Advances in Intelligent Systems and Computing 1355

Indrajit Pan
Anirban Mukherjee
Vincenzo Piuri *Editors*

Proceedings of Research and Applications in Artificial Intelligence

RAAI 2020

Springer

Advances in Intelligent Systems and Computing

Volume 1355

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia.

The publications within "Advances in Intelligent Systems and Computing" are primarily proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Indexed by DBLP, EI Compendex, INSPEC, WTI Frankfurt eG, zbMATH, Japanese Science and Technology Agency (JST).

All books published in the series are submitted for consideration in Web of Science.

More information about this series at http://www.springer.com/series/11156

Indrajit Pan · Anirban Mukherjee · Vincenzo Piuri
Editors

Proceedings of Research and Applications in Artificial Intelligence

RAAI 2020

Springer

Editors
Indrajit Pan
Department of Information Technology
RCC Institute of Information Technology
Kolkata, West Bengal, India

Anirban Mukherjee
Department of Information Technology
RCC Institute of Information Technology
Kolkata, West Bengal, India

Vincenzo Piuri
Department of Computer Science
Università degli Studi di Milano
Milan, Milano, Italy

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-981-16-1542-9 ISBN 978-981-16-1543-6 (eBook)
https://doi.org/10.1007/978-981-16-1543-6

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Dr. Indrajit Pan would like to dedicate this book to Prof. (Dr.) Siddhartha Bhattacharyya

Prof. Anirban Mukherjee would like to dedicate this book to all his research scholars

Prof. Vincenzo Piuri would like to dedicate this book to all researchers and professionals who are extraordinarily contributing to make our life better, and to students who will change our future

RAAI 2020 Committee Members

Honorary General Chairs

Dr. Kalyanmoy Deb, Michigan State University, USA
Dr. Aboul Ella Hassanien, Cairo University, Egypt
Dr. Valentina Emilia Balas, University of Arad, Romania

Chief Patron

Dr. Pranabesh Das, Director, Technical Education, Government of West Bengal, India and Chairman, BOG, RCCIIT, India

General Chair

Dr. Anirban Mukherjee, RCC Institute of Information Technology, Kolkata, India

International Advisory Chairs

Dr. Elizabeth Behrman, Wichita State University, Kansas
Dr. Xiao-Zhi Gao, University of Eastern Finland, Finland
Dr. Ricardo Baeza-Yates, Universitat Pompeu Fabra, Barcelona, Spain
Dr. Valentina Emilia Balas, University of Arad, Romania
Dr. Kalyanmoy Deb, Michigan State University, USA
Dr. Aboul Ella Hassanien, Cairo University, Egypt
Dr. Vincenzo Piuri, Universita degli Studi di Milano, Italy
Dr. Daniela Romano, University College of London, London, UK

Dr. Debotosh Bhattacharjee, Jadavpur University, India
Dr. Siddhartha Bhattacharyya, Christ University, Bengaluru, India
Dr. Debashis De, Maulana Abul Kalam Azad University of Technology, West Bengal, India
Dr. Paramartha Dutta, Visva Bharati University, Santiniketan, India
Dr. Hafizur Rahaman, Indian Institute of Engineering Science and Technology, Shibpur, India
Dr. Chiranjoy Chattopadhyay, Indian Institute of Technology, Jodhpur, India
Dr. Koushik Mondal, Indian Institute of Technology, (ISM), Dhanbad, India
Dr. Nilanjan Dey, Techno International, New Town, India
Dr. Ernesto Cuadros-Vargas, School of Computer Science, UTEC, Peru

Organizing Secretary

Dr. Indrajit Pan, RCC Institute of Information Technology, Kolkata, India
Moumita Deb, RCC Institute of Information Technology, Kolkata, India

Program Chairs

Dr. Ernesto Cuadros-Vargas, School of Computer Science, UTEC, Peru
Dr. Dipankar Majumdar, RCC Institute of Information Technology, Kolkata, India
Dr. Abhijit Das, RCC Institute of Information Technology, Kolkata, India
Dr. Shyantani Maiti, RCC Institute of Information Technology, Kolkata, India

Technical Program Committee

Dr. Rabie A. Ramadan, Cairo University, Egypt
Dr. Amlan Chatterjee, California State University, USA
Dr. Sherif Ismail, Umm Al-Qura University, Saudi Arabia
Dr. Ahmed A. Elngar, Beni-Suef University, Egypt
Dr. Pushpendu Kar, The University of Nottingham, Ningbo, China
Dr. Mohamed Abdelfattah, MET, Mansoura, Egypt
Dr. Rony Chatterjee, Microsoft Corporation, USA
Dr. Indrajit Banerjee, Indian Institute of Engineering Science and Technology, Shibpur, India
Dr. Tuhina Samanta, Indian Institute of Engineering Science and Technology, Shibpur, India
Dr. Shibakali Gupta, University of Burdwan, Burdwan, India
Dr. Subhamita Mukherjee, Techno Main, Salt Lake, India

Prof. Piyal Sarkar, Techno Main, Salt Lake, India
Dr. Sudip Ghosh, Indian Institute of Engineering Science and Technology, Shibpur, India
Dr. Arijit Ghosal, St. Thomas College of Engineering and Technology, Kolkata, India
Dr. Abhishek Bhattacharya, Institute of Engineering and Management, Kolkata, India
Dr. Sachi Nandan Mohanty, KIIT University, India
Dr. Anirban Das, University of Engineering and Management, Kolkata, India
Dr. Tiya Dey Malakar, RCC Institute of Information Technology, Kolkata, India
Dr. Arpita Ghosh, RCC Institute of Information Technology, Kolkata, India
Dr. Papia Datta, RCC Institute of Information Technology, Kolkata, India
Dr. Biswarup Neogi, JIS College of Engineering, India
Dr. Tanupriya Chourdhury, University of Petroleum and Energy Studies, India
Dr. Swarnendu Chakraborty, National Institute of Technology, Arunachal Pradesh, India
Dr. Anilesh Dey, Narula Institute of Technology, Kolkata, India
Dr. Soumi Dutta, Institute of Engineering and Management, Kolkata, India
Dr. Debashis Mondal, RCC Institute of Information Technology, Kolkata, India
Dr. Anup Kumar Kolya, RCC Institute of Information Technology, Kolkata, India
Dr. Sutirtha Kumar Guha, Meghnad Saha Institute of Technology, Kolkata, India
Dr. Baisakhi Das, Institute of Engineering and Management, Kolkata, India
Dr. Shiladitya Pujari, University Institute of Technology, Burdwan, India
Dr. Chandan Koner, Dr. B. C. Roy Engineering College, Durgapur, India
Dr. Tanmay Bhattacharya, Techno Main Salt Lake, India
Dr. Chandan Bhattacharyya, Sister Nivedita University, Kolkata, India
Dr. Sangita Agarwal, RCC Institute of Information Technology, Kolkata, India
Dr. Tathagata Deb, RCC Institute of Information Technology, Kolkata, India
Sanjib Saha, Dr. B. C. Roy Engineering College, Durgapur, India
Uddalak Chatterjee, BITM, Santiniketan, India
Rabi Narayan Behera, Institute of Engineering and Management, Kolkata, India
Sudipta Bhattacharya, Bengal Institute of Technology, Kolkata, India
Anjan Bandyopadhyay, Amity University, Kolkata, India
Anirban Bhar, Narula Institute of Technology, Kolkata, India
Mousumi Bhattacharyya, Sister Nivedita University, Kolkata, India
Sudeep Basu, North Bengal University, India
Mithun Roy, Siliguri Institute of Technology, Siliguri, India
Soumitra Sasmal, Techno Main Salt Lake, India

Institutional Advisory Chairs

Dr. Minakshi Banerjee, RCC Institute of Information Technology, Kolkata, India
Dr. Ashoke Mondal, RCC Institute of Information Technology, Kolkata, India
Dr. Alok Kole, RCC Institute of Information Technology, Kolkata, India
Dr. Abhishek Basu, RCC Institute of Information Technology, Kolkata, India

Dr. Soham Sarkar, RCC Institute of Information Technology, Kolkata, India
Dr. Srijan Bhattacharyya, RCC Institute of Information Technology, Kolkata, India
Dr. Arindam Mondal, RCC Institute of Information Technology, Kolkata, India
Dr. Shilpi Bhattacharya, RCC Institute of Information Technology, Kolkata, India

Institutional Program Committee

Arpan Deyasi, RCC Institute of Information Technology, Kolkata, India
Soumen Mukherjee, RCC Institute of Information Technology, Kolkata, India
Arup Kumar Bhattacharjee, RCC Institute of Information Technology, Kolkata, India
Rajib Saha, RCC Institute of Information Technology, Kolkata, India
Harinandan Tunga, RCC Institute of Information Technology, Kolkata, India
Biswanath Chakraborty, RCC Institute of Information Technology, Kolkata, India
Arijit Ghosh, RCC Institute of Information Technology, Kolkata, India
Nitai Banerjee, RCC Institute of Information Technology, Kolkata, India
Alokananda De, RCC Institute of Information Technology, Kolkata, India
Satabdwi Sarkar, RCC Institute of Information Technology, Kolkata, India
Pampa Debnath, RCC Institute of Information Technology, Kolkata, India
Naiwrita Dey, RCC Institute of Information Technology, Kolkata, India
Nandan Bhattacharyya, RCC Institute of Information Technology, Kolkata, India
Budhaditya Biswas, RCC Institute of Information Technology, Kolkata, India
Kalyan Biswas, RCC Institute of Information Technology, Kolkata, India
Avishek Paul, RCC Institute of Information Technology, Kolkata, India
Anindya Basu, RCC Institute of Information Technology, Kolkata, India
Sk. Mazharul Islam, RCC Institute of Information Technology, Kolkata, India
Koushik Mallick, RCC Institute of Information Technology, Kolkata, India
Parama Bagchi, RCC Institute of Information Technology, Kolkata, India
Priya Sen Purkait, RCC Institute of Information Technology, Kolkata, India
Satarupa Chatterjee, RCC Institute of Information Technology, Kolkata, India
Dr. Joyeeta Basu Pal, RCC Institute of Information Technology, Kolkata, India
Moumita Banerjee, RCC Institute of Information Technology, Kolkata, India
Deepam Ganguly, RCC Institute of Information Technology, Kolkata, India
Subhrajit Sinha Roy, RCC Institute of Information Technology, Kolkata, India
Sarbojit Mukherjee, RCC Institute of Information Technology, Kolkata, India
Nijam Ud-Din Molla, RCC Institute of Information Technology, Kolkata, India

Organizing Chairs

Hrishikesh Bhaumik, RCC Institute of Information Technology, Kolkata, India
Abantika Choudhury, RCC Institute of Information Technology, Kolkata, India
Ranjan Jana, RCC Institute of Information Technology, Kolkata, India

Jayanta Datta, RCC Institute of Information Technology, Kolkata, India
Hiranmoy Roy, RCC Institute of Information Technology, Kolkata, India
Soumyadip Dhar, RCC Institute of Information Technology, Kolkata, India
Amit Khan, RCC Institute of Information Technology, Kolkata, India
Pankaj Pal, RCC Institute of Information Technology, Kolkata, India
Sudarsan Biswas, RCC Institute of Information Technology, Kolkata, India
Shaswati Roy, RCC Institute of Information Technology, Kolkata, India

Preface

Artificial intelligence is a leading theory of computer science now. Scientists across all engineering disciplines are interested in this concept. Artificial intelligence has opened the door for all research enthusiasts and application developers to propose new-age research and application concepts. The future of artificial intelligence enabled research and application is very promising. This book will discuss the recent research trends and upcoming applications based on artificial intelligence. Many of the versatile fields of artificial intelligence will be categorically addressed in different chapters of this volume.

Over the years scientists have developed several efficient algorithms to address different real-world problems realistically and propose meaningful solution for them. However, the classical problem-solving algorithms often fall short of offering a robust solution to handle the multiple constraints encountered in real-life situations since these core methods are often uncertain and imprecise, which remain intractable to process in practice using the conventional classical methods.

Moreover, with the progress of technology, the need for advanced computational techniques is always called for addressing the complex real-life problems. The objective of such computational paradigm is to give rise to fail-safe and robust solutions to the emerging problems faced by mankind. Imparting intelligence in a machine is the need of the hour. Several intelligent techniques have been in vogue over the year in this direction. Among these techniques, the soft computing techniques stand in good stead. However, it is often noted that the soft computing techniques often fall short in offering a formidable solution. On and above, if the different components of the soft computing paradigm are conjoined together, the resultant hybrid intelligent computing paradigm is found to be more efficient and robust by design and performance in these situations.

This book aims to introduce to the prospective readers the latest trends in artificial intelligence with reference to both the classical and hybrid computational paradigms.

The editors would like to take this opportunity to express their heartfelt regards to the Management of RCC Institute of Information Technology, Kolkata, and all the

committee members of RAAI 2020, especially the technical committee members who have critically reviewed all the articles. Special thanks to Mr. Aninda Bose, Senior Editor, Springer, India, for his constant support and guidance during this book project tenure.

Kolkata, India
January 2021

Indrajit Pan
Anirban Mukherjee
Vincenzo Piuri

Contents

Editors and Contributors

About the Editors

Indrajit Pan received his B.E. in Computer Science and Engineering with Honors from The University of Burdwan (2005) and M.Tech. in Information Technology from Bengal Engineering and Science University, Shibpur (2009). He was the recipient of University Medal in his Masters. He obtained Ph.D. in Engineering from Indian Institute of Engineering Science and Technology, Shibpur, in 2015. His current research interest includes community detection, cloud computing, influence theory, and digital microfluidic biochip. He joined RCC Institute of Information Technology in 2006 and is now Associate Professor and Head of the Information Technology Department. He has research publications in different international journals, edited books, and conference proceedings. He has also co-authored some edited research volumes and international conference proceedings. He served as the guest editor in *International Journal of Hybrid Intelligence*, Inderscience, and is currently an editorial board member of Elsevier's *Applied Soft Computing Journal*. He is now a senior member of IEEE, USA, and a member of ACM, USA.

Dr. Anirban Mukherjee did his Bachelors in Civil Engineering in 1994 from Jadavpur University, Kolkata. While in service, he achieved a professional Diploma in Operations Management (PGDOM) in 1998 and completed his Ph.D. on 'Automatic Diagram Drawing based on Natural Language Text Understanding' from Indian Institute of Engineering, Science and Technology (IIEST), Shibpur, in 2014. Serving RCC Institute of Information Technology (RCCIIT), Kolkata, since inception (in 1999), he is currently Professor in the Department of Information Technology. Before joining RCCIIT, he served as Engineer in the Scientific and Technical Application Group in erstwhile RCC, Calcutta, for 6 years. He has several international journal and conference papers to his credit. He is the co-editor of one IGI Global Book and also the co-editor of two books of CRC Press and Willey. He has reviewed several journal papers and book chapters and acted as the guest editor of *Journal of Pattern Recognition Research* and *International Journal of Computers and Applications.*

Vincenzo Piuri has received his Ph.D. in Computer Engineering at Politecnico di Milano, Italy (1989). He is Full Professor in Computer Engineering at the Università degli Studi di Milano, Italy (since 2000). He has been Associate Professor at Politecnico di Milano, Italy, and Visiting Professor at the University of Texas at Austin and at George Mason University, USA. His main research interests are artificial intelligence, computational intelligence, intelligent systems, machine learning, pattern analysis and recognition, signal and image processing, biometrics, intelligent measurement systems, industrial applications, digital processing architectures, fault tolerance, and cloud computing infrastructures. Original results have been published in 400+ papers in international journals, proceedings of international conferences, books, and book chapters. He is Fellow of the IEEE, Distinguished Scientist of ACM and a senior member of INNS. He is President of the IEEE Systems Council (2020–2021) and has been IEEE Vice President for Technical Activities (2015), IEEE Director, President of the IEEE Computational Intelligence Society, Vice President for Education of the IEEE Biometrics Council, Vice President for Publications of the IEEE Instrumentation and Measurement Society and the IEEE Systems Council, and Vice President for Membership of the IEEE Computational Intelligence Society.

Contributors

Abdultaofeek Abayomi Department of Computer Science, University of Uyo, Uyo, Akwa Ibom State, Nigeria

Alimot Abdulazeez Department of Computer Science, University of Uyo, Uyo, Akwa Ibom State, Nigeria

Usha Aiyer Great Learning, Bangalore, India

Shamama Anwar Department of Computer Science and Engineering, Birla Institute of Technology, Mesra, Ranchi, India

Parama Bagchi RCC Institute of Information Technology, Kolkata, India

Debasrita Baidya Department of Geography, Kazi Nazrul University, Asansol, India

Abhishek Basu Electronics and Communication Engineering, RCC Institute of Information Technology, Kolkata, India

Soham Basu Department of Electronics and Communication Engineering, Heritage Institute of Technology, Kolkata, West Bengal, India

Subhash Bhamu Great Learning, Bangalore, India

Debotosh Bhattacharjee Jadavpur University, Kolkata, India

Vandana Bhattacharjee Department of Computer Science and Engineering, Birla Institute of Technology, Mesra, Ranchi, India

Ankit Bhattacharya Tata Consultancy Services, Kolkata, West Bengal, India

Indradeep Bhattacharya University Institute of Technology, The University of Burdwan, Bardhaman, West Bengal, India

Bhagyashri Bhosale Great Learning, Bangalore, India

Sudarsan Biswas Department of Information Technology, RCC Institute of Information Technology, Kolkata, W.B, India

Priscila V. S. Z. Capriles Federal University of Juiz de Fora, Juiz de Fora, Brazil

Ruan M. Carvalho Computational Modeling Program, Federal University of Juiz de Fora (UFJF), São Pedro, Juiz de Fora, Minas Gerais, Brazil

Susmita Chaki University of Engineering & Management, Kolkata, India

Arunava Kumar Chakraborty Department of Computer Science & Engineering, RCC Institute of Information Technology, Kolkata, India

Sujit Chakraborty Department of Information Technology, RCC Institute of Information Technology, Kolkata, W.B, India

Prapty Chanda RCC Institute of Information Technology, Kolkata, West Bengal, India

Godavarthi Charan School of Electronics Engineering, VIT-AP University, Amaravati, AP, India

Subhajit Chatterjee University of Engineering & Management, Kolkata, India

Avik Chattopadhyay Radio Physics and Electronics, University of Calcutta, Kolkata, India

Aditi Roy Chowdhury Women's Polytechnic, Kolkata, India

Fernando A. B. Colugnati Federal University of Juiz de Fora, Juiz de Fora, Brazil

Leonardo G. da Fonseca Computational Modeling Program, Federal University of Juiz de Fora, Juiz de Fora, MG, Brazil

Narayana Darapaneni Northwestern University/Great Learning, Evanston, USA

Nilormi Das RCC Institute of Information Technology, Kolkata, West Bengal, India

Sourav Das Maulana Abul Kalam Azad University of Technology, WB, Kolkata, India

Srirupa Das RCC Institute of Information Technology, Kolkata, India

Kousik Dasgupta Kalyani Government Engineering College, Kalyani, Nadia, India

Sourav Debnath Electrical Engineering Department, Camellia Institute of Technology, Kolkata, W.B, India

Arkadeep Dey Department of Computer Science and Engineering, RCC Institute of Information Technology, Kolkata, West Bengal, India

Soumyadip Dhar RCC Institute of Information Technology, Kolkata, India

Paramartha Dutta Visvabharati University, Santiniketan, India

Bishal Ghosh RCC Institute of Information Technology, Kolkata, India

Chanchal Ghosh Department of MCA, Calcutta Institute of Technology, Uluberia, Howrah, India

Leonardo Goliatt Computational Modeling, Federal University of Juiz de Fora (UFJF), São Pedro, Juiz de Fora, Minas Gerais, Brazil

Priscila V. Z. C. Goliatt Computational Modeling, Federal University of Juiz de Fora (UFJF), São Pedro, Juiz de Fora, Minas Gerais, Brazil

Diego E. B. Gomes Computational Modeling, Federal University of Juiz de Fora (UFJF), São Pedro, Juiz de Fora, Minas Gerais, Brazil

Viji Gopal School of Engineering, Cochin University of Science and Technology, Cochin, Kerala, India

Sutirtha Kumar Guha Meghnad Saha Institute of Technology, Kolkata, West Bengal, India

Shibakali Gupta University Institute of Technology, The University of Burdwan, Bardhaman, West Bengal, India

Sayak Haldar Meghnad Saha Institute of Technology, Kolkata, West Bengal, India

Joydev Hazra Heritage Institute of Technology, Kolkata, India

Molla Rashied Hussein University of Asia Pacific, Dhaka, Bangladesh

Gabriele C. Iwashima Federal University of Juiz de Fora, Juiz de Fora, Brazil

Dishani Kar RCC Institute of Information Technology, Kolkata, West Bengal, India

Reshma Kar Artificial Intelligence Laboratory, ETCE Department, Jadavpur University, Kolkata, India

Amit Khan Department of IT, RCC Institute of Information Technology, Kolkata, India

Anup Kumar Kolya Department of Computer Science & Engineering, RCC Institute of Information Technology, Kolkata, India

Nandini Kumari Department of Computer Science and Engineering, Birla Institute of Technology, Mesra, Ranchi, India

Dipankar Majumdar Department of CSE, RCC Institute of Information Technology, Kolkata, West Bengal, India

Portret Mallick Department of Computer Science and Engineering, RCC Institute of Information Technology, Kolkata, West Bengal, India

Surajit Mandal Department of ECE, B.P. Poddar Institute of Management & Technology, Kolkata, West Bengal, India

Indronil Mazumder ECE Department, RCC Institute of Information Technology, Kolkata, India

Shreyas Mishra National Institute of Technology, Rourkela, India

Diptarshi Mitra Department of Geography, Kazi Nazrul University, Asansol, India;
Salt Lake City, Kolkata, India

Arindam Mondal RCC Institute of Information Technology, Kolkata, West Bengal, India

Arpita Mondal Department of Geography, Kazi Nazrul University, Asansol, India

Bikromadittya Mondal Department of CSE, B. P. Poddar Institute of Management and Technology, Kolkata, India

Anirban Mukherjee RCC Institute of Information Technology, Kolkata, West Bengal, India

Sayantan Mukherjee Tata Consultancy Services, Kolkata, West Bengal, India

Mita Nasipuri Jadavpur University, Kolkata, India

Indrajit Pan RCC Institute of Information Technology, Kolkata, West Bengal, India

Chandra Patro RCC Institute of Information Technology, Kolkata, West Bengal, India

Varghese Paul School of Engineering, Cochin University of Science and Technology, Cochin, Kerala, India

Y. V. Pavan Kumar School of Electronics Engineering, VIT-AP University, Amaravati, AP, India

Egberto Pereira State University of Rio de Janeiro, Rio de Janeiro, RJ, Brazil

Sourav Pramanik New Alipore College, Kolkata, India

Avinash Rai Department of ECE, UIT (RGPV), Bhopal, M.P., India

K. Sandeep Rao School of Electronics Engineering, VIT-AP University, Amaravati, AP, India

Ankit Rastogi Great Learning, Bangalore, India

Anwesh Reddy Paduri Great Learning, Bangalore, India

Iago G. L. Rosa Computational Modeling Program, Federal University of Juiz de Fora (UFJF), São Pedro, Juiz de Fora, Minas Gerais, Brazil

Hiranmoy Roy RCC Institute of Information Technology, Kolkata, India

Pritam Kumar Roy Meghnad Saha Institute of Technology, Kolkata, West Bengal, India

Subhrajit Sinha Roy Electronics and Communication Engineering, RCC Institute of Information Technology, Kolkata, India;
Radio Physics and Electronics, University of Calcutta, Kolkata, India

Rajib Saha RCC Institute of Information Technology, Kolkata, India

Dasa Sampath School of Electronics Engineering, VIT-AP University, Amaravati, AP, India

Soumen Santra Department of MCA, Techno International New Town, Kolkata, West Bengal, India

Camila M. Saporetti State University of Minas Gerais, Divinópolis, MG, Brazil

Abhijit Sarkar Department of Geography, Kazi Nazrul University, Asansol, India

João P. Scoralick Federal University of Juiz de Fora, Juiz de Fora, Brazil

Anindya Sen Department of Electronics and Communication Engineering, Heritage Institute of Technology, Kolkata, West Bengal, India

Abu Salman Shaikat World University of Bangladesh, Dhaka, Bangladesh

Yash Singla Manav Rachna International Institute of Research and Studies, Faridabad, Haryana, India

Akanksha Soni Department of ECE, UIT (RGPV), Bhopal, M.P., India

Turyansu Subhadarshy Great Learning, Bangalore, India

Rumana Tasnim World University of Bangladesh, Dhaka, Bangladesh

Harinandan Tunga Department of Computer Science and Engineering, RCC Institute of Information Technology, Kolkata, West Bengal, India

Samuel Udoh Department of Information and Communication Technology, Mangosuthu University of Technology, Durban, South Africa

Uduak Umoh Department of Computer Science, University of Uyo, Uyo, Akwa Ibom State, Nigeria

Prediction and Analysis on COVID-19 Using Positive and Negative Association Rule Mining

Sujit Chakraborty, Sudarsan Biswas, and Sourav Debnath

Abstract The enormous complicated heterogeneous data can be processed easily using data mining. An attempt has been made to generate a pattern for COVID-19 disease, which can be beneficiated to detect as well as treat the affected patients. Consideration has been made to analyze and predict the most common as well as rare and hidden symptoms via applying both positive (interesting rules) and negative (uninteresting rule) data mining association rules. Thus, study has made on both frequent and infrequent itemsets of affected patients with remembering the risk levels at the pandemic situation worldwide. The extracted frequent and infrequent itemsets assist the medical professionals to make diagnostic recommendations and determine the riskiness of patients at an initial stage as test report has generated after few days of sample collection.

Keywords Data mining · COVID-19 prediction · Medical data · Frequent itemsets · Association rule mining

S. Chakraborty (✉) · S. Biswas
Department of Information Technology, RCC Institute of Information Technology, Kolkata 700015, W.B, India
e-mail: sujitchakraborty385@gmail.com

S. Debnath
Electrical Engineering Department, Camellia Institute of Technology, Kolkata 700129, W.B, India

I. Pan et al. (eds.), *Proceedings of Research and Applications in Artificial Intelligence*, Advances in Intelligent Systems and Computing 1355,
https://doi.org/10.1007/978-981-16-1543-6_1

1 Introduction

In today's world, COVID-19 is the preeminent human killer pandemic. In India, the cause of death from COVID-19 is co-morbidity in the maximum cases [1]. It claims millions of lives globally and causes thousands of diseases in India until these days. Common symptoms of COVID-19 are fever, tiredness and dry cough [2]. Other symptoms that are less familiar and may affect some patients include sore throat, aches and pains, headache, nasal congestion, loss of taste or smell, conjunctivitis, diarrhea or a rash on skin [3]. Some people who are COVID-19 positive only have very mild symptoms.

Data mining is being used in the clinical field. "Association rule mining" is the most efficient data mining approach for extracting frequent itemsets from enormous datasets. The foremost objective of association rule mining is to bring out frequent itemsets from transactional database. However, "association rule mining" algorithms neglect many valuable infrequent itemsets. These frequent itemsets with low support can give rise to important negative association rules (high confidences). The problems aimed in this paper are extracting positive(+ve) and negative(-ve) association rules from the "frequent" and "infrequent" itemsets. Application of negative rule mining along with positive rule mining in medical field is to extract rare symptoms of a particular disease.

Most research articles that have written on frequent itemsets from medical data to analyses HIV [4, 5], heart disease [6–8], cancer and tumor [9, 10], diabetes mellitus [11, 12] using apriori algorithm [13]. Little researcher work on negative association rule to find out rare features on various diseases [14–20]. There is very rare work on COVID-19 using positive(+ve) and negative(−ve) association rules, which is discussed with experiment in this paper.

2 Methodology

This proposed method focuses on both positive(+ve) and negative(−ve) association rules from frequent and infrequent itemsets with considering sample text dataset of COVID-19 symptoms. Here, "minimum support" and "minimum confidence" values are taken from pre-processed dataset to generate frequent itemsets. Finally, positive(+ve) and negative(−ve) association rules are considered from "frequent" and "infrequent" itemsets utilizing lift is the last step as elaborated in Fig. 1.

2.1 Data Source and COVID-19 Symptoms

Symptoms are simulated for COVID-19 dataset containing sample records of 1000 patients [21–24, 19, 20] have tabulated in Table 1.

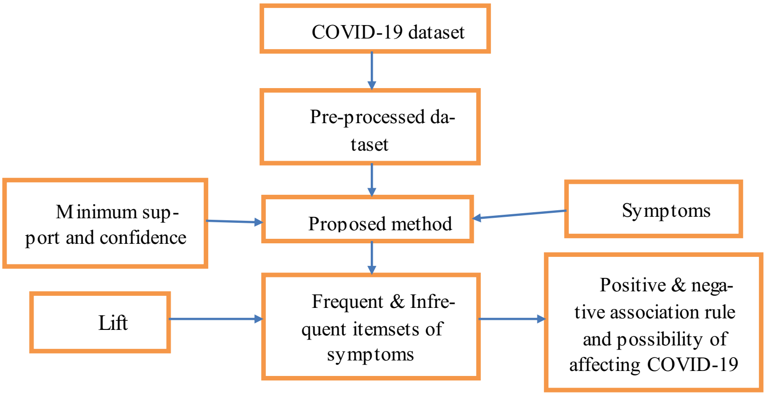

Fig. 1 Block diagram of methodology

Table 1 Symptoms for COVID-19 affected patients

Symptoms number	Symptom name	Symptoms number	Symptom name
1	Shortness of breadth	10	Cough
2	Sneezing	11	Chilblains
3	Throat pain	12	Dizziness
4	Diarrhea	13	Fatigue
5	Vomiting	14	Fever
6	Chest pain	15	Body pain
7	Nausea	16	Changes in heart rhythm
8	Headache	17	Skin rashes or blood clots
9	Conjunctivitis	18	Loss of taste or smell

2.2 Algorithms

All four algorithms are noted below.

2.2.1 Algorithm-1

Input: TRD- Covid-19 dataset; Min_support- "Minimum support" value; Min_confi- "Minimum confidence"
Output: FREQ: "frequent" itemsets; inFREQ: "infrequent" itemsets;
(1) initialization of FREQ = Φ; inFREQ = Φ;
(2) tem_1=∀ P | P ∈ Frequent symptoms
(3) $FREQ_1$ = {P | P ∈ tem_1 and support (P) ≥ min_support};
(4) $inFREQ_1 = tem_1 - FREQ_1$;
z = 2;
(5) while (tem_{z-1} !=Φ) do started
D_z = generation of (tem_{z-1}, min_support);
for individual transaction tran ∈ TRD
do started
D_{tran} = subset (D_z, tran);
for individual candidate d ∈ D_{tran}
d.counter^{++};
end of loop;
d.supp = (d.counter/|TRD|);
tem_z= {d | d ∈ d_z and (d.supp ≥ minsupport)};
(6) $FREQ_z$ = {P | P ∈ tem_z and P. supp ≥ min_support)};
(7) $inFREQ_z = tem_z - FREQ_z$;
(8) FREQ = $\cup_z$ $FREQ_z$;
(9) inFREQ = $\cup_z$ $inFREQ_z$;
(10) z^{++};
end;
(14) return FREQ and inFREQ;

Description of Algorithm-1

The present algorithm, "patients" are expressed as "Transaction" and corresponding "symptoms" are represented with "itemsets." Support is the indication of item that represents frequency of occurrence (itemsets). First of all, "FREQ" and "inFREQ" are initialized with null value. All itemsets having 1 candidate (symptom) is added to "tem_1".frequent symptoms with support count is greater than least support are assigned to "$FREQ_1$". "$FREQ_1$" is subtracted from "tem_1" to get infrequent itemsets having 1 candidate, which are assigned to "$inFREQ_1$" and "z" is initialized by 2 (>1) where z is the size of itemsets (the number of symptoms for a person). While the value of "$temp_{z-1}$" is not equal to null value, "$temp_{z-1}$" is generated with minimum support and is stored in "D_Z", which is candidate z-itemsets. Now database "TRD" is scanned for individual patient, which is represented by "trans"; "$temp_1$" candidates are generated in transaction "trans", now if itemsets exist in transaction then it is increased by 1.support of z number of itemset is calculated. An item d (symptom) is stored in $temp_z$ where support of symptom d is greater than or equal to minimum support. Frequent symptoms having support count greater than or equal to threshold support are added to "FREQz". Infrequent itemsets are added to "inFREQz" by subtracting FREQz from $temp_z$ where support is smaller than least support. Now generated frequent symptoms are added to FREQ and infrequent itemsets (symptoms) are added to "inFREQ" with z number of items. Then itemset size is incremented by 1. Return the value of "FREQ" and "inFREQ".

2.2.2 Algorithm 2

Input: minisupport: "minimum support"; minconfi: "minimum confidence"; FREQ (frequent itemsets); inFREQ (infrequent itemsets)

Output: POSR: Positive(+ve) Association Rules; NEGR: Negative(-ve) Association Rules;

(1) POSR = φ;NEGR = φ;
(2) For individual itemset S in FREQ
  do started
    for individual itemset $S_1 \cup S_2 = S, S_1 \cap S_2 = \varphi$
    do started
(3)   if confidence ($S_1 => S_2$) ≥ minconfi && lift ($S_1 => S_2$) ≥ 1
      then output of the rule ($S_1 => S_2$); POSR ∪ ($S_1 => S_2$)
    else
(4)   if confidence ($S_1 => \sim S_2$) ≥ minconfi && lift ($S_1 => \sim S_2$) ≥ 1
      output of the rule ($S_1 => \sim S_2$);NEGR ∪ ($S_1 => \sim S_2$)
    if confidence ($\sim S_1 => S_2$) ≥ minconfi && lift ($\sim S_1 => S_2$) ≥ 1
      output of the rule ($\sim S_1 => S_2$);NEGR ∪ ($\sim S_1 => S_2$)
    if confidence ($\sim S_1 => \sim S_2$) ≥ minconfi && lift ($\sim S_1 => \sim S_2$) ≥ 1
      output of the rule ($\sim S_1 => \sim S_2$);NEGR ∪ ($\sim S_1 => \sim S_2$)
    end of loop;
  end of loop ;
(5) For any itemset S in inFREQ
  do started
    For every itemset $S_1 \cup S_2 = S, S_1 \cap S_2 = \varphi$, support($S_1$) ≥ minisupport and support(S_2) ≥ minisupport
    Do started
(6)   If confidence($S_1 => S_2$) ≥ minconfi && lift($S_1 => S_2$) ≥ 1
      then output of the rule $S_1 => S_2$;
      POSR ∪ ($S_1 => S_2$)
    else
    if confidence ($S_1 => \sim S_2$) ≥ minconfi && lift ($S_1 => \sim S_2$) ≥ 1
      output of the rule ($S_1 => \sim S_2$);NEGR ∪ ($S_1 => \sim S_2$)
    if confidence ($\sim S_1 => S_2$) ≥ minconfi && lift ($\sim S_1 => S_2$) ≥ 1
      output of the rule ($\sim S_1 => S_2$);NEGR ∪ ($\sim S_1 => S_2$)
    if confidence ($\sim S_1 => \sim S_2$) ≥ minconfi && lift ($\sim S_1 => \sim S_2$) ≥ 1
      output of the rule ($\sim S_1 => \sim S_2$);NEGR ∪ ($\sim S_1 => \sim S_2$)
    end of loop;
  end of loop;
(7) return POSR and NEGR;

Description of Algorithm 2

"Confidence" is an indication of how often the rule has been found to be true and "Lift" is measurement of interestingness of a rule. At first, "PQSR" and "NEGR" are initialized by null value. After that, association rules are generated from FREQ (frequent itemsets). For an individual patient S, all symptoms belong to "FREQ". If S_1 and S_2 are any two symptoms of patient S, rules are generated for S_1 implies to S_2(S_1=>S_2). If confidence and lift of (S_1=>S_2) are higher than or equal to minimum value of confidence and value 1 respectively, then output is (+ve) rule. Else, if confidence and lift of (S_1=>~S_2) are more than or equal to least confidence and value 1 respectively, then output is (−ve) rule. If confidence and lift of (~S_1 => S_2) are higher than or equal to least confidence and value 1 respectively, then output is (−ve) rule. If confidence and lift of (S_1 =>~S_2) are higher than or equal to minimum value of confidence and value 1 respectively, the output belongs to (−ve) rule. If confidence and lift of (~S_1 =>~S_2) are more than or equal to least confidence and value 1 respectively, then output is been (−ve) rule. Now association rules are generated from "inFREQ" or infrequent itemsets where an individual patient S and corresponding symptoms are S_1 and S_2. Mentioned support S_1 and S_2 are higher than base support. Rules are generated for (S_1=>S_2). If confidence and lift of (S_1=> S_2) are more than or equal to minimum threshold of confidence and value 1 respectively, then output is (+ve) rule. Else, if confidence and lift of (S_1 => ~S_2) are higher than or equal to minimum value of confidence and value 1 then output is (−ve) rule. If confidence and lift of (~S_1 => S_2) are more than or equals to minimum threshold of confidence and value 1 respectively, then output is (−ve) rule. If confidence and lift of (~S_1 => ~S_2) are higher than or equal to minimum confidence and value 1 respectively, then output is (−ve) rule. Return the value of "PQSR" and "NEGR".

2.2.3 Algorithm 3

Given: support($S_1 \cup S_2$) ≥ minisupport

(1)if confidence (S_1 => S_2) ≥ minconfi, and lift (S_1 => S_2) > 1
    then S_1 => S_2 is a effective positive rule, S_1 and S_2 are positively correlated with minimum confidence.
  else

(2) if confidence (S_1 => S_2) < minconfi, and lift (S_1 => S_2) < 1
    then S_1 => S_2 is not a effective positive rule, S_1 and S_2 are negatively correlated with lower than minimum confidence.
  Hence, Negative association rules are generated from itemset S.

(3)if confidence (S_1 => ~S_2) ≥ minconfi, and lift (S_1 ⇒ ~S_2) > 1
    then S_1 => ~S_2 is a effective negative rule, S_1 and S_2 are positively correlated with minimum confidence.

Description of Algorithm 3

Support of ($S_1 \cup S_2$) is more than or equal to minimum support. If confidence of (S_1 => S_2) higher than or equal to minimum value of confidence, and lift of (S_1 => S_2) greater than 1 then S_1 => S_2 is an effective positive rule, S_1 and S_2 are positively correlated with minimum confidence. Else, if confidence and lift of (S_1 => S_2) less than minimum confidence and value 1 respectively, then S_1 => S_2 is not an effective positive rule, S_1 and S_2 are negatively correlated with lower than minimum confidence. Hence, negative association rules are generated from itemset S. If confidence of (S_1 => ~S_2) ≥ minimum value of confidence, and lift of (S_1 = > ~S_2) greater than 1 then S_1 => ~S_2 is an effective negative rule, S_1 and S_2 are positively correlated with minimum confidence.

2.2.4 Algorithm 4

Given: support($S_1 \cup S_2$)< minisupport, and support($S_1 \cup S_2$) != 0
support(S_1) ≥ minisupport, and support(S_2) ≥ minisupport,
(1)if confidence($S_1 \Rightarrow S_2$) ≥ minconfi, and lift($S_1 \Rightarrow S_2$) > 1
then $S_1 \Rightarrow S_2$ is a effective positive rule, X and Y is positively correlated with minimum confidence
else
(2)if confidence ($S_1 \Rightarrow$ ~S_2) ≥ minconfi and lift($S_1 \Rightarrow$ ~S_2) > 1
then $S_1 \Rightarrow$ ~S_2 is a effective negative rule, S_1 and ~S_2 is positively correlated with minimum confidence.

Description of algorithm- 4

Here, Support of ($S_1 \cup S_2$) less than minimum support, and support of ($S_1 \cup$ S2) not equal to 0, minimum support of (S_1) higher than or equal to least support, and support of (S_2) more than equal to least support. If confidence of ($S_1 \Rightarrow S_2$) more than or equal to minimum threshold of confidence, and lift of ($S_1 \Rightarrow S_2$) more than 1 then $S_1 \Rightarrow S_2$ is an effective positive rule, S_1 and S_2 is positively correlated with minimum confidence. Else, if confidence of ($S_1 \Rightarrow$ ~S_2) more than or equal to minimum value of confidence and lift of ($S_1 \Rightarrow$ ~S_2) greater than 1 then $S_1 \Rightarrow$ ~S_2 is an effective negative rule, S_1 and ~S_2 is positively correlated with minimum confidence.

3 Result and Discussions

The generated itemsets are summarized in Table 2. It has been observed that frequent itemsets are decreased with improving the minsup value. A gradual increase of the infrequent itemsets has also noticed as shown in Fig. 2. All the four types of association rules have implemented and corresponding results are recorded in Tables 3, 4, 5 and 6. These results can also be visualized with Figs. 3, 4, 5 and 6 respectively. Prediction can be made easily for the risk factor of a COVID-19 patient from these datasets. Here, {Cough ⇒ Shortness of breath} means "Shortness of breath" is experienced by a patient who is suffering from "Cough" and { ~Throat pain-Headache ⇒ ~Cough} means if "Headache" experienced by a patient is not having "Throat pain", then he may not have "Cough" with high confidence.

Table 2 Generated numbers of frequent and infrequent items

"Support"	"Frequent" itemsets	"Infrequent" itemsets
0.1	14	21
0.15	10	25
0.25	8	27
0.3	6	29
0.4	4	31

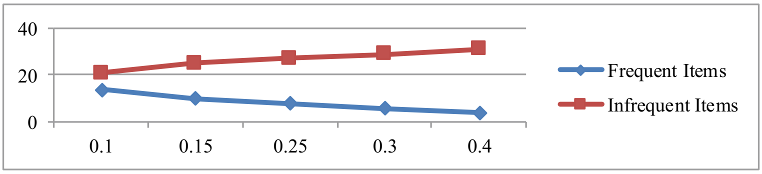

Fig. 2 Graphical representation of generated numbers of frequent and infrequent items

Table 3 [+ve] rules from frequent itemsets

Rules	"Support"	"Confidence"	"Lift"
{cough}{shortness of breath}	0.41	0.63	1.07
{shortness of breath}{fever}	0.39	0.67	1.15
{cough} {fever}	0.39	0.63	1.08
{body pain} {headache}	0.31	0.58	1.18
{cough} {body pain}	0.37	0.69	1.13