Preface

The constantly changing landscape of Data Mining makes it challenging for experts and practitioners to stay informed of the field’s most up-to-date research. That is why Information Science Reference is pleased to offer this four-volume reference collection that will empower students, researchers, and academicians with a strong understanding of critical issues within Data Mining by providing both broad and detailed perspectives on cutting-edge theories and developments. This reference is designed to act as a single reference source on conceptual, methodological, technical, and managerial issues, as well as provide insight into emerging trends and future opportunities within the discipline.

*Data Mining: Concepts, Methodologies, Tools and Applications* is organized into eight distinct sections that provide comprehensive coverage of important topics. The sections are: (1) Fundamental Concepts and Theories, (2) Development and Design Methodologies, (3) Tools and Technologies, (4) Utilization and Application, (5) Organizational and Social Implications, (6) Managerial Impact, (7) Critical Issues, and (8) Emerging Trends. The following paragraphs provide a summary of what to expect from this invaluable reference tool.

Section 1, **Fundamental Concepts and Theories**, serves as a foundation for this extensive reference tool by addressing crucial theories essential to the understanding of Data Mining. Introducing the book is “*A Study of XML Models for Data Mining*” by Richi Nayak, Sangeetha Kutty, and Tien Tran, a great foundation laying the groundwork for the basic concepts and theories that will be discussed throughout the rest of the book. Another chapter of note in Section 1 is titled “*Learning Different Concept Hierarchies and the Relations between them from Classified Data*” by Fernando Benites and Elena Sapozhnikova, which discusses concept hierarchies and classified data. Section 1 concludes, and leads into the following portion of the book with a nice segue chapter, “*Measuring Human Intelligence by Applying Soft Computing Techniques*,” by Priti Srinivas Sajja and Kunjal Mankad. Where Section 1 leaves off with fundamental concepts, Section 2 discusses architectures and frameworks in place for Data Mining.

Section 2, **Development and Design Methodologies**, presents in-depth coverage of the conceptual design and architecture of Data Mining, focusing on aspects including soft computing, association rules, data processing, spatial data mining, missing data, rule mining, and many more topics. Opening the section is “*Web Mining and Social Network Analysis*” by Roberto Marmo. This section is vital for developers and practitioners who want to measure and track the progress of digital literacy on a local, national, or international level. Through case studies, this section lays excellent groundwork for later sections that will get into present and future applications for Data Mining, including, of note: “*A Framework on Data Mining on Uncertain Data with Related Research Issues in Service Industry*” by Edward Hung, and “*A Highly Scalable and Adaptable Co-Learning Framework on Multimodal Data Mining in a Multimedia Database*” by Zhen Guo, Christos Faloutzos, and Zhongfei (Mark) Zhang. The section concludes with an excellent work by Rahime Belen and Tugba Taskaya Temizel, titled “*A Framework to Detect Disguised Missing Data.*"
Section 3, **Tools and Technologies**, presents extensive coverage of the various tools and technologies used in the implementation of Data Mining. Section 3 begins where Section 2 left off, though this section describes more concrete tools at place in the modeling, planning, and applications of Data Mining. The first chapter, “XML Mining for Semantic Web,” by Rafael Berlanga and Victoria Nebot, lays a framework for the types of works that can be found in this section, a perfect resource for practitioners looking for new ways to mine data in the burgeoning Semantic Web. Section 3 is full of excellent chapters like this one, including such titles as “Database Analysis with ANNs by means of Graph Evolution,” “Data Mining System Execution Traces to Validate Distributed System Quality-of-Service Properties,” and “A Dynamic Privacy Manager for Compliance in Pervasive Computing” to name a few. Where Section 3 described specific tools and technologies at the disposal of practitioners, Section 4 describes successes, failures, best practices, and different applications of the tools and frameworks discussed in previous sections.

Section 4, **Utilization and Application**, describes how the broad range of Data Mining efforts has been utilized and offers insight on and important lessons for their applications and impact. Section 4 includes the widest range of topics because it describes case studies, research, methodologies, frameworks, architectures, theory, analysis, and guides for implementation. Topics range from telemedicine and telehealth to DSS design, systems biology, and content-based image retrieval. The first chapter in the section is titled “Virtual Telemedicine and Virtual Telehealth,” which was written by Shazia Kareem and Imran Sarwar Bajwa. The breadth of topics covered in the chapter is also reflected in the diversity of its authors, from countries all over the globe, including Pakistan, UK, Russia, Romania, China, Australia, Italy, Slovenia, Chile, the United States, and more. Section 4 concludes with an excellent view of a case study in technology implementation and use, “Systems Biology-Based Approaches Applied to Vaccine Development” by Patricio A. Manque.

Section 5, **Organizational and Social Implications**, includes chapters discussing the organizational and social impact of Data Mining. The section opens with “From Data to Knowledge” by Tri Wijaya. Where Section 4 focused on the broad, many applications of Data Mining technology, Section 5 focuses exclusively on how these technologies affect human lives, either through the way they interact with each other, or through how they affect behavioral/workplace situations. Other interesting chapters of note in Section 5 include “A Predictive Modeling of Retail Satisfaction” by M. Hemalatha, and “Behavioral Targeting Online Advertising” by Jun Yan, Dou Shen, Teresa Mah, Ning Liu, Zheng Chen, and Ying Li. Section 5 concludes with a fascinating study of a new development in Data Mining, in “Detecting Pharmaceutical Spam in Microblog Messages.”

Section 6, **Managerial Impact**, presents focused coverage of Data Mining as it relates to effective uses of complex data warehousing, multivariate data analysis, privacy compliance, and many more utilities. This section serves as a vital resource for developers who want to utilize the latest research to bolster the capabilities and functionalities of their processes. The section begins with “Innovative Approaches for Efficiently Warehousing Complex Data from the Web,” a great look into how small colleges and companies can utilize benefits previously thought to be reserved to their larger competitors. The 14 chapters in this section offer unmistakable value to managers looking to implement new strategies that work at larger bureaucratic levels. The section concludes with “Data Mining and Economic Crime Risk Management” by Nadine Lybaert. Where Section 6 leaves off, section seven picks up with a focus on some of the more content-theoretical material of this compendium.

Section 7, **Critical Issues**, presents coverage of academic and research perspectives on Data Mining tools and applications. The section begins with “Sociocognitive Inquiry,” by Brian R. Gaines and Mildred L. G. Shaw. Other issues covered in detail in Section 7 include interoperability, developmental
psychology, legal issues, privacy expectations, and much more. The section concludes with “Tailoring FOS-ERP Packages” by Klaus Wölfel and Jean-Paul Smets, a great transitional chapter between Sections 7 and 8 because it examines an important question going into the future of the field. The last chapter manages to show a theoretical look into future and potential technologies, a topic covered in more detail in Section 8.

Section 8, Emerging Trends, highlights areas for future research within the field of Data Mining, opening with “Optimization of a Hybrid Methodology (CRISP-DM)” by José Nava and Paula Hernández. Section 8 contains chapters that look at what might happen in the coming years that can extend the already staggering amount of applications for Data Mining. Other chapters of note include “Towards Spatial Decision Support System for Animals Traceability” and “An Approach for Land-Use Suitability Assessment Using Decision Support Systems, AHP, and GIS.” The final chapter of the book looks at an emerging field within Data Mining, in the excellent contribution, “Advances of the Location Based Context-Aware Mobile Services in the Transport Sector” by Maria Giaoutzi, Vassilios Vescoukis, and Georgios Patris.

Although the primary organization of the contents in this multi-volume work is based on its eight sections, offering a progression of coverage of the important concepts, methodologies, technologies, applications, social issues, and emerging trends, the reader can also identify specific contents by utilizing the extensive indexing system listed at the end of each volume. Furthermore to ensure that the scholar, researcher, and educator have access to the entire contents of this multi volume set as well as additional coverage that could not be included in the print version of this publication, the publisher will provide unlimited multi-user electronic access to the online aggregated database of this collection for the life of the edition, free of charge when a library purchases a print copy. This aggregated database provides far more contents than what can be included in the print version, in addition to continual updates. This unlimited access, coupled with the continuous updates to the database ensures that the most current research is accessible to knowledge seekers.

As a comprehensive collection of research on the latest findings related to using technology to providing various services, Data Mining: Concepts, Methodologies, Tools and Applications, provides researchers, administrators and all audiences with a complete understanding of the development of applications and concepts in Data Mining. Given the vast number of issues concerning usage, failure, success, policies, strategies, and applications of Data Mining in countries around the world, Data Mining: Concepts, Methodologies, Tools and Applications addresses the demand for a resource that encompasses the most pertinent research in technologies being employed to globally bolster the knowledge and applications of Data Mining.