SAS® Enterprise Miner™ 13.2

Create highly accurate analytical models that enable you to predict with confidence

Overview

SAS Enterprise Miner streamlines the data mining process so you can create accurate predictive and descriptive analytical models that use vast amounts of data. You can use this software to detect fraud, minimize risk, anticipate resource demands, reduce asset downtime, increase response rates for marketing campaigns, curb customer attrition, and for customized purposes.

SAS Enterprise Miner is designed for those who need to analyze increasing volumes of data to identify and solve critical business or research issues – and help others make well-informed decisions. This includes data miners, statisticians, marketing analysts, database marketers, risk analysts, fraud investigators, engineers, scientists, and business analysts.

SAS Enterprise Miner is delivered as a distributed client/server system, providing an optimized architecture so that data scientists and business analysts can work more quickly to create accurate predictive and descriptive models, and produce results that can be shared and incorporated into business processes. To enhance the data mining process, this software is designed to work seamlessly with other SAS technologies, such as data integration, analytics, and reporting.

Why data mining?

Turning increasing volumes of data into useful information is a challenge for most organizations. Relationships and answers that identify key opportunities lie buried somewhere in all that data:

- Which customers will purchase what products and when?
- Which customers are leaving and what can be done to retain them?
- How should insurance rates be set to ensure profitability?
- How can you predict failures, reduce unnecessary maintenance, and increase uptime to optimize asset performance?

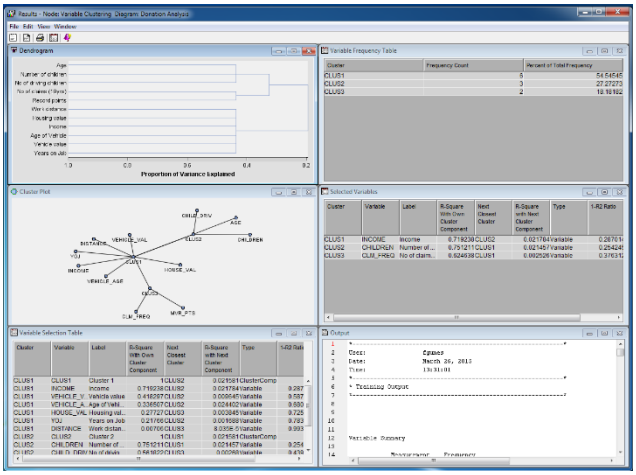
To get answers to complex questions and gain an edge in today's competitive market, powerful advanced analytic solutions are required. Discovering previously unknown patterns can help decision makers across your enterprise create effective strategies.

Benefits

- Understand key relationships and develop models intuitively and quickly
- Build better models more efficiently with a versatile data mining workbench
- Easily derive insights in a self-sufficient and automated manner
- Enhance the accuracy of predictions to ensure that the right decisions are made and the best actions are taken
- Ease model deployment and scoring processes for faster results

Statistical and Machine Learning techniques

SAS Enterprise Miner provides superior analytical depth with a suite of statistical, data mining, and machine learning algorithms. Decision trees, bagging and boosting, time series data mining, neural networks, memory-based reasoning, hierarchical clustering, linear and logistic regression, associations, sequence, and web path analysis, random forests, and support vector machine are all included. The breadth of analytical algorithms extends to industry-specific algorithms such as credit scoring and to state-of-the-art methods such as gradient boosting and least angular regression splines.

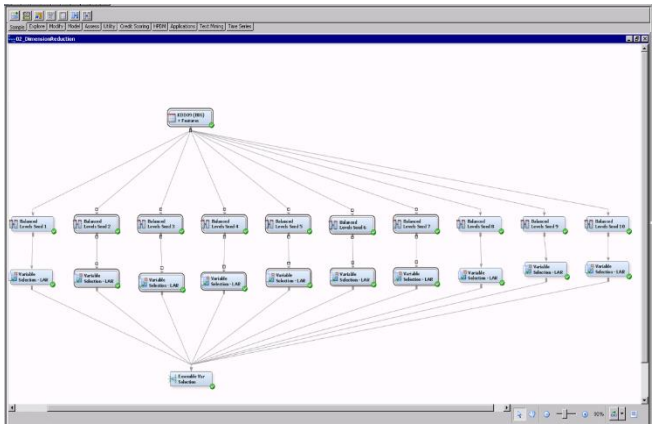


In-database and in-Hadoop scoring delivers faster results

Scoring is the process of regularly applying a model to new data for implementation into an operational environment. This can be tedious, especially when it entails manually rewriting or converting code, which delays model implementation and can introduce potentially costly errors. SAS Enterprise Miner automatically generates score code in SAS, C, Java, and PMML. The scoring code can be deployed in a variety of real-time or batch environments within SAS, on the web, or directly in relational databases or Hadoop.

Easy-to-use GUI

An easy-to-use, drag-and-drop interface is designed to appeal to analytic professionals. The advanced analytic algorithms are organized under core tasks that are performed in any successful data mining endeavor. The SAS data mining process encompasses five primary steps: sampling, exploration, modification, modelling, and assessment (SEMMA). By deploying nodes from the SEMMA toolbar, you can apply advanced statistics, identify the most significant variables, transform data elements with expression builders, develop models to predict outcomes, validate accuracy, and generate a scored data set with predicted values to deploy into your operational applications.



Sophisticated data preparation, summarization and exploration

A powerful set of interactive data preparation tools is available for addressing missing values, filtering outliers, and developing segmentation rules. Core data preparation tools include importing and appending files merging and dropping variables, and dealing with high cardinality variables. Extensive descriptive summarization features and interactive exploration tools enable even novice users to examine large amounts of data in dynamically linked, multidimensional plots.

Business-based model comparisons, reporting, and management

Models generated with different algorithms can be evaluated consistently using a highly visual assessment interface. Data miners can discuss results with business domain experts for improved collaboration and better results. An innovative *Cutoff* node examines posterior probability distributions to define the optimal actions for solving the business problem at hand.

A quick, easy, and self-sufficient way to generate models

SAS Rapid Predictive Modeler, a component of SAS Enterprise Miner, empowers business analysts and subject-matter experts with easy-to-use capabilities for quickly generating their own predictive models that are based on their specific needs and scenarios. This enables a wide range of individuals to use and benefit from predictive models without having to always rely on a potentially limited pool of advanced analytic resources.

Open, extensible design provides flexibility

You can easily integrate existing SAS models developed outside of the SAS Enterprise Miner environment into the process flow environment while maintaining full control of each syntax statement. The *Open Source* node enables integration with the R language inside a SAS Enterprise Miner flow. You can perform data transformation and exploration in addition to training and scoring supervised and unsupervised models in R. You can then seamlessly integrate the results, assess the model, and compare it to models generated by SAS Enterprise Miner. User defined *Extension* nodes enable integration using SAS code and XML logic, which opens the entire world of SAS to data miners.

Parallelized grid-enabled workbench

Scale from a single-user system to very large enterprise solutions with the Java client and SAS server architecture. Powerful servers can be dedicated to computing, while users move from office to home to remote sites without losing access to mining projects or services.

Distributable data mining system

Many process-intensive tasks, such as data sorting, summarization, variable selection and regression modeling, are multithreaded, and processes can be run in parallel for distribution and workload balancing across a grid of servers or scheduled for batch processing. SAS Enterprise Miner supports Windows servers and UNIX platforms, making it the software of choice for organizations that have large-scale data mining projects.

High-performance data mining

A select set of high-performance data mining nodes is included in SAS Enterprise Miner. Depending on the data and complexity of analysis, you might find performance gains in single-machine mode. When you need to process big data faster, a separate licensable product, SAS High-Performance Data Mining, enables you to develop timely and accurate predictive models for distributed computing envioronments.

For More Information

communities.sas.com/data-mining

sas.com/enterpriseminer