MASSEY UNIVERSITY
COLLEGE OF SCIENCES

Declaration Confirming Content of Digital Version of Thesis

I confirm that the content of the digital version of this thesis

Title: **Equity Trend Prediction with Neural Networks: An Empirical Analysis**

is the final amended version following the examination process and is identical to this hard bound paper copy.

Have you included published articles/material from your thesis? Yes / No

If yes, have you received copyright permission from the copyright holder (usually the publisher) to include this material with your thesis? Yes / No

Student's Name: **Russell James Halliday**

Student's Signature: RHalliday

Date: **25/02/2012**

Equity Trend Prediction with Neural Networks: An Empirical Analysis

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Engineering

in

Computer Systems Engineering

at Massey University, Albany,

New Zealand.

Russell James Halliday

2012

Acknowledgements

I wish to give my utmost gratitude to my supervisor Dr Mohammad Rashid and Dr Chris Messom. Dr Chris Messom gave me the guidance and input in steering the initial direction of the thesis, while Dr Rashid was instrumental in guiding it's final form.

I also wish to thank my wife Angie for her encouragement, support, patience and understanding at all times. Finally I thank my parents, brother and sisters for their unwavering support.

Abstract

This thesis presents results of neural network based trend prediction for equity markets. Despite a breadth of research which has focused on the prediction of various equity and currency exchange markets, much has focused on the use of specific techniques in such predictions. Few bodies of work have compared a wide range of equity market data preprocessing and technical analysis techniques in creating a prediction model based on feed-forward and recursive neural networks.

To achieve a broad-based prediction model, the work in this study was broken into three distinct parts. Firstly, the neural networks goal is defined as finding whether a stock will be higher or lower than the previous trading period. Subsequent to this, a variety of input data scaling and network topologies are looked at. This includes the use of Self Organising Maps (SOM) as a data classification method to limit neural network inputs and training data requirements. Feed Forward and Elman networks of various topologies are used to narrow down the best network combinations. The resulting simulation is a neural network that can predict whether the next trading period will be, on average, higher or lower than the current.

Secondly, the topology and preprocessing lessons learned during the first phase are applied to two types of neural network. Technical analysis is applied to the input data in an attempt to verify the usefulness of conventional stock indicators as inputs to neural networks. The two types of networks trained are, for the purposes of this thesis, dubbed indicator-predictive and price-predictive networks, meaning that technical analysis inputs are used to predict the next trading days technical indicator or future stock price direction respectively.

Finally, a combined network is trained which takes the inputs from the price-predictive networks in an attempt to gain better results. The hypothesis with this network is that the combined neural network should learn which of its inputs are more indicative of a stock price movement, and thus more accurately predict the future direction of the stock.

Table of Contents

ACKNOWLEDGEMENTS....................II

ABSTRACT....................III

TABLE OF CONTENTS....................IV

LIST OF ABBREVIATIONS....................VII

LIST OF MATLAB FUNCTIONS....................VIII

LIST OF TABLES....................IX

LIST OF FIGURES....................XI

CHAPTER 1 - INTRODUCTION....................15

1.1 INTRODUCTION....................15
1.2 BACKGROUND....................15
1.3 FORMULATION OF THE RESEARCH PROBLEM....................15
1.4 SCOPE AND LIMITATION....................16
1.5 METHODOLOGY....................16
1.6 THESIS ORGANISATION....................17

CHAPTER 2 - ARTIFICIAL NEURAL NETWORKS AS A PREDICTION TOOL....................18

2.1 STOCK MARKET....................21
2.2 EQUITY MARKETS....................21
2.3 INVESTMENT VERSUS TRADING....................22
2.4 TECHNICAL ANALYSIS....................22
2.5 DATA COLLECTION AND ADJUSTMENT....................23
2.5.1 Pre-processing....................24
2.5.2 Training....................26
2.6 NEURAL NETWORK DESIGN AND FEASIBILITY....................26
2.7 LITERATURE REVIEW....................27
2.8 SUMMARY....................32

CHAPTER 3 - PRELIMINARY NETWORK TESTING AND ANALYSIS....................33

3.1 DATA SELECTION....................33
3.2 DATA PRE-PROCESSING....................34
3.3 SELF ORGANISING MAP & POST PROCESSING....................36

3.4 Neural Network....36
3.5 Results....38
3.5.1 Mechanical Prediction....*40*
3.5.2 Neural Network Outputs....*40*
3.5.3 Data Pre-Processing....*40*
3.5.4 Self Organising Maps....*41*
3.5.5 Neural Network Type and Configuration....*43*
3.5.6 Mechanical Averaged-Based Prediction....*46*
3.6 Analysis....46
3.7 Summary....48
CHAPTER 4 - INDICATOR, PRICE AND COMBINED NETWORK TOPOLOGY....49
4.1 Pre-Processing - Technical Indicator Network Training....51
4.1.1 Buy / Sell / Hold....*53*
4.1.2 Top / Bottom and Trend Indicators....*55*
4.1.3 Processed Open-High-Low-Close-Volume Data....*56*
4.2 Summary....57
CHAPTER 5 - RESULTS AND ANALYSIS....58
5.1 Training Days....58
5.2 Training Epochs....63
5.3 Network Function....63
5.4 Training Goal....66
5.5 Input training period....69
5.6 Input Scaling....71
5.7 Training Step Size....74
5.8 Target Scaling....74
5.9 Test Set Size....76
5.10 Network Topology....76
5.11 Network Training Type....79
5.12 Network Type....82
5.13 Validation Set Size....84
5.14 Summary....86
CHAPTER 6 - COMBINED NETWORK....88
6.1 Network Topology....89
6.2 Results....91
6.3 Analysis....92
6.4 Summary....93

CHAPTER 7 - CONCLUSION AND FUTURE WORK......94

REFERENCES......99

APPENDIX: PREPROCESSING FUNCTION DESCRIPTIONS......108

A1. Inter-day Difference......108
A2. Intra-day Difference......109
A3. Moving Average Crossover Indicator......110
A4. Moving Average Decayed Crossover Indicator......112
A5. Moving Average Difference Indicator......114
A6. Relative Strength Index (RSI)......114
A7. Relative Strength Index (RSI) Analysis......116
A8. Bollinger Bands......117
A9. Volatility Analysis......118
A10. On Balance Volume......121
A11. Accumulation Distribution Line......122
A12. Aroon Oscillator......124
A13. Commodity Channel Index......125
A14. Bullish/Bearish Moving Average Convergence Divergence......126
A15. Zigzag Filter......128
A16. Price Channels......129
A17. Williams %R......130
A18. Ultimate Oscillator......132
A19. TRIX Data......133
A20. Standard Deviation......134
A21. Money Flow Index......134
A22. Chaikin Oscillator......136
A23. Rate of Change......137
A24. Money Envelope Analysis......137

List of Abbreviations

Various specialised abbreviations are used in this thesis as listed below:

ADL.. Accumulation Distribution Line

ANN...Artificial Neural Network

BSE...Bombay Stock Exchange

BSH..Buy Sell Hold

MLP..Multi Layer Perceptron

MACD...Moving Average Convergence Divergence

NYSE...New York Stock Exchange

OBV...On Balance Volume

OHLCV...Open High Low Close Volume

RNN...Recursive Neural Network

RMSE...Root Mean Square Error

SOM...Self Organising Map

List of Matlab Functions

BINARY..Binary Scaling

ELMAN...Elman network function

LINEARNORM..Linear Normalisation Scaling

LINEARSCALE...Linear Scaling

LINEARCLIP...Linear Clipped Scaling

LOGARITHMIC...Logarithmic Scaling

LOGSIG...Logarithmic Sigmoid

NEWFFD...Feed-forward network function

RNN...Recursive Neural Network

SOFTMAX..Softmax Scaling

TANSIG...Tangent Sigmoid function

TRAINDX...Standard Back Propagation function

TRAINLM...Levenberg-Marquardt Algorithm

TRAINRP...Resilient Back Propogation

TANH...Hyperbolic Tangent

List of Tables

Table 3.5.1: Effectiveness of Pre-processing..41
Table 3.5.2: Effectiveness f Post-Processing Normalisation on Self Organising Maps. .42
Table 3.5.3: Effect of Introduced Error on SOM Stage..42
Table 3.5.4: Effectiveness of SOM dimensions..42
Table 3.5.5: Average Success with SOM Stage Removed..43
Table 3.5.6: Effect of Neural Network Type and Number of Hidden Layers on Network Performance...44
Table 3.5.7: Network Predictive Success with 3 Hidden Neurons.................................44
Table 3.5.8: Network Predictive Success with 5 Hidden Neurons.................................45
Table 3.5.9: Network Predictive Success with 10 Hidden Neurons...............................45
Table 3.5.10: Simple Average Trend Prediction...46
Table 4.1: Weights and Training Data required for Feed Forward Network with a 0.1 Error Target...50
Table 4.2: Weights and Training Data required for Elman Network with a 0.1 Error Target...50
Table 4.1.1: Network input functions..52
Table 5.1.1: Calculation of training data required for given network types and topologies...59
Table 5.14.1: Summary or results..86
Table 6.1.1: Labelling of data to 'higher' or 'lower' based on 0.5 reference....................89
Table 6.1.2: Labelling of 'higher' and 'lower' points identified in table 13 are assigned binary values of 1 or 0 respectively...90
Table 6.2.1: Combined Network Results for Network Size Variation (hard limits 0.8 & 0.2)...91
Table 6.2.2: Combined Network Results for Network Type Variation (hard limits 0.8 & 0.2)...92
Table 6.2.3: Combined Network Results for Network Size Variation with no hard limits 92
Table 6.2.4: Combined Network Results for Network Type Variation with no hard limits 92

Table 6.3.1: Summary of Average Network Performance for Combined Network.........93
Table 7.1: Comparison of results with other studies..96

List of Figures

Figure 2.1: Simple Multilayer Perception...18
Figure 2.2: Elman Network...19
Figure 2.3: Simplified Diagram Illustrating Back Prorogation Learning (Tveter, 2001) 20
Figure 2.4: Self Organising Map Topology...20
Figure 2.5: Self Organising Map Training (KOHONEN94)..21
Figure 2.5.1: Difficulty in generalising from raw training data. Generalisation of section B from A is difficult given the entirely different trend and magnitude of the data..........25
Figure 2.5.2: Normalisation of Source Data...25
Figure 3.1: Network Architecture...33
Figure 3.4.1: Thee chunks of 20 days with a 10-day increment.....................................37
Figure 3.4.2: Elman Network..38
Figure 3.4.3: Log-Sigmoid Transfer Function..38
Figure 3.5.1: Predictive Accuracy vs. Predictive Range for Neural Network.................45
Figure 3.5.2: Predictive Accuracy vs. Predictive Reference for Neural Network...........46
Figure 4.1: Final Neural Network Architecture chosen for prediction implementation. .49
Figure 4.2: Final Neural Network Architecture chosen for prediction implementation. .49
Figure 5.1.1:Distribution of results based on Performance summary variation in training days...60
Figure 5.1.2: Number of training days required by versus network type........................60
Figure 5.1.3: Distribution of Price and Indicator performance for various numbers of training days...61
Figure 5.1.4: Distribution of results for Input period variation performance given variation in training days...62
Figure 5.1.5: Price and Indicator combination performance given variation in training days..62
Figure 5.3.1: logsig Function..63
Figure 5.3.2: transit Function..64
Figure 5.3.3: Distribution of results for variation in network function..........................64
Figure 5.3.4: Comparison of results distribution between price and indicator networks 65
Figure 5.3.5: Comparison of results distribution for different network topologies.........65

Figure 5.3.6: Comparison of results for different network topologies and type (price or indicator)....66
Figure 5.4.1: Distribution of results for training goals of 1%, 5% and 0.1%....67
Figure 5.4.2: Comparison of price and indicator network type optimal results for given training goals....67
Figure 5.4.3: Comparison of network topology optimal results for given training goals 68
Figure 5.4.4: Comparison of network topology and type for given training goals....68
Figure 5.5.1: Input period optimal results distribution for given training periods of 1, 5 and 10 days....69
Figure 5.5.2: Input period optimal results distribution for price and indicator networks given training periods of 1, 5 and 10 days....70
Figure 5.5.3: Network topology optimal results distribution for price and indicator networks given training periods of 1, 5 and 10 days....70
Figure 5.5.4: Input period and type optimal results distribution comparison for price and indicator networks given training periods of 1, 5 and 10 days....71
Figure 5.6.1: Comparison of optimal results for various types of input scaling....72
Figure 5.6.2: Distribution of optimal results for price and indicator networks given various input scaling functions Graph results....72
Figure 5.6.3: Distribution of optimal results for input period given various input scaling functions....73
Figure 5.6.4: Distribution of optimal results for variations in network type and input periods given various input scaling functions....73
Figure 5.8.1: Distribution of optimal results given various types of target scaling....74
Figure 5.8.2: Distribution of optimal results for price and indicator networks given various types of target scaling....75
Figure 5.8.3: Distribution of optimal results of various network input periods given various types of target scaling....75
Figure 5.8.4: Distribution of optimal results for different network types and input period given various types of target scaling....76
Figure 5.10.1: Distribution of optimal results given various types of network topology 77

Figure 5.10.2: Price and Indicator performance given variation in topology..................77
Figure 5.10.3: Distribution of optimal results given various input periods for different types of target scaling...78
Figure 5.10.4: Distribution of optimal results for different network types and input periods given various types of target scaling...78
Figure 5.11.1: Distribution of optimal results given various network training functions 80
Figure 5.11.2: Distribution of optimal results for different network types given various network training functions...80
Figure 5.11.3: Distribution of optimal results for different training periods given various training functions...81
Figure 5.11.4: Distribution of optimal results given various network training functions for various training functions...81
Figure 5.12.1: Distribution of optimal results given different network types (Elman and Feed Forward)...82
Figure 5.12.2: Distribution of optimal results for price and indicator networks given various network types..82
Figure 5.12.3: Distribution of optimal results for different training windows given various network types..83
Figure 5.12.4: Distribution of optimal results for different network training windows and types given various network types..83
Figure 5.13.1: Distribution of optimal results with and without validation sets...........84
Figure 5.13.2: Distribution of optimal results for price and indicator networks with and without validation sets used for training..85
Figure 5.13.3: Distribution of optimal results for different training windows with and without validation sets used for training..85
Figure 5.13.4: Distribution of optimal results for different network types and training windows, with and without validation sets used for training...86
Figure 6.1: Simplified Topology of Combined Network...88
Figure 6.3.1: Average Network Performance for Combined Network...........................93
Figure A1.1: Inter-day Differencing for AMR Corporation Jan-Feb 1985...................109
Figure A2.1: : Intra-day Difference Function applied to AMR Corporation Jan-Feb 1985

....110
Figure A3.1: Moving Averages help to reveal the underlying trend of the data...........111
Figure A3.2: Crossover Function with Buy/Sell Signal Generation.............................112
Figure A4.1: Decayed Cross Over Function with Buy/Sell Signal Generation............113
Figure A5.1: : Decayed Cross Over Function with Buy/Sell Signal Generation...........114
Figure A8.1: : Bollinger Bands...118
Figure A10.1: : On Balance Volume applied to US Steel (August 2001 - August 2002)
....122
Figure A11.1: The Application of Accumulation Distribution Line (ADL) to Exxon Mobil Corp...124
Figure A14.1: MACD Analysis Applied to Nike for May to August 2002...................128
Figure A15.1: Zigzag Analysis with 5% threshold...129
Figure A16.1: Price Channel Applied to Pepsi Co. with 10 day price channel.............130
Figure A17.1: Williams %R applied to Microsoft Corp from May to August 2002.….131
Figure A20.1: : Standard Deviation Function Output...134
Figure A24.1: : Money Envelope Analysis...138