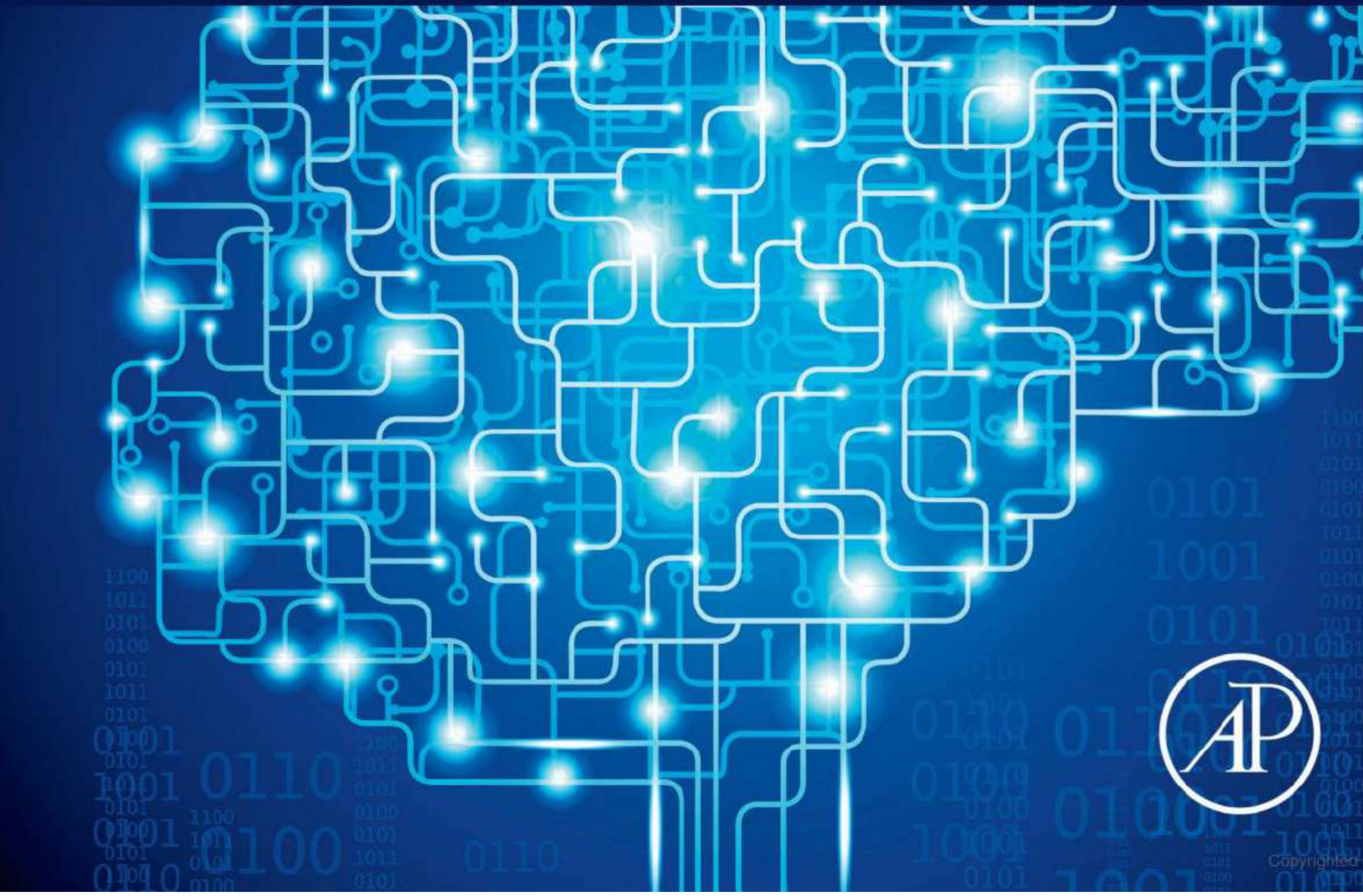
Intelligent Data Centric Systems

Series Editor Fatos Xhafa

Machine Learning, Big Data, and IoT for Medical Informatics

Edited by Pardeep Kumar, Yugal Kumar, and Mohamed A. Tawhid



Machine Learning, Big Data, and IoT for Medical Informatics

Edited by

Pardeep Kumar

Department of CSE & IT, JUIT, Solan, Himachal Pradesh, India

Yugal Kumar

Department of CSE & IT, JUIT, Solan, Himachal Pradesh, India

Mohamed A. Tawhid

Department of Mathematics and Statistics, Thompson Rivers University, Kamloops, BC, Canada

Series Editor

Fatos Xhafa

Technical University of Catalonia (UPC), Barcelona, Spain

ELSEVIER

ACADEMIC PRESS

An imprint of Elsevier

Academic Press is an imprint of Elsevier
125 London Wall, London EC2Y 5AS, United Kingdom
525 B Street, Suite 1650, San Diego, CA 92101, United States
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom

Notices
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-0-12-821777-1

For information on all Academic Press publications
visit our website at https://www.elsevier.com/books-and-journals

Publisher: Mara Conner
Acquisitions Editor: Sonnini R. Yura
Editorial Project Manager: Chiara Giglio
Production Project Manager: Nirmala Arumugam
Cover Designer: Victoria Pearson

Typeset by SPi Global, India

Contents

Contributors

Rashid Ali
School of Intelligent Mechatronics Engineering, Sejong University, Seoul, Republic of Korea

Majed Alsadhan
Department of Computer Science, Kansas State University, Manhattan, KS, United States

Emmanuel Ayodele
School of Electronic and Electrical Engineering, University of Leeds, Leeds, United Kingdom

Praveen Chakravarthy Bhallamudi
Lumirack Solutions, Chennai, India

Dilip Kumar Choubey
Department of Computer Science & Engineering, Indian Institute of Information Technology, Bhagalpur, India

Giuseppe Ciaburro
Department of Architecture and Industrial Design, Università degli Studi della Campania Luigi Vanvitelli, Aversa, Italy

Anup Das
Electrical and Computer Engineering, Drexel University, Philadelphia, PA, United States

Apurba Das
Department of CSE, PES University; Computer Vision (IoT), Tata Consultancy Services, Bangalore, India

Jaydeep Das
Advanced Technology Development Center, Indian Institute of Technology Kharagpur, West Bengal, India

Parneeta Dhaliwal
Department of Computer Science and Technology, Manav Rachna University, Faridabad, India

Haytham Elmiligi
Computing Science Department, Thompson Rivers University, Kamloops, BC, Canada

Ahmed A. Ewees
Department of Computer, Damietta University, Damietta, Egypt; Department of e-Systems, University of Bisha, Bisha, Saudi Arabia

O.K. Fasil
Department of Computer Science, Central University of Kerala, Kerala, India

Marwa A. Gaheen
Department of Computer, Damietta University, Damietta, Egypt

Pardeep Garg
Department of Electronics and Communication Engineering, Jaypee University of Information Technology, Solan, India

Isel Grau
Artificial Intelligence Lab, Free University of Brussels (VUB), Brussels, Belgium

Rahul Gupta
National Institute of Technology, Hamirpur, Himachal Pradesh, India

Sunil Kumar Hota
DIHAR, Defense Research & Development Organization, Leh, Jammu & Kashmir, India

William H. Hsu
Department of Computer Science, Kansas State University, Manhattan, KS, United States

Enas Ibrahim
Department of Computer, Damietta University, Damietta, Egypt

Muhammad Taimoor Khan
Medical Department, University of Debrecen, Debrecen, Hungary

Sung Won Kim
Department of Information and Communication Engineering, Yeungnam University, Gyeongsan, Republic of Korea

Dipesh Kumar
Department of ECE, IIT(ISM), Dhanbad, India

Vijay Kumar
National Institute of Technology, Hamirpur, Himachal Pradesh, India

Yugal Kumar
Department of CSE & IT, JUIT, Solan, Himachal Pradesh, India

Seshadri Sastry Kunapuli
Xinthe Technologies PVT LTD, Visakhapatnam, India

Deepti Lamba
Department of Computer Science, Kansas State University, Manhattan, KS, United States

Theodore V. Maliamanis
HUman-MAchines INteraction Laboratory (HUMAIN-Lab), Department of Computer Science, International Hellenic University, Kavala, Greece

Nirupama Mandal
Department of ECE, IIT(ISM), Dhanbad, India

Des McLernon
School of Electronic and Electrical Engineering, University of Leeds, Leeds, United Kingdom

Pradeep Kumar Naik
School of Life Sciences, Sambalpur University, Sambalpur, Orissa, India

Ali Nauman
Department of Information and Communication Engineering, Yeungnam University, Gyeongsan, Republic of Korea

Muhammad Hassan Nawaz
Electrical Engineering Department, University of Debrecen, Debrecen, Hungary

Ann Nowe
Artificial Intelligence Lab, Free University of Brussels (VUB), Brussels, Belgium

George A. Papakostas
HUman-MAchines INteraction Laboratory (HUMAIN-Lab), Department of Computer Science, International Hellenic University, Kavala, Greece

Advika Parthvi
School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

Yazdan Ahmad Qadri
Department of Information and Communication Engineering, Yeungnam University, Gyeongsan, Republic of Korea

Rakesh Raja
Department of Computer Science and Engineering, Birla Institute of Technology, Mesra, Ranchi, India

R. Rajesh
Department of Computer Science, Central University of Kerala, Kerala, India

Kartik Rawal
School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

Riya Sapra
Department of Computer Science and Technology, Manav Rachna University, Faridabad, India

Bikash Kanti Sarkar
Department of Computer Science and Engineering, Birla Institute of Technology, Mesra, Ranchi, India

Jane Scott
School of Architecture, Planning and Landscape, Newcastle University, Newcastle, United Kingdom

Dipankar Sengupta
PGJCCR, Queens University Belfast, Belfast, United Kingdom; Artificial Intelligence Lab, Free University of Brussels (VUB), Brussels, Belgium

Isha Sharma
National Institute of Technology, Hamirpur, Himachal Pradesh, India

Sunil Datt Sharma
Department of Electronics and Communication Engineering, Jaypee University of Information Technology, Solan, India

Vijay Kumar Sharma
DIHAR, Defense Research & Development Organization, Leh, Jammu & Kashmir, India

Rohit Shukla
Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology (JUIT), Solan, Himachal Pradesh, India

Vaibhav Shukla
Tech Mahindra Ltd., Mumbai, Maharastra, India

S.S. Shylaja
Department of CSE, PES University, Bangalore, India

Tiratha Raj Singh
Centre of Excellence in Healthcare Technologies and Informatics (CHETI), Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology (JUIT), Solan, Himachal Pradesh, India

Ravi B. Srivastava
DIHAR, Defense Research & Development Organization, Leh, Jammu & Kashmir, India

Wenhan Tan
Electrical and Computer Engineering, Drexel University, Philadelphia, PA, United States

Abdalrahman Tawhid
Computing Science Department, Thompson Rivers University, Kamloops, BC, Canada

Tanya Teotia
Computing Science Department, Thompson Rivers University, Kamloops, BC, Canada

Arvind Kumar Yadav
Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology (JUIT), Solan, Himachal Pradesh, India

Syed Ali Raza Zaidi
School of Electronic and Electrical Engineering, University of Leeds, Leeds, United Kingdom

Zhiqiang Zhang
School of Electronic and Electrical Engineering, University of Leeds, Leeds, United Kingdom

Preface

Medical informatics, also known as healthcare analytics, is a useful tool that can assess and monitor health-related behavior and conditions of individuals outside the clinic. The benefits of medical informatics are significant, including improving life expectancy, disease diagnosis, and quality of life. In many individual situations, a patient requires continuous monitoring to identify the onset of possible life-threatening conditions or to diagnose potentially dangerous diseases. Traditional healthcare systems fall short in this regard.

Meanwhile, rapid growth and advances have occurred in the digitization of information, retrieval systems, and wearable devices and sensors. Our times demand the design and development of new effective prediction systems using machine learning approaches, big data, and the Internet of Things (IoT) to meet health and life quality expectations. Furthermore, there is a need for monitoring systems that can monitor the health issues of elderly and remotely located people. In recent times, big data and IoT have played a vital role in health-related applications, mainly in disease identification and diagnosis. These techniques can provide possible solutions for healthcare analytics, in which both structured and unstructured data are collected through IoT-based devices and sensors. Machine learning and big data techniques can be applied to collected data for predictive diagnostic systems. However, designing and developing an effective diagnostic system is still challenging due to various issues like security, usability, scalability, privacy, development standards, and technologies. Therefore machine learning, big data, and IoT for medical informatics are becoming emerging research areas for the healthcare community.

Outline of the book and chapter synopses

This book presents state-of-the-art intelligent techniques and approaches, design, development, and innovative uses of machine learning, big data, and IoT for demanding applications of medical informatics. This book also focuses on different data collection methods from IoT-based systems and sensors, as well as preprocessing and privacy preservation of medical data. We have provided potential thoughts and methodologies to help senior undergraduate and graduate students, researchers, programmers, and healthcare industry professionals create new knowledge for the future to develop intelligent machine learning, big data, and IoT-based novel approaches for medical informatics applications. Further, the key roles and great importance of machine learning, big data, and IoT techniques as mathematical tools are elaborated in the book. A brief and orderly introduction to the chapters is provided in the following paragraphs. The book contains 23 chapters.

Chapter 1 presents a survey of machine learning and predictive analytics methods for medical informatics. This chapter focuses on deep neural networks with typical use cases in computational medicine, including self-supervised learning scenarios: these include convolutional neural networks for image analysis, recurrent neural networks for time series, and generative adversarial models for correction of class imbalance in differential diagnosis and anomaly detection. The authors then continue by assessing salient connections between the current state of machine learning research and data-centric healthcare analytics, focusing specifically on diagnostic imaging and multisensor integration as crucial research topics within predictive analytics. Finally, they conclude by relating open problems of machine learning for prediction-based medical informatics surveyed in this article to the impact of big data and its associated challenges, trends, and limitations of current work, including privacy and security of sensitive patient data.

Chapter 2 presents a proposed model for geolocation aware healthcare facility with IoT, Fog, and Cloud-based diagnosis in emergency cases. An end-to-end infrastructure has been modeled for the healthcare system using geolocation-enabled IoT, fog, and cloud computing technology to identify the nearest hospital or medical facility available to the patient. It has also achieved 25%–27% less delay and 27%–29% less power consumption than the cloud-only environment.

Chapter 3 aims to capture the status of medical computer vision threats and the recent defensive techniques proposed by researchers. This chapter intends to shed light on the vulnerability of machine learning models in medical image analysis, e.g., disease diagnosis, and to become a guide for any researcher working in medical image analysis toward the development of more secure machine learning-based computer-aided diagnosis systems.

Chapter 4 demonstrates a model for skull stripping and tumor detection from brain images using 3D U-Net. The demonstrated model has been tested over 373 MRIs of the LCG Segmentation Dataset, showing good standard performance over metrics of dice coefficient, and the accuracy results are competitive with the existing methods.

Chapter 5 addresses the issue of corrupted laparoscopy video by haze, noise, oversaturated illumination, etc., in minimally invasive surgery. To effectively address the issue, the authors have proposed a novel algorithm to ensure the enhancement of video with faster performance. The proposed C^2D^2A (Cross Color Dominant Deep Autoencoder) uses the strength of (a) a bilateral filter, which addresses the one-shot filtering of images both in the spatial neighborhood domain and psycho-visual range; and (b) a deep autoencoder, which can learn salient patterns. The domain-based color sparseness has further improved the performance, modulating the classical deep autoencoder to a color dominant deep autoencoder. The work has shown promise toward providing a generic framework of quality enhancement of video streams and addressing performance. This, in turn, improves the image/video analytics like segmentation, detection, and tracking the objects or regions of interest.

Chapter 6 presents an alternative way of estimating respiratory rate from ECG and PPG by using machine learning to improve estimation accuracy. The proposed methods are based on respiratory signals extracted from raw signals and use a support vector machine (SVM) and neural network (NN) to estimate respiratory rate. The proposed methods achieve comparable accuracy to current methods when the number of classes is low. Once the number of classes increases, the accuracy drops significantly.

Chapter 7 serves as an introductory guideline to address the challenges and opportunities while designing machine learning-enabled Healthcare Internet of Things (H-IoT) networks. It provides a discussion on traditional H-IoT, challenges, and opportunities in the Network 2030 paradigm. It also discusses potential machine learning techniques compatible with H-IoT and points out open issues and future research directions.

Chapter 8 presents a skin lesion segmentation approach based on the Elitist-Jaya optimization algorithm. The proposed method contains two stages: image preprocessing and edge detection. The experimental sample consists of a set of 320 images from the skin lesion dataset. The outcomes proved that the proposed approach improved the segmentation accuracy of the affected skin lesion area and outperformed the compared methods.

Chapter 9 provides its readers with an all-encompassing review that will enable a clear understanding of the current trends in glove-based gesture classification and provide new ideas for further research. The authors have analyzed deep learning approaches in terms of their current performance, advantages over classical machine learning algorithms, and limitations in specific classification scenarios. Furthermore, they present other deep learning approaches that may outperform current algorithms in glove-based gesture classification.

Chapter 10 presents an ensemble approach for evaluating the cognitive performance of the human population at high altitude. The authors identify the key multidomain cognitive screening test (MDCST) and clinical features among the lowlander ($\leq$350 m) and highlander ($\geq$1500 but <4300 m) populations, staying at an altitude $\geq$4300 m for a prolonged duration. A goodness-of-fit test was applied to the two population cohorts for identifying significant independent measures. Rule-based mining was followed to discover associative rules between the clinical, behavioral, and cognitive screening parameters. Conclusively, a unique set of association rules have been identified with at least 30% support and more than 60% confidence in behavioral and clinical features associated with the cognitive parameters.

Chapter 11 presents the role of machine learning in expert systems for disease diagnostics in human healthcare. The authors discuss essential existing expert systems for human disease diagnosis in detail. They also provide a brief evaluation of various techniques used for the development of expert systems.

Chapter 12 presents an entropy-based hybrid feature selection approach for medical datasets. A stable linear-time entropy-based ensembled feature selection approach is introduced, mainly focusing on medical datasets of several sizes. The suggested approach is validated using three state-of-the-art classifiers, namely C4.5, naïve Bayes, and JRIP, over 14 benchmark medical datasets (drawn from the UCI machine learning repository). The empirical results achieved from the datasets demonstrate that the proposed ensemble model outperforms the selected learners.

Chapter 13 shows how to utilize machine learning algorithms to create models that can predict healthcare systems' critical issues. The chapter's discussion relates to the COVID-19 pandemic and highlights the solutions offered by machine learning in such scenarios. The chapter also highlights the significance of feature engineering and its impact on machine learning models' accuracy. The chapter ends with two case studies. The first case study shows how to build a prediction model that can predict the number of diabetic patients who will visit certain hospitals in a specific geographic location in future years. The second case study analyzes health records during the COVID-19 pandemic.

Chapter 14 presents an interpretable semisupervised classifier for predicting cancer stages. Authors illustrate the self-labeling gray-box applications on the omics and clinical datasets from the cancer genome atlas. They show that the self-labeling gray-box is accurate in predicting cancer stages of rare cancers by leveraging the unlabeled instances from more common cancer types. They discuss insights, the features influencing prediction, and a global representation of the knowledge through decision trees or rule lists, which can aid clinicians and researchers.

Chapter 15 presents an overview of applications of blockchain technology in smart healthcare. The authors overviewed the fundamental blockchain concepts and applications to be used for different aspects of the smart healthcare industry and proposed a live patient monitoring system by deploying blockchain technology in the model. Keeping an eye on recent technologies in connected healthcare, they finally presented various research factors and potential challenges where blockchain technologies can play an outstanding role in realizing the concept of smart optimization in the healthcare industry.

Chapter 16 focuses on clustering and classification techniques for the prediction of leukemia. The proposed work consists of Phase I, which will be dealing with the collection of datasets and visualization of datasets, whereas Phase II will be dealing with the machine learning and data mining techniques for the prediction of leukemia disease. The authors claim that the proposed techniques would give higher performance than the existing techniques.

Chapter 17 presents a performance evaluation of fractal features toward seizure detection from electroencephalogram signals. The authors have evaluated the ability of three well-known fractal dimension feature extraction methods (the Katz fractal dimension, Higuchi fractal dimension, and Petrosian fractal dimension) to classify epileptic and nonepileptic electroencephalogram signals. The features are fed to an SVM classifier for the classification of epileptic and nonepileptic electroencephalogram signals. The SVM classifier results show that the fractal features are good measures to characterize the complex information of epileptic signals.

Chapter 18 presents an integer period discrete Fourier transform-based algorithm to identify tandem repeats in the DNA sequences. The authors have discussed the importance of tandem repeats in diverse applications. They proposed an integer period discrete Fourier transform (IPDFT)-based algorithm to detect the tandem repeats in DNA sequences. A comparison of the proposed algorithm's performance has also been made with existing methods.

Chapter 19 discusses the scope, applicability, and usage of blockchain technology to preserve patients' sensitive medical data. A framework is also proposed that allows patients and hospitals to store medical records. The framework allows patients to share the information by providing access to their data and by invoking smart contracts for automatic payments for their medical claims.

Chapter 20 presents a novel approach to securing e-health applications in the cloud environment. The authors provide an algorithm to secure data in e-health applications in the cloud environment. A new architecture for e-health applications in the cloud environment is proposed, which will provide application-level security and server-level security using certificates.

Chapter 21 presents different ensemble learning algorithms and explains how these algorithms can be used to classify health disorders. The authors have discussed an ensemble classifier approach for thyroid disease diagnosis using the AdaBoostM algorithm.

Chapter 22 presents a review of the latest artificial intelligence research in this immense medical science field, including various architectures and approaches, with special attention given to brain tumor analysis. The authors discuss various deep learning architectures used to diagnose brain tumors and compare results with existing architectures. They have examined case studies from basic clustering techniques such as K-means clustering to fuzzy and neurotrophic C-means clustering techniques and kernel graph cuts (KGC) to advanced artificial intelligence techniques such as deep convolution neural networks (DCNs), atrous convolution neural networks (ACNs), and unit architectures to find the area of interest in the coherent/incoherent regions.

Finally, Chapter 23 focuses on machine learning in precision medicine. An overview of how machine learning is used in precision medicine and its potential use in the detection, diagnosis, prognosis, risk assessment, therapy response, and discovery of new biomarkers and drug candidates is presented in this chapter.

We especially thank the Intelligent Data-Centric Systems: Sensor Collected Intelligence Series Editor, Prof. Fatos Xhafa, for his continuous support and insightful guidance.

We would also like to thank the publishers at Elsevier, in particular, Chiara Giglio, Editorial Project Manager, and Sonnini Ruiz Yura, Acquisitions Editor–Biomedical Engineering, for their helpful guidance and encouragement during this book's creation.

We are sincerely thankful to all authors, editors, and publishers whose works have been cited directly/indirectly in this manuscript.

Special acknowledgments

The first editor gratefully acknowledges the authorities of the Jaypee University of Information Technology, Waknaghat, Solan, Himachal Pradesh, India, for their kind support for this book.

The second editor gratefully acknowledges the authorities of the Jaypee University of Information Technology, Waknaghat, Solan, Himachal Pradesh, India, for their kind support for this book.

The third editor would like to acknowledge the Natural Sciences and Engineering Research Council of Canada and Thompson River University, Kamloops, Canada, for their kind support of his research on this book.

Dr. Pardeep Kumar
Solan, India

Dr. Yugal Kumar
Solan, India

Dr. Mohammad A. Tawhid
Kamloops, BC, Canada