

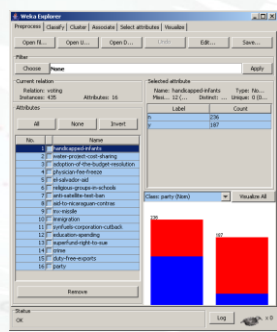
## Hands on Weka: Part III

### Classification rules and descriptive induction

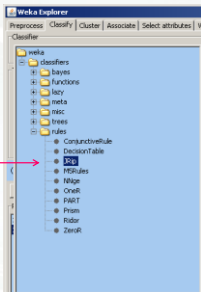
Petra Kraj Novak  
 9.12.2014

## Voting dataset

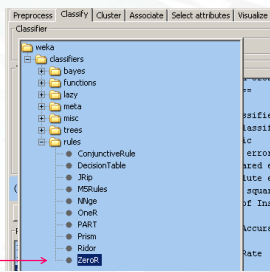
- 435 instances
- 16 attributes
  - 16 nominal attributes
  - 0 numeric attributes
- No target variable
- No missing values



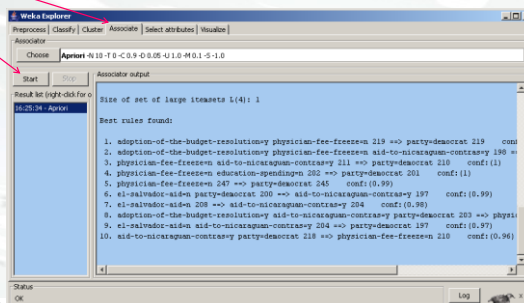
## Classification rules: Weka → classifiers → rules → JRip



## Baseline classifier: Weka → classifiers → rules → ZeroR



## Association rules Weka → associations → Apriori

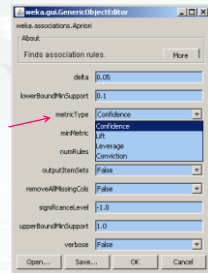


Size of set of large itemsets L[4]: 1

Best rules found:

1. adoption-of-the-budget-resolutions physician-fee-freesen 219 => party-democrat 219 conf: 1
2. adoption-of-the-budget-resolutions physician-fee-freesen aid-to-nicaragua-contra 198 => party-democrat 219 conf: 1
3. physician-fee-freesen aid-to-nicaragua-contra 211 => party-democrat 201 conf: 1
4. physician-fee-freesen education-spending 202 => party-democrat 201 conf: 1
5. physician-fee-freesen 247 => party-democrat 245 conf: (0.99)
6. el-salvador-aid 200 => aid-to-nicaragua-contra 197 conf: (0.99)
7. el-salvador-aid 208 => aid-to-nicaragua-contra 204 conf: (0.98)
8. adoption-of-the-budget-resolutions aid-to-nicaragua-contra party-democrat 203 => physi 197 conf: (0.97)
9. el-salvador-aid aid-to-nicaragua-contra 204 => party-democrat 197 conf: (0.97)
10. aid-to-nicaragua-contra party-democrat 218 => physician-fee-freesen 210 conf: (0.96)

## Quality of association rules



## Classification and Association Rules

Compare classification and association rules

- Purpose
- Format
- Quality measure
- Ruleset / List of rules
- Exhaustiveness of algorithms

INSTITUTE OF KNOWLEDGE TECHNOLOGIES

Comparison of classification and association rules

- Purpose
  - Classification rules: classification
  - Association rules: exploratory data analysis, descriptive induction
- Format
  - Both in the format  $X \rightarrow Y$
  - Classification rules: Y is a pair "target variable=class"
  - Association rules: both X and Y are itemsets = conjunctions of attribute-value pairs
- Quality measure
  - Classification rules: classification accuracy of the ruleset, precision, rule accuracy, weighted relative accuracy
  - Association rules: support, confidence, lift, leverage, conviction
- Ruleset / List of rules
  - Classification rules: can be both: unordered sets of rules or ordered list of rules
  - Association rules: unordered set of rules
- Exhaustiveness of algorithms
  - Classification rules: heuristic algorithms
  - Association rules: exhaustive algorithms, guarantee the optimal results

INSTITUTE OF KNOWLEDGE TECHNOLOGIES