Reliable Knowledge Discovery

#reliable knowledge discovery #trusted data insights #accurate information extraction #verifiable data intelligence #dependable knowledge management

Unlock the power of reliable knowledge discovery to transform raw data into actionable, trusted data insights. This process emphasizes robust methodologies for accurate information extraction, ensuring all derived intelligence is verifiable data intelligence and leads to dependable knowledge management and strategic decision-making.

Reliable Knowledge Discovery

Reliable Knowledge Discovery focuses on theory, methods, and techniques for RKDD, a new sub-field of KDD. It studies the theory and methods to assure the reliability and trustworthiness of discovered knowledge and to maintain the stability and consistency of knowledge discovery processes. RKDD has a broad spectrum of applications, especially in critical domains like medicine, finance, and military. Reliable Knowledge Discovery also presents methods and techniques for designing robust knowledge-discovery processes. Approaches to assessing the reliability of the discovered knowledge are introduced. Particular attention is paid to methods for reliable feature selection, reliable graph discovery, reliable classification, and stream mining. Estimating the data trustworthiness is covered in this volume as well. Case studies are provided in many chapters. Reliable Knowledge Discovery is designed for researchers and advanced-level students focused on computer science and electrical engineering as a secondary text or reference. Professionals working in this related field and KDD application developers will also find this book useful.

Advanced Methods for Knowledge Discovery from Complex Data

The growth in the amount of data collected and generated has exploded in recent times with the widespread automation of various day-to-day activities, advances in high-level scienti?c and engineering research and the development of e?cient data collection tools. This has given rise to the need for automa- callyanalyzingthedatainordertoextractknowledgefromit,therebymaking the data potentially more useful. Knowledge discovery and data mining (KDD) is the process of identifying valid, novel, potentially useful and ultimately understandable patterns from massive data repositories. It is a multi-disciplinary topic, drawing from s- eral ?elds including expert systems, machine learning, intelligent databases, knowledge acquisition, case-based reasoning, pattern recognition and stat- tics. Many data mining systems have typically evolved around well-organized database systems (e.g., relational databases) containing relevant information. But, more and more, one ?nds relevant information hidden in unstructured text and in other complex forms. Mining in the domains of the world-wide web,

bioinformatics, geoscienti?c data, and spatial and temporal applications comprise some illustrative examples in this regard. Discovery of knowledge, or potentially useful patterns, from such complex data often requires the - plication of advanced techniques that are better able to exploit the nature and representation of the data. Such advanced methods include, among o- ers, graph-based and tree-based approaches to relational learning, sequence mining, link-based classi?cation, Bayesian networks, hidden Markov models, neural networks, kernel-based methods, evolutionary algorithms, rough sets and fuzzy logic, and hybrid systems. Many of these methods are developed in the following chapters.

Reliable Knowledge

Reliable Knowledge offers a valuably clear account and a radically challenging investigation of the credibility of scientific knowledge.

Scalable and Accurate Knowledge Discovery in Real World Databases

Knowledge Discovery and Measures of Interest is a reference book for knowledge discovery researchers, practitioners, and students. The knowledge discovery researcher will find that the material provides a theoretical foundation for measures of interest in data mining applications where diversity measures are used to rank summaries generated from databases. The knowledge discovery practitioner will find solid empirical evidence on which to base decisions regarding the choice of measures in data mining applications. The knowledge discovery student in a senior undergraduate or graduate course in databases and data mining will find the book is a good introduction to the concepts and techniques of measures of interest. In Knowledge Discovery and Measures of Interest, we study two closely related steps in any knowledge discovery system: the generation of discovered knowledge; and the interpretation and evaluation of discovered knowledge. In the generation step, we study data summarization, where a single dataset can be generalized in many different ways and to many different levels of granularity according to domain generalization graphs. In the interpretation and evaluation step, we study diversity measures as heuristics for ranking the interestingness of the summaries generated. The objective of this work is to introduce and evaluate a technique for ranking the interestingness of discovered patterns in data. It consists of four primary goals: To introduce domain generalization graphs for describing and guiding the generation of summaries from databases. To introduce and evaluate serial and parallel algorithms that traverse the domain generalization space described by the domain generalization graphs. To introduce and evaluate diversity measures as heuristic measures of interestingness for ranking summaries generated from databases. To develop the preliminary foundation for a theory of interestingness within the context of ranking summaries generated from databases. Knowledge Discovery and Measures of Interest is suitable as a secondary text in a graduate level course and as a reference for researchers and practitioners in industry.

Knowledge Discovery and Measures of Interest

This book presents a specific and unified approach to Knowledge Discovery and Data Mining, termed IFN for Information Fuzzy Network methodology. Data Mining (DM) is the science of modelling and generalizing common patterns from large sets of multi-type data. DM is a part of KDD, which is the overall process for Knowledge Discovery in Databases. The accessibility and abundance of information today makes this a topic of particular importance and need. The book has three main parts complemented by appendices as well as software and project data that are accessible from the book's web site (http://www.eng.tau.ac.iV-maimonlifn-kdg£). Part I (Chapters 1-4) starts with the topic of KDD and DM in general and makes reference to other works in the field, especially those related to the information theoretic approach. The remainder of the book presents our work, starting with the IFN theory and algorithms. Part II (Chapters 5-6) discusses the methodology of application and includes case studies. Then in Part III (Chapters 7-9) a comparative study is presented, concluding with some advanced methods and open problems. The IFN, being a generic methodology, applies to a variety of fields, such as manufacturing, finance, health care, medicine, insurance, and human resources. The appendices expand on the relevant theoretical background and present descriptions of sample projects (including detailed results).

Knowledge Discovery and Data Mining

Digital technologies and networks are now part of everyday work in the sciences, and have enhanced access to and use of scientific data, information, and literature significantly. They offer the promise of

accelerating the discovery and communication of knowledge, both within the scientific community and in the broader society, as scientific data and information are made openly available online. The focus of this project was on computer-mediated or computational scientific knowledge discovery, taken broadly as any research processes enabled by digital computing technologies. Such technologies may include data mining, information retrieval and extraction, artificial intelligence, distributed grid computing, and others. These technological capabilities support computer-mediated knowledge discovery, which some believe is a new paradigm in the conduct of research. The emphasis was primarily on digitally networked data, rather than on the scientific, technical, and medical literature. The meeting also focused mostly on the advantages of knowledge discovery in open networked environments, although some of the disadvantages were raised as well. The workshop brought together a set of stakeholders in this area for intensive and structured discussions. The purpose was not to make a final declaration about the directions that should be taken, but to further the examination of trends in computational knowledge discovery in the open networked environments, based on the following questions and tasks: 1. Opportunities and Benefits: What are the opportunities over the next 5 to 10 years associated with the use of computer-mediated scientific knowledge discovery across disciplines in the open online environment? What are the potential benefits to science and society of such techniques? 2. Techniques and Methods for Development and Study of Computer-mediated Scientific Knowledge Discovery: What are the techniques and methods used in government, academia, and industry to study and understand these processes, the validity and reliability of their results, and their impact inside and outside science? 3. Barriers: What are the major scientific, technological, institutional, sociological, and policy barriers to computer-mediated scientific knowledge discovery in the open online environment within the scientific community? What needs to be known and studied about each of these barriers to help achieve the opportunities for interdisciplinary science and complex problem solving? 4. Range of Options: Based on the results obtained in response to items 1-3, define a range of options that can be used by the sponsors of the project, as well as other similar organizations, to obtain and promote a better understanding of the computer-mediated scientific knowledge discovery processes and mechanisms for openly available data and information online across the scientific domains. The objective of defining these options is to improve the activities of the sponsors (and other similar organizations) and the activities of researchers that they fund externally in this emerging research area. The Future of Scientific Knowledge Discovery in Open Networked Environments: Summary of a Workshop summarizes the responses to these questions and tasks at hand.

The Future of Scientific Knowledge Discovery in Open Networked Environments

The Web has emerged as a large, distributed data repository, and information on the Internet and in existing transaction databases can be analyzed for commercial gains in decision making. Therefore, how to efficiently identify quality knowledge from different data sources uncovers a significant challenge. This challenge has attracted wide interest from both academia and the industry. Knowledge Discovery in Multiple Databases provides a comprehensive introduction to the latest advancements in multi-database mining, and presents a local-pattern analysis framework for pattern discovery from multiple data sources. Based on this framework, data preparation techniques in multiple databases, an application-independent database classification for data reduction, and efficient algorithms for pattern discovery from multiple databases are described. Knowledge Discovery in Multiple Databases is suitable for researchers, professionals and students in data mining, distributed data analysis, and machine learning, who are interested in multi-database mining. It is also appropriate for use as a text supplement for broader courses that might involve knowledge discovery in databases and data mining.

Knowledge Discovery in Multiple Databases

This book addresses several knowledge discovery problems on multi-sourced data where the theories, techniques, and methods in data cleaning, data mining, and natural language processing are synthetically used. This book mainly focuses on three data models: the multi-sourced isomorphic data, the multi-sourced heterogeneous data, and the text data. On the basis of three data models, this book studies the knowledge discovery problems including truth discovery and fact discovery on multi-sourced data from four important properties: relevance, inconsistency, sparseness, and heterogeneity, which is useful for specialists as well as graduate students. Data, even describing the same object or event, can come from a variety of sources such as crowd workers and social media users. However, noisy pieces of data or information are unavoidable. Facing the daunting scale of data, it is unrealistic to expect humans to “label” or tell which data source is more reliable. Hence, it is crucial to identify trustworthy information from multiple noisy information sources, referring to the task of knowledge discovery.

At present, the knowledge discovery research for multi-sourced data mainly faces two challenges. On the structural level, it is essential to consider the different characteristics of data composition and application scenarios and define the knowledge discovery problem on different occasions. On the algorithm level, the knowledge discovery task needs to consider different levels of information conflicts and design efficient algorithms to mine more valuable information using multiple clues. Existing knowledge discovery methods have defects on both the structural level and the algorithm level, making the knowledge discovery problem far from totally solved.

Knowledge Discovery from Multi-Sourced Data

This book constitutes the refereed proceedings of the IFIP WG 8.4, 8.9, TC 5 International Cross-Domain Conference on Availability, Reliability and Security, CD-ARES 2013, held in Regensburg, Germany, in September 2013. The 21 revised papers presented were carefully reviewed and selected for inclusion in the volume. The papers concentrate on the many aspects of information systems bridging the gap between research results in computer science and the many application fields. They are organized in the following topical sections: economic, ethical, legal, multilingual, organizational and social aspects; context-oriented information integration; data/information management as a service; context-oriented information integration and location-aware computing; security and privacy; risk management and business continuity; and security and privacy and location based applications. Also included are 15 papers from a special session on Human-Computer Interaction and Knowledge Discovery (HCI-KDD 2013).

Availability, Reliability, and Security in Information Systems and HCI

This book constitutes the refereed proceedings of the 11th European Conference on Principles and Practice of Knowledge Discovery in Databases, PKDD 2007, held in Warsaw, Poland, co-located with ECML 2007, the 18th European Conference on Machine Learning. The 28 revised full papers and 35 revised short papers present original results on leading-edge subjects of knowledge discovery from conventional and complex data and address all current issues in the area.

Knowledge Discovery in Databases: PKDD 2007

This book constitutes the refereed proceedings of the joint conference on Machine Learning and Knowledge Discovery in Databases: ECML PKDD 2008, held in Antwerp, Belgium, in September 2008. The 100 papers presented in two volumes, together with 5 invited talks, were carefully reviewed and selected from 521 submissions. In addition to the regular papers the volume contains 14 abstracts of papers appearing in full version in the Machine Learning Journal and the Knowledge Discovery and Databases Journal of Springer. The conference intends to provide an international forum for the discussion of the latest high quality research results in all areas related to machine learning and knowledge discovery in databases. The topics addressed are application of machine learning and data mining methods to real-world problems, particularly exploratory research that describes novel learning and mining tasks and applications requiring non-standard techniques.

Machine Learning and Knowledge Discovery in Databases

This book is showcases recent advances in knowledge discovery enhanced with semantic and social information. It includes eight chapters that grew out of joint workshops at ECML/PKDD 2007. The contributions emphasize the vision of the Web as a social medium.

Knowledge Discovery Enhanced with Semantic and Social Information

One of the grand challenges in our digital world are the large, complex and often weakly structured data sets, and massive amounts of unstructured information. This “big data” challenge is most evident in biomedical informatics: the trend towards precision medicine has resulted in an explosion in the amount of generated biomedical data sets. Despite the fact that human experts are very good at pattern recognition in dimensions of = 3; most of the data is high-dimensional, which makes manual analysis often impossible and neither the medical doctor nor the biomedical researcher can memorize all these facts. A synergistic combination of methodologies and approaches of two fields offer ideal conditions towards unraveling these problems: Human–Computer Interaction (HCI) and Knowledge Discovery/Data Mining (KDD), with the goal of supporting human capabilities with machine learning./ppThis state-of-the-art survey is an output of the HCI-KDD expert network and features 19

carefully selected and reviewed papers related to seven hot and promising research areas: Area 1: Data Integration, Data Pre-processing and Data Mapping; Area 2: Data Mining Algorithms; Area 3: Graph-based Data Mining; Area 4: Entropy-Based Data Mining; Area 5: Topological Data Mining; Area 6 Data Visualization and Area 7: Privacy, Data Protection, Safety and Security.

Interactive Knowledge Discovery and Data Mining in Biomedical Informatics

In this work, a novel knowledge discovery framework able to analyze data produced in the Gasoline Direct Injection (GDI) context through machine learning is presented and validated. This approach is able to explore and exploit the investigated design spaces based on a limited number of observations, discovering and visualizing connections and correlations in complex phenomena. The extracted knowledge is then validated with domain expertise, revealing potential and limitations of this method.

Development of a modular Knowledge-Discovery Framework based on Machine Learning for the interdisciplinary analysis of complex phenomena in the context of GDI combustion processes

Knowledge Discovery Practices and Emerging Applications of Data Mining: Trends and New Domains introduces the reader to recent research activities in the field of data mining. This book covers association mining, classification, mobile marketing, opinion mining, microarray data mining, internet mining and applications of data mining on biological data, telecommunication and distributed databases, among others, while promoting understanding and implementation of data mining techniques in emerging domains.

Knowledge Discovery Practices and Emerging Applications of Data Mining: Trends and New Domains

Chapters “On the Current State of Reproducibility and Reporting of Uncertainty for Aspect-Based SentimentAnalysis” and “Contextualized Graph Embeddings for Adverse Drug Event Detection” are licensed under theterms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/). For further details see license information in the chapter.

Machine Learning and Knowledge Discovery in Databases

The two-volume set LNAI 8443 + LNAI 8444 constitutes the refereed proceedings of the 18th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2014, held in Tainan, Taiwan, in May 2014. The 40 full papers and the 60 short papers presented within these proceedings were carefully reviewed and selected from 371 submissions. They cover the general fields of pattern mining; social network and social media; classification; graph and network mining; applications; privacy preserving; recommendation; feature selection and reduction; machine learning; temporal and spatial data; novel algorithms; clustering; biomedical data mining; stream mining; outlier and anomaly detection; multi-sources mining; and unstructured data and text mining.

Advances in Knowledge Discovery and Data Mining

As the age of Big Data emerges, it becomes necessary to take the five dimensions of Big Data- volume, variety, velocity, volatility, and veracity- and focus these dimensions towards one critical emphasis - value. The Encyclopedia of Business Analytics and Optimization confronts the challenges of information retrieval in the age of Big Data by exploring recent advances in the areas of knowledge management, data visualization, interdisciplinary communication, and others. Through its critical approach and practical application, this book will be a must-have reference for any professional, leader, analyst, or manager interested in making the most of the knowledge resources at their disposal.

Encyclopedia of Business Analytics and Optimization

The Fuzzy Systems, Knowledge Discovery, and Natural Computation Symposium (FSKDNC 2013) was successfully held from 24 to 25 July 2013, in Shenyang, China. The Symposium was a platform for authors to present their recent development on fuzzy systems, knowledge discovery, and natural computation (i.e., intelligent techniques inspired from nature, such as neural networks, genetic algorithms, and particle swarm optimization). The Symposium attracted numerous submissions from around the globe. Each submitted paper was rigorously reviewed by the program committee and additional reviewers based on originality, significance and quality of the research, clarity of the presentation, and relevance to the Symposium theme. 60 papers are included in the Symposium proceedings after the review process. The great efforts of the authors, the Organizing Committee members, the Program

Committee members, and the additional reviewers are acknowledged here. The Symposium would not have been possible without the support from Liaoning Technical University. The professional and courteous staff from DEStech Publications, Inc also deserves special credits.

Fuzzy Systems, Knowledge Discovery and Natural Computation Symposium

"This set of books represents a detailed compendium of authoritative, research-based entries that define the contemporary state of knowledge on technology"--Provided by publisher.

Encyclopedia of Information Science and Technology, Second Edition

The growth in the amount of data collected and generated has exploded in recent times with the widespread automation of various day-to-day activities, advances in high-level scienti?c and engineering research and the development of e?cient data collection tools. This has given rise to the need for automa- callyanalyzingthedatainordertoextractknowledgefromit,therebymaking the data potentially more useful. Knowledge discovery and data mining (KDD) is the process of identifying valid, novel, potentially useful and ultimately understandable patterns from massive data repositories. It is a multi-disciplinary topic, drawing from s- eral ?elds including expert systems, machine learning, intelligent databases, knowledge acquisition, case-based reasoning, pattern recognition and stat- tics. Many data mining systems have typically evolved around well-organized database systems (e.g., relational databases) containing relevant information. But, more and more, one ?nds relevant information hidden in unstructured text and in other complex forms. Mining in the domains of the world-wide web, bioinformatics, geoscienti?c data, and spatial and temporal applications comprise some illustrative examples in this regard. Discovery of knowledge, or potentially useful patterns, from such complex data often requires the - plication of advanced techniques that are better able to exploit the nature and representation of the data. Such advanced methods include, among o- ers, graph-based and tree-based approaches to relational learning, sequence mining, link-based classi?cation, Bayesian networks, hidden Markov models, neural networks, kernel-based methods, evolutionary algorithms, rough sets and fuzzy logic, and hybrid systems. Many of these methods are developed in the following chapters.

Advanced Methods for Knowledge Discovery from Complex Data

ICNC-FSKD is a premier international forum for scientists and researchers to present the state of the art of data mining and intelligent methods inspired from nature, particularly biological, linguistic, and physical systems, with applications to computers, circuits, systems, control, communications, and more. This is an exciting and emerging interdisciplinary area in which a wide range of theory and methodologies are being investigated and developed to tackle complex and challenging problems.

International Symposium on Fuzzy Systems, Knowledge Discovery and Natural Computation (FSKD 2014)

The objective of this book is two-fold. Firstly, it is aimed at bringing to gether key research articles concerned with methodologies for knowledge discovery in databases and their applications. Secondly, it also contains articles discussing fundamentals of rough sets and their relationship to fuzzy sets, machine learning, management of uncertainty and systems of logic for formal reasoning about knowledge. Applications of rough sets in different areas such as medicine, logic design, image processing and expert systems are also represented. The articles included in the book are based on selected papers presented at the International Workshop on Rough Sets and Knowledge Discovery held in Banff, Canada in 1993. The primary methodological approach emphasized in the book is the mathematical theory of rough sets, a relatively new branch of mathematics concerned with the modeling and analysis of classification problems with imprecise, uncertain, or incomplete information. The methods of the theory of rough sets have applications in many sub-areas of artificial intelligence including knowledge discovery, machine learning, formal reasoning in the presence of uncertainty, knowledge acquisition, and others. This spectrum of applications is reflected in this book where articles, although centered around knowledge discovery problems, touch a number of related issues. The book is intended to provide an important reference material for students, researchers, and developers working in the areas of knowledge discovery, machine learning, reasoning with uncertainty, adaptive expert systems, and pattern classification.

Rough Sets, Fuzzy Sets and Knowledge Discovery

Knowledge discovery in ubiquitous environments is an emerging area of research at the intersection of the two major challenges of highly distributed and mobile systems and advanced knowledge discovery systems. It aims to provide a unifying framework for systematically investigating the mutual dependencies of otherwise quite unrelated technologies employed in building next-generation intelligent systems: machine learning, data mining, sensor networks, grids, peer-to-peer networks, data stream mining, activity recognition, Web 2.0, privacy, user modelling and others. This state-of-the-art survey is the outcome of a large number of workshops, summer schools, tutorials and dissemination events organized by KDubiq (Knowledge Discovery in Ubiquitous Environments), a networking project funded by the European Commission to bring together researchers and practitioners of this emerging community. It provides in its first part a conceptual foundation for the new field of ubiquitous knowledge discovery - highlighting challenges and problems, and proposing future directions in the area of 'smart', 'adaptive', and 'intelligent' learning. The second part of this volume contains selected approaches to ubiquitous knowledge discovery and treats specific aspects in detail. The contributions have been carefully selected to provide illustrations and in-depth discussions for some of the major findings of Part I.

Ubiquitous Knowledge Discovery

This volume presents state-of-the-art tools and techniques for automatically detecting, diagnosing, and predicting the effects of adverse events in an engineered system. It emphasizes the importance of these techniques in managing the intricate interactions within and between engineering systems to maintain a high degree of reliability. Reflecting the interdisciplinary nature of the field, the book explains how the fundamental algorithms and methods of both physics-based and data-driven approaches effectively address systems health management in application areas such as data centers, aircraft, and software systems.

Machine Learning and Knowledge Discovery for Engineering Systems Health Management

This book constitutes the refereed proceedings of the First European Symposium on Principles of Data Mining and Knowledge Discovery, PKDD '97, held in Trondheim, Norway, in June 1997. The volume presents a total of 38 revised full papers together with abstracts of one invited talk and four tutorials. Among the topics covered are data and knowledge representation, statistical and probabilistic methods, logic-based approaches, man-machine interaction aspects, AI contributions, high performance computing support, machine learning, automated scientific discovery, quality assessment, and applications.

Principles of Data Mining and Knowledge Discovery

Considers knowledge discovery, which has been defined as the extraction of implicit, previously unknown and potentially useful information from data. Early chapters examine technical issues of importance to the future development of the field, including overcoming feature interaction problems, analysis of outliers, rule discovery, and temporal processing. Later chapters describe applications in fields such as medical and health information, meteorology, organic chemistry, and the electric supply industry. The editor is a professor of information technology at the University of Portsmouth, UK. Material originated at a May 1998 colloquium. Annotation copyrighted by Book News, Inc., Portland, OR

Knowledge Discovery and Data Mining

This book outlines the core theory and practice of data mining and knowledge discovery (DM & KD) examining theoretical foundations for various methods, and presenting an array of examples, many drawn from real-life applications. Most theoretical developments are accompanied by extensive empirical analysis, offering a deep insight into both theoretical and practical aspects of the subject. The book presents the combined research experiences of 40 expert contributors of world renown.

Data Mining and Knowledge Discovery Approaches Based on Rule Induction Techniques

Recently, nature has stimulated many successful techniques, algorithms, and computational applications allowing conventionally difficult problems to be solved through novel computing systems. Nature-Inspired Informatics for Intelligent Applications and Knowledge Discovery: Implications in Business, Science, and Engineering provides the latest findings in nature-inspired algorithms and their applications for breakthroughs in a wide range of disciplinary fields. This defining reference collection

contains chapters written by leading researchers and well-known academicians within the field, offering readers a valuable and enriched accumulation of knowledge.

Nature-Inspired Informatics for Intelligent Applications and Knowledge Discovery: Implications in Business, Science, and Engineering

This comprehensive textbook on data mining details the unique steps of the knowledge discovery process that prescribes the sequence in which data mining projects should be performed, from problem and data understanding through data preprocessing to deployment of the results. This knowledge discovery approach is what distinguishes Data Mining from other texts in this area. The book provides a suite of exercises and includes links to instructional presentations. Furthermore, it contains appendices of relevant mathematical material.

Data Mining

The two-volume set LNAI 12084 and 12085 constitutes the thoroughly refereed proceedings of the 24th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2020, which was due to be held in Singapore, in May 2020. The conference was held virtually due to the COVID-19 pandemic. The 135 full papers presented were carefully reviewed and selected from 628 submissions. The papers present 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. They are organized in the following topical sections: recommender systems; classification; clustering; mining social networks; representation learning and embedding; mining behavioral data; deep learning; feature extraction and selection; human, domain, organizational and social factors in data mining; mining sequential data; mining imbalanced data; association; privacy and security; supervised learning; novel algorithms; mining multi-media/multi-dimensional data; application; mining graph and network data; anomaly detection and analytics; mining spatial, temporal, unstructured and semi-structured data; sentiment analysis; statistical/graphical model; multi-source/distributed/parallel/cloud computing.

Advances in Knowledge Discovery and Data Mining

This book constitutes the refereed proceedings of the Third Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD '99, held in Beijing, China, in April 1999. The 29 revised full papers presented together with 37 short papers were carefully selected from a total of 158 submissions. The book is divided into sections on emerging KDD technology; association rules; feature selection and generation; mining in semi-unstructured data; interestingness, surprisingness, and exceptions; rough sets, fuzzy logic, and neural networks; induction, classification, and clustering; visualization; causal models and graph-based methods; agent-based and distributed data mining; and advanced topics and new methodologies.

Methodologies for Knowledge Discovery and Data Mining

The 5-volume proceedings, LNAI 12457 until 12461 constitutes the refereed proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases, ECML PKDD 2020, which was held during September 14-18, 2020. The conference was planned to take place in Ghent, Belgium, but had to change to an online format due to the COVID-19 pandemic. The 232 full papers and 10 demo papers presented in this volume were carefully reviewed and selected for inclusion in the proceedings. The volumes are organized in topical sections as follows: Part I: Pattern Mining; clustering; privacy and fairness; (social) network analysis and computational social science; dimensionality reduction and autoencoders; domain adaptation; sketching, sampling, and binary projections; graphical models and causality; (spatio-) temporal data and recurrent neural networks; collaborative filtering and matrix completion. Part II: deep learning optimization and theory; active learning; adversarial learning; federated learning; Kernel methods and online learning; partial label learning; reinforcement learning; transfer and multi-task learning; Bayesian optimization and few-shot learning. Part III: Combinatorial optimization; large-scale optimization and differential privacy; boosting and ensemble methods; Bayesian methods; architecture of neural networks; graph neural networks; Gaussian processes; computer vision and image processing; natural language processing; bioinformatics. Part IV: applied data science: recommendation; applied data science: anomaly detection; applied data science: Web mining; applied data science: transportation; applied data science: activity recognition; applied data science: hardware

and manufacturing; applied data science: spatiotemporal data. Part V: applied data science: social good; applied data science: healthcare; applied data science: e-commerce and finance; applied data science: computational social science; applied data science: sports; demo track.

Machine Learning and Knowledge Discovery in Databases

This book constitutes the thoroughly refereed proceedings of the 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K 2018, held in Seville, Spain, in September 2018. The 12 full papers presented were carefully reviewed and selected from 167 submissions. The papers are organized in topical sections on knowledge discovery and information retrieval; knowledge engineering and ontology development; and knowledge management and information sharing.

Knowledge Discovery, Knowledge Engineering and Knowledge Management

This book integrates two areas of computer science, namely data mining and evolutionary algorithms. Both these areas have become increasingly popular in the last few years, and their integration is currently an active research area. In general, data mining consists of extracting knowledge from data. The motivation for applying evolutionary algorithms to data mining is that evolutionary algorithms are robust search methods which perform a global search in the space of candidate solutions. This book emphasizes the importance of discovering comprehensible, interesting knowledge, which is potentially useful for intelligent decision making. The text explains both basic concepts and advanced topics

Data Mining and Knowledge Discovery with Evolutionary Algorithms

The three volume proceedings LNAI 11906 – 11908 constitutes the refereed proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases, ECML PKDD 2019, held in Würzburg, Germany, in September 2019. The total of 130 regular papers presented in these volumes was carefully reviewed and selected from 733 submissions; there are 10 papers in the demo track. The contributions were organized in topical sections named as follows: Part I: pattern mining; clustering, anomaly and outlier detection, and autoencoders; dimensionality reduction and feature selection; social networks and graphs; decision trees, interpretability, and causality; strings and streams; privacy and security; optimization. Part II: supervised learning; multi-label learning; large-scale learning; deep learning; probabilistic models; natural language processing. Part III: reinforcement learning and bandits; ranking; applied data science: computer vision and explanation; applied data science: healthcare; applied data science: e-commerce, finance, and advertising; applied data science: rich data; applied data science: applications; demo track. Chapter "Heavy-tailed Kernels Reveal a Finer Cluster Structure in t-SNE Visualisations" is available open access under a Creative Commons Attribution 4.0 International License via link.springer.com.

Machine Learning and Knowledge Discovery in Databases

The multi-volume set LNAI 12975 until 12979 constitutes the refereed proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases, ECML PKDD 2021, which was held during September 13-17, 2021. The conference was originally planned to take place in Bilbao, Spain, but changed to an online event due to the COVID-19 pandemic. The 210 full papers presented in these proceedings were carefully reviewed and selected from a total of 869 submissions. The volumes are organized in topical sections as follows: Research Track: Part I: Online learning; reinforcement learning; time series, streams, and sequence models; transfer and multi-task learning; semi-supervised and few-shot learning; learning algorithms and applications. Part II: Generative models; algorithms and learning theory; graphs and networks; interpretation, explainability, transparency, safety. Part III: Generative models; search and optimization; supervised learning; text mining and natural language processing; image processing, computer vision and visual analytics. Applied Data Science Track: Part IV: Anomaly detection and malware; spatio-temporal data; e-commerce and finance; healthcare and medical applications (including Covid); mobility and transportation. Part V: Automating machine learning, optimization, and feature engineering; machine learning based simulations and knowledge discovery; recommender systems and behavior modeling; natural language processing; remote sensing, image and video processing; social media.

Machine Learning and Knowledge Discovery in Databases. Research Track

This book constitutes the refereed proceedings of the 10th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2006, held in Singapore in April 2006. The 67 revised full papers and 33 revised short papers presented together with 3 invited talks were carefully reviewed and selected from 501 submissions. The papers are organized in topical sections on Classification, Ensemble Learning, Clustering, Support Vector Machines, Text and Document Mining, Web Mining, Bio-Data Mining, and more.

Advances in Knowledge Discovery and Data Mining

This two-volume set constitutes the refereed proceedings of the workshops which complemented the 21th Joint European Conference on Machine Learning and Knowledge Discovery in Databases, ECML PKDD, held in September 2021. Due to the COVID-19 pandemic the conference and workshops were held online. The 104 papers were thoroughly reviewed and selected from 180 papers submited for the workshops. This two-volume set includes the proceedings of the following workshops:Workshop on Advances in Interpretable Machine Learning and Artificial Intelligence (AIMLAI 2021)Workshop on Parallel, Distributed and Federated Learning (PDFL 2021)Workshop on Graph Embedding and Mining (GEM 2021)Workshop on Machine Learning for Irregular Time-series (ML4ITS 2021)Workshop on IoT, Edge, and Mobile for Embedded Machine Learning (ITEM 2021)Workshop on eXplainable Knowledge Discovery in Data Mining (XKDD 2021)Workshop on Bias and Fairness in AI (BIAS 2021)Workshop on Workshop on Active Inference (IWAI 2021)Workshop on Machine Learning for Cybersecurity (MLCS 2021)Workshop on Machine Learning in Software Engineering (MLiSE 2021)Workshop on MIning Data for financial applications (MIDAS 2021)Sixth Workshop on Data Science for Social Good (SoGood 2021)Workshop on Machine Learning for Pharma and Healthcare Applications (PharML 2021)Second Workshop on Evaluation and Experimental Design in Data Mining and Machine Learning (EDML 2020)Workshop on Machine Learning for Buildings Energy Management (MLBEM 2021)

Machine Learning and Principles and Practice of Knowledge Discovery in Databases