Association Rule Software (ARS) is a basic tool which extracts association rules from attribute–value datasets (categorical or binary attributes). It is distributed with the SIPINA package which includes: a tool for the supervised learning framework, especially the decision tree induction (SIPINA RESEARCH); a tool for the linear regression (REGRESS); and thus, ARS for the association rule mining.

ARS encodes automatically the categorical attributes in dummy variables. If you want use a continuous attributes, you must discretize them before.

This tutorial describes shortly the use of the Association Rule Software (ARS).

1. Importing a dataset

First, we must import the dataset. ARS uses its proper file format (.FDM) which is optimised for the I/O processing. But we can also import dataset from other file format, especially tab–delimited text file format. We show here how to import this kind of data file.

We click on the FILE / OPEN and we select the Text Files format into the dialog box. We select the "banque.txt" data file which is available into the "Dataset" subfolder of the installation directory.
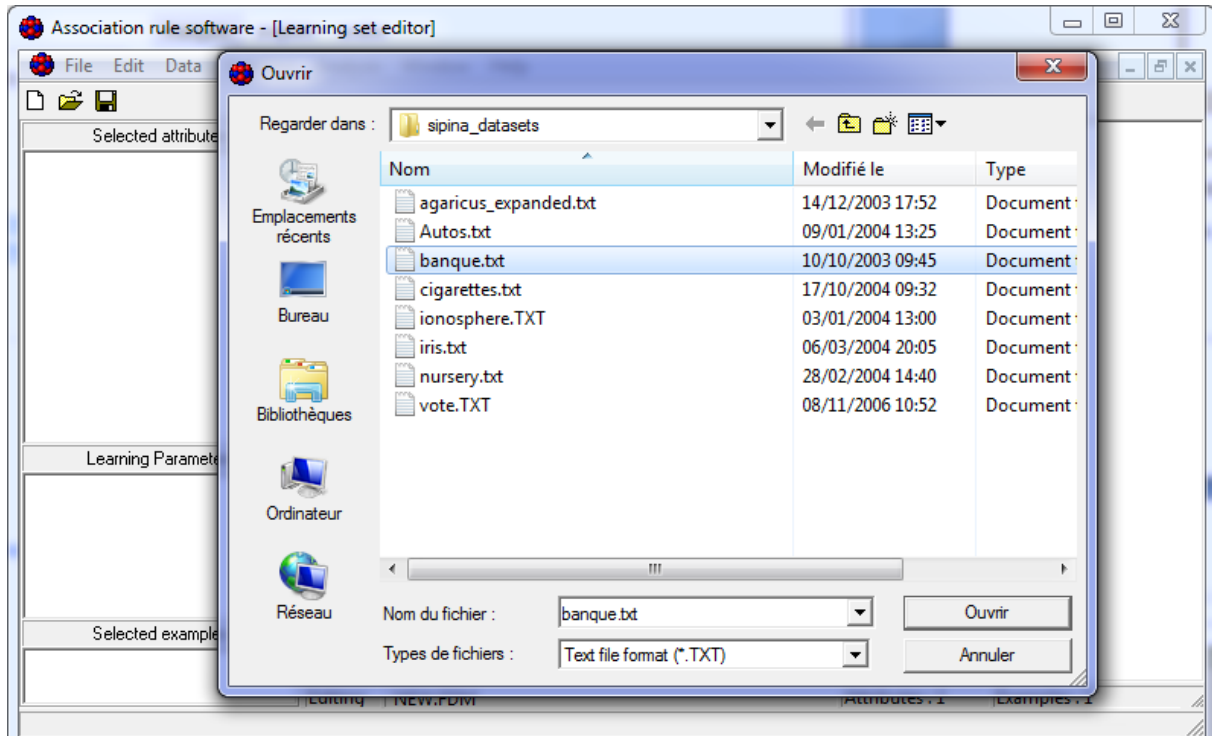

Then, we specify the organization of the text file (tab delimited, first row corresponds to the name of the attributes).
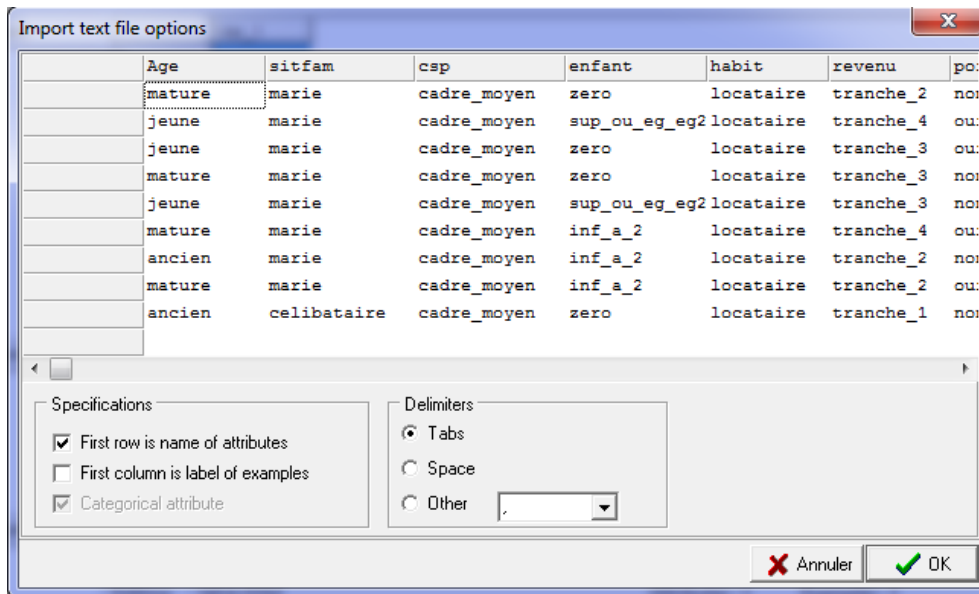
2. Specifying the role of attributes

Association Rules mining highlights the co−occurrence between items (values of attributes).
So, we must specify the attributes that we want use in our analysis.

We click on the ANALYSIS / SELECT ATTRIBUTES menu. Into the dialog box, the available attributes are showed on the left side. We select those that we want to include into the analysis by clicking on the add ▶▶ button.
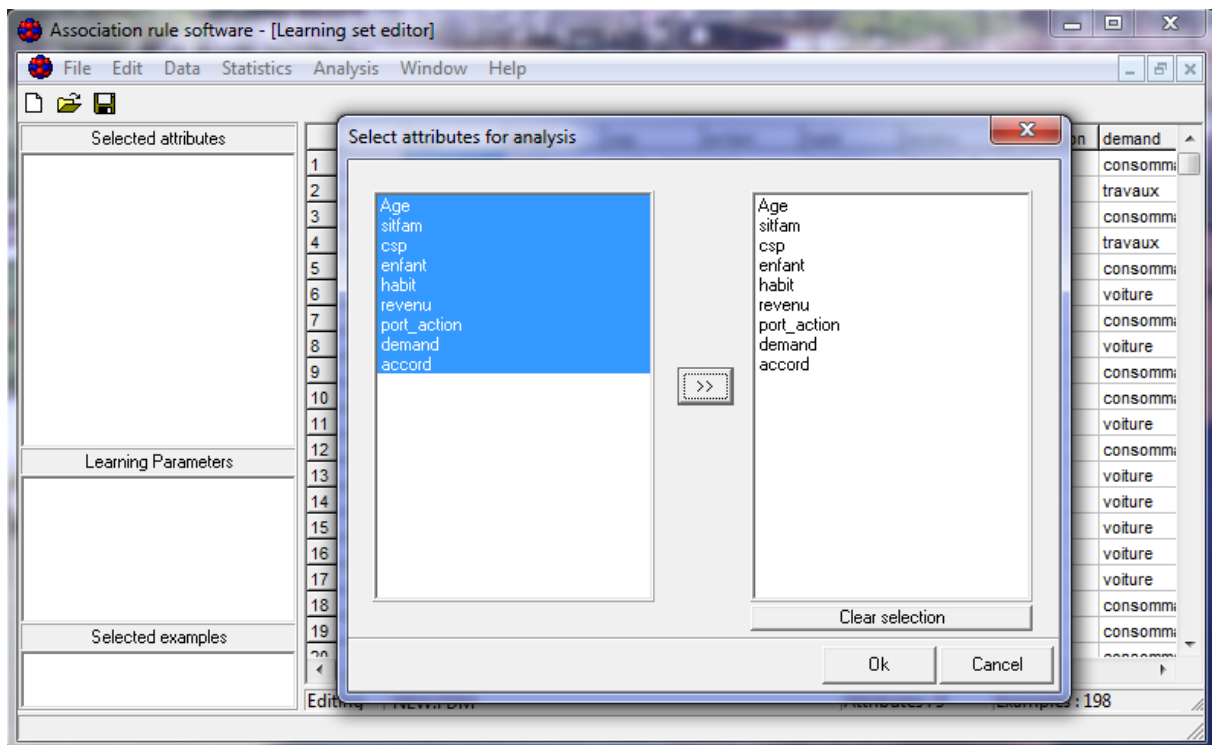
Note: The analysis can be performed on categorical attributes. But, we can use also continuous ones. In this case, they are interpreted as binary variables i.e. the values lower

or equal than 0 equates to the absence of the attribute, the values upper than 0 equates to the presence. We can also discretize the continuous attributes by selecting the appropriate method in the STATISTICS / TRANSFORM / CONTINUOUS ATTRIBUTES menu.

3. Launching the analysis

To launch the analysis, we click on the ANALYSIS / START ANALYSIS menu. A dialog box appears, asking the learning parameters. Standard settings for association rule mining are proposed:

- Support: the minimum support for the extracted rules.

- Confidence: the minimum confidence of the rules.

- Max Rule Length: the maximum length (the cardinality of the itemset) of the rules.

- Max Consequent Length: the maximum length of the consequent of the rules.
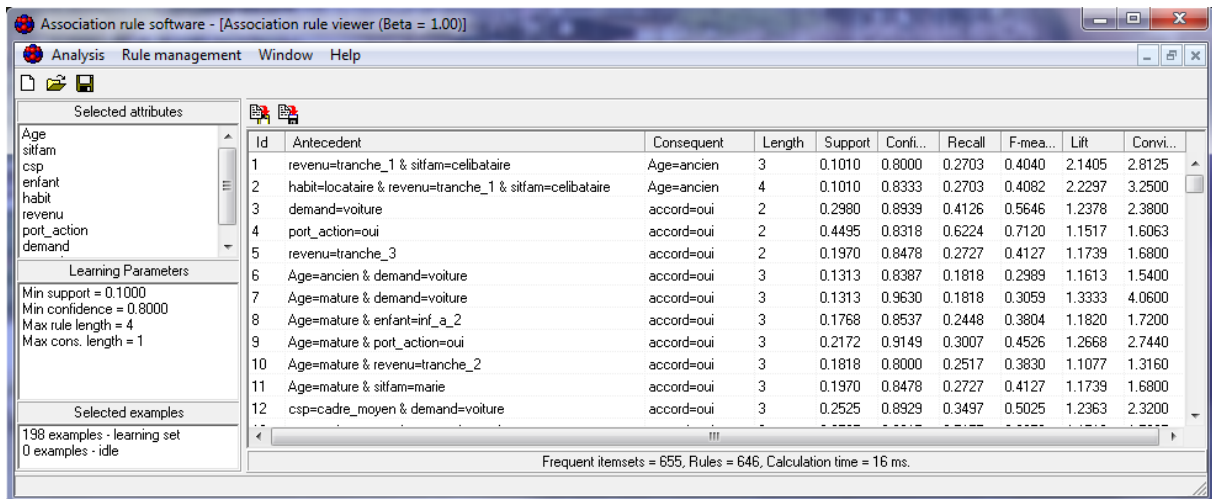
We note that ARS can extract rules with a consequent having more than one item. Of course, the length of the consequent is lower than the length of the rule.
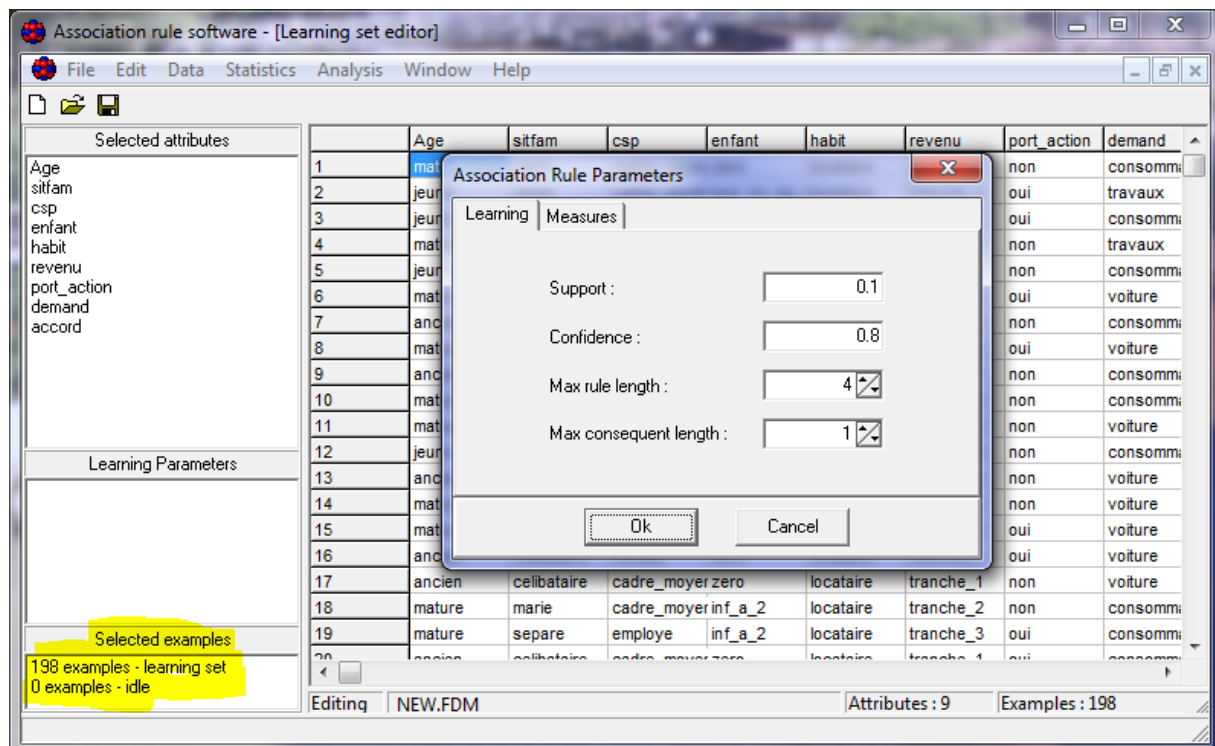
At the same time, we observe that, because we do not set explicitly the instances to treat during the analysis, ARS uses automatically all the available instances (198 cases).

We click on the OK button to start the learning process. The rules are displayed into a new visualization window.

We have the description of the rule (antecedent → consequent). We have also some standard indicators which measure the interestingness of the rules. We can sort the rules according these various criteria e.g. according the lift criterion in the screenshot bellow.

We can copy the rules into the clipboard (to paste them into a spreadsheet for instance) or save them in an output file (text file format) [RULE MANAGEMENT menu].

4. Conclusion

This short introduction describes how to use ARS in an association rules mining process.