**ARminer**: this package includes two methods for controlling false positives in association rule mining: the direct adjustment approach and the permutation-based approach

1. System requirements

* OS: Windows 7
* Software: Microsoft Visual Studio 2010 and above
* Additional library: stocc.lib

1. Include stocc.lib before compiling

* Add the path of stocc.lib to “PROJECT -> ARminer properties -> Configuration properties -> VC++ directories -> Library directories “
* Add stocc.lib to “PROJECT -> ARminer properties -> Configuration properties ->Linker -> Input -> Additional dependencies”

1. Running the program

* The datasets are stored in folder “datasets”
* Place the compiled executables under folder “run”.
* For each dataset, two .para file are created, one for direct adjustment and one for permutation test.
* Example: To run the program on dataset adult using direct adjustment, use the following command:

*ARminer adult.para*

* Example: To run the program on dataset adult using permutation test, use the following command:

*ARminer adult\_perm.para*

1. Parameters are specified in .para file.

Each parameter is specified in the format of “para\_name=para\_value”.

Lines starting with “#” are omitted.

* data\_file: full filename of the file containing the data
* names\_file: full filename of the file containing attributes and attribute types.
* min\_sup: minimum support threshold
* max\_pvalue: cut-off p-value threshold
* correction\_method: the method used to make multiple testing correction . If not specified, then the direct adjustment method is used.
* target\_attribute: the attribute appearing on the right hand side of rules. Only one attribute can be used as target\_attribute.
* target\_value: the value of the target\_attribute. Only one value can be target\_value.
* output: name of the output files

**Holdout**: This package implements the holdout method. It partitions a dataset into a mining dataset and a holdout data. It first calls ARminer to mine rules from the mining dataset and then evaluate the rules on the holdout dataset.

1. System requirements

* OS: Windows 7
* Software: Microsoft Visual Studio 2010 and above

1. Running the program

* The datasets are stored in folder “datasets”
* Place the compiled executables under folder “run”.
* Use the following command to run the program

*Holdout dataset\_name target\_attr target\_value min\_sup max\_pvalue output\_name*

* Example: to run the program on dataset mushroom using the holdout approach, use the following command:

*Holdout ..\\datasets\\mushroom class p 300 0.05 temp*

1. Parameters

* *dataset\_name : full filename of the dataset without file extension. Two files should exists: .data and .names*
* *target\_attr:* the attribute appearing on the right hand side of rules. Only one attribute can be used as target\_attribute.
* *target\_value:* the value of the target\_attribute. Only one value can be target\_value.
* *min\_sup:* minimum support threshold
* *max\_pvalue:* cut-off p-value threshold
* *output\_name:* name of the output files

If you use ARminer, please cite the following paper:

Guimei Liu, Haojun Zhang, Limsoon Wong. **Controlling False Positives in Association Rule Mining**. *Proceedings of the VLDB Endowment*, 5(2):145--156, October 2011.