Viscovery® SOMine 8.0 – What's New?

New in Viscovery® SOMine 8.0 Data Mining Suite

In **Viscovery® SOMine 8.0** a large number of features have been added, ranging from new preprocessing options and enhanced functionalities for SOM calculation to new cluster algorithms, classification and prediction techniques. Additionally, errors were corrected that have been discovered since version 7.2.

New Features

■ The map training algorithm was extended by a temperature parameter, which allows a probabilistic interpretation of data membership.

■ The topographic product is computed for every SOM, showing how well the data topology is represented by the SOM.

■ A new map picture shows the local SOM dimensionality.

■ Hedges' g* as an alternative profile value to the existing Glass' Δ. Hedges' g* is now the default.

■ Two new cluster algorithms are available: K-Means (centroid-based) and Connectivity (density-based). The latter is a completely new algorithm that uses a multi-linkage approach to generate fusion trees from which interesting clusters can be selected based on a relevance measure.

■ The cluster indicator is computed in a generalized way that better highlights good segmentations.

■ The Davies-Bouldin index is calculated for segmentations produced by algorithmic clustering.

■ A new classification technique is available in the new Create Classifier workflow.

■ Classifier models additionally compute class probabilities when classifying new data.

■ A confusion matrix and classification quality measures (accuracy, sensitivity, specificity and precision) are computed when a classifier model is tested with known classification results.

■ Logistic regressions are now available. These are implemented by use of R.

■ The Preprocessing Protocol can now accept specifications of attributes from multiple data sources.

■ Data preprocessing can now leverage the capabilities of the R scripting language.

Usability

■ Workflows have been renamed to better reflect their purpose and to be more consistent.

■ Settings made in workflow steps can be saved without the requirement to start processing of the step.

■ New function Process Now to start processing of workflow steps without entering the dialogs.

■ The workflow step that was active when a project is saved is restored when the project is reopened.

■ Workflow step placeholders (arrows) show more information because they now extend horizontally to occupy the available space.

■ Workflow Automation heeds the user's permissions and settings.

■ Data marts can be shrunk by removing data that is not used anymore in a project.

■ Names of files specified in the Import Data step can now be edited in-place; it is no longer necessary to add a new file import and to remove the old one.

■ Data processing happens in the background in more cases while the user interface is operated, most notably in the Tune Histograms step.

Viscovery® SOMine 8.0 – What's New?

- Attribute values captured by nominal value expressions are listed and counted to permit early correctness checks.
- In the Tune Histograms step, the removal of values from a nominal attribute can now be done using a list of occurring values.
- Attribute priorities can now be imported from a file, for example, the Preprocessing Protocol.
- Names of nominal values can be changed on the fly in the SOM for presentation purposes.
- Attribute lists can now be rearranged in original order.
- Editing names and descriptions of attributes can be initiated from the Cluster Characteristics window. Attribute order in the Statistics and Data Records windows are now kept track of in Arrangements.
- The labels on scatter plot chart axes are now more frequent and are located at more pleasing (round) numbers.
- The size of thumbnail images over the map can now be customized.
- Names of new segmentations mention the number of clusters.
- Segmentation pictures can now be rearranged.
- Regression results that use a confidence setting are now updated on the fly when the confidence is changed.

Modules and Licensing

- Some licensable modules have been renamed: Visual Clusters to **Visual Explorer** and Explore and Classify to **Cluster and Classify**.
- The basic module **Visual Explorer** is now available for free.
- User interface elements pertaining to features that are not covered by the current license are visible, but disabled.
- Entering a license key does not require administrative privileges anymore.
- Different users using the same computer (with different user accounts) can use **Viscovery SOMine** provided each user has their own license. (Alternatively, they can also use a floating network license.)
- License validity is checked on start-up. An Internet connection is required. If validation fails for 30 days, the license becomes invalid, but becomes valid again when validation is successful.
- The Viscovery License Server checks validity of the license daily; it becomes invalid after 10 days of absent Internet connectivity.

Compatibility Notes

- **Viscovery SOMine** files created with version 7.2 or earlier can be imported, but files created with version 8 cannot be imported in earlier versions.
- Opening SOM files that have been last saved with an earlier version may be slow. To accelerate loading times, it is recommended to save such files once with version 8.0.
- The cluster method SOM-Single-Linkage has been discontinued. Use Connectivity clustering with a multi-linkage K of 1 instead to achieve a similar clustering.
- In the Export Classification step, cluster statistics are no longer exported. Users who wish to associate this information with each data record have to use the cluster names that are exported to extract the respective information from the SOM.