Make Better Decisions Through Predictive Intelligence

Highlights

- Easily access, prepare and model structured data with this intuitive, visual data mining workbench

- Rapidly build and validate models using the most advanced statistical and machine-learning techniques available

- Support the entire data mining process with a broad set of tools based on CRISP-DM methodology

- Efficiently deploy insight and predictive models on a scheduled basis or in real time

Data mining provides organizations with a clearer view of current conditions and deeper insight into future events. With IBM SPSS Modeler Professional, your organization can conduct data mining that takes advantage of many types of data, resulting in more revealing analysis and in-depth understanding of your customers or constituents. This comprehensive data mining workbench has powerful data preparation, visualization and predictive modeling capabilities to help you solve any business challenge faster, with more accurate results.

Modeler is popular worldwide with analysts and business users alike. Its unique features enable non-analysts to produce accurate models quickly and easily without sophisticated analytical skills, while professional analysts can take advantage of the software's advanced predictive modeling capabilities.

Streamline the data mining process

Modeler's intuitive graphical interface makes it easy for users to visualize every step of the data mining process as part of a "stream." By interacting with streams, analysts and business users can collaborate in adding business knowledge to the data mining process. Because data miners can focus on knowledge discovery rather than on technical tasks like writing code, they can pursue "train-of-thought" analysis, explore the data more deeply and uncover additional hidden relationships.

From this visual interface, you can easily access and integrate data from a number of sources, including IBM® SPSS® Data Collection products, and data in virtually any type of database, spreadsheet or flat file – including IBM® SPSS® Statistics, SAS® and Microsoft® Excel® files. No other data mining solution offers this versatility.

Modeler's powerful automation tools, such as automated data preparation and auto modeling, make it easy to prepare data for analysis, find the best model based on hidden patterns in data and quickly produce consistent and accurate results.
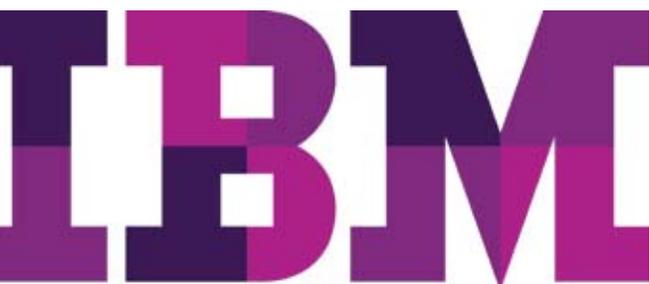
Leverage more types of data for better results

Our customers have found that incorporating all available types of data increases the "lift" or accuracy of predictive models, leading to more useful recommendations and improved outcomes.

If your organization collects large amounts of text data, the interactive text mining workbench available in IBM SPSS Modeler Premium will enable you to extract concepts and opinions from any type of text – such as the text captured in operationl sources, call center notes, customer e-mails, media or journal articles, blogs, RSS feeds and more. Direct access to survey data in Data Collection products makes it easy to include demographic, attitudinal and behavioral information in your models – rounding out your understanding of the people or organizations you serve.

Choose from an unparalleled breadth of techniques

Modeler offers an array of advanced data mining techniques that are designed to meet the needs of every data mining application, including all of the following algorithms.

- Classification algorithms – Make predictions or forecasts based on historical data using techniques such as Decision Tree, Neural Networks, Logistic Regression, Time- Series, Support Vector Machines, Cox regression and more. Leverage automatic classification modeling for both binary and numeric outcomes to streamline model creation.
- Segmentation algorithms – Group people or detect unusual patterns with automatic clustering, anomaly detection and clustering neural network techniques. Use automatic classification to apply multiple algorithms with a single step and take the guesswork out of selecting the right technique.
- Association algorithms – Discover associations, links or sequences using Apriori, CARMA and sequential association.

Optimize your current information technologies

Modeler's open and scalable architecture makes the best use of your existing IT infrastructure. It integrates with your existing systems, both when accessing data and when deploying results, so you don't need to move data into and out of a proprietary format. And techniques such as in-database mining, multithreading, server clustering and SQL pushback help you conserve resources, deliver results faster and reduce overall IT costs.

Follow a proven, repeatable process

During every phase of the data mining process, Modeler supports the de facto industry standard, the CRoss-Industry Standard Process for Data Mining (CRISP-DM). This means your company can focus on solving business problems through data mining, rather than on reinventing a new process for every project. Individual Modeler projects can be efficiently organized using the CRISP-DM project manager.
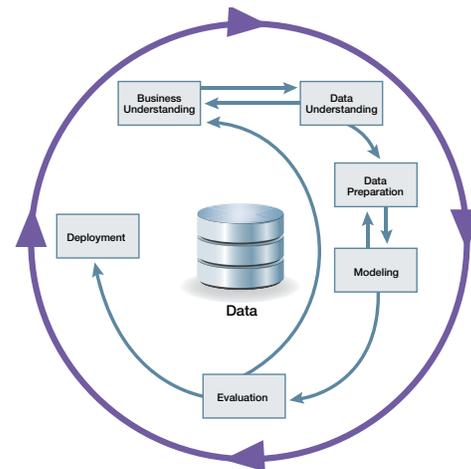
The CRISP-DM process, as shown in this diagram, enables data miners to implement efficient data mining projects that yield measurable business results.

Deploy predictive modeling across the enterprise

Modeler can efficiently analyze the amounts of data typically generated by small to mid-sized organizations. Organizations with high-volume or complex data mining requirements leverage IBM SPSS Modeler Server. Using client/server architecture, Modeler Server allows many data analysts to work simultaneously without straining computing resources. You can take advantage of in-database mining on leading information platforms and efficiently process large amounts of data. Modeler Server also offers additional deployment options to help you extend the benefits of data mining across geographic or functional lines and put results in the hands of decision makers quickly.

What's new in Modeler Professional 14

This release includes new features and enhancements that will enable you to create and interpret models easily using cutting-edge techniques, integrate seamlessly with other IBM SPSS software and third-party technologies and embed predictive modeling into your organization's business processes.

Performance improvements

- Improve the stability and accuracy of your models by leveraging large dataset optimization techniques, including boosting and bagging, for Neural Net, Linear and Decision Tree algorithms
- Run many models at once and interact with them using a new visualization tool that enables you to better understand the results of ensemble models and share them with others in your organization
- Enhance scalability and performance when using Modeler Professional Server by leveraging the new large database processing optimization capability for key algorithms. Build and refresh models on databases of unlimited size for enterprise-scale processes

Updated algorithms

- The new Neural Net algorithm supports new analysis methods and includes multilayer perceptron and radial basis functions. It includes innovative, interactive visualization that makes it easy to understand and communicate results.
- Build better linear models using a new Linear Regression method that leverages large database processing, has built-in automatic data preparation options and produces rich visualizations that make it easy to interpret model results interactively

Data enhancements

- Extended support for enterprise data sources with the ability to read and write data to and from XML
- Leverage the strength of your operational databases better with additional in database mining options and push Modeler results back to operational database tables from the interface
- Simplify the re-use of modeling streams across users and ensure proper parameter settings with runtime parameter prompts
- Gain more control when exporting results to Excel by adding to an existing workbook or specifying where results should be placed in a spreadsheet

Improved platform support and deployment

- Improve deployment and scoring using a visual deployment definition that includes automatic model rebuilding, branching and model refresh capabilities
- Manage enterprise-wide login standards with new support for standard single sign-on technology (SSO)

Features
Data understanding
- Create a wide range of interactive graphs with automatic assistance
- Use visual link analysis to see the associations in your data
- Interact with data by selecting regions or items on a graph and viewing the selected information; or select key data for use in analysis
- Access Statistics graphs and reporting tools directly from Modeler

Data preparation
- Access operational data from IBM DB2®, Oracle®, Microsoft SQL Server™, Informix®, Neoview, Netezza, mySQL (Sun) and Teradata data sources
- Import delimited and fixed-width text files, Statistics files, SAS, Data Collection data sources or XML
- Use multiple data-cleaning options to remove or replace invalid data, automatically impute missing values and mitigate for outliers and extremes
- Apply automatic data preparation to interrogate and condition data for analysis in a single step
- Export data to delimited text files, Excel, Statistics, SAS and operational databases
- Field filtering, naming, derivation, binning, re-categorization, value replacement and field reordering
- Record selection, sampling, merging and concatenation; sorting, aggregation and balancing
- Data restructuring, partitioning and transposition
- Extensive string functions: string creation, substitution, search and matching, whitespace removal and truncation
- Access data management and transformations performed in Statistics directly from Modeler

- RFM scoring: aggregate customer transactions to provide Recency, Frequency, and Monetary value scores and combine these to produce a complete RFM analysis

Modeling and evaluation
- Employ advanced data mining algorithms to get the best results from your data
- Use interactive model and equation browsers and view advanced statistical output
- Show relative impact of data attributes on predicted outcomes with variable importance graphs
- Combine multiple models (ensemble models) or use one model to analyze a second model
- Use automatic (binary and numeric) classification and automatic clustering in place of individual algorithms
- Use Modeler's Component-Level Extension Framework (CLEF) to integrate custom algorithms
- Through the integration of Statistics, use R to extend analysis options

Modeling algorithms included
- C&RT, C5.0, CHAID & QUEST – Decision tree algorithms including interactive tree building
- Decision List – Interactive rule-building algorithm
- K-Means, Kohonen, Two Step, Discriminant, Support Vector Machine (SVM) – Clustering and segmentation algorithms
- Factor/PCA, Feature Selection – Data reduction algorithms
- Regression, Linear, GenLin (GLM) – Linear equation modeling
- Self-learning response model (SLRM) – Bayesian model with incremental learning
- Time-series – Generate and automatically select time-series forecasting models

- Neural Networks – Multi-layer perceptrons with back-propagation learning, and radial basis function networks
- Support Vector Machines – Advanced algorithm for wide datasets
- Bayesian Networks – Graphical probabilistic models
- Cox regression – Calculate likely time to an event
- Anomaly Detection – Cluster-based algorithm for detecting unusual results
- KNN – Nearest neighbor modeling and scoring algorithm
- Apriori – Popular association discovery algorithm with advanced evaluation functions
- CARMA – Association algorithm which supports multiple consequents
- Sequence – Sequential association algorithm for order-sensitive analyses

Deployment
- Export models using SQL or PMML (the XML-based standard format for predictive models)
- Leverage IBM SPSS Collaboration and Deployment Services for innovative analytics management, process automation and deployment capabilities

Modeler server (optional*)
- Use in-database mining to build models in the database using leading database technologies and leverage high-performance database implementations
- Leverage high-performance hardware, experience quicker time-to solution, and achieve greater ROI through parallel execution of streams and multiple models
- Transmit sensitive data securely between Modeler Client and Modeler Server through secure sockets layer (SSL) encryption