

PORTFOLIO OPTIMIZATION IN INDIAN STOCK MARKET: A STUDY USING FINANCIAL ANALYTICS

DHANYA JOTHIMANI



DEPARTMENT OF MANAGEMENT STUDIES
INDIAN INSTITUTE OF TECHNOLOGY DELHI
INDIA

NOVEMBER 2017

PORTFOLIO OPTIMIZATION IN INDIAN STOCK MARKET: A STUDY USING FINANCIAL ANALYTICS

by

DHANYA JOTHIMANI
Department of Management Studies

Submitted

in fulfillment of the requirements of the degree of Doctor of Philosophy

to the



INDIAN INSTITUTE OF TECHNOLOGY DELHI
INDIA

NOVEMBER 2017

“The greatest enemy of knowledge is not ignorance, it is the illusion of knowledge.”

-Stephen Hawking

Dedicated to my parents
Shri. R. Jothimani and Smt. T. Eswari

Certificate

The thesis entitled “**Portfolio Optimization in Indian Stock Market: A Study Using Financial Analytics**”, being submitted by **Ms. Dhanya Jothimani** to the Indian Institute of Technology Delhi, for the award of the degree of “**Doctor of Philosophy**” is a record of *bona fide* research work carried out by her. She has worked under our supervision in conformity with rules and regulations of the Indian Institute of Technology Delhi. The research reports and results presented in the thesis have not been submitted in part or full for the award of any degree or diploma in any other University or Institute.

Date:

Place:

Prof. Surendra S. Yadav
Professor
Department of Management Studies,
Indian Institute of Technology Delhi,
Hauz Khas, New Delhi

Prof. Ravi Shankar
Professor
Department of Management Studies,
Indian Institute of Technology Delhi,
Hauz Khas, New Delhi

Acknowledgements

“Gratitude is the attitude that takes you to your altitude.”

First and foremost, I would like to thank the Almighty for providing me necessary strength and support during this part of my (academic) life.

I take this opportunity to express my deepest gratitude to most revered supervisors (*gurus*) Prof. Surendra S. Yadav and Prof. Ravi Shankar. No words can express how thankful and blessed I am to be associated with Prof. Surendra S. Yadav. He has been more than a father figure to me and helping me grow in both personal and research life. His constant encouragement to me and confidence in me helped me to evolve in this research area. I express my sincere gratitude to Prof. Ravi Shankar for enabling me to explore the area of Financial Analytics. His encouraging words at right times helped me to target the work on high impact journals and to attend well-reputed conferences.

I am thankful to my Student Research Committee (SRC) members: Prof. Pramod Kumar Jain, Dr. Seema Sharma and Prof. S. Dharmaraja (Department of Mathematics) for their suggestions and comments that helped me to improve my work. I would like to thank Dr. P. Vignesh and Dr. Shveta Singh for allowing me to attend their lectures during formative stage of my research. Special thanks are due to Prof. Prithpal Singh Sir and Jaswinder aunty for their blessings and affection always.

I am grateful to few researchers including Prof. Ross Barmish (University of Wisconsin-Madison), Prof. Hirsohi Morita (Osaka University), Prof. Nicole Adler (The Hebrew University of Jerusalem), Prof. Rajiv Banker (Temple University), Prof. Dessi Pachamano (Babson College), Prof. Phillip Ernst (Rice University), Dr. Sander Gerber (Hudson Bay Capital Management) and Prof. Frank Fabozzi (EDHEC Business School), for sparing their valuable time and promptly responding to my queries over e-mail.

I am extremely lucky to have few close friends, especially, Abhay Sir, Rajeev Sir, Saptha, Tapan, Bishal, Megha, Burhan Sir, Gurmeet, Isha, Amrutha Das-Mithun, Amrutha Das-Viswanath, Lavanya and Jai, who extended their help and support in their own capacity and being there by my side always. There are many colleagues,

like Arun Ji, Anushruti, Aakriti, Krishnendu Sir, Sumant Sir, Nakul, Fahima, Ashish Rathore, Ajay Sir, Vijayta, Divya, Rachita, Harshita, Monika Sharma, Monika Singla, Veepan, Ashish Kaushal, Devender, Sushil, Nisha Thomas, Vipulesh, Rishi, Amit ji, Sonali, Mahamaya mam and many others who should not be forgotten for their unconditional support and encouragement at various points. I would like to thank Saptha, Bhagyalakshmi and Shreyas for tirelessly sending across the research articles that were not accessible at IIT Delhi. The support from staff members of DMS (Amit ji, Amit bhaiya, Ditpal ji, Jagadish ji, Parikshit ji, Prakash ji, Ravi, Anita, Sushil and Vijender ji) and PG section (Rajender ji and Chopra ji) is duly acknowledged for their help with the official processes without any pain. In addition, I am obliged to all the respondents of the survey, who took out time from their busy schedules to respond to the survey questionnaire.

As always, Amma and Appa have been a great source of inspiration and support throughout this journey by understanding the constraints that I have not been able to spend time with them and that I was not able to fulfil my responsibilities towards them during this period.

Dhanya Jothimani

Abstract

Portfolio optimization is an investment decision-making process to hold a set of assets based on various criteria such as risk and return. Availability of several financial instruments makes it complex. Portfolio optimization can be carried out in three stages, namely, asset selection, asset weighting and asset management. Previous researches have invariably focussed on either first two stages of portfolio optimization or on each stage separately. This study aims to provide a comprehensive end-to-end framework for portfolio optimization focussing on all three stages. Further, it also addresses the issues related to improving the covariance matrix and improving the forecasting accuracy of stock price. The scope of the work is limited to analysis of non-financial stocks, listed on National Stock Exchange.

First phase of the study proposes an analytical framework for portfolio optimization. The historical performance of the firms is evaluated using data envelopment analysis, thus, aiding in selection of potential stocks for portfolio formation.

Hierarchical risk parity model based on Gerber statistics (HRP-GS) is proposed for determining the weights to be assigned for each asset in the portfolio. The performance of the proposed model is compared with three models, namely, global minimum variance model based on historical correlation, global minimum variance model based on Gerber statistics, and hierarchical risk parity model based on historical correlation. The HRP-GS model is proposed to address instability issues in covariance matrix and improving the estimates of covariance matrix.

Six ensemble forecasting models (Empirical Mode Decomposition (EMD)-Artificial neural network (ANN), EMD-Support Vector Regression (SVR), Ensemble EMD (EEMD)-ANN, EEMD-SVR, Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN)-ANN and CEEMDAN-SVR) are developed for predicting prices of stocks constituting the portfolio. Trading rules are designed to guide the investors to buy/sell/hold stocks according to varying market conditions. The experimental analysis carried out on the data from Indian stock market data shows the strength of the proposed framework.

In second phase of the study, a survey is carried out to understand the investment practices adopted by portfolio managers in Indian stock market. Survey results suggest that implementation of Financial theories in Indian stock market is very minimal.

सार

पोर्टफोलियो ऑप्टिमाइजेशन विभिन्न मानदंडों जैसे रिस्क और रिटर्न के आधार पर परिसंपत्तियों (assets) का एक समूह रखने के लिए एक निवेश निर्णय प्रक्रिया है। कई वित्तीय साधनों की उपलब्धता इसको जटिल बना देता है। पोर्टफोलियो ऑप्टिमाइजेशन तीन चरणों में किया जा सकता है: एसेट सिलेक्शन, एसेट वेटिंग (weighting) और एसेट मैनेजमेंट। पिछले शोध कार्य में हमेशा पोर्टफोलियो ऑप्टिमाइजेशन के पहले दो चरणों या प्रत्येक चरण पर अलग-अलग रूप से ध्यान केंद्रित किया गया है। इस अध्ययन का उद्देश्य सभी तीन चरणों पर पोर्टफोलियो ऑप्टिमाइजेशन के लिए व्यापक एंड -टू-एंड फ्रेमवर्क प्रदान करना है। इसके अलावा, यह कोवेरियन्स (covariance) मैट्रिक्स को बेहतर बनाने और स्टॉक की कीमत की भविष्यवाणी की सटीकता में सुधार करने से संबंधित मुद्दों को भी संबोधित करता है। नेशनल स्टॉक एक्सचेंज में सूचीबद्ध नॉन-फाइनेंसियल शेयरों के विश्लेषण के लिए काम का दायरा सीमित है।

अध्ययन का पहला चरण पोर्टफोलियो ऑप्टिमाइजेशन के लिए एक विश्लेषणात्मक रूपरेखा का प्रस्ताव करता है। इस प्रकार, पोर्टफोलियो के गठन के लिए संभावित शेयरों के चयन में सहायता के लिए डाटा इनवेलपमेंट एनालिसिस (data envelopment analysis) का उपयोग करके फर्मों (firms) के ऐतिहासिक प्रदर्शन का मूल्यांकन किया जाता है।

पोर्टफोलियो में प्रत्येक परिसंपत्ति के लिए आवंटित किए जाने वाले वजन (weight) का निर्धारण करने के लिए गेरबेर स्टेटिस्टिक्स (Gerber Statistics) (एचआरपी-जीएस – HRP-GS) पर आधारित हिरार्चिकल रिस्क पैरिटी (Hierarchical Risk Parity) मॉडल प्रस्तावित किया गया है। प्रस्तावित मॉडल के प्रदर्शन की तुलना तीन मॉडलों के साथ की गयी है, अर्थात्, ग्लोबल मिनिमम वेरियंस मॉडल बेस्ड ऑन हिस्टोरिकल कॉरलेशन (Global Minimum Variance based on historical correlation), ग्लोबल मिनिमम वेरियंस मॉडल बेस्ड ऑन गेरबेर स्टेटिस्टिक्स (Global minimum variance based on Gerber Statistics), और हिरार्चिकल रिस्क पैरिटी मॉडल बेस्ड ऑन हिस्टोरिकल कॉरलेशन (Hierarchical Risk Parity model based on historical correlation)। एचआरपी-जीएस मॉडल को मैट्रिक्स के अस्थिरता मुद्दों को संबोधित करने और कोवेरियन्स मैट्रिक्स के अनुमानों में सुधार करने का प्रस्ताव है।

छह एन्सेम्बल पूर्वानुमान मॉडलों का उपयोग किया है जो इस प्रकार हैं: एम्पिरिकल मोड डिक्मोज़िशन () - आर्टिफिशियल न्यूरल नेटवर्क (ईएमडी-एनएन-EMD-ANN), ईएमडी- सपोर्ट वेक्टर रिग्रेसन (ईएमडी - एसवीआर-EMD-SVR), एन्सबलबल ईएमडी - एनएन (EEMD-ANN), ईईएमडी-एसवीआर (EEMD-SVR), कम्पलीट एन्सेम्बल मोड डिक्मोज़िशन विथ अडाप्टिव नॉइज़ (सीईएमडीएन) -एनएनएन (CEEMDAN-ANN) और सीईएमडीएन-एसवीआर (CEEMDAN-SVR)। यह सारे मॉडल पोर्टफोलियो का गठन करने वाले शेयरों की कीमतों की भविष्यवाणी के लिए विकसित किए गए हैं। व्यापारिक नियमों

(trading rules) को डिज़ाइन किया गया है ताकि बाज़ारों की अलग-अलग स्थितियों के अनुसार शेयरों को खरीदने / बेचने / रखने के लिए निवेशकों को निर्देशित किया जा सके। भारतीय शेयर बाजार के आंकड़ों के डेटा पर किया गया प्रयोगात्मक विश्लेषण प्रस्तावित ढांचे की ताकत को दर्शाता है।

अध्ययन के दूसरे चरण में भारतीय शेयर बाजार में पोर्टफोलियो प्रबंधकों द्वारा अपनाई गई निवेश प्रक्रिया को समझने के लिए एक सर्वेक्षण किया गया है। सर्वेक्षण के परिणाम बताते हैं कि भारतीय शेयर बाजार में वित्तीय सिद्धांतों का कार्यान्वयन बहुत कम है।

Contents

Certificate	i
Acknowledgements	iii
Abstract	v
List of Figures	xiii
List of Tables	xv
Abbreviations	xvii

1	Introduction	1
1.1	Introduction	1
1.2	Motivation for the Research	4
1.3	Indian Stock Market	5
1.4	Research Objectives	10
1.5	Research Methodology and Scope	10
1.5.1	Research Methodology	10
1.5.2	Scope of the Study	13
1.5.3	Use of Programming Language and Statistical Software	14
1.6	Organization of Thesis	14
1.7	Chapter Summary	15
2	Literature Review	17
2.1	Introduction	17
2.2	Stock Selection	17
2.3	Portfolio Construction	22
2.3.1	Markowitz Mean-Variance Framework	23
2.3.1.1	Practical Factors	23
2.3.1.2	Estimation Error of Sample Parameters	26
2.3.2	Other Models for Portfolio Construction	29
2.4	Portfolio Management	31

2.4.1	Stock Price Prediction	31
2.4.2	Trading Decisions	38
2.5	Investment Practices Adopted by Practitioners	46
2.6	Gaps in Contemporary Literature	49
2.7	Concluding Observations	50
3	Selection of Profitable Stocks	51
3.1	Introduction	51
3.2	Data Envelopment Analysis	52
3.3	Data	55
3.4	Input and Output Parameters	56
3.5	Stock Selection Using DEA	59
3.6	Results and Discussion	59
3.7	Concluding Observations	62
4	Portfolio Construction	63
4.1	Introduction	63
4.2	Global Minimum Variance (GMV) Model	68
4.3	Gerber Statistics	69
4.4	Hierarchical Risk Parity Model	70
4.5	Data	74
4.6	Analytical Models	75
4.7	Results and Discussion	79
4.8	Concluding Observations	88
5	Construction of Stock Price Forecasting Model	91
5.1	Introduction	91
5.2	Ensemble Framework	93
5.2.1	Classification of Ensemble Forecasting Models	93
5.2.2	Proposed Ensemble Framework	96
5.3	Non-classical Decomposition Models	98
5.3.1	Empirical Mode Decomposition	98
5.3.2	Ensemble Empirical Mode Decomposition	100
5.3.3	Complete Ensemble Empirical Mode Decomposition with Adaptive Noise	101
5.4	Machine Learning Algorithms	103
5.4.1	Artificial Neural Network	103
5.4.2	Support Vector Regression	105
5.5	Performance Measures	107
5.5.1	Error Measure	107
5.5.2	Wilcoxon Signed Rank Test	108
5.5.3	Friedmann Test	109
5.6	Analytical Model	111
5.6.1	Data	111
5.6.2	Phase I: Non-classical Decomposition of Stock Price	112

5.6.3	Phase II: Prediction of Subseries Using Machine Learning Models	113
5.7	Results and Discussion	119
5.7.1	Performance Comparison of Ensemble Models with Traditional Models	125
5.7.2	Comparison of Performance among EMD-based ANN models and EMD-based SVR models	126
5.7.3	Comparison of Performance among EMD-based Ensemble Models	127
5.8	Concluding Observations	130
6	Determination of Timing of Trading	133
6.1	Introduction	133
6.2	Trading Rules	134
6.3	Illustration	136
6.4	Portfolio Performance	138
6.5	Concluding Observations	141
7	Survey of Investment Practices in Indian Stock Market	143
7.1	Introduction	143
7.2	Survey Methodology	144
7.2.1	Questionnaire Development and Identification of Target Respondents	144
7.2.2	Pre-Testing of Questionnaire (Reliability and Validity)	145
7.2.3	Sample Size Determination	146
7.2.4	Questionnaire Administration and Data Collection	147
7.3	Profile of Respondents	148
7.4	Analysis and Empirical Results	149
7.4.1	Stock Selection	150
7.4.2	Portfolio Construction	153
7.4.3	Portfolio Management	156
7.4.4	Algorithmic Trading	158
7.5	Concluding Observations	160
8	Summary and Conclusions	169
8.1	Introduction	169
8.2	Summary of the Work Done	170
8.3	Major Findings from the Research	172
8.3.1	Findings from Analytical Models	173
8.3.2	Findings from the Survey	174
8.4	Contributions of Research	175
8.4.1	Implications to Academics	176
8.4.2	Managerial Implications	177
8.5	Limitations of the Study and Scope for Future Research	178
8.6	Concluding Observations	179

References	181
Index	217
Appendix A Condition Number of Matrix	219
Appendix B Correlation-based Distance Metric	223
Appendix C A Survey on Investment Practices in Indian Stock Market	225
Publications	237
Short Biography	241

List of Figures

1.1	Timeline of Various Models Developed for Portfolio Optimization . . .	3
1.2	Flowchart Illustrating Research Design	11
2.1	Themes for Stock Selection	18
2.2	Themes for Portfolio Construction	22
2.3	Themes for Portfolio Management	31
3.1	Flowchart Showing Data Selection Process	57
4.1	Hierarchical Clustering of Stocks based on Covariance Obtained Using Gerber Statistics	78
4.2	Gerber Statistics Matrix Before Clustering (Year 2009)	79
4.3	Gerber Statistics Matrix After Clustering (Year 2009)	80
5.1	Classification of Ensemble Forecasting Techniques	94
5.2	Flow Chart of Ensemble Forecasting Framework	97
5.3	Architecture of ANN Model	104
5.4	Training and Testing of a ANN Model Using BP algorithm	105
5.5	Schematic Diagram of SVR	105
5.6	Daily Stock Price of PersonalProd_5 for the Year 2011	112
5.7	Decomposed Components of PersonalProd_5 Using EMD (Year 2011)	113
5.8	Decomposed Components of PersonalProd_5 Using EEMD (Year 2011)	114
5.9	Decomposed Components of PersonalProd_5 Using CEEMDAN (Year 2011)	115
5.10	ACF and PACF of IMF_3 Obtained Using EMD for PersonalProd_5	116
5.11	Predicted Values Using ANN and Ensemble Models (Year 2016)	120
5.12	Predicted values using SVR and ensemble models (Year 2016)	121
6.1	Plot Showing the Varying Number of Stocks in the Portfolio (Year 2016)	140
7.1	Business Turnover of the Participating Firms	148
7.2	Area of Specialization	149
7.3	Experience of the Respondents	149
7.4	Professional Education of the Respondents	150
7.5	Seniority in the Firm	150

7.6	Percentage of Responses on Usage of Markowitz's Mean-Variance Framework	153
7.7	Consideration of Impact of Financial News on Prediction of Stock Price/Returns	157
7.8	Implementation of Algorithmic Trading	159
8.1	Flowchart Showing Summary of the Work Done	171

List of Tables

1.1	Trends of Equity Segment of BSE.	8
1.2	Trends of Equity Segment of NSE.	9
2.1	Summary of literature review on portfolio optimization	41
2.2	Summary of Studies on Survey of Investment Practices.	49
3.1	Sectors and Number of Firms.	56
3.2	Input and Output Parameters for DEA Analysis.	58
3.3	Efficient Firms under Automotive Sector (Years 2008-15)	60
3.4	Number of Efficient Firms under Different Sectors (Years 2008-15)	62
4.1	Weights of MV-HC and MV-GS Portfolios (Year 2012)	80
4.2	Weights of HRP-HC and HRP-GS Portfolios (Year 2012)	81
4.3	Various Measures of Portfolio (Years 2008-15).	88
5.1	Network Structure of ANN Used in All Three Ensemble Models	117
5.2	Error Measures (Year 2016)	122
5.3	WSRT Results between Ensemble Models and Traditional ANN Model	126
5.4	WSRT Results between Ensemble Models and Traditional SVR Model	126
5.5	WSRT Results of Ensemble EMD-based ANN Vs Ensemble EMD-based SVR Models	128
5.6	Friedman Test and post hoc Nemeyi Test	129
6.1	Trading Rules for Investment Decisions	135
6.2	Illustration of Trading Rules for PersonalProd_5 (Year 2016)	136
6.3	Returns Obtained Using Trading Rules and Buy-and-Hold Strategy	140
7.1	Statistics of World Stock Exchange (2016).	163
7.2	Primary Business of the Participant's Firm.	164
7.3	Methods Adopted for Selection of Stocks.	164
7.4	Factors Considered for Stock Selection.	164
7.5	Cross-tabulation on Awareness and Usage of DEA for Stock Selection Problem.	165
7.6	Approaches Adopted for Portfolio Formation Process.	165
7.7	Observations on Markowitz Mean Variance framework.	165
7.8	Factors Considered During Markowitz MV Optimization.	165
7.9	Error Mitigation During Portfolio Construction Process.	166

7.10 Risk Measures Considered During Portfolio Formation Process. . . .	166
7.11 Approaches Adopted for Calculating Extreme Risk Measures. . . .	166
7.12 Measures Used for Performance Measurement.	166
7.13 Methods Used for Rebalancing the Portfolio.	167
7.14 Frequency of Rebalancing the Portfolio.	167
7.15 Methods Adopted for Prediction of Stock Prices/Returns.	167
7.16 Frequency of Data Used for Prediction of Stock Prices/Returns. . .	168
7.17 Error Measures Used for Measuring the Performance of Prediction Models.	168
7.18 Reasons for Implementing Algorithmic Trading.	168
7.19 Reasons for Not Implementing Algorithmic Trading.	168

Abbreviations

ABC	Artificial Bee Colony
AHP	Analytic Hierarchy Process
ANN	Artificial Neural Network
ANP	Analytic Network Process
ARCH	AutoRegressive Conditional Heteroskedasticity
ARMA	AutoRegressive Moving Average
ARIMA	AutoRegressive Integrated Moving Average
BCC	Banker Charnes Cooper
BSE	Bombay Stock Exchange
CCR	Charnes Cooper Rhodes
CEEMDAN	Complete Ensemble Empirical Mode Decomposition with Adaptive Noise
CRS	Constant Returns to Scale
DA	Directional Accuracy
DE	Differential Evolution
DEA	Data Envelopment Analysis
DJIA	Dow Jones Industrial Average
DMU	Decision Making Units
DWT	Discrete Wavelet Transform
EOB	Electronic Order Book
EMD	Empirical Mode Decomposition
EEMD	Ensemble Empirical Mode Decomposition
FA	Firefly Algorithm
GA	Genetic Algorithm

GARCH	G eneralised A uto R egressive C onditional H eteroskedasticity
GMV	G lobal M inimum V ariance
GP	G enetic P rogramming
GS	G erber S tatistics
HRP	H ierarchical R isk P arity
IMF	I ntrinsic M ode F unction
MAD	M ean A bsolute D eviation
MAE	M ean A bsolute E rror
MCDA	M ulti C riteria D ecision A nalysis
MSE	M ean S quare E rror
NSE	N aional S tock E xchange
NMSE	N ormalised M ean S quare E rror
NYSE	N ew Y ork S tock E xchange
PSO	P article S warm O ptimization
OWA	O rdered W eighted A veraging
RMSE	R oot M ean S quare E rror
ROA	R eturn O n A sset
ROE	R eturn O n E quity
ROI	R eturn O n I nterest
RSE	R egional S tock E xchange
SA	S imulated A nnual
SEBI	S ecurities and E xchange B oard of I ndia
SVM	S upport V ector M achine
SVR	S upport V ector R egression
TS	T abu S earch
VRS	V ariable R eturns to S cale
WSRT	W ilcoxon S igned R ank T est