Data Mining Orange Documentation

With the recent growth of big data and the internet of things (IoT), individuals can now upload, retrieve, store, and collect massive amounts of information to help drive decisions and optimize processes. Due to this, a new age of predictive computing is taking place, and data can now be harnessed to predict unknown occurrences or probabilities based on data collected in real time. Predictive Intelligence Using Big Data and the Internet of Things highlights state-of-the-art research on predictive intelligence using big data, the IoT, and related areas to ensure quality assurance and compatible IoT systems. Featuring coverage on predictive application scenarios to discuss these breakthroughs in real-world settings and various methods, frameworks, algorithms, and security concerns for predictive intelligence, this book is ideally designed for academicians, researchers, advanced-level students, and technology developers. This book A Beginner's Guide to Learning Analytics is designed to meet modern educational trends' needs. It is addressed to readers who have no prior knowledge of learning analytics and functions as an introductory text to learning analytics for those who want to do more with evaluation/assessment in their organizations. The book is useful to all who need to evaluate their learning and teaching strategies. It aims to bring greater efficiency and deeper engagement to individual students, learning communities, and educators. Covered here are the key concepts linked to learning analytics for researchers and practitioners interested in learning analytics. This book helps those who want to apply analytics to learning and development programs and helps educational institutions to identify learners who require support and provide a more personalized learning experience. Like chapters show diverse uses of learning analytics to enhance student and faculty performance. It presents a coherent framework for the effective translation of learning analytics research for educational practice to its practical application in different educational domains. This book provides educators and researchers with the tools and frameworks to effectively make sense of and use data and analytics in their everyday practice. This book will be a valuable addition to researchers' bookshelves.

Artificial Intelligence and Innovations (AIAI) will interest researchers, IT professionals and consultants by examining technologies and applications of demonstrable value. The conference focused on profitable intelligent systems and technologies. AIAI focuses on real world applications; therefore authors should highlight the benefits of AI technology for industry and services. Novel approaches solving business and industrial problems, using AI, will emerge from this conference.

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

This book constitutes the refereed proceedings of the 19th EPIA Conference on Artificial Intelligence, EPIA 2019, held in Funchal, Madeira, Portugal, in September 2019. The 119 revised full papers and 6 short papers presented were carefully reviewed and selected from a total of 252 submissions. The papers are organized in 18 tracks devoted to the following topics: AIEd - Artificial Intelligence in Education, AI4G - Artificial Intelligence for Games, AIoTA - Artificial Intelligence and IoT in Agriculture, AIL - Artificial Intelligence and Law, AIM - Artificial Intelligence in Medicine, AICPDES - Artificial Intelligence in Cyber-Physical and Distributed Embedded Systems, AIPES - Artificial Intelligence in Power and Energy Systems, AITS - Artificial Intelligence in Transportation Systems, ALEA - Artificial Life and Evolutionary Algorithms, AmIA - Ambient Intelligence and Affective Environments, BAAI - Business Applications of Artificial Intelligence, GAI- General AI, IROBOT - Intelligent Robotics, KDBI - Knowledge Discovery and Business Intelligence, KRR - Knowledge Representation and Reasoning, MASTA - Multi-Agent Systems: Theory and Applications, SSM - Social Simulation and Modelling, TeMA - Text Mining and Applications.

The Sixth SIAM International Conference on Data Mining continues the tradition of presenting approaches, tools, and systems for data mining in fields such as science, engineering, industrial processes, healthcare, and medicine. The datasets in these fields are large, complex, and often noisy. Extracting knowledge requires the use of sophisticated, high-performance, and principled analysis techniques and algorithms, based on sound statistical foundations. These techniques in turn require powerful visualization technologies; implementations that must be carefully tuned for performance; software systems that are usable by scientists, engineers, and physicians as well as researchers; and infrastructures that support them.

This book presents Proceedings of the 2021 Intelligent Systems Conference which is a remarkable collection of chapters covering a wider range of topics in areas of intelligent systems and artificial intelligence and their applications to the real world. The conference attracted a total of 496 submissions from many academic pioneering researchers, scientists, industrial engineers, and students from all around the world. These submissions underwent a double-blind peer-review process. Of the total submissions, 180 submissions have been selected to be included in these proceedings. As we witness exponential growth of computational intelligence in several directions and use of intelligent systems in everyday applications, this book is an ideal resource for reporting latest innovations and future of AI. The chapters include theory and application on all aspects of artificial intelligence, from classical to intelligent scope. We hope that readers find the book interesting and valuable; it provides the state-of-the-art intelligent methods and techniques for solving real-world problems along with a vision of the future research. .

Data Preparation for Data Mining addresses an issue unfortunately ignored by most authorities on data mining: data preparation. Thanks largely to its perceived difficulty, data preparation has traditionally taken a backseat to the more alluring question of how best to extract meaningful knowledge. But without adequate preparation of your data, the return on the resources invested in mining is certain to be disappointing. Dorian Pyle corrects this imbalance. A twenty-five-year veteran of what has become the data mining industry, Pyle shares his own successful data preparation methodology, offering both a conceptual overview for managers and complete technical details for IT professionals. Apply his techniques and watch your mining efforts pay off-in the form of improved performance, reduced distortion, and more valuable results. On the enclosed CD-ROM, you'll find a suite of programs as C source code and compiled into a command-line-driven toolkit. This code illustrates how the author's techniques can be

applied to arrive at an automated preparation solution that works for you. Also included are demonstration versions of three commercial products that help with data preparation, along with sample data with which you can practice and experiment. * Offers in-depth coverage of an essential but largely ignored subject. * Goes far beyond theory, leading you-step by step-through the author's own data preparation techniques. * Provides practical illustrations of the author's methodology using realistic sample data sets. * Includes algorithms you can apply directly to your own project, along with instructions for understanding when automation is possible and when greater intervention is required. * Explains how to identify and correct data problems that may be present in your application. * Prepares miners, helping them head into preparation with a better understanding of data sets and their limitations.

Catalog of reports, decisions and opinions, testimonies and speeches.

Data mining is often referred to by real-time users and software solutions providers as knowledge discovery in databases (KDD). Good data mining practice for business intelligence (the art of turning raw software into meaningful information) is demonstrated by the many new techniques and developments in the conversion of fresh scientific discovery into widely accessible software solutions. This book has been written as an introduction to the main issues associated with the basics of machine learning and the algorithms used in data mining. Suitable for advanced undergraduates and their tutors at postgraduate level in a wide area of computer science and technology topics as well as researchers looking to adapt various algorithms for particular data mining tasks. A valuable addition to the libraries and bookshelves of the many companies who are using the principles of data mining (or KDD) to effectively deliver solid business and industry solutions. Provides an introduction to the main issues associated with the basics of machine learning and the algorithms used in data mining A valuable addition to the libraries and bookshelves of companies using the principles of data mining (or KDD) to effectively deliver solid business and industry solutions

Integrates social media, social network analysis, and data mining to provide an understanding of the potentials of social media mining.

This book constitutes the second part of refereed proceedings of the 5th Computational Methods in Systems and Software 2021 (CoMeSySo 2021) proceedings. The real-world problems related to data science and algorithm design related to systems and software engineering are presented in this papers. Furthermore, the basic research' papers that describe novel approaches in the data science, algorithm design and in systems and software engineering are included. The CoMeSySo 2021 conference is breaking the barriers, being held online. CoMeSySo 2021 intends to provide an international forum for the discussion of the latest high-quality research results

This volume showcases the presentations and discussions delivered at the 2018 POMS International Conference in Rio. Through a collection of selected papers, it is possible to review the impact and application of operations management for social good, with contributions across a wide range of topics, including: humanitarian operations and crisis management, healthcare operations management, sustainable operations, artificial intelligence and data analytics in operations, product innovation and technology in operations management, marketing and operations management, service operations and servitization, logistics and supply chain management, resilience and risk in

operations, defense, and tourism among other emerging Operations Management issues. The Production and Operations Management Society (POMS) is one of the most important and influential societies in the subject of Production Engineering and, as an international professional and academic organization, represents the interests of professionals and academics in production management and operations around the world.

Written for drug developers rather than computer scientists, this monograph adopts a systematic approach to mining scientifi c data sources, covering all key steps in rational drug discovery, from compound screening to lead compound selection and personalized medicine. Clearly divided into four sections, the first part discusses the different data sources available, both commercial and non-commercial, while the next section looks at the role and value of data mining in drug discovery. The third part compares the most common applications and strategies for polypharmacology, where data mining can substantially enhance the research effort. The final section of the book is devoted to systems biology approaches for compound testing. Throughout the book, industrial and academic drug discovery strategies are addressed, with contributors coming from both areas, enabling an informed decision on when and which data mining tools to use for one's own drug discovery project.

This book highlights recent research on bio-inspired computing and its various innovative applications in information and communication technologies. It presents 38 high-quality papers from the 10th International Conference on Innovations in Bio-Inspired Computing and Applications (IBICA 2019) and 9th World Congress on Information and Communication Technologies (WICT 2019), which was held at GIET University, Gunupur, India, on December 16-18, 2019. As a premier conference, IBICA-WICT brings together researchers, engineers and practitioners whose work involves bio-inspired computing, computational intelligence and their applications in information security, real-world contexts, etc. Including contributions by authors from 18 countries, the book offers a valuable reference guide for all researchers, students and practitioners in the fields of Computer Science and Engineering.

This book constitutes the refereed proceedings of the 15th International Conference on Text, Speech and Dialogue, TSD 2012, held in Brno, Czech Republic, in September 2012. The 82 papers presented together with 2 invited talks were carefully reviewed and selected from 173 submissions. The papers are organized in topical sections on corpora and language resources, speech recognition, tagging, classification and parsing of text and speech, speech and spoken language generation, semantic processing of text and speech, integrating applications of text and speech processing, machine translation, automatic dialogue systems, multimodal techniques and modeling.

This book constitutes the refereed proceedings of the Second International Conference on Data Mining and Big Data, DMBD 2017, held in Fukuoka, Japan, in July/August 2017. The 53 papers presented in this volume were carefully reviewed and selected from 96 submissions. They were organized in topical sections named: association analysis; clustering; prediction; classification; schedule and sequence analysis; big data; data analysis; data mining; text mining; deep learning; high performance computing; knowledge base and its framework; and fuzzy control.

With a pedigree going back over ten years, The DefinitiveHandbook of Business Continuity Management can rightly claim tobe a classic guide to business risk

management and contingencyplanning, with a style that makes it accessible to all businessmanagers. Some of the original underlying principles remain thesame – but much has changed. This is reflected in thisradically updated third edition, with exciting and helpful newcontent from new and innovative contributors and new case studiesbringing the book right up to the minute. This book combines over 500 years of experience from leadingBusiness Continuity experts of many countries. It is presented inan easy-to-follow format, explaining in detail the core BCactivities incorporated in BS 25999, Business ContinuityGuidelines, BS 25777 IT Disaster Recovery and other standards andin the body of knowledge common to the key business continuityinstitutes. Contributors from America, Asia Pacific, Europe, China, Indiaand the Middle East provide a truly global perspective, bringingtheir own insights and approaches to the subject, sharing bestpractice from the four corners of the world. We explore and summarize the latest legislation, guidelines andstandards impacting BC planning and management and explain theirimpact. The structured format, with many revealing case studies,examples and checklists, provides a clear roadmap, simplifying andde-mystifying business continuity processes for those new to itsdisciplines and providing a benchmark of current best practice forthose more experienced practitioners. This book makes a massive contribution to the knowledge base ofBC and risk management. It is essential reading for all businesscontinuity, risk managers and auditors: none should be withoutit.

Open Source Data Warehousing and Business Intelligence is an all-in-one reference for developing open source based data warehousing (DW) and business intelligence (BI) solutions that are business-centric, cross-customer viable, cross-functional, cross-technology based, and enterprise-wide. Considering the entire lifecycle of an open source DW &

Molecular Pathology: The Molecular Basis of Human Disease provides a current and comprehensive view of the molecular basis and mechanisms of human disease. Combining accepted principles with broader theoretical concepts and with contributions from a group of experts, the book looks into disease processes in the context of traditional pathology and their implications for translational molecular medicine. It also discusses concepts in molecular biology and genetics, recent scientific and technological advances in modern pathology, the concept of ""molecular pathogenesis"" of disease, and how disease evolves from normal cells and tissues due to perturbations in molecular pathways. The book describes the integration of molecular and cellular pathogenesis using a bioinformatics approach and a systems biology approach to disease pathogenesis. It also discusses current and future strategies in molecular diagnosis of human disease, and the impact of molecular diagnosis on treatment decisions and the practice of personalized medicine. This book is a valuable resource for students, biomedical researchers, practicing physician-scientists who undertake disease-related basic science and translational research, and pathology residents and other postdoctoral fellows. * Exam Master® web site will host "Self-assessment" questions that students can use to study for the molecular section of the board exam * Companion Web Site – will host a complete set of PowerPoint slides: to include images from the book and additional images for teaching; course materials; lecture materials * Teaches from the perspective of "integrative systems biology, which encompasses the intersection of all molecular aspects of biology, as applied to

understanding human disease * Outlines the principles and practice of molecular pathology * Explains the practice of "molecular medicine and the translational aspects of molecular pathology

This book constitutes the refereed proceedings of the Third International Workshop on Mining Complex Data, MCD 2007, held in Warsaw, Poland, in September 2007, co-located with ECML and PKDD 2007. The 20 revised full papers presented were carefully reviewed and selected; they present original results on knowledge discovery from complex data. In contrast to the typical tabular data, complex data can consist of heterogenous data types, can come from different sources, or live in high dimensional spaces. All these specificities call for new data mining strategies.

The field of data mining is receiving significant attention in today's information-rich society, where data is available from different sources and formats, in large volumes, and no longer constitutes a bottleneck for knowledge acquisition. This rich information has paved the way for novel areas of research, particularly in the crime data analysis realm. Data Mining Trends and Applications in Criminal Science and Investigations presents scientific concepts and frameworks of data mining and analytics implementation and uses across various domains, such as public safety, criminal investigations, intrusion detection, crime scene analysis, and suspect modeling. Exploring the diverse ways that data is revolutionizing the field of criminal science, this publication meets the research needs of law enforcement professionals, data analysts, investigators, researchers, and graduate-level students.

"This book provides an overall view of recent solutions for mining, and explores new patterns,offering theoretical frameworks and presenting challenges and possible solutions concerning pattern extractions, emphasizing research techniques and real-world applications. It portrays research applications in data models, methodologies for mining patterns, multi-relational and multidimensional pattern mining, fuzzy data mining, data streaming and incremental mining"--Provided by publisher.

This three-volume set, LNAI 10937, 10938, and 10939, constitutes the thoroughly refereed proceedings of the 22nd Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining, PAKDD 2018, held in Melbourne, VIC, Australia, in June 2018. The 164 full papers were carefully reviewed and selected from 592 submissions. The volumes present papers focusing on new ideas, original research results and practical development experiences from all KDD related areas, including data mining, data warehousing, machine learning, artificial intelligence, databases, statistics, knowledge engineering, visualization, decision-making systems and the emerging applications.

Intelligent Systems and ApplicationsProceedings of the 2021 Intelligent Systems Conference (IntelliSys).. Volume 3Springer Nature

Language—that is, oral or written content that references abstract concepts in subtle ways—is what sets us apart as a species, and in an age defined by such content, language has become both the fuel and the currency of our modern

information society. This has posed a vexing new challenge for linguists and engineers working in the field of language-processing: how do we parse and process not just language itself, but language in vast, overwhelming quantities? Modern Computational Models of Semantic Discovery in Natural Language compiles and reviews the most prominent linguistic theories into a single source that serves as an essential reference for future solutions to one of the most important challenges of our age. This comprehensive publication benefits an audience of students and professionals, researchers, and practitioners of linguistics and language discovery. This book includes a comprehensive range of topics and chapters covering digital media, social interaction in online environments, text and data mining, language processing and translation, and contextual documentation, among others.

This book examines the impact of digital innovation on organizations. It reveals how the digital revolution is redefining traditional levels of analysis while at the same time blurring the internal and external boundaries of the organizational environment. It presents a collection of research papers that examine the interaction between Information and Communication Technology (ICT) and behavior from a threefold perspective: First, they analyze individual behavior in terms of specific organizational practices like learning, collaboration and knowledge transfer, as well as the use of ICT within the organization. Second, they explore the dynamics at work on the border between the internal and the external environments by analyzing the organizational impact of ICT usage outside the company, as can be seen in employer branding, consumer behavior and organizational image. Third, they investigate how ICT is being adopted to help face societal challenges outside the company like waste and pollution, smart cities, and e-government. The diversity of views offered makes this book particularly relevant for users, companies, scientists and governments. The content is based on a selection of the best papers – original double-blind peer reviewed contributions –presented at the annual conference of the Italian chapter of AIS, which took place in Rome, Italy, in October 2015, or in other reputable international peer-refereed Information Systems conferences.

The progress of data mining technology and large public popularity establish a need for a comprehensive text on the subject. The series of books entitled by "Data Mining" address the need by presenting in-depth description of novel mining algorithms and many useful applications. In addition to understanding each section deeply, the two books present useful hints and strategies to solving problems in the following chapters. The contributing authors have highlighted many future research directions that will foster multi-disciplinary collaborations and hence will lead to significant development in the field of data mining.

This streamlined "essential" version of the Molecular Pathology (2009) textbook extracts key information, illustrations and photographs from the main textbook in the same number and organization of chapters. It is aimed at teaching students in courses where the full textbook is not needed, but the concepts included are desirable (such as

graduate students in allied health programs or undergraduates). It is also aimed at students who are enrolled in courses that primarily use a traditional pathology textbook, but need the complementary concepts of molecular pathology (such as medical students). Further, the textbook will be valuable for pathology residents and other postdoctoral fellows who desire to advance their understanding of molecular mechanisms of disease beyond what they learned in medical/graduate school. Offers an essential introduction to molecular genetics and the "molecular" aspects of human disease Teaches from the perspective of "integrative systems biology," which encompasses the intersection of all molecular aspects of biology, as applied to understanding human disease In-depth presentation of the principles and practice of molecular pathology: molecular pathogenesis, molecular mechanisms of disease, and how the molecular pathogenesis of disease parallels the evolution of the disease using histopathology. "Traditional" pathology section provides state-of-the-art information on the major forms of disease, their pathologies, and the molecular mechanisms that drive these diseases. Explains the practice of "molecular medicine" and the translational aspects of molecular pathology: molecular diagnostics, molecular assessment, and personalized medicine Each chapter ends with Key Summary Points and Suggested Readings

This book comprises high-quality refereed research papers presented at the 2021 International Conference on Artificial Intelligence and Logistics Engineering (ICAILE2021), held in Kyiv, Ukraine, on 22-24 January 2021, organized jointly by Wuhan University of Technology, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" and the International Research Association of Modern Education and Computer Science. The topics discussed in the book include state-of-the-art papers in artificial intelligence and logistics engineering. It is an excellent source of references for researchers, graduate students, engineers, management practitioners and undergraduate students interested in artificial intelligence and their applications in logistics engineering.

Data Mining and Data Visualization focuses on dealing with large-scale data, a field commonly referred to as data mining. The book is divided into three sections. The first deals with an introduction to statistical aspects of data mining and machine learning and includes applications to text analysis, computer intrusion detection, and hiding of information in digital files. The second section focuses on a variety of statistical methodologies that have proven to be effective in data mining applications. These include clustering, classification, multivariate density estimation, tree-based methods, pattern recognition, outlier detection, genetic algorithms, and dimensionality reduction. The third section focuses on data visualization and covers issues of visualization of high-dimensional data, novel graphical techniques with a focus on human factors, interactive graphics, and data visualization using virtual reality. This book represents a thorough cross section of internationally renowned thinkers who are inventing methods for dealing with a new data paradigm. Distinguished contributors who are international experts in aspects of data mining Includes data mining approaches to non-numerical data mining including text data, Internet traffic data, and geographic data Highly topical discussions reflecting current thinking on contemporary technical issues, e.g. streaming data Discusses taxonomy of dataset sizes, computational complexity, and scalability usually ignored in most discussions Thorough discussion of data visualization issues

blending statistical, human factors, and computational insights