

How To Start

On this page you will find the most important information, how to start **knowledge discovery in databases** or **data mining** with **LISp-Miner** tool.

To run LISp-Miner you need to have:

- Analysed data (i.e. the data you want to analyse) in any ODBC accessible format (we provide sample data sets for free download),
- Some of LISp-Miner KDD procedures (free for download),
- Metabase (metadata, free for download).

Read detail description for more information:

Analysed data

The data, you want to analyse have to be in any **ODBC accessible** format. However, if you are interested in the knowledge discovery and you have not any data to analyse, we provide two sample data set. There are free to download sample financial data and data from the medical project STULONG.

LISp-Miner: software tool for data mining

The system LISp-Miner consists of several **modules** (subsystems – **KDD procedures**) and is running on Windows 95 or higher operating system. There are several modules (subsystems), prepared for free download:

Elementary

Elementary subsystem consists of **empty metabase**, **LMAdmin.exe** and **LMDataSource.exe** modules (files). That is all you need to set up connection of analysed data with metabase and LISp-Miner system – the LMAdmin.exe is to set up the connection. The LMDataSource.exe is for data preparation – you can prepare the data before you will do more sophisticated analyses (eg. mining for asociational rules by 4ft-Miner or other KDD analysis).

4ft-Miner

The 4ft-Miner module (KDD procedure) can be used for mining for association rules and conditional association rules. This module (KDD procedure) can mine for much general set of association rules than the apriori algorithm mines – in the result 4ft-Miner can discover association rules, which apriori cannot discover. By the parameters of the 4ft-Miner task you can specify what rules may be interesting for you, so you do not get useless rules.

4ft-Miner consists of 4ftTask.exe and 4ftResult.exe files. 4ftTask.exe is to set up the task of mining for association rules and for the run of the task. The 4ftResult.exe is for viewing discov-

ered rules, sorting them according several parameters and to filter discovered rules. 4ftResult.exe is for simple interpretation of discovered association rules.

KEX

The KEX module (KDD procedure) is to mine for decision rules. The KEX procedure generates automatically all the potentially usefull implications and tests each of them in the analyzed data to see if the impication can improve the set of decision rules learned so far. The resulting set of decision rules consists of several percent of all the generated implications.

Metadata (Metabase)

Metabase is an MS Access database, where all the *data* about processed task (i.e. *metadata*) are stored. Into the metabase **all the KDD procedures** (modules, LISp-Miner subsystems) save all the parameters of the tasks, results, etc. For example, if you want to work with results of your work on another computer, you should copy both **analysed data** and **metabase**. So, metabase is the file, where all the results, all the done task, etc. are stored!

You can download empty metabase as a part of **elementary subsystem**. This means, that this empty metabase contains all the tables, which it must have for storing LISp-Miner task, results, etc. This metabase contains no one task or result. You can start using this empty metabase for your own work with LISp-Miner.

First steps

Overview

The process of KDD (knowledge discovery in databases) with LISp-Miner tool consists of following steps:

- 1 Connect data and metabase (use LMAdmin.exe)
- 2 Mark primary key and check it for duplicites (use LMDataSource.exe)
- 3 Create attributes (use LMDataSource.exe)
- 4 More sophisticated KDD analyse (use 4ftTask.exe, KexTask.exe, CFTask.exe, KLTask.exe)

Files preparation

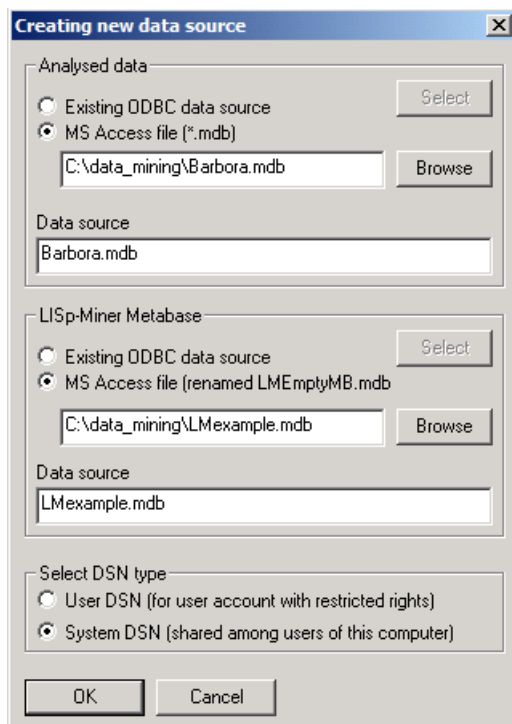
Start with downloading of **elementary** and **4ft-Miner KDD procedures**. Assume, the data you want to analyse are in any ODBC accessible format. Follow this text to get first experiences with mining for association rules with 4ft-Miner.

Unzipp both elementary and 4ft-Miner downloaded ZIPPED packages. Rename the LMEEmpty.mdb metabase.

In this example we use as **analysed data** the Barbora.mdb file. We renamed the LMEEmpty.mdb file, so its' name is LMexample.mdb now.

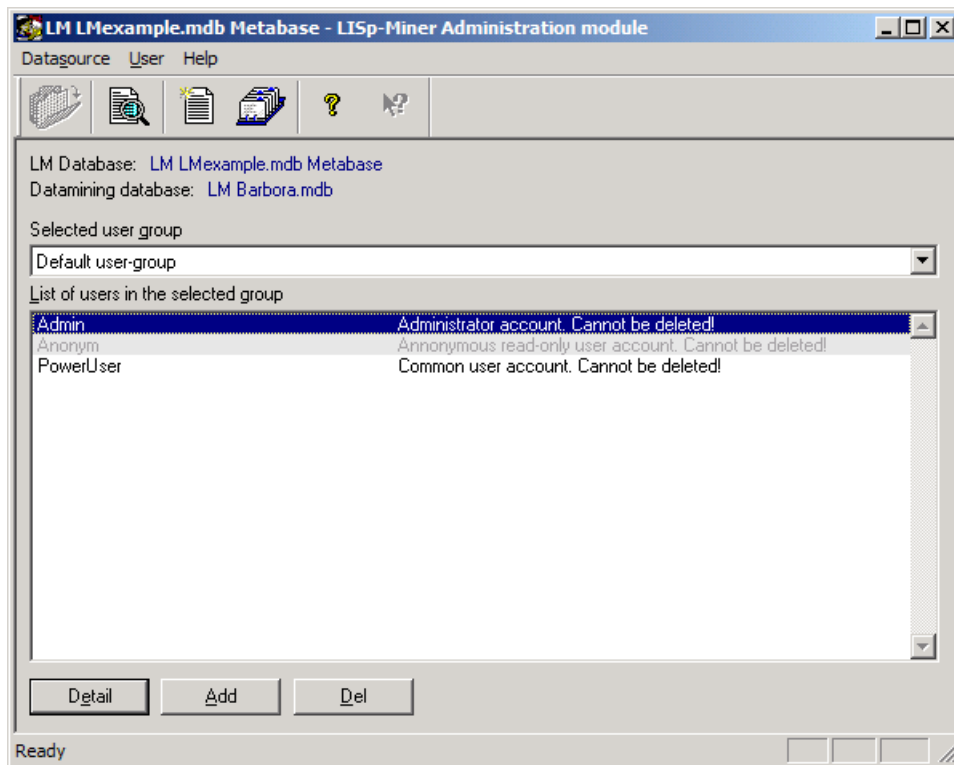
Start the **LMAdmin.exe** file.

Select (click on) **Create new data source** button. You will get window Creating new datasource. Select analysed data and LISp-Miner Metabase: in our example we select Barbora.mdb file and LMexample.mdb.



Anyway you can select any ODBC data source with your data, you want to analyse or any MS Access file with your data, you want to analyse. As LISp-Miner metabase you can select any renamed LMEEmpty.mdb file.

Continue with the *OK* button. After confirmation, you will get the LMAdmin main window:



As you will no more work with LMAdmin.exe, you can finish it by *Exit* in menu *Datasource*.

You can see video (AVI file) instructions: in this video we:

- 1 rename LMEmty.mdb into LMexample.mdb,
- 2 start LMAdmin.exe,
- 3 create new ODBC datasource,
- 4 close LMAdmin.exe module.

Attributes preparation

See Demonstration of the DataSource procedure to study how to prepare attributes using DataSource module.

Sophisticated data mining

See Demonstration of the 4ft-Miner to study how to mine for association rules and demonstration of the KL-Miner to study how to do the KL-analysis, demonstartion of the CF-Miner to study how to do the CG-analysis.