Make Better Decisions Through Predictive Intelligence

Highlights

- Easily access, prepare and model structured data with this intuitive, visual data mining workbench
- Expand the benefits of business analytics through integration with Cognos® Business Intelligence software and InfoSphere™ Warehouse
- Analyze data stored in legacy systems with IBM Classic Federation Server and zDB2® support
- Rapidly build and validate models using the most advanced statistical and machine-learning techniques available
- Efficiently deploy insight and predictive models
- More deployment options through zLinux, SuSE Linux® Enterprise Server, and inclusion into IBM Smart Analytics System for Power

Data mining provides organizations with a clearer view of current conditions and deeper insight into future events.

IBM SPSS Modeler Professional* is a data mining workbench for the analysis of structured numerical data to model outcomes and to make predictions that inform business decisions with predictive intelligence. Predictive intelligence creates more effective strategies because it allows organizations not only to evaluate trends – benchmarking results, plans and performance but also to look into the future – evaluating likely outcomes and understanding how the interplay of factors affects those outcomes.

The latest release of IBM SPSS Modeler Professional strengthens its integration with IBM's Cognos® 8 Business Intelligence and InfoSphere Warehouse solutions – allowing users to create a best-of -breed solution from a single vendor that extends from the storage and management of data to the deployment of predictive intelligence to decision makers. Now the thousands of organizations worldwide who trust Cognos applications to monitor their vital signs will be able to incorporate predictive intelligence into their reports and dashboards – for greater insight into the future and an improved ability to take advantage of evolving opportunities.

The reach of predictive intelligence will stretch even further now that IBM SPSS Modeler Professional software can be hosted in mainframe computing environments. Read more about these enhancements on page 2.

Streamline the data mining process

IBM SPSS Modeler is popular worldwide with analysts and business users alike. Its automated data preparation and modeling 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 data preparation and predictive modeling capabilities.

** IBM SPSS Modeler Professional was formerly called PASW® Modeler.*

Business Benefits:

Organizations of all types have found that they can benefit from using IBM SPSS Modeler Professional:

- Businesses can attract customers, strengthen their loyalty, reduce customer attrition more cost effectively and also reduce risk
- Public sector organizations can predict workforce capacity, evaluate program effectiveness and proactively respond to public safety issues
- Educational institutions can manage the student lifecycle, improve classroom performance and address many other operational challenges

IBM SPSS 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 these 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 – and, as mentioned above, you now can access data directly from Cognos 8 Business Intelligence.

No other data mining solution offers this versatility. IBM SPSS 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.

What's new in Modeler Professional 14.1

The enhancements to Modeler Professional in this release bring predictive intelligence to a broader range of organizations, enabling them to be more focused and agile in their planning and daily decision-making because they will have a more nuanced understanding of their own enterprises, the environment in which they operate, and of their customers and other stakeholders. New capabilities include:

- Integration with Cognos. Analysts now can access data from their Cognos 8 Business Intelligence environment directly within the Modeler interface. Cognos software organizes and delivers a complete and consistent view of information for enterprise-wide decision-making. By adding Modeler's analytic capabilities, organizations can quickly and reliably evaluate the likelihood of specific outcomes, using their familiar enterprise-wide data view. Also – since Modeler can write results into Cognos 8 Business Intelligence – they can make predictive intelligence available to business users and all information stakeholders who rely on Cognos as their informational portal into enterprise analytics.

*** IBM SPSS Data Collection and IBM SPSS Statistics were formerly called PASW® Data Collection and PASW® Statistics.*

- **Enhanced InfoSphere™ and DB2® integration.** Organizations with an InfoSphere data warehouse can now use the full range its data mining algorithms from within Modeler. This gives users the advantage of an intuitive, graphical interface without compromising data mining performance. Supported algorithms now include logistic regression, Naïve Bayes, time-series and radial basis function (RBF). Additional DB2 access options, such as compression and partitioning, ensures that organizations can fully leverage large-scale data sources easily.
- **Mining against mainframe data.** Mainframe computers typically hold a wealth of data about an organization's operational history. With support for zDB2 and IBM Classic Federation server, these organizations can see what effect past events will have on the future. This will help them understand current activity more clearly, evaluate changing business realities and base their planning on reliable predictive intelligence.
- **Support for Linux® on System z®.** Running Linux on System z offers many advantages to organizations seeking to simplify complex information systems while meeting today's requirements for security, transparency and cost-control. IBM SPSS Modeler Professional Server is now supported in these environments.

Include 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 operational sources, call center notes, customer emails, media or journal articles, blogs, RSS feeds and more. Direct access to survey data in IBM SPSS 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

IBM SPSS 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.

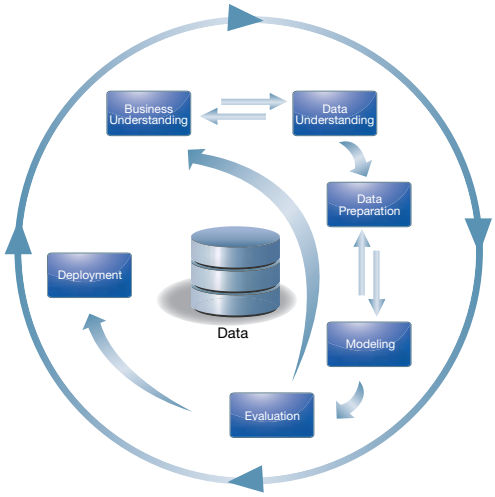
- Segmentation algorithms – Group people or detect unusual patterns with automatic clustering, anomaly detection and clustering neural network techniques. Use automatic classification to apply multiple algorithms with a single step and take the guesswork out of selecting the right technique.
- Association 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, multi-threading, 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

IBM SPSS 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.

About IBM Business Analytics

IBM Business Analytics software delivers complete, consistent and accurate information that decision-makers trust to improve business performance. A comprehensive portfolio of business intelligence, predictive analytics, financial performance and strategy management, and analytic applications provides clear, immediate and actionable insights into current performance and the ability to predict future outcomes. Combined with rich industry solutions, proven practices and professional services, organizations of every size can drive the highest productivity, confidently automate decisions and deliver better results.

As part of this portfolio, IBM SPSS Predictive Analytics software helps organizations predict future events and proactively act upon that insight to drive better business outcomes. Commercial, government and academic customers worldwide rely on IBM SPSS technology as a competitive advantage in attracting, retaining and growing customers, while reducing fraud and mitigating risk. By incorporating IBM SPSS software into their daily operations, organizations become predictive enterprises – able to direct and automate decisions to meet business goals and achieve measurable competitive advantage. For further information or to reach a representative visit www.ibm.com/spss.

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 Cognos® Business Intelligence, IBM DB2®, Oracle®, Microsoft SQL Server™, Informix®, Neoview, Netezza, mySQL (Sun) and Teradata data sources, as well as mainframe data through zDB2 and IBM Classic Federation Server support
- 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, Cognos Business Intelligence packages and operational databases
- 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 Machine – 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

- IBM InfoSphere in-database mining algorithms supported: Decision Tree, Association, Sequence, Regression, Logistic Regression, Clustering, Naïve Bayes, Time-Series and Radial Basis Function (RBF)
- Microsoft SQL Server in-database mining algorithms supported: Decision Tree, Association Rules, Linear Regression, Clustering, Sequence Clustering, Naïve Bayes, Time-Series and Neural Network
- Oracle in-database mining algorithms supported: Decision Tree, General Linear Model (GLM), O-Cluster (Orthogonal Partitioning Clustering), KMeans, Apriori, Minimum Description
  Length (MDL), Support Vector Machine, Naïve Bayes, Adaptive Bayes, Non-Negative Matrix nFactorization and Artificial Intelligence (AI)

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
- Use SQL-pushback to push data transformations and select modeling algorithms directly into your operational databases
- Leverage high-performance hardware, including IBM System z machines, to 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

IBM Corporation
Route 100
Somers, NY 10589

Produced in the United States of America
May 2010

Please Recycle

Business Analytics software

IMD14303USEN-02