Ibm Spss Modeler Algorithms Guide

Interpreting Quantitative Data with IBM SPSS StatisticsHandbook on Constructing Composite Indicators: Methodology and User GuideSPSS Statistics 17.0Machine Learning for Data MiningDecision Trees and Applications with IBM SPSS ModelerIBM Predictive Maintenance and Quality 2.0 Technical OverviewOptimization and Decision Support Design Guide: Using IBM ILOG Optimization Decision ManagerRisk Scoring for a Loan Application on IBM System z: Running IBM SPSS Real-Time AnalyticsImplementing an Optimized Analytics Solution on IBM Power SystemsEnabling Real-time Analytics on IBM z Systems PlatformUsing IBM® SPSS® Statistics for Research Methods and Social Science StatisticsAdvanced Analytics MethodologiesIBM Software for SAP SolutionsPredictive Analytics and Data MiningPredictive Analytics For DummiesIBM SPSS Modeler CookbookData Mining and Statistics for Decision MakingPerforming Data Analysis Using IBM SPSS資料探勘: 人工智慧與機器學習發展以SPSS Modeler為範例IBM SPSS Modeler EssentialsData Mining with SPSS ModelerA Concise Guide to Market ResearchWorkload Optimized Systems: Tuning POWER7 for AnalyticsSPSS For DummiesData Mining With ConfidenceEffective CRM using Predictive AnalyticsIBM Software Defined Infrastructure for Big Data Analytics WorkloadsHandbook of Statistical Analysis and Data Mining ApplicationsData Mining with Rattle and RProceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013)Harness the Power of Big Data The IBM Big Data PlatformMachine Learning for Decision MakersDiscovering Statistics Using IBM SPSS StatisticsAccelerating Data Transformation with IBM DB2 Analytics Accelerator for z/OSSystems of Insight for Digital Transformation: Using IBM Operational Decision Manager Advanced and Predictive AnalyticsLearning IBM Watson AnalyticsReal-time Fraud Detection Analytics on IBM System zMethodologies of Pattern RecognitionApplied Predictive AnalyticsData Mining Techniques in CRM

Interpreting Quantitative Data with IBM SPSS Statistics

SPSS (Statistical Package for the Social Sciences) is a data management and analysis software that allows users to generate solid, decision-making results by performing statistical analysis This book provides just the information needed: installing the software, entering data, setting up calculations, and analyzing data Covers computing cross tabulation, frequencies, descriptive ratios, means, bivariate and partial correlations, linear regression, and much more Explains how to output information into striking charts and graphs For ambitious users, also covers how to program SPSS to take their statistical analysis to the next level

Handbook on Constructing Composite Indicators: Methodology and User Guide

This IBM® Redbooks® publication addresses topics to use the virtualization strengths of the IBM POWER8® platform to solve clients' system resource utilization challenges and maximize systems' throughput and capacity. This book addresses performance tuning topics that will help answer clients' complex analytic workload requirements, help maximize systems' resources, and provide expert-level documentation to transfer the how-to-skills to the worldwide teams.

This book strengthens the position of IBM Analytics and Big Data solutions with a well-defined and documented deployment model within a POWER8 virtualized environment, offering clients a planned foundation for security, scaling, capacity, resilience, and optimization for analytics workloads. This book is targeted toward technical professionals (analytics consultants, technical support staff, IT Architects, and IT Specialists) who are responsible for providing analytics solutions and support on IBM Power SystemsTM.

SPSS Statistics 17.0

A wide range of applications, such as R, SAS, MATLAB, and SPSS Statistics, provide a huge toolbox of methods to analyze large data and can be used by experts to find patterns and interesting structures in the data. Many of these tools are mainly programming languages, which assumes the analyst has deeper programming skills and an advanced background in IT and mathematics. Since this field is becoming more important, graphic user-interfaced data analysis software is starting to enter the market, providing "drag and drop" mechanisms for career changers and people who are not experts in programming or statistics.One of these easy to handle, data analytics applications is the IBM SPSS Modeler. This book is dedicated to the introduction and explanation of its data analysis power and focused in decision trees. The more important topics are the next: Decision Tree Models General Uses of Tree-Based Analysis C&RT Algorithms CHAID Algorithms QUEST Algorithms C5.0 Algorithms Decision Trees with IM SPSS Modeler Building a Decision Tree with the C5.0 Node Building a decision tree with the CHAID node The C&R Tree node and variable generation The QUEST node-Boosting & Imbalanced data Detection of diabetes-comparison of decision tree nodes Rule set and cross-validation with C5.0 The Auto Classifier Node Building a Stream with the Auto Classifier Node The Auto Classifier Model Nugget Models for credit rating with the Auto Classifier node SVM classifier Interactive decision Trees with IBM SPSS Modeler The Interactive Tree Builder Growing and Pruning the Tree Defining Custom Splits Customizing the Tree View Gains Risks The Growing Directives Generation Filter and Select Nodes Building a Tree Model Directly C&R Tree, CHAID, QUEST, and C 5.0 Models Nuggets Model Nuggets for Boosting, Bagging and Very Large Datasets

Machine Learning for Data Mining

Regarding online transaction processing (OLTP) workloads, IBM® z SystemsTM platform, with IBM DB2®, data sharing, Workload Manager (WLM), geoplex, and other high-end features, is the widely acknowledged leader. Most customers now integrate business analytics with OLTP by running, for example, scoring functions from transactional context for real-time analytics or by applying machine-learning algorithms on enterprise data that is kept on the mainframe. As a result, IBM adds investment so clients can keep the complete lifecycle for data analysis, modeling, and scoring on z Systems control in a cost-efficient way, keeping the qualities of services in availability, security, reliability that z Systems solutions offer. Because of the changed architecture and tighter integration, IBM has shown, in a customer proof-of-concept, that a particular client was able to achieve an orders-of-magnitude improvement in performance, allowing that client's data scientist to investigate the data in a more interactive process. Open technologies, such as

Predictive Model Markup Language (PMML) can help customers update single components instead of being forced to replace everything at once. As a result, you have the possibility to combine your preferred tool for model generation (such as SAS Enterprise Miner or IBM SPSS® Modeler) with a different technology for model scoring (such as Zementis, a company focused on PMML scoring). IBM SPSS Modeler is a leading data mining workbench that can apply various algorithms in data preparation, cleansing, statistics, visualization, machine learning, and predictive analytics. It has over 20 years of experience and continued development, and is integrated with z Systems. With IBM DB2 Analytics Accelerator 5.1 and SPSS Modeler 17.1, the possibility exists to do the complete predictive model creation including data transformation within DB2 Analytics Accelerator. So, instead of moving the data to a distributed environment, algorithms can be pushed to the data, using cost-efficient DB2 Accelerator for the required resource-intensive operations. This IBM Redbooks® publication explains the overall z Systems architecture, how the components can be installed and customized, how the new IBM DB2 Analytics Accelerator loader can help efficient data loading for z Systems data and external data, how in-database transformation, in-database modeling, and in-transactional real-time scoring can be used, and what other related technologies are available. This book is intended for technical specialists and architects, and data scientists who want to use the technology on the z Systems platform. Most of the technologies described in this book require IBM DB2 for z/OS®. For acceleration of the data investigation, data transformation, and data modeling process, DB2 Analytics Accelerator is required. Most value can be achieved if most of the data already resides on z Systems platforms, although adding external data (like from social sources) poses no problem at all.

Decision Trees and Applications with IBM SPSS Modeler

This is a practical cookbook with intermediate-advanced recipes for SPSS Modeler data analysts. It is loaded with step-by-step examples explaining the process followed by the experts.If you have had some hands-on experience with IBM SPSS Modeler and now want to go deeper and take more control over your data mining process, this is the guide for you. It is ideal for practitioners who want to break into advanced analytics.

IBM Predictive Maintenance and Quality 2.0 Technical Overview

Data mining is the process of automatically searching large volumes of data for models and patterns using computational techniques from statistics, machine learning and information theory; it is the ideal tool for such an extraction of knowledge. Data mining is usually associated with a business or an organization's need to identify trends and profiles, allowing, for example, retailers to discover patterns on which to base marketing objectives. This book looks at both classical and recent techniques of data mining, such as clustering, discriminant analysis, logistic regression, generalized linear models, regularized regression, PLS regression, decision trees, neural networks, support vector machines, Vapnik theory, naive Bayesian classifier, ensemble learning and detection of association rules. They are discussed along with illustrative examples throughout the book to explain the theory of these methods, as well as their strengths and limitations. Key Features: Presents a comprehensive introduction to all techniques used in data mining and statistical learning, from classical to latest techniques. Starts from basic principles up to advanced concepts. Includes many step-by-step examples with the main software (R, SAS, IBM SPSS) as well as a thorough discussion and comparison of those software. Gives practical tips for data mining implementation to solve real world problems. Looks at a range of tools and applications, such as association rules, web mining and text mining, with a special focus on credit scoring. Supported by an accompanying website hosting datasets and user analysis. Statisticians and business intelligence analysts, students as well as computer science, biology, marketing and financial risk professionals in both commercial and government organizations across all business and industry sectors will benefit from this book.

Optimization and Decision Support Design Guide: Using IBM ILOG Optimization Decision Manager

Boost your Big Data IQ! Gain insight into how to govern and consume IBM's unique in-motion and at-rest Big Data analytic capabilities Big Data represents a new era of computing—an inflection point of opportunity where data in any format may be explored and utilized for breakthrough insights—whether that data is in-place, in-motion, or at-rest. IBM is uniquely positioned to help clients navigate this transformation. This book reveals how IBM is infusing open source Big Data technologies with IBM innovation that manifest in a platform capable of "changing the game." The four defining characteristics of Big Data—volume, variety, velocity, and veracity—are discussed. You'll understand how IBM is fully committed to Hadoop and integrating it into the enterprise. Hear about how organizations are taking inventories of their existing Big Data assets, with search capabilities that help organizations discover what they could already know, and extend their reach into new data territories for unprecedented model accuracy and discovery. In this book you will also learn not just about the technologies that make up the IBM Big Data platform, but when to leverage its purpose-built engines for analytics on data in-motion and data at-rest. And you'll gain an understanding of how and when to govern Big Data, and how IBM's industry-leading InfoSphere integration and governance portfolio helps you understand, govern, and effectively utilize Big Data. Industry use cases are also included in this practical guide.

Risk Scoring for a Loan Application on IBM System z: Running IBM SPSS Real-Time Analytics

A step-by-step guide to data mining applications in CRM. Following a handbook approach, this book bridges the gap between analytics and their use in everyday marketing, providing guidance on solving real business problems using data mining techniques. The book is organized into three parts. Part one provides a methodological roadmap, covering both the business and the technical aspects. The data mining process is presented in detail along with specific guidelines for the development of optimized acquisition, cross/ deep/ up selling and retention campaigns, as well as effective customer segmentation schemes. In part two, some of the most useful data mining algorithms are explained in a simple and comprehensive way for business users with no technical expertise. Part three is packed with real world case studies which employ the use of three leading data mining tools: IBM SPSS Modeler, RapidMiner and Data Mining for Excel. Case studies from industries including banking, retail and telecommunications are presented in detail so as to serve as templates for developing similar applications. Key Features: Includes numerous real-world case studies which are presented step by step, demystifying the usage of data mining models and clarifying all the methodological issues. Topics are presented with the use of three leading data mining tools: IBM SPSS Modeler, RapidMiner and Data Mining for Excel. Accompanied by a website featuring material from each case study, including datasets and relevant code. Combining data mining and business knowledge, this practical book provides all the necessary information for designing, setting up, executing and deploying data mining techniques in CRM. Effective CRM using Predictive Analytics will benefit data mining practitioners and consultants, data analysts, statisticians, and CRM officers. The book will also be useful to academics and students interested in applied data mining.

Implementing an Optimized Analytics Solution on IBM Power Systems

Introducing the IBM SPSS Modeler, this book guides readers through data mining processes and presents relevant statistical methods. There is a special focus on step-by-step tutorials and well-documented examples that help demystify complex mathematical algorithms and computer programs. The variety of exercises and solutions as well as an accompanying website with data sets and SPSS Modeler streams are particularly valuable. While intended for students, the simplicity of the Modeler makes the book useful for anyone wishing to learn about basic and more advanced data mining, and put this knowledge into practice.

Enabling Real-time Analytics on IBM z Systems Platform

Get efficient in performing data mining and machine learning using IBM SPSS Modeler Key Features Learn how to apply machine learning techniques in the field of data science Understand when to use different data mining techniques, how to set up different analyses, and how to interpret the results A step-by-step approach to improving model development and performance Book Description Machine learning (ML) combined with data mining can give you amazing results in your data mining work by empowering you with several ways to look at data. This book will help you improve your data mining techniques by using smart modeling techniques. This book will teach you how to implement ML algorithms and techniques in your data mining work. It will enable you to pair the best algorithms with the right tools and processes. You will learn how to identify patterns and make predictions with minimal human intervention. You will build different types of ML models, such as the neural network, the Support Vector Machines (SVMs), and the Decision tree. You will see how all of these models works and what kind of data in the dataset they are suited for. You will learn how to combine the results of different models in order to improve accuracy. Topics such as removing noise and handling errors will give you an added edge in model building and optimization. By the end of this book, you will be able to build predictive models and extract information of interest from the dataset What you will learn Hone your model-building skills and create the most accurate models Understand how predictive machine learning models work Prepare your data to acquire the best possible results Combine models in order to suit the requirements of different types of data Analyze single and multiple models and understand their combined results Derive worthwhile insights from your data using histograms and graphs Who this book is for If you are a data scientist, data analyst, and data mining professional and are keen to achieve a 30% higher salary by adding machine learning to your skillset, then this is the ideal book for you. You will learn to apply machine learning techniques to various data mining challenges. No prior knowledge of machine learning is assumed.

Using IBM® SPSS® Statistics for Research Methods and Social Science Statistics

Advanced Analytics Methodologies is today's definitive guide to analytics implementation for MBA and university-level business students and sophisticated practitioners. Its expanded, cutting-edge coverage helps readers systematically "jump the gap" between their organization's current analytical capabilities and where they need to be. Step by step, Michele Chambers and Thomas Dinsmore help readers customize a complete roadmap for implementing analytics that supports unique corporate strategies, aligns with specific corporate cultures, and serves unique customer and stakeholder communities. Drawing on work with dozens of leading enterprises, Michele Chambers and Thomas Dinsmore provide advanced applications and examples not available elsewhere, describe high-value applications from many industries, and help you systematically identify and deliver on your company's best opportunities. They show how to: Go beyond the Analytics Maturity Model: power your unique business strategy with an equally focused analytics strategy Link key business objectives with core characteristics of your organization, value chain, and stakeholders Take advantage of game changing opportunities before competitors do Effectively integrate the managerial and operational aspects of analytics Measure performance with dashboards, scorecards, visualization, simulation, and more Prioritize and score prospective analytics projects Identify "Quick Wins" you can implement while you're planning for the long-term Build an effective Analytic Program Office to make your roadmap persistent Update and revise your roadmap for new needs and technologies This advanced text will serve the needs of students and faculty studying cutting-edge analytics techniques, as well as experienced analytics leaders and professionals including Chief Analytics Officers; Chief Data Officers; Chief Scientists; Chief Marketing Officers; Chief Risk Officers; Chief Strategy Officers; VPs of Analytics or Big Data; data scientists; business strategists; and many line-of-business executives.

Advanced Analytics Methodologies

The second edition of Interpreting Quantitative Data with IBM SPSS Statistics is an invaluable resource for students analysing quantitative data for the first time. The book clearly sets out a range of statistical techniques and their common applications, explaining their logic and links to the research process. It also shows how SPSS can be used as a tool to aid analysis. Key features of the second edition include: - new chapters on one-way and two-way ANOVA, the Chi-square test and linear regression. - SPSS lab sessions following each chapter which demonstrate how SPSS can be used in practice - sets of exercises and 'real-life' examples to aid teaching and learning - lists of key terms to aid revision and further reading to enhance students' understanding - an improved text design making the book easier to navigate - a companion website with answers to the labs and exercises, along with additional data sets and powerpoint slides

IBM Software for SAP Solutions

This IBM® Redbooks® publication addresses topics to help clients to take advantage of the virtualization strengths of the POWER® platform to solve system resource utilization challenges and maximize system throughput and capacity. This publication examines the tools, utilities, documentation, and other resources available to help technical teams provide business solutions and support for Cognos® Business Intelligence (BI) and Statistical Package for the Social Sciences (SPSS®) on Power SystemsTM virtualized environments. This book addresses topics to help address complex high availability requirements, help maximize the availability of systems, and provide expert-level documentation to the worldwide support teams. This book strengthens the position of the Cognos and SPSS solutions with a well-defined and documented deployment model within a POWER system virtualized environment. This model provides clients with a planned foundation for security, scaling, capacity, resilience, and optimization. This book is targeted toward technical professionals (BI consultants, technical support staff, IT Architects, and IT Specialists) who are responsible for providing Smart Analytics solutions and support for Cognos and SPSS on Power Systems.

Predictive Analytics and Data Mining

Make the most advanced predictive analytical processes easy using Watson Analytics with this easy-to-follow practical guide About This Book This is the first and the only book on IBM Watson Analytics, and it shows you how to leverage Watson in an enterprise environment through rich use cases Incorporate Watson Analytics into your business strategy and confidently add this cutting edge expertise to your resume This book is written by James D Miller, IBM-certified expert and accomplished Director and Sr. Project Leader Who This Book Is For If you want to perform data discovery and analysis and make sense of data you have, this book for you. Data scientists can also use this book to explore a new way to perform data analysis tasks on cloud with ease. This book does not require a programming background. What You Will Learn Study the language of Watson while you discover how easy it is to access and configure Review what a Watson use case is, why it's important, and how to identify one Design Watson Analytical solutions based upon your use cases Understand the basic concepts behind the content analysis cycle and where Watson fits in Explore all the features of Watson, such as Explore, Predict, and Assemble Customize and extend your Watson solutions Use Watson at the Enterprise level Integrate Watson with other toolsets In Detail Today, only a small portion of businesses actually use a real analytical tool as part of routine decision making. IBM Watson Analytics is changing that making the most advanced and predictive analytical techniques understandable and usable for any industry. This book will be the vital tour guide for your trip, starting with what IBM Watson Analytics is. We'll start off with introduction to Watson Analytics and then quickly move on to various use cases under which one can use the different analytics functionalities offered by Watson. During the course of the book, you will learn how to design solutions, and customize and extend Watson analytics. We will conclude by taking Watson Analytics to enterprise and integrating it with other solutions (other IBM solutions and analytics). Now is the time for you to learn IBM Watson to compete in the world. Style and approach Watson provides individuals with the ability to perform sophisticated data discovery and analysis without all of the complexity that usually goes along with it. This book will get you started with Watson analytics and how you can use it in day-to-day data analysis. The book introduces the key concepts and terminology and then uses practical use case examples to reinforce your understanding.

Predictive Analytics For Dummies

Today many organizations face challenges when developing a realistic plan or schedule that provides the best possible balance between customer service and revenue goals. Optimization technology has long been used to find the best solutions to complex planning and scheduling problems. A decision-support environment that enables the flexible exploration of all the trade-offs and sensitivities needs to provide the following capabilities: Flexibility to develop and compare realistic planning and scheduling scenarios Quality sensitivity analysis and explanations Collaborative planning and scenario sharing Decision recommendations This IBM® Redbooks® publication introduces you to the IBM ILOG® Optimization Decision Manager (ODM) Enterprise. This decision-support application provides the capabilities you need to take full advantage of optimization technology. Applications built with IBM ILOG ODM Enterprise can help users create, compare, and understand planning or scheduling scenarios. They can also adjust any of the model inputs or goals, and fully understanding the binding constraints, trade-offs, sensitivities, and business options. This book enables business analysts, architects, and administrators to design and use their own operational decision management solution.

IBM SPSS Modeler Cookbook

Data mining is the art and science of intelligent data analysis. By building knowledge from information, data mining adds considerable value to the ever increasing stores of electronic data that abound today. In performing data mining many decisions need to be made regarding the choice of methodology, the choice of data, the choice of tools, and the choice of algorithms. Throughout this book the reader is introduced to the basic concepts and some of the more popular algorithms of data mining. With a focus on the hands-on end-to-end process for data mining, Williams guides the reader through various capabilities of the easy to use, free, and open source Rattle Data Mining Software built on the sophisticated R Statistical Software. The focus on doing data mining rather than just reading about data mining is refreshing. The book covers data understanding, data preparation, data refinement, model building, model evaluation, and practical deployment. The reader will learn to rapidly deliver a data mining project using software easily installed for free from the Internet. Coupling Rattle with R delivers a very sophisticated data mining environment with all the power, and more, of the many commercial offerings.

Data Mining and Statistics for Decision Making

Key Message:SPSS®17.0: Advanced Statistical Procedures Companioncontains valuable tips, warnings, and examples that will help you take advantage of SPSS and better analyze data. This book offers clear and concise explanations and examples of advanced statistical procedures in the SPSS Advanced and Regression modules. Key Topics: Model Selection Loglinear Analysis; Logit Loglinear Analysis; Multinomial Logistic Regression; Ordinal Regression; Probit Regression; Kaplan-Meier Survival Analysis; Life Tables; Cox Regression; Variance Components; Linear Mixed Models; Generalized Linear Models; Generalized Estimating Equations; Nonlinear Regression; Two-Stage Least-Squares Regression; Weighted Least-Squares Regression; Multidimensional Scaling Market: for all readers interested in SPSS.

Performing Data Analysis Using IBM SPSS

This IBM® RedpaperTM publication updated technical overview provides essential details about the data processing steps, message flows, and analytical models that power IBM Predictive Maintenance and Quality (PMQ) Version 2.0. The new version of PMQ builds on the first one, released in 2013, to help companies efficiently monitor and maintain production assets and improve their overall availability, utilization, and performance. It analyzes various types of data to detect failure patterns and poor quality parts earlier than traditional quality control methods, with the goal of reducing unscheduled asset downtime and improving quality metrics. Version 2.0 includes an improved method of interacting with the solution's analytic data store using an API from the new Analytics Solution Foundation, a reusable, configurable, and extensible component that supports a number of the solution's analytic functions. The new version also changes the calculation of profiles and KPIs, which is now done using orchestrations that are defined in XML. This updated technical overview provides details about these new orchestration definitions.

資料探勘：人工智慧與機器學習發展以SPSS Modeler為範例

The proceeding is a collection of research papers presented at the International Conference on Data Engineering 2013 (DaEng-2013), a conference dedicated to address the challenges in the areas of database, information retrieval, data mining and knowledge management, thereby presenting a consolidated view to the interested researchers in the aforesaid fields. The goal of this conference was to bring together researchers and practitioners from academia and industry to focus on advanced on data engineering concepts and establishing new collaborations in these areas. The topics of interest are as follows but are not limited to: • Database theory • Data management • Data mining and warehousing • Data privacy & security • Information retrieval, integration and visualization • Information system • Knowledge discovery in databases • Mobile, grid and cloud computing • Knowledge-based • Knowledge management • Web data, services and intelligence

IBM SPSS Modeler Essentials

Put Predictive Analytics into Action Learn the basics of Predictive Analysis and Data Mining through an easy to understand conceptual framework and immediately practice the concepts learned using the open source RapidMiner tool. Whether you are brand new to Data Mining or working on your tenth project, this book will show you how to analyze data, uncover hidden patterns and relationships to aid important decisions and predictions. Data Mining has become an essential tool for any enterprise that collects, stores and processes data as part of its operations. This book is ideal for business users, data analysts, business analysts, business intelligence and data warehousing professionals and for anyone who wants to learn Data Mining. You'll be able to: 1. Gain the necessary knowledge of different data mining techniques, so that you can select the right technique for a given data problem and create a general purpose analytics process. 2. Get up and running fast with more than two dozen commonly used powerful algorithms for predictive analytics using practical use cases. 3. Implement a simple step-by-step process for predicting an outcome or discovering hidden relationships from the data using RapidMiner, an open source GUI based data mining tool Predictive analytics and Data Mining techniques covered: Exploratory Data Analysis, Visualization, Decision trees, Rule induction, k-Nearest Neighbors, Naïve Bayesian, Artificial Neural Networks, Support Vector machines, Ensemble models, Bagging, Boosting, Random Forests, Linear regression, Logistic regression, Association analysis using Apriori and FP Growth, K-Means clustering, Density based clustering, Self Organizing Maps, Text Mining, Time series forecasting, Anomaly detection and Feature selection. Implementation files can be downloaded from the book companion site at www.LearnPredictiveAnalytics.com Demystifies data mining concepts with easy to understand language Shows how to get up and running fast with 20 commonly used powerful techniques for predictive analysis Explains the process of using open source RapidMiner tools Discusses a simple 5 step process for implementing algorithms that can be used for performing predictive analytics Includes practical use cases and examples

Data Mining with SPSS Modeler

A guide for constructing and using composite indicators for policy makers, academics, the media and other interested parties. In particular, this handbook is concerned with indicators which compare and rank country performance.

A Concise Guide to Market Research

This is an applied handbook for the application of data mining techniques in the CRM framework. It combines a technical and a business perspective to cover the needs of business users who are looking for a practical guide on data mining. It focuses on Customer Segmentation and presents guidelines for the development of actionable segmentation schemes. By using non-technical language it guides readers through all the phases of the data mining process.

Workload Optimized Systems: Tuning POWER7 for Analytics

Transforming data from operational data models to purpose-oriented data structures has been commonplace for the last decades. Data transformations are heavily used in all types of industries to provide information to various users at different levels. Depending on individual needs, the transformed data is stored in various different systems. Sending operational data to other systems for further processing is then required, and introduces much complexity to an existing information technology (IT) infrastructure. Although maintenance of additional hardware and software is one component, potential inconsistencies and individually managed refresh cycles are others. For decades, there was no simple and efficient way to perform data transformations on the source system of operational data. With IBM® DB2® Analytics Accelerator, DB2 for z/OS is now in a unique position to complete these transformations in an efficient and well-performing way. DB2 for z/OS completes these while connecting to the same platform as for operational transactions, helping you to minimize your efforts to manage existing IT infrastructure. Real-time analytics on incoming operational transactions is another demand. Creating a comprehensive scoring model to detect specific patterns inside your data can easily require multiple iterations and multiple hours to complete. By enabling a first set of analytical functionality in DB2 Analytics Accelerator, those dedicated mining algorithms can now be run on an accelerator to efficiently perform these modeling tasks. Given the speed of query processing on an accelerator, these modeling tasks can now be performed much quicker compared to traditional relational database management systems. This speed enables you to keep your scoring algorithms more up-to-date, and ultimately adapt more quickly to constantly changing customer behaviors. This IBM Redbooks® publication describes the new table type that is introduced with DB2 Analytics Accelerator V4.1 PTF5 that enables more efficient data transformations. These tables are called accelerator-only tables, and can exist on an accelerator only. The tables benefit from the accelerator performance characteristics, while maintaining access through existing DB2 for z/OS application programming interfaces (APIs). Additionally, we describe the newly introduced analytical capabilities with DB2 Analytics Accelerator V5.1, putting you in the position to efficiently perform data modeling for online analytical requirements in your DB2 for z/OS environment. This book is intended for technical decision-makers who want to get a broad understanding about the analytical capabilities and accelerator-only tables of DB2 Analytics Accelerator. In addition, you learn about how these capabilities can be used to accelerate in-database transformations and in-database analytics in various environments and scenarios, including the following scenarios: Multi-step processing and reporting in IBM DB2 Query Management FacilityTM, IBM Campaign, or Microstrategy environments In-database transformations using IBM InfoSphere® DataStage® Ad hoc data analysis for data scientists In-database analytics using IBM SPSS® Modeler

SPSS For Dummies

SAP is a market leader in enterprise business application software. SAP solutions provide a rich set of composable application modules, and configurable functional capabilities that are expected from a comprehensive enterprise business application software suite. In most cases, companies that adopt SAP software remain heterogeneous enterprises running both SAP and non-SAP systems to support their business processes. Regardless of the specific scenario, in heterogeneous enterprises most SAP implementations must be integrated with a variety of non-SAP enterprise systems: Portals Messaging infrastructure Business process management (BPM) tools Enterprise Content Management (ECM) methods and tools Business analytics (BA) and business intelligence (BI) technologies Security Systems of record Systems of engagement The tooling included with SAP software addresses many needs for creating SAP-centric environments. However, the classic approach to implementing SAP functionality generally leaves the business with a rigid solution that is difficult and expensive to change and enhance. When SAP software is used in a large, heterogeneous enterprise environment, SAP clients face the dilemma of selecting the correct set of tools and platforms to implement SAP functionality, and to integrate the SAP solutions with non-SAP systems. This IBM® Redbooks® publication explains the value of integrating IBM software with SAP solutions. It describes how to enhance and extend pre-built capabilities in SAP software with best-in-class IBM enterprise software, enabling clients to maximize return on investment (ROI) in their SAP investment and achieve a balanced enterprise architecture approach. This book describes IBM Reference Architecture for SAP, a prescriptive blueprint for using IBM software in SAP solutions. The reference architecture is focused on defining the use of IBM software with SAP, and is not intended to address the internal aspects of SAP components. The chapters of this book provide a specific reference architecture for many of the architectural domains that are each important for a large enterprise to establish common strategy, efficiency, and balance. The majority of the most important architectural domain topics, such as integration, process optimization, master data management, mobile access, Enterprise Content Management, business intelligence, DevOps, security, systems monitoring, and so on, are covered in the book. However, there are several other architectural domains which are not included in the book. This is not to imply that these other architectural domains are not important or are less important, or that IBM does not offer a solution to address them. It is only reflective of time constraints, available resources, and the complexity of assembling a book on an extremely broad topic. Although more content could have been added, the authors feel confident that the scope of architectural material that has been included should provide organizations with a fantastic head start in defining their own enterprise reference architecture for many of the important architectural domains, and it is hoped that this book provides great value to those reading it. This IBM Redbooks publication is targeted to the following audiences: Client decision makers and solution architects leading enterprise transformation projects and wanting to gain further insight so that they can benefit from the integration of IBM software in large-scale SAP projects. IT architects and consultants integrating IBM technology with SAP solutions.

Data Mining With Confidence

資料探勘是一門結合統計學與資訊科學相關理論的方法學，藉由各種功能與模式的導入與實踐，使得資料探勘的應用遍及各個領域，成為研究與實務工作者重要的研究方法，尤其是運用在人工智慧及機器學習的未來發展。

再者，隨著知識經濟的發展，以資料探勘為基礎，創造個人、組織競爭優勢、與經營績效的管理理論及工具，也就成為資料探勘發展及應用的趨勢。故資料探勘理論與工具方法的學習與導入於組織、企業，就成為知識探勘、運用與管理的重要工作。因此，我們也可以說資料探勘，對於學術界與實務界而言，是一門兼具問題、理論、與方法的學科。

這本書所要提供給讀者的內容，即嘗試以不同資料探勘的理論為經，演算方法為緯，在經、緯的架構中，藉著個案實例，以及SPSS Modeler系統實際的操作，來說明資料探勘模式與功能所能提供問題解決的方法，以及在人工智慧及機器學習未來的發展。

【本書範例檔案請至博碩官網下載】

Effective CRM using Predictive Analytics

Take a deep dive into the concepts of machine learning as they apply to contemporary business and management. You will learn how machine learning techniques are used to solve fundamental and complex problems in society and industry. Machine Learning for Decision Makers serves as an excellent resource for establishing the relationship of machine learning with IoT, big data, and cognitive and cloud computing to give you an overview of how these modern areas of computing relate to each other. This book introduces a collection of the most important concepts of machine learning and sets them in context with other vital technologies that decision makers need to know about. These concepts span the process from envisioning the problem to applying machine-learning techniques to your particular situation. This discussion also provides an insight to help deploy the results to improve decision-making. The book uses case studies and jargon busting to help you grasp the theory of machine learning quickly. You'll soon gain the big picture of machine learning and how it fits with other cutting-edge IT services. This knowledge will give you confidence in your decisions for the future of your business. What You Will Learn Discover the machine learning, big data, and cloud and cognitive computing technology stack Gain insights into machine learning concepts and practices Understand business and enterprise decision-making using machine learning Absorb machine-learning best practices Who This Book Is For Managers tasked with making key decisions who want to learn how and when machine learning and related technologies can help them.

IBM Software Defined Infrastructure for Big Data Analytics Workloads

Methodologies of Pattern Recognition is a collection of papers that deals with the two approaches to pattern recognition (geometrical and structural), the Robbins-Monro procedures, and the implications of interactive graphic computers for pattern recognition methodology. Some papers describe non-supervised learning in statistical pattern recognition, parallel computation in pattern recognition, and statistical analysis as a tool to make patterns emerge from data. One paper points out the importance of cluster processing in visual perception in which proximate points of similar brightness values form clusters. At higher levels of mental activity humans are efficient in clumping complex items into clusters. Another paper suggests a recognition method which combines versatility and an efficient noise-proofness in dealing with the two main problems in the field of recognition. These difficulties are the presence of a large variety of observed signals and the presence of interference. One paper reports on a possible feature selection for pattern recognition systems employing the minimization of population entropy. Electronic engineers, physicists, physiologists, psychologists, logicians, mathematicians, and philosophers will find great rewards in reading the above collection.

Handbook of Statistical Analysis and Data Mining Applications

Handbook of Statistical Analysis and Data Mining Applications, Second Edition, is a comprehensive professional reference book that guides business analysts, scientists, engineers and researchers, both academic and industrial, through all stages of data analysis, model building and implementation. The handbook helps users discern technical and business problems, understand the strengths and weaknesses of modern data mining algorithms and employ the right statistical methods for practical application. This book is an ideal reference for users who want to address massive and complex datasets with novel statistical approaches and be able to objectively evaluate analyses and solutions. It has clear, intuitive explanations of the principles and tools for solving problems using modern analytic techniques and discusses their application to real problems in ways accessible and beneficial to practitioners across several areas—from science and engineering, to medicine, academia and commerce. Includes input by practitioners for practitioners Includes tutorials in numerous fields of study that provide step-by-step instruction on how to use supplied tools to build models Contains practical advice from successful real-world implementations Brings together, in a single resource, all the information a beginner needs to understand the tools and issues in data mining to build successful data mining solutions Features clear, intuitive explanations of novel analytical tools and techniques, and their practical applications

Data Mining with Rattle and R

Using IBM SPSS for Social Statistics and Research Methods supports the use of SPSS for social statistics and research methods classes and is an excellent companion to any undergraduate statistics or research methods textbook. The book covers a wide range of data analysis topics to help students working on papers, research projects, and proposals. Using examples, tables, and actual SPSS screen captures, along with current data sets from the General Social Survey, it guides users through several different kinds of SPSS files including data files, output files, and syntax files.

Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013)

When ricocheting a solution that involves analytics, the mainframe might not be the first platform that comes to mind. However, the IBM® System z® group has developed some innovative solutions that include the well-respected mainframe benefits. This book describes a workshop that demonstrates the use of real-time advanced analytics for enhancing core banking decisions using a loan origination example. The workshop is a live hands-on experience of the entire process from analytics modeling to deployment of real-time scoring services for use on IBM z/OS®. In this IBM Redbooks® publication, we include a facilitator guide chapter as well as a participant guide chapter. The facilitator guide includes information about the preparation, such as the needed material, resources, and steps to set up and run this workshop. The participant guide shows step-by-step the tasks for a successful learning experience. The goal of the first hands-on exercise is to learn how to use IBM SPSS® Modeler for Analytics modeling. This provides the basis for the next exercise "Configuring risk assessment in SPSS Decision Management". In the third exercise, the participant experiences how real-time scoring can be implemented on a System z. This publication is written for consultants, IT architects, and IT administrators who want to become familiar with SPSS and analytics solutions on the System z.

Harness the Power of Big Data The IBM Big Data Platform

Combine business sense, statistics, and computers in a new and intuitive way, thanks to Big Data Predictive analytics is a branch of data mining that helps predict probabilities and trends. Predictive Analytics For Dummies explores the power of predictive analytics and how you can use it to make valuable predictions for your business, or in fields such as advertising, fraud detection, politics, and others. This practical book does not bog you down with loads of mathematical or scientific theory, but instead helps you quickly see how to use the right algorithms and tools to collect and analyze data and apply it to make predictions. Topics include using structured and unstructured data, building models, creating a predictive analysis roadmap, setting realistic goals, budgeting, and much more. Shows readers how to use Big Data and data mining to discover patterns and make predictions for tech-savvy businesses Helps readers see how to shepherd predictive analytics projects through their companies Explains just enough of the science and math, but also focuses on practical issues such as protecting project budgets, making good presentations, and more Covers nuts-and-bolts topics including predictive analytics basics, using structured and unstructured data, data mining, and algorithms and techniques for analyzing data Also covers clustering, association, and statistical models; creating a predictive analytics roadmap; and applying predictions to the web, marketing, finance, health care, and elsewhere Propose, produce, and protect predictive analytics projects through your company with Predictive Analytics For Dummies.

Machine Learning for Decision Makers

Systems of record (SORs) are engines that generates value for your business. Systems of engagement (SOE) are always evolving and generating new customer-centric experiences and new opportunities to capitalize on the value in the systems of record. The highest value is gained when systems of record and systems of engagement are brought together to deliver insight. Systems of insight (SOI) monitor and analyze what is going on with various behaviors in the systems of engagement and information being stored or transacted in the systems of record. SOIs seek new opportunities, risks, and operational behavior that needs to be reported or have action taken to optimize business outcomes. Systems of insight are at the core of the Digital Experience, which tries to derive insights from the enormous amount of data generated by automated processes and customer interactions. Systems of Insight can also provide the ability to apply analytics and rules to real-time data as it flows within, throughout, and beyond the enterprise (applications, databases, mobile, social, Internet of Things) to gain the wanted insight. Deriving this insight is a key step toward being able to make the best decisions and take the most appropriate actions. Examples of such actions are to improve the number of satisfied clients, identify clients at risk of leaving and incentivize them to stay loyal, identify patterns of risk or fraudulent behavior and take action to minimize it as early as possible, and detect patterns of behavior in operational systems and transportation that lead to failures, delays, and maintenance and take early action to minimize risks and costs. IBM® Operational Decision Manager is a decision management platform that provides capabilities that support both event-driven insight patterns, and business-rule-driven scenarios. It also can easily be used in combination with other IBM Analytics solutions, as the detailed examples will show. IBM Operational Decision Manager Advanced, along with complementary IBM software offerings that also provide capability for systems of insight, provides a way to deliver the greatest value to your customers and your business. IBM Operational Decision Manager Advanced brings together data from different sources to recognize meaningful trends and patterns. It empowers business users to define, manage, and automate repeatable operational decisions. As a result, organizations can create and shape customer-centric business moments. This IBM Redbooks® publication explains the key concepts of systems of insight and how to implement a system of insight solution with examples. It is intended for IT architects and professionals who are responsible for implementing a systems of insights solution requiring event-based context pattern detection and deterministic decision services to enhance other analytics solution components with IBM Operational Decision Manager Advanced.

Discovering Statistics Using IBM SPSS Statistics

Accelerating Data Transformation with IBM DB2 Analytics Accelerator for z/OS

Learn the art and science of predictive analytics — techniques that get results Predictive analytics is what translates big data into meaningful, usable business information. Written by a leading expert in the field, this guide examines the science of the underlying algorithms as well as the principles and best practices that govern the art of predictive analytics. It clearly explains the theory behind predictive analytics, teaches the methods, principles, and techniques for conducting predictive analytics projects, and offers tips and tricks that are essential for successful predictive modeling. Hands-on examples and case studies are included. The ability to successfully apply predictive analytics enables businesses to effectively interpret big data; essential for competition today This guide teaches not only the principles of predictive analytics, but also how to apply them to achieve real, pragmatic solutions Explains methods, principles, and techniques for conducting predictive analytics projects from start to finish Illustrates each technique with hands-on examples and includes as series of in-depth case studies that apply predictive analytics to common business scenarios A companion website provides all the data sets used to generate the examples as well as a free trial version of software Applied Predictive Analytics arms data and business analysts and business managers with the tools they need to interpret and capitalize on big data.

Systems of Insight for Digital Transformation: Using IBM Operational Decision Manager Advanced and Predictive Analytics

Get to grips with the fundamentals of data mining and predictive analytics with IBM SPSS Modeler Key Features Get up–and-running with IBM SPSS Modeler without going into too much depth. Identify interesting relationships within your data and build effective data mining and predictive analytics solutions A quick, easy–to-follow guide to give you a fundamental understanding of SPSS Modeler, written by the best in the business Book Description IBM SPSS Modeler allows users to quickly and efficiently use predictive analytics and gain insights from your data. With almost 25 years of history, Modeler is the most established and comprehensive Data Mining workbench available. Since it is popular in corporate settings, widely available in university settings, and highly compatible with all the latest technologies, it is the perfect way to start your Data Science and Machine Learning journey. This book takes a detailed, step-by-step approach to introducing data mining using the de facto standard process, CRISP-DM, and Modeler's easy to learn "visual programming" style. You will learn how to read data into Modeler, assess data quality, prepare your data for modeling, find interesting patterns and relationships within your data, and export your predictions. Using a single case study throughout, this intentionally short and focused book sticks to the essentials. The authors have drawn upon their decades of teaching thousands of new users, to choose those aspects of Modeler that you should learn first, so that you get off to a good start using proven best practices. This book provides an overview of various popular data modeling techniques and presents a detailed case study of how to use CHAID, a decision tree model. Assessing a model's performance is as important as building it; this book will also show you how to do that. Finally, you will see how you can score new data and export your predictions. By the end of this book, you will have a firm understanding of the basics of data mining and how to effectively use Modeler to build predictive models. What you will learn Understand the basics of data mining and familiarize yourself with Modeler's visual programming interface Import data into Modeler and learn how to properly declare metadata Obtain summary statistics and audit the quality of your data Prepare data for modeling by selecting and sorting cases, identifying and removing duplicates, combining data files, and modifying and creating fields Assess simple relationships using various statistical and graphing techniques Get an overview of the different types of models available in Modeler Build a decision tree model and assess its results Score new data and export predictions Who this book is for This book is ideal for those who are new to SPSS Modeler and want to start using it as quickly as possible, without going into too much detail. An understanding of basic data mining concepts will be helpful, to get the best out of the book.

Learning IBM Watson Analytics

Lecturers/instructors - request a free digital inspection copy here With a little help from his weird band of characters the Fourth Edition of the award-winning book continues, with its unique blend of humour and collection of bizarre examples, to bring statistics - from first principles to advanced concepts - well and truly to life using IBM SPSS Statistics. Lecturers: with WebAssign® you can manage and monitor your students' progress quickly and easily online or give them more opportunities to practise! Ideal for short courses, choose to use WebAssign® alongside the Fourth Edition of Andy Field's textbook to quickly set up courses and schedule assignments (using the 2159 questions available) and track individual performance so you can spot in an instant where more instruction or practice is needed. If not using for fomal assessment, WebAssign® still lets you set questions for your students to practise over and over again. They get instant feedback and also links to the relevant chapter or section in the integral ebook to help them work out the correct solution. For more information on how to integrate WebAssign® into a forthcoming course or to arrange a class test please contact your local SAGE representative for more details. (Students please note: access to WebAssign® is dependent not only on the purchase of a student access code (ISBN: 9781446273043) but also a username, institution code and password supplied by your course leader/instructor). SAGE MobileStudy - study where and when you like Scan any QR code within the book to access revision material on a smartphone or tablet such as Cramming Sam's Study tips, flashcard glossaries, interactive multiple choice questionnaires and more. Click here to take a look (if you're accessing the site from a desktop you'll be taken to the Companion Website instead; look out for the MobileStudy icon to show you which pages are also available on the MobileStudy site). See how Andy's book is changing the landscape for textbooks through the use of technology! Support materials for a wide range of disciplines Education and Sport Sciences lecturer support materials with enhanced ones for Psychology, Business and Management and the Health Sciences on the enhanced Companion Website make the book even more relevant to a wider range of subjects across the social sciences and where statistics is taught to a cross-disciplinary audience. Other major new updates include: Now fully compatible with recent IBM SPSS Statistics releases. Two new characters! Statistical cult leader Oditi provides students with access to video clips via his Lantern to help further understanding of statistical/SPSS concepts, while Confusius helps students to make better sense of statistical terms. The enhanced Companion Website offers plenty of lecturer and student material to use in conjunction with the textbook. These include PowerPoints and subject-specific testbanks for lecturers as well as answers to the Smart Alex tasks at the end of the each chapter; datafiles for testing problems in SPSS; flashcards of key concepts; self-assessment multiple-choice questions; and online videos of key statistical and SPSS procedures discussed in the textbook for students. Video Links Go behind the scenes of the Fourth Edition, and find out about the man behind the book Watch Andy introduce SAGE MobileStudy Ask Andy Anything: Teaching stats and Robbie Williams' head Ask Andy Anything: Gibson or Fender Ask Andy Anything: The one part of the book Andy hated writing Available with Perusall—an eBook that makes it easier to prepare for class Perusall is an award-winning eBook platform featuring social annotation tools that allow students and instructors to collaboratively mark up and discuss their SAGE textbook. Backed by research and supported by technological innovations developed at Harvard University, this process of learning through collaborative annotation keeps your students engaged and makes teaching easier and more effective. Learn more.

Real-time Fraud Detection Analytics on IBM System z

Each chapter of Performing Data Analysis Using IBM SPSS covers a particular statistical procedure and offers the following: an example problem or analysis goal, together with a data set; IBM SPSS analysis with step-by-step analysis setup and accompanying screen shots; and IBM SPSS output with screen shots and narrative on how to read or interpret the results of the analysis.

Methodologies of Pattern Recognition

Payment fraud can be defined as an intentional deception or misrepresentation that is designed to result in an unauthorized benefit. Fraud schemes are becoming more complex and difficult to identify. It is estimated that industries lose nearly $1 trillion USD annually because of fraud. The ideal solution is where you avoid making fraudulent payments without slowing down legitimate payments. This solution requires that you adopt a comprehensive fraud business architecture that applies predictive analytics. This IBM® Redbooks® publication begins with the business process flows of several industries, such as banking, property/casualty insurance, and tax revenue, where payment fraud is a significant problem. This book then shows how to incorporate technological advancements that help you move from a post-payment to pre-payment fraud detection architecture. Subsequent chapters describe a solution that is specific to the banking industry that can be easily extrapolated to other industries. This book describes the benefits of doing fraud detection on IBM System z®. This book is intended for financial decisionmakers, consultants, and architects, in addition to IT administrators.

Applied Predictive Analytics

This IBM® Redbooks® publication documents how IBM Platform Computing, with its IBM Platform Symphony® MapReduce framework, IBM Spectrum Scale (based Upon IBM GPFSTM), IBM Platform LSF®, the Advanced Service Controller for Platform Symphony are work together as an infrastructure to manage not just Hadoop-related offerings, but many popular industry offeringsm such as Apach Spark, Storm, MongoDB, Cassandra, and so on. It describes the different ways to run Hadoop in a big data environment, and demonstrates how IBM Platform Computing solutions, such as Platform Symphony and Platform LSF with its MapReduce Accelerator, can help performance and agility to run Hadoop on distributed workload managers offered by IBM. This information is for technical professionals (consultants, technical support staff, IT architects, and IT specialists) who are responsible for delivering cost-effective cloud services and big data solutions on IBM Power SystemsTM to help uncover insights among client's data so they can optimize product development and business results.

Data Mining Techniques in CRM

This accessible, practice-oriented and compact text provides a hands-on introduction to market research. Using the market research process as a framework, it explains how to collect and describe data and presents the most important and frequently used quantitative analysis techniques, such as ANOVA, regression analysis, factor analysis and cluster analysis. The book describes the theoretical choices a market researcher has to make with regard to each technique, discusses how these are converted into actions in IBM SPSS version 22 and how to interpret the output. Each chapter concludes with a case study that illustrates the process using real-world data. A comprehensive Web appendix includes additional analysis techniques, datasets, video files and case studies. Tags in the text allow readers to quickly access Web content with their mobile device. The new edition features: Stronger emphasis on the gathering and analysis of secondary data (e.g., internet and social networking data) New material on data description (e.g., outlier detection and missing value analysis) Improved use of educational elements such as learning objectives, keywords, self-assessment tests, case studies, and much more Streamlined and simplified coverage of the data analysis techniques with more rules-of-thumb Uses IBM SPSS version 22

ROMANCE ACTION & ADVENTURE MYSTERY & THRILLER BIOGRAPHIES & HISTORY CHILDREN’S YOUNG ADULT FANTASY HISTORICAL FICTION HORROR LITERARY FICTION NON-FICTION SCIENCE FICTION