Save...
OK
Cancel
Status
OK
Log
x 0


Weka Explorer
Preprocess
Classify
Cluster
Associate
Select attributes
Visualize
Associator
Choose
Apriori -N 10 -T 0 -C 0.9 -D 0.05 -U 1.0 -M 0.1 -S -1.0 -c -1
Start
Stop
Result list (right-click fo
16:23:55 - Apriori
Associator output
Apriori
=======
Minimum support: 0.45 (196 instances)
Minimum metric <confidence>: 0.9
Number of cycles performed: 11
Generated sets of large itemsets:
Size of set of large itemsets L(1): 20
Size of set of large itemsets L(2): 17
Size of set of large itemsets L(3): 6
Size of set of large itemsets L(4): 1
Best rules found:
1. adoption-of-the-budget-resolution=y physician-fee-freeze=n 219 ==> Class=democr
2. adoption-of-the-budget-resolution=y physician-fee-freeze=n aid-to-nicaraguan-co
3. physician-fee-freeze=n aid-to-nicaraguan-contras=y 211 ==> Class=democrat 210
4. physician-fee-freeze=n education-spending=n 202 ==> Class=democrat 201 conf:
5. physician-fee-freeze=n 247 ==> Class=democrat 245 conf:(0.99)
6. el-salvador-aid=n Class=democrat 200 ==> aid-to-nicaraguan-contras=y 197 cor
7. el-salvador-aid=n 208 ==> aid-to-nicaraguan-contras=y 204 conf:(0.98)
8. adoption-of-the-budget-resolution=y aid-to-nicaraguan-contras=y Class=democrat
9. el-salvador-aid=n aid-to-nicaraguan-contras=y 204 ==> Class=democrat 197 cor
10. aid-to-nicaraguan-contras=y Class=democrat 218 ==> physician-fee-freeze=n 210
Status
OK
Log
x 0



Choose Predictive Apriori algorithm (without parameter)

Occurrences of Outlook = overcast

Support for Outlook = overcast AND Play = yes

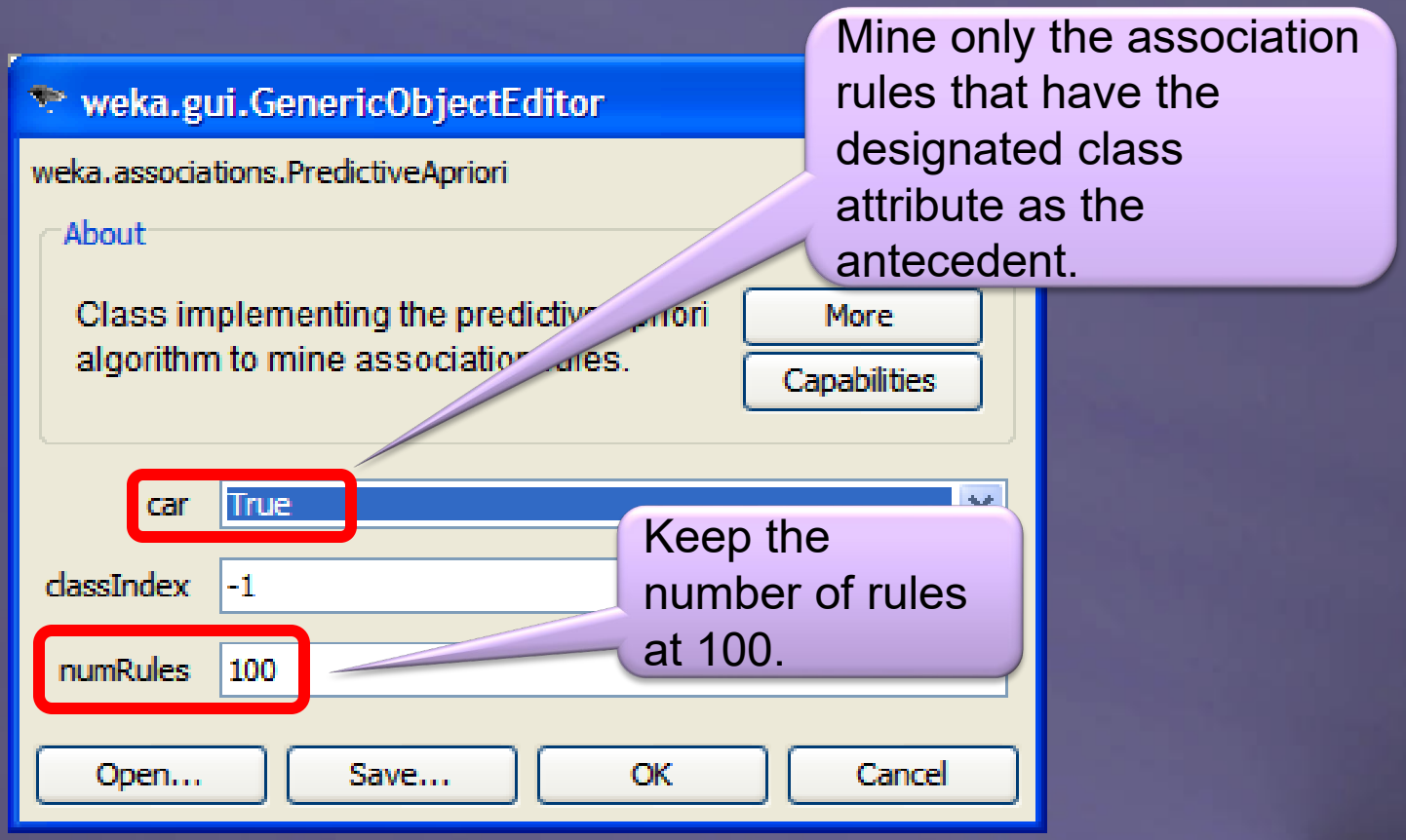
Predictive accuracy = 95.6%

Confidence of If Outlook = overcast, then Play = yes is 4/4 = 100%

Associator output

1. Outlook=overcast 4 ==> Play=yes 4 acc:(0.95583)
2. Temp.=cool 4 ==> Humidity=normal 4
3. Humidity=normal Windy=false 4 ==> Play=yes 4
4. Outlook=sunny Humidity=high 3 ==> Play=no 3
5. Outlook=sunny Play=no 3 ==> Humidity=high 3
6. Outlook=rainy false Play=yes 3
7. Outlook= lse 3
8. Outlook= =high Play=no 2 acc:(0.86872)
9. Outlook= lay=yes 2 acc:(0.86872)
10. Outlook= =normal 2 acc:(0.86872)
11. Outlook= =false Play=yes 2 acc:(0.86872)
12. Outlook= Temp.=hot Play=yes 2 acc:(0.86872)
13. Outlook=rainy Humidity=high 2 ==> Temp.=mild 2 acc:(0.86872)
14. Outlook=rainy Windy=true 2 ==> Play=no 2 acc:(0.86872)
15. Outlook=rainy Play=no 2 ==> Windy=true 2 acc:(0.86872)
16. Temp.=hot Play=no 2 ==> Outlook=sunny Humidity=high 2 acc:(0.86872)
17. Temp.=hot Play=yes 2 ==> Outlook=overcast Windy=false 2 acc:(0.86872)
18. Temp.=mild Humidity=normal 2 ==> Play=yes 2 acc:(0.86872)
19. Temp.=mild Play=no 2 ==> Humidity=high 2 acc:(0.86872)
20. Temp.=cool Windy=false 2 ==> Humidity=normal Play=yes 2 acc:(0.86872)

Class Association Rules



Associator output

```
 1. Outlook=overcast 4 ==> Play=yes 4    acc:(0.97484)
 2. Humidity=normal Windy=false  4 ==> Play=yes 4    acc:(0.97484)
 3. Outlook=sunny Humidity=high 3 ==> Play=no 3    acc:(0.96246)
 4. Outlook=rainy Windy=false  3 ==> Play=yes 3    acc:(0.96246)
 5. Outlook=sunny Temp.=hot 2 ==> Play=no 2    acc:(0.94114)
 6. Outlook=sunny Humidity=normal 2 ==> Play=yes 2    acc:(0.94114)
 7. Outlook=rainy Windy=true  2 ==> Play=no 2    acc:(0.94114)
 8. Temp.=mild Humidity=normal 2 ==> Play=yes 2    acc:(0.94114)
 9. Temp.=cool Windy=false  2 ==> Play=yes 2    acc:(0.94114)
10. Humidity=normal 7 ==> Play=yes 6    acc:(0.78559)
11. Temp.=cool 4 ==> Play=yes 3    acc:(0.64518)
12. Windy=false  8 ==> Play=yes 6    acc:(0.62995)
13. Outlook=sunny Windy=false  3 ==> Play=no 2    acc:(0.60023)
14. Outlook=rainy Temp.=mild 3 ==> Play=yes 2    acc:(0.60023)
15. Temp.=hot Humidity=high 3 ==> Play=no 2    acc:(0.60023)
16. Temp.=mild Windy=true  3 ==> Play=yes 2    acc:(0.60023)
17. Humidity=high Windy=true  3 ==> Play=no 2    acc:(0.60023)
18. Temp.=mild 6 ==> Play=yes 4    acc:(0.58517)
19. Outlook=sunny 5 ==> Play=no 3    acc:(0.56036)
20. Outlook=rainy 5 ==> Play=yes 3    acc:(0.56036)
21. Humidity=high 7 ==> Play=no 4    acc:(0.55274)
22. Temp.=hot 4 ==> Play=no 2    acc:(0.53273)
23. Temp.=hot 4 ==> Play=yes 2    acc:(0.53273)
24. Windy=true  6 ==> Play=no 3    acc:(0.52874)
25. Windy=true  6 ==> Play=yes 3    acc:(0.52874)
26. Humidity=high 7 ==> Play=yes 3    acc:(0.50181)
27. Outlook=sunny 5 ==> Play=yes 2    acc:(0.50124)
28. Outlook=rainy 5 ==> Play=no 2    acc:(0.50124)
29. Temp.=mild 6 ==> Play=no 2    acc:(0.4736)
30. Windy=false  8 ==> Play=no 2    acc:(0.42451)
```

Class Association Rules