SHARE PRICE ACTION ANALYSIS USING NATURAL LANGUAGE PROCESSING

D M G C M Nalinga

219373D

Master of Science in Computer Science

Department of Computer Science and Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

April 2025

SHARE PRICE ACTION ANALYSIS USING NATURAL LANGUAGE PROCESSING

D M G C M Nalinga

219373D

Thesis/Dissertation submitted in partial fulfillment of the requirements for the degree

Master of Science in Computer Science

Department of Computer Science and Engineering

Faculty of Engineering

University of Moratuwa

Sri Lanka

April 2025

DECLARATION

I declare that this is my own work and this thesis/dissertation does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books).

8th July 2025

Signature: Date:

The above candidate has carried out research for the PhD/MPhil/Masters thesis/dissertation under my supervision. I confirm that the declaration made above by the student is true and correct.

Prof. G I U S Perera

Name of Supervisor:

Signature of the Supervisor: Date:

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to all those who provided support to make my research on “SHARE PRICE ACTION ANALYSIS USING NATURAL LANGUAGE PROCESSING” successful.

First, I would like to express my gratitude to my project supervisor Prof. Indika Perera, Senior Lecturer, Department of Computer Science and Engineering. I am highly indebted to him for his guidance and for providing necessary information regarding the project and for his support in completing the project successfully.

I am sincerely thankful to Dr. Chathuranga Hettiarachchi, Senior Lecturer, Department of Computer Science and Engineering for the support given throughout the project period. Further, I would like to extend my gratitude to Dr. Shehan Perera for participating in meetings and providing me with very useful guidance in the beginning to make my research successful.

Finally, I wish to thank the academic and non-academic staff of the Department of Computer Science and Engineering and colleagues for the support and encouragement given.

ABSTRACT

Stock price prediction has been a widely researched topic, primarily through technical and fundamental analysis. While technical analysis relies on historical stock data and mathematical indicators, its effectiveness diminishes in illiquid stock markets such as the Colombo Stock Exchange (CSE) due to low trading volumes and irregular price movements. Fundamental analysis, on the other hand, focuses on intrinsic company value but does not fully capture short-term market reactions to external events.

This research explores an alternative approach by applying Natural Language Processing (NLP) techniques to conduct an event study analysis. The study examines how news articles influence stock price movements in the CSE by transforming textual data into numerical representations using Large Language Model (LLM)-based embeddings. The extracted feature vectors are then analysed using machine learning algorithms to identify correlations between news representation and stock price fluctuations.

By leveraging NLP-based vectorization and predictive modelling, this research provides new insights into price action analysis in illiquid stock markets, where traditional prediction methods often fail. The findings contribute to the field of financial analytics by demonstrating the feasibility of using textual data to enhance stock price forecasting in under-researched market conditions.

Keywords: Stock Market Prediction, Natural Language Processing, Event Study Analysis, Colombo Stock Exchange, Machine Learning

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES