Advances In Knowledge Discovery And Data Mining 17th Pacific Asia Conference Pakdd 2013 Gold Coast Australia April 14 17 2013 Proceedings Part I Lecture Notes In Computer Science

During the last decade, the French-speaking scientific community developed a very strong research activity in the field of Knowledge Discovery and Management (KDM or EGC for " Extraction et Gestion des Connaissances" in French), which is concerned with, among others, Data Mining, Knowledge Discovery, Business Intelligence, Knowledge Engineering and Semantic Web. The recent and novel research contributions collected in this book are extended and reworked versions of a selection of the best papers that were originally presented in French at the EGC 2009 Conference held in Strasbourg, France on January 2009. The volume is organized in four parts. Part I includes five papers concerned by various aspects of supervised learning or information retrieval. Part II presents five papers concerned with unsupervised learning issues. Part III includes two papers on data streaming and two on security while in Part IV the last four papers are concerned with ontologies and semantic.

This book presents a collection of representative and novel work in the field of data mining, knowledge discovery, clustering and classification, based on expanded and reworked versions of a selection of the best papers originally presented in French at the EGC 2014 and EGC 2015 conferences held in Rennes (France) in January 2014 and Luxembourg in January 2015. The book is in three parts. The first four chapters discuss optimization considerations in data mining. The second part explores specific quality measures, dissimilarities and ultrametrics. The final chapters focus on semantics, ontologies and social networks. Written for PhD and MSc students, as well as researchers working in the field, it addresses both theoretical and practical aspects of knowledge discovery and management.

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.

Eight sections of this book span fundamental issues of knowledge discovery, classification and clustering, trend and deviation analysis, dependency derivation, integrated discovery systems, augmented database systems and application case studies. The appendices provide a list of terms used in the literature of the field of data mining and knowledge discovery in databases, and a list of online resources for the KDD researcher.

Volume 6

6th Pacific-Asia Conference, PAKDD 2002, Taipei, Taiwan, May 6-8, 2002. Proceedings

PAKDD 2018 Workshops, BDASC, BDM, ML4Cyber, PAISI, DaMEMO, Melbourne, VIC, Australia, June 3, 2018, Revised Selected Papers

Volume 9

Proceedings

This two-volume set, LNAI 10234 and 10235, constitutes the thoroughly refereed proceedings of the 21st Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining, PAKDD 2017, held in Jeju, South Korea, in May 2017. The 129 full papers were carefully reviewed and selected from 458 submissions. They are organized in topical sections named: classification and deep learning; social network and graph mining; privacy-preserving mining and security/risk applications; spatio-temporal and sequential data mining; clustering and anomaly detection; recommender system; feature selection; text and opinion mining; clustering and matrix factorization; dynamic, stream data mining; novel models and algorithms; behavioral data mining; graph clustering and community detection; dimensionality reduction.

The two-volume set LNAI 7818 + LNAI 7819 constitutes the refereed proceedings of the 17th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2013, held in Gold Coast, Australia, in April 2013. The total of 98 papers presented in these proceedings was carefully reviewed and selected from 363 submissions. They cover the general fields of data mining and KDD extensively, including pattern mining, classification, graph mining, applications, machine learning, feature selection and dimensionality reduction, multiple information sources mining, social networks, clustering, text mining, text classification, imbalanced data, privacy-preserving data mining, recommendation, multimedia data mining, stream data mining, data preprocessing and representation.

This book is a collection of representative and novel works in the field of data mining, knowledge discovery, clustering and classification. Discussing both theoretical and practical aspects of "Knowledge Discovery and Management" (KDM), it is intended for researchers interested in these fields, including PhD and MSc students, and researchers from public or private laboratories. The contributions included are extended and reworked versions of six of the best papers that were originally presented in French at the EGC'2016 conference held in Reims (France) in January 2016. This was the 16th edition of this successful conference, which takes place each year, and also featured workshops and other events with the aim of promoting exchanges between researchers and companies concerned with KDM and its applications in business, administration, industry and public organizations. For more details about the EGC society, please consult egc.asso.fr.

This book constitutes the refereed proceedings of the 11th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2007, held in Nanjing, China, May 2007. It covers new ideas, original research results and practical development experiences from all KDD-related areas including data mining, machine learning, data warehousing, data visualization, automatic scientific discovery, knowledge acquisition and knowledge-based systems.

Advances in Knowledge Discovery and Data Mining, Part I

Knowledge Mining Using Intelligent Agents

Advances in Distributed and Parallel Knowledge Discovery

... Pacific-Asia Conference, PAKDD ..., Proceedings

Trends and Applications in Knowledge Discovery and Data Mining

This book presents introductions to DKD and PKD, extensive reviews of the field, and state-of-the-art techniques. Foreword by Vipin Kumar Knowledge discovery and data mining (KDD) deals with the problem of extracting interesting associations, classifiers, clusters, and other patterns from data. The emergence of network-based distributed computing environments has introduced an important new dimension to this problem--distributed sources of data. Traditional centralized KDD typically requires central aggregation of distributed data, which may not always be feasible because of limited network bandwidth, security concerns, scalability problems, and other practical issues. Distributed knowledge discovery (DKD) works with the merger of communication and computation by analyzing data in a distributed fashion. This technology is particularly useful for large heterogeneous distributed environments such as the Internet, intranets, mobile computing environments, and sensor-networks. When the data sets are large, scaling up the speed of the KDD process is crucial. Parallel knowledge discovery (PKD) techniques addresses this problem by using high-performance multiprocessor machines. This book presents introductions to DKD and PKD, extensive reviews of the field, and state-of-the-art techniques. Contributors Rakesh Agrawal, Khaled AlSabti, Stuart Bailey, Philip Chan, David Cheung, Vincent Cho, Joydeep Ghosh, Robert Grossman, Yi-ke Guo, John Hale, John Hall, Daryl Hershberger, Ching-Tien Ho, Erik Johnson, Chris Jones, Chandrika Kamath, Hillol Kargupta, Charles Lo, Balinder Malhi, Ron Musick, Vincent Ng, Byung-Hoon Park, Srinivasan Parthasarathy, Andreas Prodromidis, Foster Provost, Jian Pun, Ashok Ramu, Sanjay Ranka, Mahesh Sreenivas, Salvatore Stolfo, Ramesh Subramonian, Janjao Sutiwaraphun, Kagan Tummer, Andrei Turinsky, Beat Wüthrich, Mohammed Zaki, Joshua Zhang

Knowledge discovery and data mining have become areas of growing significance because of the recent increasing demand for KDD techniques, including those used in machine learning, databases, statistics, knowledge acquisition, data visualization, and high performance computing. In view of this, and following the success of the five previous PAKDD conferences, the sixth Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2002) aimed to provide a forum for the sharing of original research results, innovative ideas, state-of-the-art developments, and implementation experiences in knowledge discovery and data mining among researchers in academic and industrial organizations. Much work went into preparing a program of high quality. We received 128 submissions. Every paper was reviewed by 3 program committee members, and 32 were selected as regular papers and 20 were selected as short papers, representing a 25% acceptance rate for regular papers. The PAKDD 2002 program was further enhanced by two keynote speeches, delivered by Vipin Kumar from the Univ. of Minnesota and Rajeev Rastogi from AT&T. In addition, PAKDD 2002 was complemented by three tutorials, XML and data mining (by Kyuseok Shim and Surajit Chadhuri), mining customer data across various customer touchpoints at- commerce sites (by Jaideep Srivastava), and data clustering analysis, from simple groupings to scalable clustering with constraints (by Osmar Zaiane and Andrew Foss).

This book presents recent advances in Knowledge discovery in databases (KDD) with a focus on the areas of market basket database, time-stamped databases and multiple related databases. Various interesting and intelligent algorithms are reported on data mining tasks. A large number of association measures are presented, which play significant roles in decision support applications. This book presents, discusses and contrasts new developments in mining time-stamped data, the identification of temporal patterns, the mining of multiple related databases, as well as local patterns analysis.

This book is a collection of representative and novel works done in Data Mining, Knowledge Discovery, Clustering and Classification that were originally presented in French at the EGC'2012 Conference held in Bordeaux, France, on January 2012. This conference was the 12th edition of this event, which takes place each year and which is now successful and well-known in the French-speaking community. This community was structured in 2003 by the foundation of the French-speaking EGC society (EGC in French stands for "Extraction et Gestion des Connaissances" and means "Knowledge Discovery and Management", or KDM). This book is intended to be read by all researchers interested in these fields, including PhD or MSc students, and researchers from public or private laboratories. It concerns both theoretical and practical aspects of KDM. The book is structured in two parts called "Knowledge Discovery and Data Mining" and "Classification and Feature Extraction or Selection". The first part (6 chapters) deals with data clustering and data mining. The three remaining chapters of the second part are related to classification and feature extraction or feature selection.

Volume 7

15th Pacific-Asia Conference, PAKDD 2011, Shenzhen, China, May 24-27, 2011, Proceedings

14th Pacific-Asia Conference, PAKDD 2010, Hyderabat, India, June 21-24, 2010, Proceedings

Volume 2

Advances in Algorithms, Theory, and Applications

Since the initial work on constrained clustering, there have been numerous advances in methods, applications, and our understanding of the theoretical properties of constraints and constrained clustering algorithms. Bringing these developments together, Constrained Clustering: Advances in Algorithms, Theory, and Applications presents an extensive collection of the latest innovations in clustering data analysis methods that use background knowledge encoded as constraints. Algorithms The first five chapters of this volume investigate advances in the use of instance-level, pairwise constraints for partitional and hierarchical clustering. The book then explores other types of constraints for clustering, including cluster size balancing, minimum cluster size, and cluster-level relational constraints. Theory It also describes variations of the traditional clustering under constraints problem as well as approximation algorithms with helpful performance guarantees. Applications The book ends by applying clustering with constraints to relational data, privacy-preserving data publishing, and video surveillance data. It discusses an interactive visual clustering approach, a distance metric learning approach, existential constraints, and automatically generated constraints. With contributions from industrial researchers and leading academic experts who pioneered the field, this volume delivers thorough coverage of the capabilities and limitations of constrained clustering methods as well as introduces new types of constraints and clustering algorithms.

This book is a collection of representative and novel works done in Data Mining, Knowledge Discovery, Clustering and Classification that were originally presented in French at the EGC'2013 (Toulouse, France, January 2013) and EGC'2014 Conferences (Rennes, France, January 2014). These conferences were respectively the 13th and 14th editions of this event, which takes place each year and which is now successful and well-known in the French-speaking community. This community was structured in 2003 by the foundation of the French-speaking EGC society (EGC in French stands for "Extraction et Gestion des Connaissances" and means "Knowledge Discovery and Management", or KDM). This book is aiming at all researchers interested in these fields, including PhD or MSc students, and researchers from public or private laboratories. It concerns both theoretical and practical aspects of KDM. The book is structured in two parts called "Applications of KDM to real datasets" and "Foundations of KDM". 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.

During the last decade, Knowledge Discovery and Management (KDM or, in French, EGC for Extraction et Gestion des connaissances) has been an intensive and fruitful research topic in the French-speaking scientific community. In 2003, this enthusiasm for KDM led to the foundation of a specific French-speaking association, called EGC, dedicated to supporting and promoting this topic. More precisely, KDM is concerned with the interface between knowledge and data such as, among other things, Data Mining, Knowledge Discovery, Business Intelligence, Knowledge Engineering and Semantic Web. The recent and novel research contributions collected in this book are extended and reworked versions of a selection of the best papers that were originally presented in French at the EGC 2010 Conference held in Tunis, Tunisia in January 2010. The volume is organized in three parts. Part I includes four chapters concerned with various aspects of Data Cube and Ontology-based representations. Part II is composed of four chapters concerned with Efficient Pattern Mining issues, while in Part III the last four chapters address Data Preprocessing and Information Retrieval.

This book constitutes the thoroughly refereed post-workshop proceedings at PAKDD Workshops 2017, held in conjunction with PAKDD, the 21st Pacific-Asia Conference on Knowledge Discovery and Data Mining in May 2017 in Jeju, South Korea. The 17 revised papers presented were carefully reviewed and selected from 38 submissions. The workshops affiliated with PAKDD 2017 include: Workshop on Machine Learning for Sensory Data Analysis (MLSDA), Workshop on Biologically Inspired Data Mining Techniques (BDM), Pacific Asia Workshop on Intelligence and Security Informatics (PAISI), and Workshop on Data Mining in Business Process Management (DM-BPM).

This book constitutes the refereed proceedings of the 5th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2001, held in Hong Kong, China in April 2001. The 38 revised full papers and 22 short papers presented were carefully reviewed and selected from a total of 152 submissions. The book offers topical sections on Web mining, text mining, applications and tools, concept hierarchies, feature selection, interestingness, sequence mining, spatial and temporal mining, association mining, classification and rule induction, clustering, and advanced topics and new methods.

The three-volume set LNAI 11439, 11440, and 11441 constitutes the thoroughly refereed proceedings of the 23rd Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2019, held in Macau, China, in April 2019. The 137 full papers presented were carefully reviewed and selected from 542 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: classification and supervised learning; text and opinion mining; spatio-temporal and stream data mining; factor and tensor analysis; healthcare, bioinformatics and related topics; clustering and anomaly detection; deep learning models and applications; sequential pattern mining; weakly supervised learning; recommender system; social network and graph mining; data pre-processing and featureselection; representation learning and embedding; mining unstructured and semi-structured data; behavioral data mining; visual data mining; and knowledge graph and interpretable data mining.

Advances in Machine Learning and Data Mining for Astronomy documents numerous successful collaborations among computer scientists, statisticians, and astronomers who illustrate the application of state-of-the-art machine learning and data mining techniques in astronomy. Due to the massive amount and complexity of data in most scientific disciplines

Data Mining and Knowledge Discovery for Process Monitoring and Control

Advances in Knowledge Discovery and Data Mining

11th Pacific-Asia Conference, PAKDD 2007, Nanjing, China, May 22-25, 2007, Proceedings

Volume 5

23rd Pacific-Asia Conference, PAKDD 2019, Macau, China, April 14-17, 2019, Proceedings, Part III

Knowledge Mining Using Intelligent Agents explores the concept of knowledge discovery processes and enhances decision-making capability through the use of intelligent agents like ants, termites and honey bees. In order to provide readers with an integrated set of concepts and techniques for understanding knowledge discovery and its practical utility, this book blends two distinct disciplines data mining and knowledge discovery process, and intelligent agents-based computing (swarm intelligence and computational intelligence). For the more advanced reader, researchers, and decision/policy-makers are given an insight into emerging technologies and their possible hybridization, which can be used for activities like dredging, capturing, distributions and the utilization of knowledge in their domain of interest (i.e. business, policy-making, etc.). By studying the behavior of swarm intelligence, this book aims to integrate the computational intelligence paradigm and intelligent distributed agents architecture to optimize various engineering problems and efficiently represent knowledge from the large gamut of data.

The two-volume set LNAI 6634 and 6635 constitutes the refereed proceedings of the 15th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2011, held in Shenzhen, China in May 2011. The total of 32 revised full papers and 58 revised short papers were carefully reviewed and selected from 331 submissions. The papers present new ideas, original research results, and practical development experiences from all KDD-related areas including data mining, machine learning, artificial intelligence and pattern recognition, data warehousing and databases, statistics, knowledge engineering, behavior sciences, visualization, and emerging areas such as social network analysis.

This book is a collection of high scientific novel contributions addressing several of these challenges. These articles are extended versions of a selection of the best papers that were initially presented at the French-speaking conferences EGC'2019 held in Metz (France, January 21-25, 2019). These extended versions have been accepted after an additional peer-review process among papers already accepted in long format at the conference. Concerning the conference, the long and short papers selection were also the result of a double blind peer review process among the hundreds of papers initially submitted to each edition of the conference (acceptance rate for long papers is about 25%.

This book constitutes the thoroughly refereed post-workshop proceedings at PAKDD Workshops 2018, held in conjunction with the 22nd Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2018, in Melbourne, Australia, in June 2018. The 32 revised papers presented were carefully reviewed and selected from 46 submissions. The workshops affiliated with PAKDD 2018 include: Workshop on Big Data Analytics for Social Computing, BDASC, Australasian Workshop on Machine Learning for Cyber-security, ML4Cyber, Workshop on Biologically-inspired Techniques for Knowledge Discovery and Data Mining, BDM, Pacific Asia Workshop on Intelligence and Security Informatics, PAISI, and Workshop on Data Mining for Energy Modeling and Optimization, DaMEMO.

22nd Pacific-Asia Conference, PAKDD 2018, Melbourne, VIC, Australia, June 3-6, 2018, Proceedings, Part I

5th Pacific-Asia Conference, PAKDD 2001 Hong Kong, China, April 16-18, 2001. Proceedings

22nd Pacific-Asia Conference, PAKDD 2018, Melbourne, VIC, Australia, June 3-6, 2018, Proceedings, Part II

Volume 4

The two-volume set LNAI 7301 and 7302 constitutes the refereed proceedings of the 16th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2012, held in Kuala Lumpur, Malaysia, in May 2012. The total of 20 revised full papers and 66 revised short papers were carefully reviewed and selected from 241 submissions. The papers present new ideas, original research results, and practical development experiences from all KDD-related areas. The papers are organized in topical sections on supervised learning: active, ensemble, rare-class and online; unsupervised learning: clustering, probabilistic modeling in the first volume and on data mining: networks, graphs, time-series and outlier detection, and data manipulation: pre-processing and dimension reduction in the second volume.

Mohamed Medhat Gaber " It is not my aim to surprise or shock you – but the simplest way I can summarise is to say that there are now in the world machines that think, that learn and that create. Moreover, their ability to do these things is going to increase rapidly until – in a visible future – the range of problems they can handle will be coextensive with the range to which the human mind has been applied" by Herbert A. Simon (1916-2001) 1 Overview This book suits both graduate students and researchers with a focus on discovering knowledge from scienti c data. The use of computational power for data analysis and knowledge discovery in scienti c disciplines has found its roots with the re- lution of high-performance computing systems. Computational science in physics, chemistry, and biology represents the rst step towards automation of data analysis tasks. The rational behind the development of computational science in different - eas was automating mathematical operations performed in those areas. There was no attention paid to the scienti c discovery process. Automated Scienti c Disc- ery (ASD) [1–3] represents the second natural step. ASD attempted to automate the process of theory discovery supported by studies in philosophy of science and cognitive sciences. Although early research articles have shown great successes, the area has not evolved due to many reasons. The most important reason was the lack of interaction between scientists and the automating systems.

This book constitutes the proceedings of the 14th Pacific-Asia Conference, PAKDD 2010, held in Hyderabad, India, in June 2010.

This text surveys research from the fields of data mining and information visualisation and presents a case for techniques by which information visualisation can be used to uncover real knowledge hidden away in large databases.

PAKDD 2017 Workshops, MLSDA, BDM, DM-BPM Jeju, South Korea, May 23, 2017, Revised Selected Papers

Scientific Data Mining and Knowledge Discovery

Advances in Knowledge Discovery in Databases

Advances in Knowledge Discovery and Data Mining, Part II

Constrained Clustering

Advances in Knowledge Discovery and Data Mining5th Pacific-Asia Conference, PAKDD 2001 Hong Kong, China, April 16-18, 2001. ProceedingsSpringer Science & Business Media

The two-volume set LNAI 6634 and 6635 constitutes the refereed proceedings of the 15th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2011, held in Shenzhen, China in May 2011. The total of 32 revised full papers and 58 revised short papers were carefully reviewed and selected from 331 submissions. The papers present new ideas, original research results, and practical development experiences from all KDD-related areas including data mining, machine learning, artificial intelligence and pattern recognition, data warehousing and databases, statistics, knoweldge engineering, behavior sciences, visualization, and emerging areas such as social network analysis.

This book highlights novel research in Knowledge Discovery and Management (KDM), gathering the extended, peer-reviewed versions of outstanding papers presented at the annual conferences EGC'2017 & EGC'2018. The EGC conference cycle was founded by the International French-speaking EGC society ("Extraction et Gestion des Connaissances") in 2003, and has since become a respected fixture among the French-speaking community. In addition to the annual conference, the society organizes various other events in order to promote exchanges between researchers and companies concerned with KDM and its applications to business, administration, industry and public organizations. Addressing novel research in data science, semantic Web, clustering, and classification, the content presented here will chiefly benefit researchers interested in these fields, including Ph.D./M.Sc. students, at public and private laboratories alike.

14th Pacific-Asia Conference, PAKDD 2010, Hyderabad, India, June 21-24, 2010, Proceedings

Advances in Knowledge Discovery and Management

16th Pacific-Asia Conference, PAKDD 2012, Kuala Lumpur, Malaysia, May 29-June 1, 2012, Proceedings, Part II

Volume 8

Principles and Foundations