LEIC – ALAMEDA 2017/18

Sistemas de Apoio à Decisão

Lab 2 – Mining Association Rules

1. **Loading Transactions**
    a. Load *example* and discover the set of frequent itemsets (support≥20%).
    b. Load *adult* and discover the set of rules (support≥50%, confidence≥90%)

2. Analyze the *vote* data, discovering all the association rules with 100% of confidence, and support above 10%.
    a. How does the number of discovered rules change with the minimum confidence and support thresholds?
    b. How many rules do you find with lift > 1?
    c. And for conviction > 1?
    d. How useful are the rules selected with the different measures?

3. Analyze the *supermarket* data and compare the results against the previous ones, according to
    a. The number of discovered rules
    b. The quality and usefulness of the rules
    c. The time required to discover the rules

4. Load the *glass* data. Since the attributes are numeric, discretize them using both the **equal-width** and **equal-frequency** strategies.
    a. What do you observe when you compare the distributions obtained?
    b. Compare the rules discovered.

R packages

- arules
- dplyr