Faculty of Information and Communication Technology

THREE-POINT MOVING GRADIENT WITH LONG SHORT-TERM MEMORY FOR STOCK PRICES TIME SERIES FORECASTING

Wong Pee Shoong

Doctor Of Philosophy

2025

THREE-POINT MOVING GRADIENT WITH LONG SHORT-TERM MEMORY FOR STOCK PRICES TIME SERIES FORECASTING

WONG PEE SHOONG

A thesis submitted
in fulfillment of the requirements for the degree of Doctor of Philosophy

Faculty of Information and Communication Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2025

DECLARATION

I declare that this thesis entitled "Three-Point Moving Gradient With Long Short-Term Memory For Stock Prices Time Series Forecasting" is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature: ……………………………

Name: WONG PEE SHOONG

Date: 19 MAY 2025

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of Doctor of Philosophy.

Signature: ………………………………………………………

Supervisor Name: Ts. Dr. SITI AZIRAH ASMAI

Date: 20 MAY 2025

DEDICATION

To my lovely and beloved wife, Jasmine Wong Leng Syn.

ABSTRACT

The forecasting the stock market has always been a fascination among investors and speculators. They wish to know what the stock market is like in the future. This research is on the use of price gradients in forecasting. The significance of this research is that it is to provide an alternative of forecasting time series prices technique apart from the standard techniques such as Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Convolutional Neural Network (CNN) etc. The objectives of this research is to propose Three-Point Moving Gradient (TPMG) with Long Short Term Memory (LSTM) technique in stock price forecasting to have linkage between the immediate past, the present and the immediate future values in standard LSTM, and to evaluate the performance accuracy of the proposed TPMG with the standard LSTM and CNN for the actual Kuala Lumpur Stock Exchange (KLSE) stock price. The main aim of this research is to propose an alternative technique that is Three-Point Moving Gradient with LSTM (TPMG-LSTM) to forecasting time series stock prices. The key hypothesis of this study is that TPMG-LSTM technique can provide more accurate forecasting than the standard LSTM or CNN. This research feeds daily market stock prices into LSTM, CNN and price gradients into TPMG-LSTM networks and then compare the forecasting results of each network to find out which is closest to the observed price. The main data source is the webpage https://www.malaysiastock.biz/Market-Watch.aspx and the sample size is 245 counters out of the population of 2845 counters in KLSE. The sampling technique involves counters with minimum 80 price points. From the forecasting testing data collected, TPMG-LSTM Residual Analysis (RA) (-0.0200) is closer to the real data than LSTM RA (-0.0306) or CNN RA (0.4480). TPMG-LSTM Coefficient of Determination (R^2) (-0.4782) is closer to the real data than LSTM-LSTM R^2 (-4.1052) or CNN R^2 (-12.4956). TPMG-LSTM Mean Absolute Error (MAE) (0.0391) is closer to the real data than LSTM MAE (0.1179) or CNN MAE (0.4831). TPMG-LSTM Root Mean Square Error (RMSE) (0.0486) is closer to the real data than LSTM RMSE (0.1399) or CNN RMSE (0.5190). All results show that TPMG-LSTM produces forecasting closest to the observed data compared to standard LSTM or CNN. The main conclusion of this research is that TPMG-LSTM provides a more accurate stock price forecasting than standard LSTM or CNN.

ABSTRAK

GRADIEN BERGERAK TIGA TITIK BERSAMA INGATAN JANGKA PANJANG PENDEK UNTUK RAMALAN SIRI MASA HARGA SAHAM

Meramal pasaran saham selalunya merupakan satu daya tarikan di kalangan pelabur-pelabur dan spekulator-spekulator. Mereka ingin tahu bagaimana keadaan pasaran saham di masa hadapan. Kajian ini bertumpu kepada penggunaan gradien harga dalam ramalan. Kepentingan kajian ini adalah ia membekalkan satu alternatif dalam teknik ramalan siri masa selain dari teknik-teknik yang standard seperti Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Convolutional Neural Network (CNN) dan lain-lain. Objektif kajian ini adalah untuk mencadangkan Gradien Bergerak Tiga Titik (TPMG) Bersama Ingatan Jangka Panjang Pendek (LSTM) dalam famalan harga saham yang mempunyai kaitan antara masa lalu serta merta, masa kini serta merta dan masa depan serta merta, dan menilai kejituan prestasi TPMG yang dicadangkan dengan LSTM dan CNN standard untuk harga saham Pasaran Saham Kuala Lumpur (KLSE). Tujuan utama kajian ini ialah untuk mencadangkan satu teknik alternatif iaitu teknik TPMG bersama LSTM (TPMG-LSTM) meramal harga saham siri masa. Hipotesis utama dalam kajian ini ialah teknik TPMG-LSTM boleh memberikan satu ramalan yang lebih jitu dari LSTM atau CNN standard. Kajian ini memasukkan harga saham pasaran harian ke dalam LSTM, CNN dan gradien harga ke dalam rangkaian TPMG-LSTM

dan membandingkan keputusan ramalan setiap rangkaian untuk mengetahui mana satu yang paling dekat dengan harga yang diperhatikan. Sumber data utama ialah laman web https://www.malaysi-astock.biz/Market-Watch.aspx dan saiz sampel ialah 245 kaunter dari populasi sebanyak 2845 kaunter di KLSE. Teknok persampelan melibatkan kaunter-kaunter dengan minima 80 titik harga. Dari data ujian ramalan terkumpul, TPMG-LSTM Residual Analysis (RA) (-0.0200) adalah lebih dekat kepada data sebenar dari LSTM RA (-0.0306) atau CNN RA (0.4480). TPMG-LSTM Coefficient of Determination (R^2) (-0.4782) adalah lebih dekat kepada data sebenar dari LSTM R2 (-4.1052) atau CNN R2 (-12.4956). TPMG-LSTM Mean Absolute Error (MAE) (0.0391) adalah lebih dekat kepada data sebenar dari LSTM MAE (0.1179) atau CNN MAE (0.4831). TPMG-LSTM Root Mean Square Error (RMSE) (0.0486) adalah lebih dekat daripada data sebenar dari LSTM RMSE (0.1399) atau CNN RMSE (0.5190). Semua keputusan menunjukkan bahawa TPMG-LSTM menghasilkan ramalan yang paling dekat dengan data yang diperhatikan berbanding dengan LSTM atau CNN standard. Kesimpulan utama kajian ini adalah TPMG-LSTM memberikan satu ramalan lebih jitu berbanding dengan LSTM atau CNN standard.

ACKNOWLEDGEMENTS

First and foremost, I would like to take this opportunity to express my sincere acknowledgement to my supervisor Ts. Dr. Siti Azirah Asmai from the Faculty of Artificial Intelligence and Cyber Security (FAIX), Universiti Teknikal Malaysia Melaka (UTeM) for her essential supervision, support and encouragement towards the completion of this thesis.

I would also like to express my greatest gratitude to Dr. Tay Choo Chuan from Faculty of Electrical Engineering, co-supervisor of this research for his advice and suggestions in RMSE test.

My gratitude also goes to Dr. Norhazwani Binti Md Yunos and Dr. Nur Zareen Binti Zulkarnain for supervising my Proposal Defence while Dr. Zahriah Binti Sahri for supervising my Progress Report Presentation.

Special thanks to all my colleagues and my beloved wife, Jasmine Wong Leng Syn for their sacrifices, material and moral support in completing this degree. Lastly, thank you to everyone who had been associated to the crucial parts of realization of this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

LSTM	-	Long Short-Term Memory
TPMG	-	Three-Point Moving Gradient
ANN	-	Artificial Neural Network
RNN	-	Recurrent Neural Network
GRU	-	Gated Recurrent Unit
NLP	-	Natural Language Processsing
RA	-	Residual Analysis
MAE	-	Mean Absolute Error
RMSE	-	Root Mean Square Error
EOD	-	End Of Day
FVA	-	Forecast Value Added
MA	-	Moving Average
VAR	-	Vector Auto Regression
ARMA	-	AutoRegressive Moving Average
ARIMA	-	AutoRegressive Integrated Moving Average

LIST OF PUBLICATIONS

The following is the list of publications related to the work in this thesis:

Shoong W.P., Asmai S. A., and Chuan, T. C., 2020. Stock Prices Time Series Forecasting by Deep Learning Using Three-Point Moving Gradient. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(4), pp. 6622 - 6630.

Shoong W. P., Asmai S. A., Zulkarnain N. Z., and Chuan, T. C., 2022. An Improved LSTM Technique Using Three-Point Moving Gradient for Stock Price Forecasting. *International Journal of Computer Information Systems and Industrial Management Applications*, 14 (2022), pp. 338 - 346.

CHAPTER 1

INTRODUCTION

1.1 Background of the Problem

In recent years, Deep Learning (DL) has transformed artificial intelligence, leading to cutting-edge performance on a range of tasks. (Al-lQubaydhi et al., 2024). When processing sequential data, such as text or time series, Recurrent Neural Networks (RNNs), are better at capturing these sequence links than classic feed-forward networks, which have trouble with sequence processing (Beltozar-Clemente and Iparraguirre-Villanueva, 2024). RNN transfers data back into itself and has cycles Schmidt (2019). RNNs have gained a lot of traction in research domains that deal with sequential input, such as text, audio, and video. However, traditional RNNs with sigmoid or tanh activation functions struggle to grasp long-term dependencies when the input data contains a significant gap (Yu, 2019). RNN has problems of recalling older information and long-term dependency due to vanishing and exploding gradients (Waqas and Humphries, 2024). In addition, RNN also has slow training speed (Zhang, 2023).

Nevertheless, there is a special class of RNN called LSTM (Sherstinsky, 2020). LSTM has been used for a long time to estimate stock prices, and its impressive performance is well recognized. Studies show that LSTM performs well in stock market forecasting. incorporating gate functions into the cell structure, LSTM could efficiently handle the problem of long-term dependency (Zhao, 2020). The ability of LSTMs to process sequential data and retain information from earlier steps in the sequence allows them to efficiently anticipate subsequent steps. They can efficiently forecast future steps by retaining information from earlier steps in the sequence (Al-Selwi et al., 2024). In short, a conventional RNN can process long sequence data more effectively because due to the gating mechanism introduced by the LSTM neural network. This is advantageous for simulating nonlinear systems. This type of RNN also successfully resolves the issue of gradient vanishing and exploding (Kang et al., 2023). When learning long-term dependencies, exploding and vanishing gradient problems usually occur. This is addressed by the powerful recurrent neural system known as the LSTM model, even in cases when the minimal time lags are quite lengthy (Houdt, 2020). Sequence modelling benefits greatly from the use of RRNs. They are extensively versatile and can contain a range of information kinds, including temporal order (Zargar, 2021). These qualities make them perfect for generating a list of recommendations. In essence, LSTM is a RNN that is trained to forecast a value by finding the greatest fit between the past and present data. To best fit its forecast, LSTM makes use of its historical data. The LSTM achieves this by using internal gates to regulate its internal memory and backpropagating previously collected data as input (Elsayed et al., 2020). The use of internal gates helps LSTM to overcome the problem of vanishing and exploding gradient.

In time series stock price forecasting, LSTM is not a novel or obscure technique. Deep recurrent neural networks, such as LSTM, perform better than deep feedforward neural networks and efficiently use model parameters. They also accumulate quickly (Ayitey Junior et al., 2022). An LSTM-based method for forecasting stock returns was proposed by Chen et al. (2015), as cited in Gao (2024) and it was validated on the mainland stock market of China. They verified that the performance of the LSTM model was better than that of the random forecasting method. Using characteristics like open, close, high, and low prices, Roondiwala et al. (2015) also implemented an LSTM network to forecast India's Nifty prices with some degree of success, as cited in Patil et al. (2021). They were reasonably accurate in forecasting the open, close, high, and low prices. Di Persio and Honchar (2016) evaluated the relative performance of three distinct RNN models - the RNN, the LSTM, and the GRU - on the price of Google stock in 2016, as cited in Phuoc et al. (2024). It appeared that LSTM performed better than the others. To anticipate the starting price of ten distinct firms, Akita et al. (2016) gathered recorded data from a Japanese newspaper to be used as LSTM input in conjunction with market time-series data. They were able to accurately forecast the opening prices of each of the ten distinct companies. Based on price history and technical analysis indications, Nelson, D.M., as cited in Hu et al. (2021), suggested an LSTM model. The suggested model outperformed the MLP and Random Forest models, according to the findings. The implementation's efficacy has been evaluated in comparison to multi-layered perceptron baselines, random forest, and pseudo-random models. Bao et al. (2017) applied the Haar wavelet transformation to denoise the financial time series data and implemented the stacked autoencoders to learn the deep features of the data and then used LSTM to forecast the closing price of stock indices, as cited in Bhandari et al. (2022). To forecast CSI300 and sentiment, Li et al. (2017) used investor opinion from forum postings and combined it with historical market data into an LSTM network, as cited in Karlis et al. (2021). The research found that LSTMs outperformed benchmark models and that adding sentiment characteristics resulted in notable improvements in accuracy (from 78.57% to 87.86%) when forecasting the open price the following day. This demonstrated the superior accuracy of LSTM over other models. To forecasting NSEI listed equities, Selvin et al compared three distinct Deep Learning architectures: CNN-sliding window models, LSTM, and RNN, as cited in Fathali et al. (2022). LSTM appeared to do better than the rest. Yan and Ouyang (2017) demonstrated that LSTM is a better strategy to estimate the daily returns of the shanghai composite index as cited in Arslan et al. (2023). Kim and Won (2018) built a hybrid model to forecast the volatility of the Korean stock price index (KOPSI 200) by fusing GARCH-type models with the LSTM model, as cited in Amirshahi and Lahmiri (2023).

A hybrid LSTM and GRU Deep Learning model was suggested by Hossain et al (2019) to forecast the American S and P 500 time series using data from 1950 to 2016. To produce the first-level forecasting, the procedure, as mentioned by Dael et al. (2023), entails feeding the input data into an LSTM network and then shifting the LSTM layer's output to the GRU layer. With a focus on LSTM forecasting, Sezer et al. (2020) wrote a study titled "Financial Time Series Forecasting with Deep Learning: A Systematic Literature Review: 2005-2019". Many Deep Learning model developers are found to favour LSTM networks for solving challenging tasks like handwritten character recognition and automatic speech. Time-series data are typically utilized with these LSTM models. They can be used in financial time series analysis, sentiment analysis, speech recognition, language modelling, natural language processing (NLP), speech recognition, and predictive analysis. An article titled "Forecasting stock prices with Long Short-Term Memory neural network based on attention mechanism" was published by Qiu et al. (2020). Text processing and speech recognition are two areas where LSTM neural networks excel. In addition, LSTM neural networks are appropriate for random nonstationary sequences like stock-price time series due to their selectivity and memory cells. Lin et al. (2021) use LSTM combined with Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) to forecast the stock index price of Standard and Poor’s 500 Index (SandP500) and the China Security 300 Index (CSI300). This model performed well with great accuracy. Wu et al. (2021) proposed a new framework structure combining Convolutional Neural Network with LSTM Neural Network to forecast 10 stocks in USA and Taiwan. Jabeen et al. (2021) proposed a LSTM method of forecasting for major stock sectors using COVID sentiment.

For LSTM, there are a few success stories. A recurrent LSTM network was suggested by Wiese (2007) as a method for identifying fraudulent credit card transactions. Recurrent networks should be adept at identifying fraudulent patterns in card transaction sequences, according to the authors, as they were intended to handle sequences. The authors were able to show how the LSTM method is superior to the SVM algorithm (Afanasiev, 2024). Using LSTM, Roondiwala et al. (2015) were able to forecast India's NIFTY50 with remarkable accuracy, as cited in Patil et al. (2021). Gunduz (2021) used ensemble learning and linear-based Deep Learning (LSTM) models to forecast the hourly direction of eight banking stocks in Borsa Istanbul with remarkable accuracy. Nawaz et al. (2021) found that LSTM outperformed other models in sentiment analysis of eWOM from a women's clothing company, according to performance assessment factors. Joseph et al. (2023) proposes a novel hybrid bidirectional LSTM (BiLSTM) model for near real-time wind speed forecasting. Gulmez created an optimized deep LSTM network with the ARO model (LSTM-ARO) created to forecast stock prices (Gulmez, 2023). Xu et al. forecasted crude oil futures based on Bi-LSTM-Attention model (Xu et al., 2024). Shi et al. developed an integrated GCN-LSTM stock price movement forecasting based on knowledge-incorporated graph construction (Shi et al., 2024). Arepalli and Naik devised a Deep Learning-enabled IoT framework for early hypoxia detection in aqua water using light weight spatially shared attention-LSTM network (Arepalli and Naik, 2024). Mamdouh et al. improved flight delays forecasting by developing attention-based bidirectional LSTM network (Mamdouh et al., 2024). Because LSTM was able to forecast the stock market rather well, it is now used more often in stock markets across the globe. Globally, the application of LSTM in stock market forecasting is growing quickly.

However, using statistical tests, Greaves-Tunnell & Harchaoui (2019) came to the conclusion that LSTM is unable to produce long memory sequences from white noise inputs or accurately represent the long memory effect in the input. Furthermore, one shortcoming of typical LSTM networks is that they cannot handle irregular time intervals (Baytas et al., 2017).

This research is important because it explores the possibility of an alternate way using LSTM to forecast the stock prices.