

Data Science in Theory and Practice

Data Science in Theory and Practice

Edited by

Jaydip Sen and Sayantani Roy Choudhury

Cambridge
Scholars
Publishing



Data Science in Theory and Practice

Edited by Jaydip Sen and Sayantani Roy Choudhury

This book first published 2024

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

Copyright © 2024 by Jaydip Sen, Sayantani Roy Choudhury
and contributors

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

ISBN: 978-1-0364-0898-5

ISBN (Ebook): 978-1-0364-0899-2

TABLE OF CONTENTS

| | |
|---|-------|
| Preface | vii |
| List of Figures..... | xi |
| List of Tables..... | xxiii |
| Chapter 1 | 1 |
| A Comparative Study of Portfolio Optimization Methods for the Indian Stock Market | |
| <i>Jaydip Sen, Arup Dasgupta and Sayantani Roy Choudhury</i> | |
| Chapter 2 | 71 |
| A Portfolio Rebalancing Approach for the Indian Stock Market | |
| <i>Jaydip Sen, Arup Dasgupta, Subhasis Dasgupta and Sayantani Roy Choudhury</i> | |
| Chapter 3 | 146 |
| Technical Analysis of Indian Stocks: A Triad of Bollinger Bands, MACD, and RSI Strategies | |
| <i>Jaydip Sen, Sayantani Roy Choudhury and Anshuman Pathak</i> | |
| Chapter 4 | 198 |
| Marital Status and Workforce Participation Decision of Indian Women | |
| <i>Alina Guha, Papiya Ghosh, Sujay Kapat and Sayantani Roy Choudhury</i> | |
| Chapter 5 | 209 |
| Interstate Inequalities in LPG Consumption: A Study in the Indian Context | |
| <i>Moupiyali Koley and Devleena Majumdar</i> | |
| Chapter 6 | 231 |
| MGNREGP of India and the Poor States: A Quest for Reality | |
| <i>Harsha Tiwary and Kshitish Kumar Khuntia</i> | |

| | |
|---|-----|
| Chapter 7 | 249 |
| Ethical Practices in the Banking Sector: A Comparative Analysis of the State Bank of India and HDFC Bank | |
| <i>Sumee Dastidar and Minhaz Sabiha Osmani</i> | |
| Chapter 8 | 258 |
| A Comparative Study of Hyperparameter Tuning Methods | |
| <i>Subhasis Dasgupta and Jaydip Sen</i> | |
| Chapter 9 | 289 |
| The Backpropagation Algorithm in Employee Competency Assessment | |
| <i>Kankana Mukhopadhyay</i> | |
| Chapter 10 | 306 |
| Cloud Cover Forecasting for Solar Power Plants Using Sequential Sky Images | |
| <i>Gourab Nath, Kandarp Chaudhary, Kurumaddali Sakhet, Sahil Gupta, Anjita Merin Alex and Adarsh Srivastava</i> | |
| Chapter 11 | 333 |
| Enhancing Crop Health Monitoring: The Use of Convolutional Neural Networks for Early Blight and Late Blight Identification | |
| <i>Subasish Roy, Rahul P and Hitaishi Roy</i> | |
| Chapter 12 | 345 |
| Use of Artificial Intelligence for Detection of Data Mining Errors in Real Time Customer Databases | |
| <i>Sandeep Bhattacharjee</i> | |

PREFACE

The volume titled *Data Science in Theory and Practice* serves as a comprehensive exploration of the dynamic and rapidly evolving field of data science. In today's digital age, the proliferation of data from various sources has led to unprecedented opportunities and challenges. Organizations across industries are increasingly relying on data-driven insights to inform decision-making, drive innovation, and gain a competitive edge in the marketplace. At the heart of this data revolution lies the interdisciplinary field of data science, which combines principles from statistics, computer science, machine learning, and domain expertise to extract meaningful insights from vast and complex datasets. From financial markets to social dynamics, from agricultural landscapes to technological advancements, data science plays a pivotal role in unraveling insights and driving informed decision-making.

This edited volume aims to provide readers with a comprehensive understanding of both the theoretical foundations and practical applications of data science. Through a collection of chapters authored by experts in the field, the volume covers a wide range of topics, spanning from fundamental concepts to advanced methodologies and emerging trends. Each chapter offers a unique perspective on various aspects of data science, providing readers with valuable insights, practical guidance, and cutting-edge research findings. The book is organized into twelve chapters, each addressing various domains, techniques, and applications of data science.

In Chapter 1 titled *A Comparative Study of Portfolio Optimization Methods for the Indian Stock Market Sector Index*, Sen, Dasgupta, and Roy Choudhury present a comparative analysis of three portfolio optimization methods, mean-variance portfolio (MVP), hierarchical risk parity (HRP) portfolio, and hierarchical equal risk contribution (HERC) portfolio, in the Indian stock market across 15 sectors of stocks listed on the National Stock Exchange (NSE) of India. Top stocks in each sector are chosen based on their free-float market capitalization. Three portfolios per sector are constructed using data from July 1, 2019, to June 30, 2022, and evaluated from July 1, 2022, to June 30, 2023. Performance metrics include cumulative returns, volatility, and Sharpe ratios, identifying the portfolios with superior performance characteristics.

In Chapter 2 titled *A Portfolio Rebalancing Approach for the Indian Stock Market*, Sen, Dasgupta, Dasgupta, and Roy Choudhury introduce a

calendar-based rebalancing approach to stock portfolios in the Indian market across ten sectors listed on the NSE of India. It selects the top ten stocks from each sector based on their market capitalization values. Using historical stock data from January 4, 2021, to September 20, 2023, portfolios are constructed and trained from January 4, 2021, to June 30, 2022. Performance evaluation is done from July 1, 2022, to September 20, 2023, with a focus on analyzing the effectiveness of the calendar rebalancing method across sectors.

In Chapter 3 titled *Technical Analysis of Indian Stocks: A Triad of Bollinger Bands, MACD, and RSI Strategies*, Sen, Roy Choudhury, and Pathak explore three widely-used technical indicators - Bollinger Bands, Moving Average Convergence Divergence (MACD), and Relative Strength Index (RSI) - and their efficacy in the Indian stock market. The proposed approach of the authors involves selecting 14 sectors listed on the National Stock Exchange (NSE) and identifying 10 top stocks from each sector based on their market capitalization. Trading activities are conducted from July 1, 2022, to June 30, 2023, with an initial capital of INR 100,000, using the three technical indicators. The performance of each indicator is evaluated, and a comparative analysis across all sectors is conducted to assess their effectiveness.

In Chapter 4 *Marital Status and Workforce Participation Decision of Indian Women*, Guha, Ghosh, Kapat, and Roy Choudhury examine the statistical significance of 'marital status' and several other associated factors influencing female workforce participation, with the Worker Population Ratio (WPR) serving as the primary variable of interest. Additionally, this study seeks to identify factors contributing to the proportion of NEET (Not in Employment, Education, and Training). The work shows that an analysis of these aspects provides insights into the factors affecting the employment status of women, encompassing both participation in the workforce and unemployment.

In Chapter 5 *Interstate Inequalities in LPG Consumption: A Study in Indian Context*, Koley and Majumdar present a study that focuses on focus on the state-wise LPG consumption status in India from 1955 until now. According to the authors, the state-wise LPG consumption trend in India reveals a massive inequality. While some states exhibit higher per capita consumption, there are states where the per capita LPG consumption has been significantly low. The work attempts to identify and focus on the factors responsible for interstate differences in LPG consumption in India over the last twenty years.

In Chapter 6 *MGNREGP of India and the Poor States: A Quest for Reality*, Tiwary and Khuntia examine the effectiveness of the social welfare

scheme, Mahatma Gandhi National Rural Employment Guarantee Program (MGNREGP) in terms of person-days created and funding distributed to the states in India by the central government based on the needs of the states. The study also investigates how much benefits this scheme has been able to bring to the poorest in society.

In Chapter 7 *Ethical Practices in the Banking Sector: A Comparative Analysis of State Bank of India and HDFC Bank*, Dastidar and Osmani carry out a study on a comparative analysis of the perceptions of bank employees regarding ethical issues at two of the largest banks in India, the State Bank of India and the HDFC Bank. The authors investigate whether employees at these banks adhere to their respective codes of ethics in their job performance and analyze potential strategies for enhancing the ethical standards of the banks.

In Chapter 8 *A Comparative Study of Hyperparameter Tuning Methods*, Dasgupta and Sen present three well-known algorithms for hyperparameter tuning of models. Linear and non-linear models for regression and classification tasks based on some public datasets are built and their performances are evaluated. The non-linear models are found to outperform the linear models if their hyperparameters are tuned accurately.

In Chapter 9 *The Backpropagation Algorithm in Employee Competency Assessment*, Mukhopadhyay presents an implementation of the backpropagation algorithm for employee competency assessment. Competencies are categorized into various areas and defined in a way that job incumbents could easily relate to them. Positional competencies, comprising intrinsic and extrinsic elements are also considered and benchmark levels are established for positional competencies.

In Chapter 10 *Cloud Cover Forecasting for Solar Plants Using Sequential Sky Images*, Nath, Chaudhary, Sakhet, Gupta, Alex, and Srivastava investigate image forecasting using both classical time series models and deep learning methods. The study finds VAR model is effective in predicting pixel values based on nearby pixels, and it outperforms models that incorporate weather variables as exogenous, and a deep learning-based LSTM model in image population forecasting. Additionally, an unsupervised technique utilizing the K-means algorithm with RB-ratios is proposed that is proven superior to grayscale image analysis.

In Chapter 11 *Enhancing Crop Health Monitoring: The Use of Convolutional Neural Networks for Early Blight and Late Blight Identification*, Roy, Rahul, and Roy examine the utilization of Convolutional Neural Networks (CNNs) to detect late and early blight diseases in potato crops via leaf optical images. The work assesses CNN's capability in accurately identifying various blight syndromes and healthy

potato leaves. Employing a dataset of potato leaf images, the study applies deep learning techniques for model training. Evaluation of the model's performance focuses on its accuracy in distinguishing between different blight diseases and healthy foliage.

In Chapter 12 *Use of Artificial Intelligence for Detection of Data Mining Errors in Real-Time Customer Databases*, Bhattacharjee presents two error detection techniques, namely Isolation Forest and one-class SVM, each applied to measure distinct types of errors. The work compares the performance of these techniques using sixteen parameters. Additionally, it explores two application areas for each method.

Although the chapters in this volume do not delve into the fundamental theories of the discussed topics, they provide concise discussions of relevant principles and basics to ensure comprehensive coverage. Therefore, while some background knowledge in data science may be beneficial, readers are not required to possess advanced expertise in these areas. We anticipate that this volume will serve as a valuable resource for individuals interested in exploring various applications of data science across diverse domains. The primary audience for this book includes advanced postgraduate and doctoral students in finance, econometrics, management, data science, computer science, and information technology. Additionally, faculty members at graduate schools and universities, as well as data science practitioners in the industry, are likely to find the content highly beneficial.

We extend our heartfelt gratitude to all the contributors who have dedicated their time and expertise to the chapters included in this volume. Their invaluable contributions have been instrumental in making this project a success. We would also like to express our sincere appreciation to Cambridge Scholars Publishing for granting us the opportunity to publish our work under their esteemed publishing banner. Special recognition is owed to Adam Rummens and Sophie Edminson from Cambridge Scholars Publishing for their unwavering patience, collaboration, and support throughout the extensive publishing process. Additionally, we would like to acknowledge the indispensable support and cooperation received from our esteemed faculty colleagues at Praxis Business School, Kolkata, India. Their cooperation has been indispensable, and without their involvement, the publication of this volume would not have been achievable. We extend our heartfelt thanks to every one of them for their invaluable contributions and support.

Jaydip Sen & Sayantani Roy Choudhury

LIST OF FIGURES

| | |
|--|----|
| Fig. 1-1 The efficient frontier of 10,000 candidate portfolios | 6 |
| Fig. 1-2 The dendrogram of the agglomerative clustering is created by the hierarchical risk parity method | 8 |
| Fig. 1-3 The dendrogram of the agglomerative clustering of the stocks from the auto sector | 13 |
| Fig. 1-4 The allocation of weights done by the MVP, HRP, and HERC algorithms for the auto sector stocks..... | 13 |
| Fig. 1-5 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the auto sector stocks on training data. | 14 |
| Fig. 1-6 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the auto sector stocks on the test data | 15 |
| Fig. 1-7 The dendrogram of the agglomerative clustering of the banking sector stocks | 16 |
| Fig. 1-8 The allocation of weights done by the MVP, HRP, and HERC algorithms for the banking sector stocks..... | 16 |
| Fig. 1-9 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the banking sector stocks on the training data | 17 |
| Fig. 1-10 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the banking sector stocks on the test data | 18 |
| Fig. 1-11 The dendrogram of the agglomerative clustering of the financial services sector stocks | 19 |
| Fig. 1-12 The allocation of weights done by the MVP, HRP, and HERC algorithms for the financial services sector stocks..... | 19 |
| Fig. 1-13 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the financial services sector stocks on the training data 20 | |
| Fig. 1-14 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the financial services sector stocks on the test data | 21 |
| Fig. 1-15 The dendrogram of the agglomerative clustering of the consumer durables sector stocks | 22 |
| Fig. 1-16 The allocation of weights done by the MVP, HRP, and HERC algorithms for the consumer durables sector stocks..... | 22 |
| Fig. 1-17 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the consumer durables sector stocks on the training data | 23 |

| | |
|--|----|
| Fig. 1-18 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the consumer durables sector stocks on the test data | 24 |
| Fig. 1-19 The dendrogram of the agglomerative clustering of the FMCG sector stocks | 25 |
| Fig. 1-20 The allocation of weights done by the MVP, HRP, and HERC algorithms for the FMCG sector stocks | 25 |
| Fig. 1-21 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the FMCG sector stocks on the training data | 26 |
| Fig. 1-22 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the FMCG sector stocks on the test data | 27 |
| Fig. 1-23 The dendrogram of the agglomerative clustering of the IT sector stocks | 28 |
| Fig. 1-24 The allocation of weights done by the MVP, HRP, and HERC algorithms for the IT sector stocks | 28 |
| Fig. 1-25 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the IT sector stocks on the training data | 29 |
| Fig. 1-26 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the IT sector stocks on the test data | 30 |
| Fig. 1-27 The dendrogram of the agglomerative clustering of the media sector stocks | 31 |
| Fig. 1-28 The allocation of weights done by the MVP, HRP, and HERC algorithms for the media sector stocks | 31 |
| Fig. 1-29 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the media sector stocks on the training data | 32 |
| Fig. 1-30 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the media sector stocks on the test data | 33 |
| Fig. 1-31 The dendrogram of the agglomerative clustering of the metal sector stocks | 34 |
| Fig. 1-32 The allocation of weights done by the MVP, HRP, and HERC algorithms for the metal sector stocks | 34 |
| Fig. 1-33 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the metal sector stocks on the training data | 35 |
| Fig. 1-34 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the metal sector stocks on the test data | 36 |
| Fig. 1-35 The dendrogram of the agglomerative clustering of the mid- small IT & telecom sector stocks | 37 |

| | |
|--|----|
| Fig. 1-36 The allocation of weights done by the MVP, HRP, and HERC algorithms for the mid-small IT & telecom sector stocks | 37 |
| Fig. 1-37 The cum. ret. yielded by the MVP, HRP, and HERC portfolios for the mid-small IT sector stocks on the training data | 38 |
| Fig. 1-38 The cum. ret. yielded by the MVP, HRP, and HERC portfolios for the mid-small IT & telecom sector stocks on the test data | 39 |
| Fig. 1-39 The dendrogram of the agglomerative clustering of the oil & gas sector stocks | 40 |
| Fig. 1-40 The allocation of weights done by the MVP, HRP, and HERC algorithms for the oil & gas sector stocks..... | 40 |
| Fig. 1-41 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the oil & gas sector stocks on the training data..... | 41 |
| Fig. 1-42 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the oil & gas sector stocks on the test data | 42 |
| Fig. 1-43 The dendrogram of the agglomerative clustering of the pharma sector stocks..... | 43 |
| Fig. 1-44 The allocation of weights done by the MVP, HRP, and HERC algorithms for the pharma sector stocks | 43 |
| Fig. 1-45 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the pharma sector stocks on the training data | 44 |
| Fig. 1-46 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the pharma sector stocks on the test data | 45 |
| Fig. 1-47 The dendrogram of the agglomerative clustering of the private banks sector stocks | 46 |
| Fig. 1-48 The allocation of weights done by the MVP, HRP, and HERC algorithms for the private banks sector stocks..... | 46 |
| Fig. 1-49 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the private banks sector stocks on the training data | 47 |
| Fig. 1-50 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the private banks sector stocks on the test data | 48 |
| Fig. 1-51 The dendrogram of the agglomerative clustering of the PSU banks sector stocks..... | 49 |

| | |
|---|----|
| Fig. 1-52 The allocation of weights done by the MVP, HRP, and HERC algorithms for the PSU banks sector stocks | 49 |
| Fig. 1-53 The cum. ret. yielded by the MVP, HRP, and HERC portfolios for the PSU banks sector stocks on the training data | 50 |
| Fig. 1-54 The cum. ret. yielded by the MVP, HRP, and HERC portfolios for the PSU banks sector stocks on the test data | 51 |
| Fig. 1-55 The dendrogram of the agglomerative clustering of the realty sector stocks | 52 |
| Fig. 1-56 The allocation of weights done by the MVP, HRP, and HERC algorithms for the realty sector stocks | 52 |
| Fig. 1-57 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the realty sector stocks on the training data | 53 |
| Fig. 1-58 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the realty sector stocks on the test data | 54 |
| Fig. 1-59 The dendrogram of the agglomerative clustering of the NIFTY 50 stocks | 55 |
| Fig. 1-60 The allocation of weights done by the MVP, HRP, and HERC algorithms for the NIFTY 50 stocks | 56 |
| Fig. 1-61 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the NIFTY 50 stocks on the training data | 57 |
| Fig. 1-62 The cumulative returns yielded by the MVP, HRP, and HERC portfolios for the NIFTY 50 stocks on the test data | 57 |
| Fig. 2-1 The daily number of shares of each stock in the auto sector portfolio from January 4, 2021, to September 20, 2023 | 91 |
| Fig. 2-2 The daily allocation of weights to each stock of the auto sector portfolio from January 4, 2021, to September 20, 2023 | 91 |
| Fig. 2-3 The cumulative return of the auto sector portfolio and the cumulative return of the benchmark index of NIFTY 50 | 92 |
| Fig. 2-4 The statistical distribution and box plots of the returns of the rebalanced portfolio of the auto sector | 93 |
| Fig. 2-5 The daily number of shares of each stock in the banking sector portfolio from January 4, 2021, to September 20, 2023 | 95 |

| | |
|---|-----|
| Fig. 2-6 The daily allocation of weights to each stock of the banking sector portfolio from January 4, 2021, to September 20, 2023 | 95 |
| Fig. 2-7 The cumulative return of the banking sector portfolio and the cumulative return of the benchmark index of NIFTY 50 | 96 |
| Fig. 2-8 The statistical distribution and box plots of the returns of the rebalanced portfolio of the banking sector..... | 97 |
| Fig. 2-9 The daily number of shares of each stock in the consumer durables sector portfolio from January 4, 2021, to September 20, 2023 | 99 |
| Fig. 2-10 The daily allocation of weights to each stock of the consumer durables sector portfolio from Jan 4, 2021, to Sep 20, 2023 | 99 |
| Fig. 2-11 The cumulative return of the consumer durables sector portfolio and the cumulative return of the benchmark index of NIFTY 50 | 100 |
| Fig. 2-12 The statistical distribution and box plots of the returns of the rebalanced portfolio of the consumer durables sector. | 101 |
| Fig. 2-13 The daily number of shares of each stock in the FMCG sector portfolio from January 4, 2021, to September 20, 2023 | 103 |
| Fig. 2-14 The daily allocation of weights to each stock of the FMCG sector portfolio from Jan 4, 2021, to Sep 20, 2023 | 103 |
| Fig. 2-15 The cumulative return of the FMCG sector portfolio and the cumulative return of the benchmark index of NIFTY 50 | 104 |
| Fig. 2-16 The statistical distribution and box plots of the returns of the rebalanced portfolio of the FMCG sector | 105 |
| Fig. 2-17 The daily number of shares of each stock in the IT sector portfolio from January 4, 2021, to September 20, 2023 | 107 |
| Fig. 2-18 The daily allocation of weights to each stock of the IT sector portfolio from Jan 4, 2021, to Sep 20, 2023 | 107 |
| Fig. 2-19 The cumulative return of the IT sector portfolio and the cumulative return of the benchmark index of NIFTY 50 | 108 |
| Fig. 2-20 The statistical distribution and box plots of the returns of the rebalanced portfolio of the IT sector..... | 109 |

| | |
|--|-----|
| Fig. 2-21 The daily number of shares of each stock in the metal sector portfolio from January 4, 2021, to September 20, 2023 | 111 |
| Fig. 2-22 The daily allocation of weights to each stock of the metal sector portfolio from Jan 4, 2021, to Sep 20, 2023 | 111 |
| Fig. 2-23 The cumulative return of the metal sector portfolio and the cumulative return of the benchmark index of NIFTY 50..... | 112 |
| Fig. 2-24 The statistical distribution and box plots of the returns of the rebalanced portfolio of the metal sector..... | 113 |
| Fig. 2-25 The daily number of shares of each stock in the pharma sector portfolio from January 4, 2021, to September 20, 2023 | 115 |
| Fig. 2-26 The daily allocation of weights to each stock of the pharma sector portfolio from Jan 4, 2021, to Sep 20, 2023 | 115 |
| Fig. 2-27 The cumulative return of the pharma sector portfolio and the cumulative return of the benchmark index of NIFTY 50 | 116 |
| Fig. 2-28 The statistical distribution and box plots of the returns of the rebalanced portfolio of the pharma sector..... | 117 |
| Fig. 2-29 The daily number of shares of each stock in the private banks sector portfolio from January 4, 2021, to September 20, 2023 | 119 |
| Fig. 2-30 The daily allocation of weights to each stock of the private banks sector portfolio from Jan 4, 2021, to Sep 20, 2023 | 119 |
| Fig. 2-31 The cumulative return of the private banks sector portfolio and the cumulative return of the benchmark index of NIFTY 50 | 120 |
| Fig. 2-32 The statistical distribution and box plots of the returns of the rebalanced portfolio of the private banks sector..... | 121 |
| Fig. 2-33 The daily number of shares of each stock in the PSU banks sector portfolio from January 4, 2021, to September 20, 2023 | 123 |
| Fig. 2-34 The daily allocation of weights to each stock of the PSU banks sector portfolio from Jan 4, 2021, to Sep 20, 2023 | 123 |

| | |
|--|-----|
| Fig. 2-35 The cumulative return of the PSU banks sector portfolio and the cumulative return of the benchmark index of NIFTY 50 | 124 |
| Fig. 2-36 The statistical distribution and box plots of the returns of the rebalanced portfolio of the PSU banks sector | 125 |
| Fig. 2-37 The daily number of shares of each stock in the realty sector portfolio from January 4, 2021, to September 20, 2023 | 127 |
| Fig. 2-38 The daily allocation of weights to each stock of the realty sector portfolio from Jan 4, 2021, to Sep 20, 2023 | 127 |
| Fig. 2-39 The cumulative return of the realty sector portfolio and the cumulative return of the benchmark index of NIFTY 50 | 128 |
| Fig. 2-40 The statistical distribution and box plots of the returns of the rebalanced portfolio of the realty sector | 129 |
| Fig. 2-41 A graphical representation of the performance of the auto sector portfolio | 133 |
| Fig. 2-42 A graphical representation of the performance of the banking sector portfolio | 133 |
| Fig. 2-43 A graphical representation of the performance of the consumer durables sector portfolio | 133 |
| Fig. 2-44 A graphical representation of the performance of the FMCG sector portfolio | 134 |
| Fig. 2-45 A graphical representation of the performance of the IT sector portfolio | 134 |
| Fig. 2-46 A graphical representation of the performance of the metal sector portfolio | 134 |
| Fig. 2-47 A graphical representation of the performance of the pharma sector portfolio | 135 |
| Fig. 2-48 A graphical representation of the performance of the private banks sector portfolio | 135 |
| Fig. 2-49 A graphical representation of the performance of the PSU banks sector portfolio | 135 |
| Fig. 2-50 A graphical representation of the performance of the realty sector portfolio | 136 |
| Fig. 3-1 The Bollinger Bands plot of Mahindar & Mahindra stock with the trading signal points identified | 156 |
| Fig. 3-2 The MACD plot of Mahindra & Mahindra stock with the trading signal points identified | 156 |

| | |
|---|-----|
| Fig. 3-3 The RSI plot of Mahindra & Mahindra stock with the trading signal points identified..... | 157 |
| Fig. 3-4 The Bollinger Bands plot of ICICI Bank stock with the trading signal points identified..... | 158 |
| Fig. 3-5 The MACD plot of ICICI Bank stock with the trading signal points identified..... | 159 |
| Fig. 3-6 The RSI plot of ICICI Bank stock with the trading signal points identified..... | 159 |
| Fig. 3-7 The Bollinger Bands plot of Bajaj Finance stock with the trading signal points identified..... | 161 |
| Fig. 3-8 The MACD plot of Bajaj Finance stock with the trading signal points identified..... | 161 |
| Fig. 3-9 The RSI plot of Bajaj Finance stock with the trading signal points identified..... | 162 |
| Fig. 3-10 The Bollinger Bands plot of Titan Company stock with the trading signal points identified..... | 163 |
| Fig. 3-11 The MACD plot of Titan Company stock with the trading signal points identified..... | 163 |
| Fig. 3-12 The RSI plot of Titan Company stock with the trading signal points identified..... | 164 |
| Fig. 3-13 The Bollinger Bands plot of Hindustan Unilever stock with the trading signal points identified | 165 |
| Fig. 3-14 The MACD plot of Hindustan Unilever stock with the trading signal points identified..... | 165 |
| Fig. 3-15 The RSI plot of Hindustan Unilever stock with the trading signal points identified..... | 166 |
| Fig. 3-16 The Bollinger Bands plot of Coforge stock with the trading signal points identified..... | 167 |
| Fig. 3-17 The MACD plot of Coforge stock with the trading signal points identified..... | 167 |
| Fig. 3-18 The RSI plot of Coforge stock with the trading signal points identified..... | 168 |
| Fig. 3-19 The Bollinger Bands plot of Zee Ent. Enterprise stock with the trading signal points identified | 169 |
| Fig. 3-20 The MACD plot of Zee Ent. Enterprise stock with the trading signal points identified..... | 170 |
| Fig. 3-21 The RSI plot of Zee Ent. Enterprise stock with the trading signal points identified..... | 170 |
| Fig. 3-22 The Bollinger Bands plot of Hindalco Industries stock with the trading signal points identified | 171 |

| | |
|--|-----|
| Fig. 3-23 The MACD plot of Hindalco Industries stock with the trading signal points identified..... | 172 |
| Fig. 3-24 The RSI plot of Hindalco Industries stock with the trading signal points identified..... | 172 |
| Fig. 3-25 The Bollinger Bands plot of L&T Tech. Services stock with the trading signal points identified | 174 |
| Fig. 3-26 The MACD plot of L&T Tech. Services stock with the trading signal points identified..... | 174 |
| Fig. 3-27 The RSI plot of L&T Tech. Services stock with the trading signal points identified..... | 175 |
| Fig. 3-28 The Bollinger Bands plot of Hindustan Petroleum Corporation stock with the trading signal points identified..... | 176 |
| Fig. 3-29 The MACD plot of Hindustan Petroleum Corporation stock with the trading signal points identified | 176 |
| Fig. 3-30 The RSI plot of Hindustan Petroleum Corporation stock with the trading signal points identified | 177 |
| Fig. 3-31 The Bollinger Bands plot of Sun Pharma Industries stock with the trading signal points identified | 178 |
| Fig. 3-32 The MACD plot of Sun Pharma Industries stock with the trading signal points identified..... | 178 |
| Fig. 3-33 The RSI plot of Sun Pharma Industries stock with the trading signal points identified..... | 179 |
| Fig. 3-34 The Bollinger Bands plot of Axis Bank stock with the trading signal points identified..... | 180 |
| Fig. 3-35 The MACD plot of Axis Bank stock with the trading signal points identified..... | 180 |
| Fig. 3-36 The RSI plot of Axis Bank stock with the trading signal points identified..... | 181 |
| Fig. 3-37 The Bollinger Bands plot of State Bank of India stock with the trading signal points identified | 182 |
| Fig. 3-38 The MACD plot of State Bank of India stock with the trading signal points identified..... | 183 |
| Fig. 3-39 The RSI plot of State Bank of India stock with the trading signal points identified..... | 183 |
| Fig. 3-40 The Bollinger Bands plot of DLF stock with the trading signal points identified..... | 184 |
| Fig. 3-41 The MACD plot of DLF stock with the trading signal points identified..... | 185 |
| Fig. 3-42 The RSI plot of DLF stock with the trading signal points identified | 185 |

| | |
|--|-----|
| Fig. 4-1 The decision tree classification model for WPR (male)..... | 204 |
| Fig. 4-2 The decision tree classification model for WPR (female)..... | 204 |
| Fig. 4-3 The decision tree classification model for NEET (male)..... | 206 |
| Fig. 4-4 The decision tree classification model for NEET (female)..... | 206 |
| Fig. 5-1 The Energy Ladder Model..... | 210 |
| Fig. 5-2 The historical pattern of the LPG consumption in India..... | 215 |
| Fig. 5-3 The phases of the LPG consumption in India. | 216 |
| Fig. 5-4 The map of India showing the intensity of per capita LPG consumption in different states in 2001 and 2021..... | 217 |
| Fig. 5-5 The per capita LPG consumption in India in 2021, 2011, and 2021..... | 218 |
| Fig. 6-1 Number of job cards issued under MGNREGP in poor and non-poor states..... | 239 |
| Fig. 6-2 Number of person-days generated under MGNREGP in poor and non-poor states..... | 240 |
| Fig. 6-3 Number of households employed under MGNREGP in both poor and non-poor states..... | 240 |
| Fig. 6-4 The spread of MGNREGP in both poor and non-poor states..... | 241 |
| Fig. 6-5 The depth of MGNREGP in both poor and non-poor states..... | 242 |
| Fig. 8-1 The bias-variance trade-off..... | 260 |
| Fig. 8-2 The scatter plot of kWh and other numerical variables in the training and the test dataset..... | 268 |
| Fig. 8-3 The distribution of data in training and test set for the steel industry dataset..... | 269 |
| Fig. 8-4 The distribution of the predictor variables of the gas turbine data in both training and test data set..... | 272 |
| Fig. 8-5 The scatter plot of CO and other numerical variables in the training and the test dataset of the gas turbine data..... | 274 |
| Fig. 8-6 The distribution of the numerical variables in the training and test dataset of the adult data..... | 276 |
| Fig. 8-7 The distribution of the categorical variables in the training and test datasets of the adult data..... | 278 |

| | |
|---|-----|
| Fig. 8-8 The distribution of data in the training and test set of the dry bean dataset..... | 280 |
| Fig. 9-1 Illustration of backpropagation in a neural network | 292 |
| Fig. 9.2 The training results – Quality competencies (Mgr) I..... | 299 |
| Fig. 9.3 The coefficient of correlation: Quality competencies (Mgr) I..... | 300 |
| Fig. 9.4 The training results: Quality competencies (Mgr) II | 301 |
| Fig. 9.5 The coefficient of correlation: Quality competencies (Mgr) II..... | 301 |
| Fig. 9.6 The competency assessment form (Quality) | 304 |
| Fig. 10.1 A sample image of the sky was captured using a fish-eye lens | 309 |
| Fig. 10.2 A sketch to represent some neighborhood pixels at a given time..... | 311 |
| Fig. 10.3 A sample of 6 sky images with different cloud coverages | 315 |
| Fig. 10.4 An illustration of a resized image, data segmentation for batch processing, and sliding window..... | 319 |
| Fig. 10.5 Actual vs. forecasted pixel values using LSTM model with the Adam optimizer..... | 321 |
| Fig. 10.6 Actual and forecasted pixel values using the LSTM model with the RMSProp optimizer | 321 |
| Fig. 10-7 K-means clustering on two different images – (a) a clear sky, and (b) a cloudy sky for different values of K..... | 325 |
| Fig. 10-8 K-means clustering for a sky image with a 29% cloud cover percentage for different values of K..... | 327 |
| Fig. 10-9 K-means clustering for a sky image with a 33% cloud cover percentage for different values of K..... | 328 |
| Fig. 10-10 K-means clustering for a sky image with a 67% cloud cover percentage for different values of K..... | 328 |
| Fig. 10-11 K-means clustering for a sky image with a 71% cloud cover percentage for different values of K..... | 329 |
| Fig. 11-1 K-means clustering for a sky image with a 71% cloud cover percentage for different values of K..... | 337 |
| Fig. 11-2 The stepwise flowchart of the proposed methodology | 338 |
| Fig. 11-3 The architecture of the CNN model | 339 |
| Fig. 11-4 The training and validation loss and accuracy plots for the CNN model..... | 341 |

| | |
|--|-----|
| Fig. 11-5 The figure exhibits various predicted outputs for some sample potato leaves and classifies leaf images into their classes..... | 342 |
| Fig. 12-1 Anomaly detection using an isolation forest..... | 350 |
| Fig. 12-2 Anomaly detection using a one-class SVM | 353 |

LIST OF TABLES

| | |
|---|----|
| Table 1.1 The portfolio compositions of the auto sector (Period: July 1, 2019 – June 30, 2022) | 14 |
| Table 1.2 The performance results of the portfolios of the auto sector stocks | 15 |
| Table 1.3 The portfolio compositions of the banking sector (Period: July 1, 2019 – June 30, 2022) | 16 |
| Table 1.4 The performance results of the banking sector portfolios | 18 |
| Table 1.5 The portfolio compositions of the financial services sector (Period: July 1, 2019 – June 30, 2022) | 20 |
| Table 1.6 The performance results of the financial services sector portfolios | 21 |
| Table 1.7 The portfolio compositions of the consumer durables sector (Period: July 1, 2019 – June 30, 2022)..... | 23 |
| Table 1.8 The performance results of the consumer durables sector portfolios | 24 |
| Table 1.9 The portfolio compositions of the FMCG sector (Period: July 1, 2019 – June 30, 2022)..... | 26 |
| Table 1.10 The performance results of the FMCG sector portfolios | 27 |
| Table 1.11 The portfolio compositions of the IT sector (Period: July 1, 2019 – June 30, 2022)..... | 29 |
| Table 1.12 The performance results of the portfolios of the IT sector | 30 |
| Table 1.13 The portfolio compositions of the media sector | 32 |
| Table 1.14 The performance results of the portfolios of the media sector | 33 |
| Table 1.15 The portfolio compositions of the metal sector (Period: July 1, 2019 – June 30, 2022)..... | 35 |
| Table 1.16 The performance results of the portfolios of the metal sector..... | 36 |
| Table 1.17 The portfolio of the mid-small IT & telecom sector..... | 38 |
| Table 1.18 The performance results of the mid-small IT & telecom portfolios | 39 |

| | |
|--|-----|
| Table 1.19 The portfolio compositions of the oil & gas sector (Period: July 1, 2019 – June 30, 2022) | 41 |
| Table 1.20 The performance results of the oil & gas sector portfolios | 42 |
| Table 1.21 The portfolio compositions of the pharma sector (Period: July 1, 2019 – June 30, 2022) | 44 |
| Table 1.22 The performance results of the pharma sector portfolios | 45 |
| Table 1.23 The portfolio compositions of the private banks sector (Period: July 1, 2019 – June 30, 2022) | 47 |
| Table 1.24 The performance results of the private banks sector portfolios | 48 |
| Table 1.25 The portfolio compositions of the PSU banks sector (Period: July 1, 2019 – June 30, 2022) | 50 |
| Table 1.26 The performance results of the PSU banks sector portfolios | 51 |
| Table 1.27 The portfolio compositions of the realty sector (Period: July 1, 2019 – June 30, 2022) | 53 |
| Table 1.28 The performance results of the realty sector portfolios | 54 |
| Table 1.29 The portfolio compositions of the NIFTY 50 stocks (Period: July 1, 2019 – June 30, 2022) | 56 |
| Table 1.30 The performance results of the NIFTY 50 portfolios | 58 |
| Table 1.31 The summary of the performances of the portfolios on the training data (Period: July 1, 2019 – June 30, 2022) | 58 |
| Table 1.32 The summary of the performances of the portfolios on the test data (Period: July 1, 2022 – June 30, 2023) | 58 |
| Table 2.1 The performance of the auto sector portfolio on the in-sample and out-of-sample data | 94 |
| Table 2.2 The performance of the banking sector portfolio on the in-sample and out-of-sample data | 98 |
| Table 2.3 The performance of the consumer durables sector portfolio on the in-sample and out-of-sample data | 102 |
| Table 2.4 The performance of the FMCG sector portfolio on the in-sample and out-of-sample data | 106 |
| Table 2.5 The performance of the IT sector portfolio on the in-sample and out-of-sample data | 110 |

| | |
|--|-----|
| Table 2.6 The performance of the metal sector portfolio on the in-sample and out-of-sample data | 114 |
| Table 2.7 The performance of the pharma sector portfolio on the in-sample and out-of-sample data | 118 |
| Table 2.8 The performance of the private banks sector portfolio on the in-sample and out-of-sample data | 121 |
| Table 2.9 The performance of the PSU banks sector portfolio on the in-sample and out-of-sample data | 125 |
| Table 2.10 The performance of the realty sector portfolio on the in-sample and out-of-sample data | 129 |
| Table 2.11 The summary of the performances of the portfolios on the in-sample data (Period: Jan 1, 2021 – Jun 30, 2022)..... | 130 |
| Table 2.12 The summary of the performances of the portfolios on the out-of-sample data (Period: Jul 1, 2022 – Sept 20, 2023)..... | 131 |
| Table 2.13 The summary of the performances of the portfolios on the overall data (Period: Jan 4, 2021 – Sept 20, 2023) | 131 |
| Table 3.1 The annual returns yielded by the BB, MACD, and RSI methods for the auto sector stocks (Period: July1, 2022 – June 30, 2023) | 157 |
| Table 3.2 The annual returns yielded by the BB, MACD, and RSI methods for the banking sector stocks (Period: July1, 2022 – June 30, 2023) | 160 |
| Table 3.3 The annual returns yielded by the BB, MACD, and RSI methods for the financial services sector stocks (Period: July1, 2022 – June 30, 2023) | 162 |
| Table 3.4 The annual returns yielded by the BB, MACD, and RSI methods for the consumer durables sector stocks (Period: July1, 2022 – June 30, 2023)..... | 164 |
| Table 3.5 The annual returns yielded by the BB, MACD, and RSI methods for the FMCG sector stocks (Period: July1, 2022 – June 30, 2023) | 166 |
| Table 3.6 The annual returns yielded by the BB, MACD, and RSI methods for the IT sector stocks (Period: July1, 2022 – June 30, 2023) | 168 |
| Table 3.7 The annual returns yielded by the BB, MACD, and RSI methods for the media sector stocks (Period: July1, 2022 – June 30, 2023) | 171 |

| | |
|--|-----|
| Table 3.8 The annual returns yielded by the BB, MACD, and RSI methods for the metal sector stocks (Period: July1, 2022 – June 30, 2023) | 173 |
| Table 3.9 The annual returns yielded by the BB, MACD, and RSI methods for the mid-small IT & telecom sector stocks (Period: July1, 2022 – June 30, 2023) | 175 |
| Table 3.10 The annual returns yielded by the BB, MACD, and RSI methods for the oil & gas sector stocks (Period: July1, 2022 – June 30, 2023) | 177 |
| Table 3.11 The annual returns yielded by the BB, MACD, and RSI methods for the pharma sector stocks (Period: July1, 2022 – June 30, 2023) | 179 |
| Table 3.12 The annual returns yielded by the BB, MACD, and RSI methods for the private banks sector stocks (Period: July1, 2022 – June 30, 2023) | 181 |
| Table 3.13 The annual returns yielded by the BB, MACD, and RSI methods for the PSU banks sector stocks (Period: July1, 2022 – June 30, 2023) | 184 |
| Table 3.14 The annual returns yielded by the BB, MACD, and RSI methods for the realty sector stocks (Period: July1, 2022 – June 30, 2023) | 186 |
| Table 3.15 The summary of the performances of the three technical indicators on the stocks from fourteen sectors..... | 186 |
| Table 4.1 The output of the logistic regression model for WPR (Male & Female) – Part I | 202 |
| Table 4.2 The output of the logistic regression model for WPR (Male & Female) – Part II | 202 |
| Table 4.3 The codes for marital status | 203 |
| Table 4.4 The performance results of the decision tree classification for WPR (Male & Female)..... | 203 |
| Table 4.5 The output of the logistic regression model for NEET (Male & Female) – Part I | 205 |
| Table 4.6 The output of the logistic regression model for NEET (Male & Female) – Part II..... | 205 |
| Table 4.7 The performance results of the decision tree classification for NEET (Male & Female) | 205 |
| Table 5.1 The fixed effect panel regression results (Period: 2001-2021) | 220 |
| Table 5.2 The correlation among the independent variables | 220 |
| Table 6.1 The list of determinants | 236 |
| Table 6.2 The basic indicators of MGNREGP | 238 |

| | |
|---|-----|
| Table 6.3 The correlation of two important indicators | 243 |
| Table 6.4 The regression of two important indicators | 244 |
| Table 7.1 Employees' perception of banking ethics | 252 |
| Table 7.2 Employees' opinion in context to the importance of ethics in the banking sector..... | 252 |
| Table 7.3 The independent samples t-test for the difference between SBI and HDFC Bank in the promotion of ethics | 253 |
| Table 7.4 The chi-square test for the association of compliance of ethical practices by bank employees | 254 |
| Table 7.5 The response to the question: Whether business plans are analyzed before loans are sanctioned..... | 254 |
| Table 7.6 Ethical issues in the selected banks | 255 |
| Table 7.7 Improving ethics in the selected banks | 255 |
| Table 8.1 The description of the basic dataset | 265 |
| Table 8.2 The description of the steel industry data | 267 |
| Table 8.3 The description of the gas turbine data | 270 |
| Table 8.4 The correlation matrix for the gas turbine data..... | 273 |
| Table 8.5 The variable descriptions of the adult dataset..... | 274 |
| Table 8.6 A brief description of the variables of the dry bean data | 279 |
| Table 8.7 The set of hyperparameters of different models | 281 |
| Table 8.8 The performance of the fine-tuned models on the steel industry dataset (Metric: RMSE) | 282 |
| Table 8.9 The performance of the models based on fine-tuned hyperparameters on the gas turbine dataset (Metric: RMSE) | 283 |
| Table 8.10 The performance of the fine-tuned models on the adult dataset (Metric: Cohen Kappa) | 283 |
| Table 8.11 The performance of the fine-tuned models on the bean dataset (Metric: Cohen Kappa) | 283 |
| Table 9.1 The quality competencies – input values | 297 |
| Table 9.2 The quality competencies (MGR 1) | 299 |
| Table 9.3 The quality competencies – input/output (MGR II)..... | 300 |
| Table 9.4 The quality competencies – benchmark table (MGR) | 303 |
| Table 10.1 The descriptions of the variables in the weather dataset..... | 315 |
| Table 10.2 The results of the VAR (3) model for sky image forecasting | 319 |
| Table 10.3 The results of the LSTM model for sky image forecasting .. | 320 |

| | |
|---|-----|
| Table 10.4 The VAR (3) model's results in cloud cover forecasting | 322 |
| Table 10.5 The best K-means results for grey-scaled images..... | 324 |
| Table 10.6 The best K-means results for RB-ratio images..... | 327 |
| Table 11.1 The performance results of the CNN model..... | 342 |
| Table 12.1 A comparative analysis of isolation forest and one-class SVM methods for anomaly detection..... | 355 |

CHAPTER 1

A COMPARATIVE STUDY OF PORTFOLIO OPTIMIZATION METHODS FOR THE INDIAN STOCK MARKET

JAYDIP SEN, ARUP DASGUPTA AND
SAYANTANI ROY CHOUDHURY

Introduction

The design of optimized portfolios has remained a research topic of broad and intense interest among researchers of quantitative and statistical finance for a long time. An optimum portfolio allocates the weights to a set of capital assets in a way that optimizes the return and risk of those assets. Markowitz in his seminal work proposed the *mean-variance optimization* approach which is based on the mean and covariance matrix of returns (Markowitz, 1952). The *mean-variance portfolio* (MVP) design works on an algorithm, known as the *critical line algorithm* (CLA). The CLA algorithm, despite the elegance of its theoretical framework, has some major limitations. One of the major problems is the adverse effects of the estimation errors in its expected returns and covariance matrix on the performance of the portfolio. Since it is extremely challenging to accurately estimate the expected returns of an asset from its historical prices, it is a popular practice to use either a minimum variance portfolio or an optimized risk portfolio with the maximum Sharpe ratio as better proxies for the expected returns. However, due to the inherent complexity, several factors have been used to explain the expected returns.

The *hierarchical risk parity* (HRP) algorithm attempts to address three major shortcomings of quadratic optimization methods which are particularly relevant to the CLA used in the MVP approach to portfolio design (de Prado, 2016). These problems are instability, concentration, and underperformance. Unlike the CLA, the HRP algorithm does not require the

covariance matrix of return values to be invertible. Hence, the HRP portfolios can deliver good results even if the covariance matrices are ill-degenerated or singular, which is an impossibility for a quadratic optimizer like CLA. Interestingly, even though MVP's objective is to minimize the variance, HRP is proven to have a lower probability of yielding lower out-of-sample variance than MVP. Given the fact that future returns cannot be forecasted with sufficient accuracy, many researchers have proposed risk-based asset allocation using the covariance matrix of the returns. However, this approach brings in another problem of instability that arises because the quadratic programming methods require the inversion of a covariance matrix whose all eigenvalues must be positive (Baily & de Prado, 2012). HRP is a new portfolio design method that addresses the pitfalls of the quadratic optimization-based MVP approach using techniques of graph theory and machine learning (de Prado, 2016). This method exploits the features of the covariance matrix without the requirement of its invertibility.

The HERC portfolio optimization uses an integrated approach to machine learning and a top-down recursive bisection method of the HRP portfolio method (Raffinot, 2018). The proponents of the HERC method identified several shortcomings of the HRP portfolio optimization. The single linkage-based cluster trees constructed in the HRP method usually led to deep and wide trees and suboptimal allocation of weights to the clusters. The HRP algorithm usually involves higher computations. Finally, the recursive bijection approach used in HRP bisects the cluster tree before the weight allocation instead of directly allocating the weights based on the dendrogram of clustering. This makes the computed weights inaccurate. The HERC algorithm avoids these problems using a top-down recursive bisection and a naive risk parity within the clusters.

This chapter presents a comparative study of the three portfolio optimization methods, MVP, HRP, and HERC, on the Indian stock market, particularly focusing on the stocks chosen from 15 sectors listed on the National Stock Exchange (NSE) of India. The top stocks of each cluster are identified based on their free-float market capitalization from the NSE's report published on July 1, 2022 (NSE Website). For each sector, three portfolios are designed on stock prices from July 1, 2019, to June 30, 2022, following three portfolio optimization approaches. The portfolios are tested over the period from July 1, 2022, to June 30, 2023. For evaluation of the performances of the portfolios, three metrics are used (i) cumulative returns, (ii) annual volatilities, and (iii) Sharpe ratios. based on their cumulative returns. For each sector, the portfolios that yield the highest cumulative return, the lowest volatility, and the maximum Sharpe Ratio over the training and the test periods are identified.