Preface

It is a pleasure to present an eclectic collection of papers on Big Data, Data Mining, and Data Science within the Intelligent Computing (De Gruyter) book series. In an era where data reigns supreme, harnessing its value has become necessary in research, business, government, and military. The insights from analyzing large datasets have revolutionized industries, driven innovation, and transformed how we understand and interact with the world and technology ecosystems.

From a computing perspective in our data-rich world, Big Data, Data Mining, and Data Science collectively leverage data to uncover hidden knowledge and solve complex problems. Big Data deals with the vast volumes, velocity, and variety of structured and unstructured data. Data Mining focuses on extracting meaningful patterns and insights from large datasets for predictive modeling and decision support. Data Science aims to extract actionable insights using various techniques to solve complex problems and drive decision-making. These techniques are applied to diverse problems and domains, such as the financial sector, healthcare, e-commerce, and cybersecurity.

The work presented in this book can be loosely categorized into two distinct themes. The first theme is "methods and instrumentation," where authors provide insight into systematic methods, procedures, and techniques within a research or experimental framework and the tooling to measure, observe, or manipulate variables of interest. In this thematic collection, papers explore a range of topics such as *hypergraph databases*, *automated determination of cluster numbers for high-dimensional big data*, *centrality metrics for identifying dominant factors in datasets*, *machine learning-based data preprocessing approaches*, *estimation of time-series outliers using multi-objective optimization with non-stationary means*, and the *development of languages for generating random data to facilitate random testing of hardware and software applications.*

The second theme is "applications and case studies," where authors apply and implement theories, techniques, methodologies, and technologies in specific contexts, showcasing their practical relevance and effectiveness. In this thematic collection, papers explore a range of topics such as using *high-volume dynamic ensemble-based model computations in e-commerce,* deploying *explainable artificial intelligence (AI) to explain an assessment analytics algorithm for free text exams*, using *graph neural networks (NN) and gene interaction data*, applying *recurrent neural network (RNN) models to examine the volatility in financial markets during a global pandemic*, using *skill-centered qualification ontologies to support data mining of human resources in knowledge-based enterprise process-representations*, extracting *information from vibration sensor data using topological data analysis*, leveraging *generative AI (GenAI) and table arrangement techniques to analyze newspaper stories for stock price insight*, creating *metadata schemas for data reservoirs*, and exploring *the discrimination capabilities of a set of features for road surface classification.*

https://doi.org/10.1515/9783111344553-202

The book is mainly composed of selected papers that were accepted for the 2022 and 2023 International Conferences on Computational Science and Computational Intelligence (CSCI: December, Las Vegas, USA) and the 2023 International Conference on Data Science (CSCE/ICDATA: July, Las Vegas, USA).

Selected authors were given the opportunity to submit the extended versions of their conference papers for publication consideration in this book. An important mission of CSCI and CSCE annual conferences includes "*Providing a unique platform for a diverse community of constituents composed of scholars, researchers, developers, educators, and practitioners. The Congress makes a concerted effort to reach out to participants affiliated with diverse entities (such as universities, institutions, corporations, government agencies, and research centers/labs) worldwide. The Congress also attempts to connect participants from institutions that have teaching as their main mission with those who are affiliated with institutions that have research as their main mission. The Congress uses a quota system to achieve its institution and geography diversity objectives*." Since this book comprises the extended versions of the accepted papers of CSCI and CSCE annual conferences, it is no surprise that it has chapters from a highly qualified and diverse group of authors.

Aside from recognizing the authors who provided their research contributions, we are also grateful to the many colleagues who offered their time and effort in organizing the CSCI and CSCE conferences. Their help was instrumental in the formation of this book. The editorial committee members appear on the CSCI and CSCE's websites. Finally, we want to thank Steve Elliot (De Gruyter Editor) and Aleksandra Ślosarczyk (De Gruyter Editorial Project Manager) for their continuous support throughout the development and production of the book.

We hope our readers find as much value and gain insight from this book as we have.

George Dimitoglou, Ph.D.
Department of Computer Science & IT,
Hood College,
401 Rosemont Ave, Frederick, MD 21701, USA

Leonidas Deligiannidis, Ph.D.
School of Computing and Data Science,
Wentworth Institute of Technology,
550 Huntington Ave, Boston, MA 02115, USA

Hamid R. Arabnia, Ph.D.
University of Georgia,
415 GSRC, Computer Science,
Athens, GA 30602-7404, USA