Advances in Sustainability Science and Technology

Series Editors

Robert J. Howlett, Bournemouth University & KES International, Shoreham-by-sea, UK

John Littlewood, School of Art & Design, Cardiff Metropolitan University, Cardiff, UK

Lakhmi C. Jain, University of Technology Sydney, Broadway, NSW, Australia

The book series aims at bringing together valuable and novel scientific contributions that address the critical issues of renewable energy, sustainable building, sustainable manufacturing, and other sustainability science and technology topics that have an impact in this diverse and fast-changing research community in academia and industry.

The areas to be covered are

- Climate change and mitigation, atmospheric carbon reduction, global warming
- Sustainability science, sustainability technologies
- Sustainable building technologies
- Intelligent buildings
- Sustainable energy generation
- Combined heat and power and district heating systems
- Control and optimization of renewable energy systems
- Smart grids and micro grids, local energy markets
- Smart cities, smart buildings, smart districts, smart countryside
- Energy and environmental assessment in buildings and cities
- Sustainable design, innovation and services
- Sustainable manufacturing processes and technology
- Sustainable manufacturing systems and enterprises
- Decision support for sustainability
- Micro/nanomachining, microelectromechanical machines (MEMS)
- Sustainable transport, smart vehicles and smart roads
- Information technology and artificial intelligence applied to sustainability
- Big data and data analytics applied to sustainability
- Sustainable food production, sustainable horticulture and agriculture
- Sustainability of air, water and other natural resources
- Sustainability policy, shaping the future, the triple bottom line, the circular economy

High quality content is an essential feature for all book proposals accepted for the series. It is expected that editors of all accepted volumes will ensure that contributions are subjected to an appropriate level of reviewing process and adhere to KES quality principles.

The series will include monographs, edited volumes, and selected proceedings.

More information about this series at http://www.springer.com/series/16477

Neha Sharma · Santanu Ghosh ·
Monodeep Saha

Open Data for Sustainable Community

Glocalized Sustainable Development Goals

Springer

Neha Sharma
Analytics and Insights
Tata Consultancy Services Ltd.
Pune, Maharashtra, India

Santanu Ghosh
Analytics and Insights
Tata Consultancy Services Ltd.
Kolkata, West Bengal, India

Monodeep Saha
Analytics and Insights
Tata Consultancy Services Ltd.
Bengaluru, Karnataka, India

ISSN 2662-6829 ISSN 2662-6837 (electronic)
Advances in Sustainability Science and Technology
ISBN 978-981-33-4311-5 ISBN 978-981-33-4312-2 (eBook)
https://doi.org/10.1007/978-981-33-4312-2

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

Foreword

I am pleased to write this foreword, because of my close association with the themes of 'Data Centricity' and 'Sustainability Ecosystems' that is relevant for businesses, communities and individuals. In my current role as the Global Head of Analytics and Insights, Tata Consultancy Services, I come across various customer execs, policy makers, government officials and industry leaders who are aspiring to make their organizations more purpose driven and adaptive, attain business growth but at the same time make a strong positive contribution towards environment and communities. My association with industry bodies, academia and strategic think tanks has led me to a view that growth, development and transformations are only meaningful if they are sustainable. The power of data, analytics and AI and the related digital technologies play a major role in this journey which is also vetted by the guidance provided by the United Nations Sustainable Development Goals. The world is seeing democratization of technologies and data resulting in newer possibilities. Open data is one of the levers for achieving those.

In order to further the sustainability agenda, there needs to be a strong collaboration among governments, corporates, academia and citizens with resulting ecosystem capabilities. This publication stands at the intersection of such an ecosystem and an excellent compilation of research projects based on open data with outcomes having a bearing on the community. The next generation is a very important stakeholder of this ecosystem as they would foster innovation to shape up a better tomorrow. Such a publication will encourage the students and academic institutions to focus on sustainability as a theme and contribute to the ecosystem. Each of the work outlines interesting possibilities which will be of interest to business enterprises.

Open Data for Sustainable Community is a result of joint interests between the students, data science communities and the TCS Analytics and Insights team. They have been instrumental in terms of topic selection, open data resources and researcher orientation. It explores and presents a part of the open datasets from government institutions to achieve the sustainable objectives at local level, in turn contributing towards global mission. Reading through the book, you will find some

of the specific issues in the areas of environment, Indian agriculture and health care seen through the lens of data science, which has deep relevance in today's world.

I would encourage students, researchers and practitioners to contribute to the sustainability ecosystem and build further on the good work that has been done.

Dinanath Kholkar
Vice President and Global Head
Analytics and Insights
Tata Consultancy Services Ltd.
Pune, India

Preface

By and large, we are at the crossroad of the 4th Industrial Revolution, where phy-gital systems are going to play a massive role. This transformation is cutting across every known sphere to mankind. The world will become globally localized marketplace. COVID-19 has convoluted the entire space-time fabric, and there is a massive paradigm shift. We are looking the world through the new lenses where technological transformation via machine learning and artificial intelligence is the new norm. We are at the cusp of the future where AI/ML will be imbibed in day-to-day activities via cloud platforms. Doing business in a greener way is going to be norm for us if we intend to sustain life force on this planet. Fighting against natural calamities like drought, pandemics and pollution needs proactive intervention, clear vision, ground-level implementable and scalable technology. The transformation of thought process at an individual level will help to achieve the same. Even the policies and strategies have to be top down and the implementation needs to start bottom up and most importantly at grass-root level.

Since the last 300 years, industrial revolutions have game changing impact on societies. As our topic suggests, we are looking at some of these like health care, agriculture and environment through the lens of AI. Demographic-level analysis and GIS analysis are novel methods that are used in this field. The intent is to explore into these areas and identify cracks through which deeper in roads can be made. With the above background given, we will go in the deeper waters and explore the content of this book.

Health care—This chapter intends to dig on the Sustainable Development Goal 3. With that bigger picture as a vision, a study is conducted on COVID-19. The kind of impact coronavirus has on our society is at a massive scale and there is no geography which is untouched. There is no geography which is untouched. A cohort study on the indeterminate contagion pattern of COVID-19 is done, and effort is made to map the same with the potential features. The demand and supply sides both are mapped and identified with some of the state-of-the-art ML techniques which are used on the data to create analysis to derive insights from them. Though these are early days, these are initial steps in that direction.

Medical fraternity along with researchers are working to fast track a vaccine for the COVID-19. Nations and pharmaceutical organizations are joining forces and trying to gain a handle on situation carrying joint clinical trials. This book is a first-hand attempt to provide a consultative approach of looking at the demand and supply sides using federated data using open-source technology adoption. The data-driven thought leadership shown in the book ensures a detailed outlook. While pharmaceutical and healthcare organizations across the globe are in the process of doing deep research around application of machine learning on federated data, the opportunities are limitless. This approach of looking at catastrophic events, which has the ability to shock both demand and supply metrics, is a perfect experimental set-up and would be leveraged in future as framework for analysis. This study can further be utilized in the financial services industry to analyse the perturbation effect and how individual demographics are impacted. Utilities industries also can benefit by bringing in another angle of pricing. This base framework in the chapters provides the detail set-up of this experiment, data staging, analysing the data, steps involved, peeling various layers of data, using algorithms to derive insights. Apart from the results, the setting of the process will result in value.

Agriculture—The study is being done on farmer contact centre by Government of India queries. The intent of the study is to reduce the false positives and identify, define and create recommendations to remove these process deficiencies which can reduce time debt on the government and in turn resolve queries for the farmers faster. This solution provides and opens up the doors for usage of natural language process and use of automating. The idea is novel and can be used in setting up of a command centre which is in line with the revolution of Industry 4.0. Smart farming alarm system can be used to build early warning systems, which can map the grievances with solutions. One potential use case which can be tapped in is creating a database for common grievances and mapping them demography-wise, which when mapped with the loan data provided to farmers can become strong indicators. Those who have operated with micro-finance institutions will find this information extremely valuable. What this book aims is to create a methodology in terms of how to stage the data and convert it into a goldmine.

Environment—This section highlights the use of data democratization to identify correlations and patterns between air pollution and green cover for Pune City. Different techniques are fused together to create analysis which are technically intensive and extremely data driven to derive key insights from the data. They not only establish the age-old fact that the trees are key to the societal development but also help in forming strategy regarding where and how these plantations need to be done for it to have maximum impact. Reduction in carbon footprint is one major goal of all the corporate houses. Automobile sector and industries using fossil fuel for their energy requirements have to relook from a different lens. The Pune City is taken as a use case to highlight the fact that in case we want to develop technology solutions to take us on the road of creating smart cities, then it can not happen without a green cover. The book attempts not only using the available data but also looking at environment barometer by using open-source technology.

The three domains mentioned above are the intended social sample that weaves the social fabric by enabling interaction between individuals, technology and governance. Our study of sustainable development by means of artificial intelligence will help us understand these dynamics. The intent with which this book is written is very close to our hearts and is to provoke the thought process which draws out various possibilities in which our society becomes a better place for living. The caveat to be drawn here is that all the three frameworks in finer way are monetizable. The three domains and pillars of this book call out the essence of **process excellence**, use of **federated data** and **open-source technology**.

In the new normal, these three trends are not going to go anywhere soon.

The book will cater to the wide range of readers including professionals from sustainable development goals, social scientists, data scientists and machine learning experts.

This book is an early attempt towards that process. The book presents a prototype of thought process about taking the initial steps of harnessing data openly available and how to craft a solution. The authors of this book have come together from different walks of life with one common goal—a strong sense of commitment and a burning desire for betterment of society utilizing their technical skills. A sincere effort from the entire team which includes the authors, publishing house and the students. A special mention to our family members, friends and colleagues who kept us sane and focused during the entire journey.

Pune, India — Neha Sharma
Kolkata, India — Santanu Ghosh
Bengaluru, India — Monodeep Saha

Contents

Part I Environment—A Fact-Based Study using Tree Census and Air Pollution Data

1 Inching Towards Sustainable Smart Cities—Literature Review and Data Preparation 3
1.1 Introduction 3
1.2 Literature Review 5
1.3 Data Preparation: Tree Census Data 8
1.3.1 Understanding the Importance of Tree Census 8
1.3.2 Introduction to Tree Census Open Data 12
1.3.3 Data Profiling of Tree Census Open Data 13
1.3.4 Data Cleaning and Wrangling for Analysis 19
1.4 Data Preparation: Air Pollution Data 30
1.4.1 Understanding Air Pollution 31
1.4.2 Introduction to Air Pollution Open Data 36
1.4.3 Data Profiling of Air Pollution Open Data 37
1.4.4 Data Cleaning and Wrangling for Analysis 38
1.5 Conclusion 45
References 45

2 Exploring Air Pollution and Green Cover Dataset—A Quantitative Approach 49
2.1 Data Exploration: Tree Census 49
2.1.1 Hexbin Plot Representing Tree Density 50
2.1.2 Pie Chart Representing the Categories of Tree Condition 51
2.1.3 Violin Plots to Study Health-Category-Wise Distribution of Girth 52
2.1.4 Scatter Plot to Spot Clusters of Poor Quality and Dead Trees 54

2.1.5 Mode of All Categorical Variables 56
2.1.6 Top 10 Wards with Highest Number of Trees 57
2.1.7 In-Depth Analysis of Tree Condition and Ownership Type 57
2.1.8 Box Plot of Tree Girth and Canopy by Tree Condition and Rarity 63
2.1.9 Count Plot of Trees by Their Yield Type 65
2.1.10 Counts of Top 10 Most Commonly Occurring Trees in Pune Which Yield Timber Wood 67
2.1.11 Counts of Balanced and Unbalanced Trees 68
2.1.12 Count of Trees with Respect to the Reported Signs of Stress/Damage on the Tree 69
2.1.13 Count Plot of Trees by the Rarity of Their Occurrence 69
2.1.14 Count Plot of Trees by Their Phenology Category 70
2.1.15 Flowering Season of the Trees 70
2.1.16 Pair Plot of All the Numerical Variables 71
2.1.17 K-Means Clustering 72
2.2 Data Exploration: Air Pollution Data 77
2.2.1 Descriptive Statistics of Pollutants for the Five Locations 78
2.2.2 Visualizing Air Quality Index (AQI) 81
2.2.3 Visualizing Individual Pollutant Levels 84
2.2.4 Interrelationships Between AQI, SO_2 and NOx (in µg/m^3) Concentration 87
2.2.5 Pollutant Concentration for the Months of 2018 91
2.2.6 AQI Variation for the Months of 2018 93
2.3 Conclusion 103
References 104

3 Application of Statistical Analysis in Uncovering the Spatio-Temporal Relationships Between the Environmental Datasets 105
3.1 Introduction 105
3.2 Data Sketch for Air Pollution and Tree Census Dataset 107
3.3 Methodology to Find Correlation 109
3.3.1 Measuring Correlation 109
3.3.2 Air Quality Index 110
3.3.3 Time-Series Analysis 111
3.3.4 Haversine Formula 111
3.4 Analysis of Datasets to Find Correlation 111
3.4.1 Exploratory Data Analysis 111
3.4.2 GIS Analysis 114
3.4.3 Time-Series Analysis 119
3.4.4 Statistical Analysis 120

3.5 Results . . . 125
3.6 Conclusion and Future Work . . . 126
References . . . 127

Part II Resilient Agriculture—A War Against Hunger

4 Farmer Call Centre Literature Review and Data Preparation . . . 131
4.1 Introduction . . . 131
4.2 Literature Review . . . 134
4.3 Understanding the Operations of Kisan Call Centre . . . 140
4.4 Data Preparation: Kisan Call Centre Queries . . . 141
4.4.1 Data of Kisan Call Centre Queries . . . 142
4.4.2 Preparation of Kisan Call Centre Queries . . . 144
4.4.3 Pre-processing of Kisan Call Centre Queries . . . 146
4.5 Conclusion . . . 148
References . . . 149

5 Analysis and Visualization of Farmer Call Center Data . . . 151
5.1 Introduction . . . 151
5.2 Material Methods Used for Analysis . . . 152
5.2.1 Check and Confirm the Pre-processed Data . . . 152
5.2.2 Form an Objective and Acquire Domain Knowledge . . . 152
5.2.3 Data Visualization Criteria . . . 153
5.2.4 Libraries Used for Visualization . . . 154
5.2.5 Visualization Charts Used . . . 154
5.3 Data Exploration and Visualization . . . 155
5.3.1 Donut Pie Chart Presenting Overview of Query Types . . . 156
5.3.2 Radar Chart and Stacked Bar Graph to Analyse District-Wise Query Type . . . 157
5.3.3 Radar Chart to Present Queries According to Seasons . . . 160
5.3.4 Radar Chart and Plot Chart to Present Category-Wise Query Type . . . 163
5.4 Conclusion and Future Scope . . . 165
References . . . 167

6 An Approach for Exploring New Frontiers for Optimizing Query Volume for Farmer Call Centre—KCC Query Pattern . . . 169
6.1 Introduction . . . 169
6.2 Different Approaches for Query Text to Query Type Classification . . . 172
6.2.1 Text Similarity and Clustering . . . 172
6.2.2 Word-Based Encodings . . . 173

6.2.3 Text to Sequences . . . 175
6.2.4 Out of Vocabulary (OOVs) . . . 176
6.2.5 Padding . . . 176
6.2.6 Visualization . . . 178
6.3 Conclusions . . . 179
References . . . 180

Part III Demand and Supply Study of Healthcare Human Resource and Infrastructure—Through the Lens of COVID 19

7 Sustainable Healthcare in COVID-19 Pandemic—Literature Survey and Data Lifting . . . 183
7.1 Introduction . . . 183
7.2 Literature Review . . . 185
7.3 Data Preparation: COVID-19, Infrastructure, Human Resource, State Population Data . . . 191
7.3.1 Data Source Identification and Data Acquisition . . . 192
7.3.2 Data Profiling: COVID-19, Infrastructure, Human Resource, State Population Data . . . 195
7.3.3 Data Cleaning and Wrangling for Analysis . . . 196
7.4 Exploratory Data Analysis (EDA) . . . 201
7.5 Conclusion . . . 207
References . . . 208

8 COVID-19 and Indian Healthcare System—A Race Against Time . . . 211
8.1 Introduction . . . 211
8.2 Material Methods Used for Analysis . . . 212
8.3 Data Analysis and Visualization . . . 213
8.3.1 Progression of COVID-19 in India . . . 214
8.3.2 Healthcare Infrastructure in India . . . 225
8.3.3 Healthcare Human Resource in India . . . 242
8.4 Conclusion . . . 250
References . . . 253

9 Estimating Cases for COVID-19 in India . . . 255
9.1 Introduction . . . 255
9.2 Methodology . . . 256
9.2.1 The Database . . . 256
9.2.2 The Models . . . 257
9.3 Polynomial Regression . . . 258
9.3.1 Why Polynomial Regression Model? . . . 258
9.3.2 Polynomial Regression—Model A . . . 258
9.3.3 Polynomial Regression—Model B . . . 260

9.4 Long Short-Term Memory (LSTM) 262
9.4.1 Why LSTM? 262
9.4.2 LSTM Model 262
9.5 Autoregressive Integrated Moving Average (ARIMA) 264
9.5.1 Why Is ARIMA Preferred to Exponential Smoothing? 264
9.5.2 ARIMA Model A 265
9.5.3 ARIMA Model B 267
9.5.4 ARIMA Model C 270
9.5.5 ARIMA Model D 272
9.5.6 Validation of Models B, C and D 274
9.6 Prophet 275
9.6.1 Why Prophet? 277
9.6.2 Prophet Model 277
9.7 Conclusions 280
References 282

10 Multifacet Impact of Pandemic on Society 283
10.1 Introduction 283
10.2 Literature Review 284
10.3 Economy During the Time of Pandemic 287
10.3.1 Impact of Lockdown on Banking Sector and Insurers 288
10.3.2 Layoffs in Various Industries 289
10.3.3 Migration and Livelihood 289
10.3.4 Reverse Migration 291
10.4 Supply Chain Management 292
10.4.1 Demand of Essential Commodities 293
10.4.2 What Are Decentralized Supply Chains? 293
10.4.3 Framework of Decentralized Supply Chain 294
10.4.4 Technologies Used 295
10.4.5 Quality Assurance 296
10.4.6 Salient Features 297
10.5 Conclusion 297
References 298

Authors and Contributors

About the Authors

Neha Sharma is working with Tata Consultancy Services and is a Founder Secretary, Society for Data Science. Prior to this she has worked as Director of premier Institute of Pune, that run post-graduation courses like MCA and MBA. She is an alumnus of a premier College of Engineering and Technology, Bhubaneshwar and completed her PhD from prestigious Indian Institute of Technology, Dhanbad. She is an ACM Distinguished Speaker, a Senior IEEE member and Executive Body member of IEEE Pune Section. She is an astute academician and has organized several national and international conferences and published several research papers. She is the recipient of "Best PhD Thesis Award" and "Best Paper Presenter at International Conference Award" at National Level. She is a well-known figure among the IT circle, and well sought over for her sound knowledge and professional skills. Neha Sharma has been instrumental in integrating teaching with the current needs of the Industry and steering students towards their bright future.

Santanu Ghosh is a IT veteran with over 26 years of experience and a Sustainability Evangelist. He leads the Sustainability Ecosystem for TCS A&I unit. He is responsible for the exploration and launch of Data & Analytics-driven sustainability initiatives in the areas of Environment, Climate, Health and Biodiversity for both commercial and community enterprises. His endeavors align with UN's Sustainable Development Goals and its mapping to Enterprise and Community KPIs. He opines that data plays a critical role in creating data-driven environment-friendly solutions and creating a circular economy by creation of a knowledge ecosystem. He has been part of many fora promoting his views on Environment, Climate, Community Impacts and their relationships with the SDGs.

Mr. Monodeep Saha is currently working as Project Lead in TCS for BFSI clients. He has a rich and diverse experience of over a decade in Financial services Analytics industry, with consulting experience of major financial institutions/FinTech's in Asia pacific/US/EMEA region. Analytics/Machine Learning expert with industry expertise in Risk and Marketing Analytics space. Mr. Saha has master's in physics from IIT Madras with focus on advanced mathematics and computing. In his spare time, he spends time with family, loves to read and write. He authors short writeup at various forums covering fiction and non-fictional topics. This book is an attempt to converge his professional skills for greater good of humanity and society at large.

Contributors

Sudhanshu Bhoi is a computer engineering student who loves to think through problems and build creative solutions and an enthusiast in the field of AI; currently, he is studying machine learning and applying models to create a better world.

Sanket Bijawe is an undergraduate student at VIT, Pune, and an Exe-com member at Student Branch, VIT, Pune. He is interested in the field of deep learning and NLP.

Arjun Ghose is pursuing M.Sc. in applied statistics from Symbiosis Statistical Institute, Pune. He has graduated with a B.Sc. degree in statistics from St. Xavier's College, Mumbai.

Chaitanyasuma Jain is a computer engineering undergraduate student. She has previously worked on software development and data science projects. She is an active participant in women-based STEM communities and has a keen interest in machine learning. She believes in ideation and wishes to use her technical skills for meaningful social causes.

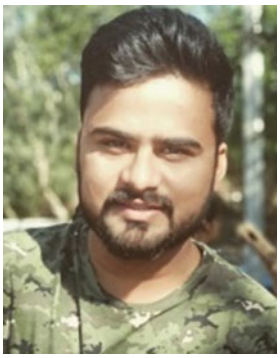
Monisha Jhamb is currently pursuing masters in applied statistics at Symbiosis Statistical Institute, Pune. She is a data science engineer with a post-graduate programme in data science and engineering from Great Lakes Institute of Management. She is also Six Sigma Green Belt certified by KPMG.

Shreyas Joshi is currently pursuing his graduation in electronics and telecommunication engineering at P. E. S. Modern College of Engineering, Pune. His area of interest includes embedded systems and Python for speech recognition.

Anirudh Kolwadkar is a dynamic, team-spirited and performance-driven computer engineering student. He is CS sophomore passionate about technology, Android development and artificial intelligence. His interests include coding and writing technical blogs.

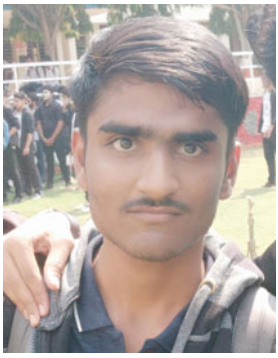
Aboli Marathe is a third-year student studying computer engineering at Pune Institute of Computer Technology. She is a machine learning researcher and is working towards transforming the idea of artificial intelligence for social good. Her research interests include causal inference, explainable artificial intelligence and deep learning.

Vaishnavi Nair is currently pursuing her undergraduate course in electronics and telecommunication engineering from P. E. S Modern College of Engineering, Pune-05. She is currently serving as the CoSSR of the IEEE Pune Section.



Aman Pande is pursuing M.Sc. in applied statistics from Symbiosis Statistical Institute. He has graduated with a B.Sc. degree in statistics from Sir Purshurambhau College of Pune.

Manoj Patil, is currently pursuing his graduation in electronics and telecommunication engineering at P. E. S. Modern College of Engineering, Pune. His area of interest includes IoT and Python programming.



Shivraj Patil is currently pursuing bachelor's degree and has developed a keen interest in the field of machine learning, especially in computer vision. He wishes to work on more amazing projects in the future. As a beginner in reinforcement learning, his current goal is to train an agent (bot) to play tennis.

Atreyee Saha is pursuing M.Sc. in applied statistics from Symbiosis Statistical Institute. She has completed B.Sc. in statistics honours from Bethune College affiliated by Calcutta University, Kolkata.

Siddharth Saoji is pursuing CS Undergrad at VIT, Pune. He is CoSSR of IEEE Pune section and interested in finding solutions to real-world problems using data science and full stack development.

Madhavi Shamkuwar is currently working as an assistant professor in Zeal Education Society, Zeal Institute of Business Administration, Computer Application and Research **(ZIBACAR)**, Pune, Maharashtra, India. She is pursuing Ph.D. from Department of Management Sciences (PUMBA), Savitribai Phule Pune University, Pune, in wildlife management using information technology tools. She had published several research papers and participated in conferences, seminar and FDPs. She had 11 years of teaching experience, and her areas of interest are data mining, system analysis and design, algorithms and big data.

Mishita Sharma is currently pursuing master of science in applied statistics at the Symbiosis Statistical Institute, Pune. She is having a strong background in mathematics; her interests lie in data mining, statistical analysis and machine learning techniques.

Arpan Sil is currently a student of applied statistics in Symbiosis Statistical Institute. He is a bronze medallist in electrical engineering from Maulana Abul Kalam Azad University of Technology for his outstanding academic achievements at undergraduate level. Besides taking an avid interest in data sciences and allied fields, he is passionate about football and writing and has published several articles in reputed sports websites. A through and through optimist, he envisages a better future for humankind and is dedicated to make it happen.

Vineet Tambe is a budding E&TC engineer with a keen interest in robotics and AI. He works with embedded systems and robotics and aims to build advanced secure robotic technology and hardware systems for the future using various artificial intelligence techniques.