Viscovery SOMine | Data Mining Suite

Workflow-oriented software suite based on self-organizing maps (SOM) and multivariate statistics for explorative data mining and predictive modeling

General Features of the System

Project environment

- ◼ Analytical work organized in projects, providing focused navigation through the application
- ◼ Goal-oriented operation, shielding the user from the technology core and the statistical algorithms
- ◼ Projects, consisting of up to 5 workflows, provided for data import and preprocessing, SOM creation and exploration, classification, local SOM prediction, and scoring
- ◼ Project workflows, optionally processed with minimal user interaction
- ◼ Administration of projects, performed in directories with "clean directories" function

Workflow orientation

- ◼ Dedicated workflows, each consisting of 4 steps, with clear tasks
- ◼ Steps providing proven default settings
- ◼ Workflow branching allowing generation of model variations
- ◼ Functions allowing parameterizations to be copied between workflow steps
- ◼ Workflow branch sorting with drag-and-drop function

Reporting and documentation

- ◼ Instant, on-demand reports for each workflow step as well as the entire workflow
- ◼ Dynamic reporting of setting differences between steps in a branching
- ◼ Integrated documentation for projects
- ◼ Production journal of created models
- ◼ Functions for adding descriptions and detailed comments to attributes, clusters, and models

Usability

- ◼ Informative pop-ups throughout the system, as well as supportive context menus in numerous places
- ◼ Multiple handling tools, such as quick search for strings and sortable tables
- ◼ Transfer of attribute selections and retention of attribute order between lists with copy and paste
- ◼ Function for saving combinations of visible windows with selected properties as "Arrangements"
- ◼ Multiple options for setting preferences for charts, tables, statistics, modeling, and visualization
- ◼ Online help function and comprehensive user manual

Proprietary Data Formats

Viscovery projects

- ◼ *.visdm files: proprietary format for Viscovery data mining project files, which contain all relevant information of a Viscovery SOMine analysis project, including all preprocessing and other settings
- ◼ Earlier Viscovery project file formats (*.vsp, *.csp, and *.vpp from version 4, 5, and 6 of Viscovery SOMine, Viscovery Profiler, and Viscovery Predictor) can be imported

Viscovery data marts

- *.dms files: Viscovery data mart files, containing preprocessing and other meta information
- *.dmd files: associated Viscovery data mart files, containing portions of imported data, optimized for rapid access

Viscovery models

- *.som files: Viscovery SOMine model files, which contain the SOM data representation, including preprocessing settings, formulas, clusters, and predictive modeling definitions
- Older Viscovery model file formats (*.som of versions 4, 5, and 6 of Viscovery SOMine, Viscovery Profiler, and Viscovery Predictor) can be imported

Technical Requirements and Configuration

Minimum workstation configuration (recommended configuration)

- 2 GHz CPU (3 GHz or higher)
- 2 GB RAM (8 GB or more)
- Windows 7 or newer
- 32-bit OS (64-bit for voluminous data sets)
- 24 bit color graphics, 1280 x 800 (full HD or higher)
- Required disk space is 6 times the size of the analyzed data file (10 times for comfortable project work)

Modular software configuration

- The minimum configuration of Viscovery SOMine consists of the **Visual Clusters** core module.
- For advanced exploration and modeling, the **Explore and Classify** and **Predict and Score** extension modules are available.
- The **Enterprise Data** extension module provides enhanced features for data import and handling, as well as for the analyses of very high-dimensional and voluminous data.
- Each workflow can be automated using the corresponding **Workflow Automation Service**, which is available as part of an additional package.

Licensing

- Each modular configuration is available with a perpetual license or, alternatively, with a term-license for a specified period of time.
- Each modular configuration is available as a single-user license or as a network license.
- Single-user licenses are bound to the user account on a single computer and may be transferred to a different computer system once per year.
- Network licenses allow operation of the software for a defined maximum of concurrent users and require the additional installation of the Viscovery License Server.
- Each licensed user may operate one interactive instance of the software at a time.
- All configuration and validity information regarding licenses is coded in unique license keys.
- Software installation and license activation require administrator privileges and Internet connectivity.