What's New in SAS Enterprise Miner 12.1

Overview

SAS Enterprise Miner 12.1 is the version of SAS Enterprise Miner that supersedes SAS Enterprise Miner 7.1M1, released in December 2011. The updated version numbering is a result of synchronizing analytical content among SAS data mining software releases. SAS Enterprise Miner 12.1 provides improvements and enhancements to the product's core user interface, as well as updates to the SAS Enterprise Miner Credit Scoring nodes, the SAS Enterprise Miner Application nodes, the Rapid Predictive Modeler (RPM) SAS Enterprise Miner add-on to SAS Enterprise Guide, and the SAS Enterprise Miner High Performance Data Mining nodes.

SAS Enterprise Miner Core

- Several important and common SAS Enterprise Miner project properties have been promoted from macro variables that were deployed via SAS code or project start code to property selections that can be defined via the Enterprise Miner GUI properties panel.

- The SAS Enterprise Miner client can now be opened directly into a specific project or diagram, or from the most recent project and diagram.

- The %AA_MODEL_REGISTER macro has been added to enable users to directly register models that were developed in SAS code into the SAS Metadata Server. Once in SAS Metadata Server, the common model data can be accessed by SAS products such as SAS Model Manager, SAS Enterprise Guide, and SAS Data Integration Studio.

- The %AA_MODEL_EVAL macro has been added in order to compute lift and Receiver Operating Characteristic-type measures on any data set that contains probabilities and events. The %AA_MODEL_EVAL macro uses a two-stage approximation algorithm to calculate the model performance measures.

- Batch code from the SAS Enterprise Miner client can be used more easily with input tables of different names and locations. The SAS Enterprise Miner 12.1 batch code now integrates project start code that you can use to define libraries and options.

- The SAS Enterprise Miner SAS Code node for version 12.1 contains enhanced support for score code that contains SAS Procedure steps.

- PMML scoring for the Decision Tree, Regression, Neural Network, and Clustering nodes has been promoted to production status. New experimental functions have been introduced for general regression and scorecards in the SAS Enterprise Miner 12.1 release.

SAS Enterprise Miner Credit Scoring

- The Interactive Grouping node user interface has been redesigned to provide improved usability, performance, and computational scalability.

- The Interactive Grouping node has added a new **Calculated** variable role.

- The Scorecard node features an output variable that counts the number of adverse characteristics, and users can select named input variables for adverse characteristic reporting.

- The Scorecard node now processes indeterminate outcome values.

SAS Enterprise Miner Applications

- The Gradient Boosting node now provides users with the capability to disable the H statistic calculation, resulting in improved run-time performance.

- The Incremental Response node has been promoted from experimental to production status.

- The Time Series Data Mining nodes have been promoted from experimental to production status. The production Time Series Data Mining nodes have been redesigned to provide greater ease of use as well as improved performance and scalability. It is no longer necessary to precede Time Series Data Mining nodes with a Time Series Data Preparation node. Time series data can now be processed with or without a numeric TimeID variable. Time Series Data Mining now supports sequence data. Season and Trend information is now extracted and included in Time Series Data Mining results..

- The Decision Tree output displays have been enhanced to display variable precision values in the split branches and nodes. This change improves the usability of the decision tree tool when mining with extremely large and extremely small values.

Rapid Predictive Modeler (RPM)

- The neural network functions in the SAS Enterprise Miner Rapid Predictive Modeler (RPM) add-on for SAS Enterprise Guide have been revised. The new neural network functions provide a simplified architecture that is appropriate for business problems. The updated changes also improve the run-time performance of the RPM tool.

Support for High Performance Data Mining

SAS is developing a key set of statistical and data mining tasks that execute on a dedicated high performance appliance. SAS High Performance (HP) software distributes data, memory, and computations over a grid of systems that produces dramatic improvements in large data scalability and run times. Enterprise Miner 12.1 uses the SAS High Performance system for building predictive models.

- All SAS Enterprise Miner High Performance data mining procedures have been enhanced with new features. The HPFOREST and HP4SCORE procedures have been promoted from experimental to production status. For more information about SAS High Performance Data Mining procedures, see the SAS Enterprise Miner documentation page at `http://support.sas.com/documentation/onlinedoc/miner/`.

- The score code generation macros have been deprecated, in favor of CODE statements on the relevant HP data mining procedures. The High Performance Impute node now supports Winsorized and Trimmed data calculations.

- A High Performance data mining Data Validation node has been added. The new High Performance Data Validation node provides users with the ability to assign data rows into training or validation partitions when training models for generalization.

- A High Performance Forest node has been added to facilitate modeling of highly nonlinear data.

- The High Performance Neural node has been enhanced with new architecture and generalization options.

- All High Performance data mining model nodes support missing values in class variables as a distinct level. This technique improves both model accuracy and model deployment.

- All High Performance data mining model nodes now report both training and validation data set fit statistics, as well as lift and Receiver Operating Characteristic (ROC) values.

- The set of High Performance data mining nodes can now be connected to more of the traditional SAS Enterprise Miner nodes.

- SAS Enterprise Miner High Performance procedures and data mining nodes have been enhanced with support for multi-byte international data sources.