

Opisna indukcija v WEKI

23.3.2007

Petra Kralj

Petra.Kralj@ijs.si

Descriptive induction

Voting dataset
Iris dataset

Voting dataset

- 435 instances
- 16 attributes
 - 16 nominal attributes
 - 0 numeric attributes
- No target variable
- No missing values

Weka Explorer

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

Open fil... | Open U... | Open D... | Undo | Edit... | Save...

Filter: Choose **None** Apply

Current relation: Relation: voting, Instances: 435, Attributes: 16

Selected attribute: Name: handicapped-infants, Type: No..., Missi... 12 (...), Distinct: ..., Unique: 0 (0...)

Label	Count
n	236
y	187

Class: party (Nom) Visualize All

Attributes: All | None | Invert

No.	Name
<input checked="" type="checkbox"/>	1 handicapped-infants
<input type="checkbox"/>	2 water-project-cost-sharing
<input type="checkbox"/>	3 adoption-of-the-budget-resolution
<input type="checkbox"/>	4 physician-fee-freeze
<input type="checkbox"/>	5 el-salvador-aid
<input type="checkbox"/>	6 religious-groups-in-schools
<input type="checkbox"/>	7 anti-satellite-test-ban
<input type="checkbox"/>	8 aid-to-nicaraguan-contras
<input type="checkbox"/>	9 mx-missile
<input type="checkbox"/>	10 immigration
<input type="checkbox"/>	11 synfuels-corporation-cutback
<input type="checkbox"/>	12 education-spending
<input type="checkbox"/>	13 superfund-right-to-sue
<input type="checkbox"/>	14 crime
<input type="checkbox"/>	15 duty-free-exports
<input type="checkbox"/>	16 party

Remove

Status: OK Log x 0

Association rules

1

2

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

Start Stop

Result list (right-click for o

16:25:34 - Apriori

Associator output

```
Size of set of large itemsets L(4): 1

Best rules found:

1. adoption-of-the-budget-resolution=y physician-fee-freeze=n 219 ==> party=democrat 219  conf: (1)
2. adoption-of-the-budget-resolution=y physician-fee-freeze=n aid-to-nicaraguan-contras=y 198 ==> party=democrat 198  conf: (1)
3. physician-fee-freeze=n aid-to-nicaraguan-contras=y 211 ==> party=democrat 210  conf: (1)
4. physician-fee-freeze=n education-spending=n 202 ==> party=democrat 201  conf: (1)
5. physician-fee-freeze=n 247 ==> party=democrat 245  conf: (0.99)
6. el-salvador-aid=n party=democrat 200 ==> aid-to-nicaraguan-contras=y 197  conf: (0.99)
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 party=democrat 203 ==> physician-fee-freeze=n 203  conf: (1)
9. el-salvador-aid=n aid-to-nicaraguan-contras=y 204 ==> party=democrat 197  conf: (0.97)
10. aid-to-nicaraguan-contras=y party=democrat 218 ==> physician-fee-freeze=n 210  conf: (0.96)
```

Status

OK

Log x 0

Iris dataset

- 150 instances
- 4 attributes
 - 0 nominal attributes
 - 4 numeric attributes
- Nominal target variable
 - 3 values:
 - Iris-setosa (30%)
 - Iris-versicolor (30%)
 - Iris-virginica (30%)
- No missing values

Weka Explorer

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

Ope... | Ope... | Ope... | Undo | Edit... | Sav...

Filter: Choose **None** Apply

Current relation
Relation: iris
Instances: 150 Attributes: 5

Attributes
All None Inv...

No.	Name
1	sepal length
2	sepal width
3	petal length
4	petal width
5	iris

Selected attribute
Name: iris Type: N...
Missing: ... Distinct: Unique: 0...

Label	Count
Iris-setosa	50
Iris-versicolor	50
Iris-virginica	50

Class: iris (Nom) Visualize All

50 50 50

Status: OK Log x 0

Clustering

1



2



Weka Explorer

Preprocess | **Cluster** | Associate | Select attributes | Visualize

Clusterer

- weka
 - clusters
 - Cobweb
 - EM
 - FarthestFirst
 - MakeDensityBasedClusterer
 - SimpleKMeans**

Clusterer output

Number of iterations: 6
Within cluster sum of squared errors: 6.99811400482676

Cluster centroids:

weka.gui.GenericObjectEditor

weka.clusterers.SimpleKMeans

About

Cluster data using the k means algorithm More

numClusters

seed

Open... Save... OK Cancel

0	61	(41%)
1	50	(33%)
2	39	(26%)

17:12:13 - SimpleKMeans
17:12:22 - SimpleKMeans

Status: OK Log x 0

3



Clustering visualization

The image shows the Weka Clusterer Visualize interface. On the left, the main Weka window is partially visible, showing the 'Cluster' tab and a 'Result list' with a context menu open. A red arrow points to the 'Visualize cluster assignments' option in the menu. The main window displays a scatter plot titled 'Weka Clusterer Visualize: 17:12:22 - SimpleKMeans (iris-weka.filters.uns...'. The plot shows 'X: sepal length (Num)' on the horizontal axis and 'Y: petal width (Num)' on the vertical axis. The data points are colored according to their cluster assignment: red for cluster0, blue for cluster1, and green for cluster2. The plot includes a 'Jitter' slider and buttons for 'Reset', 'Clear', and 'Save'. Below the plot, the 'Class colour' section shows the mapping: cluster0 (red), cluster1 (blue), and cluster2 (green).