CMP7158 Project Management Coursework

Peer Evaluation and Self-reflection Forms

Directions: In the space below, honestly evaluate the work of other students in your group by answering yes or no and by using a scale from 1 to 3, **1 being poor, 2 being average, 3 being above average.**

Evaluator's Name: Inderpal

Date: 1st May 2023

Group Member 1/Student ID: **Afshin Barjaneh (22184198)**

1. Did this group member complete his/her assigned tasks for the group? Yes
2. How would you rate the quality of this person's work? 3
3. How would you rate the timeliness of the completion of the work? 3
4. How would you rate the accuracy of the work? 3
5. Overall, how would you rank this group member's performance in the group? 3
6. Would you want to work with this person again? Yes

Explain why in the space below.

Everything was perfect about this guy he helped in everything from articles to editing.

Group Member 2/Student ID: **Dansteve Adekanbi (22178806)**

1. Did this group member complete his/her assigned tasks for the group? Yes
2. How would you rate the quality of this person's work? 3
3. How would you rate the timeliness of the completion of the work? 3
4. How would you rate the accuracy of the work? 3
5. Overall, how would you rank this group member's performance in the group? 3
6. Would you want to work with this person again? Yes

Explain why in the space below.

Dan was up to date everytime regarding his work.

Group Member 3/Student ID: **Ibim Mercy Miller Braide (22166610)**

1. Did this group member complete his/her assigned tasks for the group? Yes

2. How would you rate the quality of this person's work? 3

3. How would you rate the timeliness of the completion of the work? 3

4. How would you rate the accuracy of the work? 3

5. Overall, how would you rank this group member's performance in the group? 3

6. Would you want to work with this person again? Yes

Explain why in the space below.

We can say that this man was the leader of the group. So, everything is best about this man.

Group Member 4/Student ID: **Vijaya Kumar Padavala (22167683)**

1. Did this group member complete his/her assigned tasks for the group? Yes

2. How would you rate the quality of this person's work? 2

3. How would you rate the timeliness of the completion of the work? 2

4. How would you rate the accuracy of the work? 3

5. Overall, how would you rank this group member's performance in the group? 3

6. Would you want to work with this person again? No

Explain why in the space below.

I rated him average because he was slow at first then he increased his speed and he did good in the last.

Group Member 5/Student ID: **Walawe Dolamulla Kankanamge Hiran Tharinda (22190026)**

1. Did this group member complete his/her assigned tasks for the group? Yes

2. How would you rate the quality of this person's work? 3

3. How would you rate the timeliness of the completion of the work? 3

4. How would you rate the accuracy of the work? 3

5. Overall, how would you rank this group member's performance in the group? 3

6. Would you want to work with this person again? Big Yes

Explain why in the space below.

As you can see his name is so big, like that his work was big in the group he handled the editing part.

Self-reflection

- **Benefits of the group work**

Enhanced Learning: Group work can facilitate the exchange of ideas and knowledge between members. It also helps learners to develop a deeper understanding of the subject by hearing different perspectives.
Improved Critical Thinking: Group work provides a platform for discussion and debate, which can help learners to develop their critical thinking skills. It also encourages them to question their assumptions and consider alternative viewpoints.
Improved Communication Skills: Group work requires effective communication, which can help learners to develop their listening, speaking, and writing skills. It also helps them to learn how to work collaboratively with others.

- **The challenges in the group**

Communication Issues: One of the most common challenges in group work is communication issues. Misunderstandings can arise when members have different communication styles or when there is a language barrier.
Unequal Participation: Unequal participation can lead to resentment among group members. This can happen when some members are more engaged than others or when some members dominate the conversation.

- **Benefits and challenges your group faced**

Benefits:

Enhanced learning through the exchange of ideas and knowledge
Improved critical thinking through discussion and debate
Increased motivation and engagement
Shared responsibility and workload distribution
Improved communication and social skills

Challenges:

Communication issues such as misunderstandings or language barriers
Unequal participation or engagement among group members
Conflicting goals or expectations for the project
Time management difficulties due to varying schedules or availability
Personality clashes or conflicts within the group
Ineffective leadership or a lack of direction within the group

- **Discuss what will you do differently next time**

1. Clearly define goals and objectives: Ensure that all members of the group understand the goals and objectives of the project. This will help to ensure that everyone is working towards a common goal.
2. Establish roles and responsibilities: Assign specific roles and responsibilities to each member of the group based on their skills and strengths. This will help to ensure that everyone is contributing and taking ownership of their tasks.
3. Set deadlines and milestones: Establish clear deadlines and milestones for the project. This will help to ensure that the group is making progress and staying on track.
4. Communicate effectively: Communication is key in group work. Establish a regular communication schedule and make sure that everyone has the opportunity to share their thoughts and ideas.
5. Foster a positive group dynamic: Encourage a positive group dynamic by being respectful and supportive of each other. This will help to create a positive working environment and increase motivation and engagement.

6. Use technology to collaborate: Utilize technology tools to facilitate collaboration such as Google Drive, Trello, or Asana. These tools can help to streamline communication and collaboration, and keep everyone on the same page.

SCHOOL OF COMPUTING AND DIGITAL TECHNOLOGY

MSc Computer Science

Predicting Stock Prices Using Machine Learning

Module: Research Methods and Project Management

Module Code: CMP7158

Module Leader: Dr. Mariam Adedoyin-Olowe

Group Members:

Afshin Barjaneh - 22184198

Dansteve Adekanbi – 22178806

Ibim Mercy Miller Braide – 22166610

Inderpal – 22195795

Vijaya Kumar Padavala – 22167683

Walawe Dolamulla Kankanamge Hiran Tharinda – 22190026

Word Count (excluding references) = 1959

Glossary

Term	**Definition**
ARIMA	autoregressive integrated moving average
CNN	Convolutional Neural Networks
DL	Deep Learning
DTR	Decision Tree Regressor
EMH	Efficient market hypothesis
GFNN	Gaussian-Fuzzy-Neural network
KNN	K-Nearest Neighbors
LASSO	Least Absolute Shrinkage and Selection Operator
LR	Logistic regression
LSTM	Long short-term memory
MAPE	Mean absolute percentage error.
NLP	Natural language processing
OLS	Ordinary least squares
RMSE	Root Mean Square Error
SMAPE	Symmetric Mean Absolute Percentage Error
SVM	Support Vector Machine
SVR	Support Vector Regressor
VAR	Vector Autoregression
XGBoost	Extreme Gradient Boosting

Table Of Content

Table of Tables

Table of Figures

Aim

To investigate the most efficient Machine Learning Model for Stock Price Prediction, which can aid investors in making informed and profitable stock trading decisions.

Objectives

- Perform a comparative analysis of suitable machine learning models and their corresponding algorithms to determine the optimal approach for stock price prediction.
- Analyse the performance of machine learning models by using appropriate indicators.
- Visualize and interpret the outcomes to gain insights into the variables driving the predicted stock prices.
- Assess the best-performing machine learning models.

Abstract

This review examines the effectiveness of 15 different machine learning algorithms and 1 decision tree algorithm across two research studies published in academic journals. The first study evaluated twelve classifiers and their performance in predicting stock prices using semantic data. The study conducted a thorough and concise assessment of the algorithms' performances.

The research study evaluated the effect of social media and financial data on the accuracy of machine learning algorithms in predicting stock prices. The study included hybrid algorithms that combined stock market, economic, and social news with deep learning. However, this literature review focuses on analysing the performance of non-hybrid machine learning algorithms to achieve its objectives. The second case study critically reviews five machine learning algorithms in predicting stock prices for twelve leading companies in the Indian Stock Market. Both studies conducted a detailed and comparative analysis of the algorithms' performances.

Keywords: Machine learning algorithms, Stock Market prediction, Financial Market, Social media, Sentiments, Technical analysis, Fundamental analysis

Scope and Delimits

This analysis focuses on stock market prediction, efficient market hypothesis, and machine learning algorithms. It only analyses existing works and draws conclusions based on reviewed literature.

This analysis does not provide investment advice, cover other market analysis forms, or conduct empirical research. It also excludes recommendations on other investment instruments besides stocks.

Introduction

It is crucial to predict stock prices to examine financial systems, as the stock market reflects the economic condition of a country or even the world. As data acquisition, machine learning, and big data techniques continue to advance, researchers are exploring machine learning methods to identify non-linear patterns in financial systems. Fischer and Krauss (2018) prognosticated the S&P 500 trend via the up-down signals produced by Long Short-Term Memory (LSTM). Precise anticipation of the stock market enables investors and other interested parties to comprehend the trends in the financial market (Lu et al., 2020).

Stock market prediction uses two main analytical approaches: technical analysis and fundamental analysis (Nti et al., 2019). Technical analysis studies market prices and trading volumes to predict future movements based on past patterns (Han et al., 2021). On the other hand, fundamental analysis focuses on a company's financial data to forecast future cash flows and stock prices (Dhafer et al., 2022). However, this information is typically unstructured, which poses a significant challenge for the analyst. (Lohrmann et al., 2019; Javed Awan et al., 2020; Leigh et al., 2002; Fama ,1965)

The Efficient Market Hypothesis (EMH) proposes that market prices behave like random walks. Therefore, current information cannot predict future changes (Fama, 1965; Fama, 1970). EMH categorises market efficiency into weak, semi-strong, and strong forms, with semi-strong efficiency implying that publicly available information is fully incorporated into stock prices (Aminimehr et al., 2022). If investors can regularly generate a risk-adjusted return greater than that of the market, it may contradict the EMH. Machine learning algorithms can help make predictions but require careful selection based on the dataset's characteristics and intended purpose (Htun et al., 2023).

Related works

Khan et al. (2020) investigated the effect of financial news and social media on stock prediction spanning a period of ten consecutive days using machine learning models. The random forest classifier model yielded the highest accuracy of 83.22%.

Vignesh (2020) applied machine learning to stock prediction and analysed past stock data. The LSTM model achieved better results with an accuracy of 66.83% compared to SVM and backpropagation.

Daul et al. (2022) compared the performance of linear models (OLS and LASSO) and non-linear models (boosted trees and neural networks) in predicting stock returns. The study found that non-linear methods, especially neural networks, outperformed linear models regarding the Sharpe Ratio.

Yun et al. (2021) conducted a study exploring the efficacy of the GA-XGBoost algorithm in predicting the direction of stock prices. The researchers used an enhanced 3-stage feature engineering process to train the model, yielding superior results and greater flexibility than benchmark studies. Notably, the model expanded feature set two in a specific dataset achieved a remarkable prediction accuracy of 93.28%, marking a significant increase of 32.79% from the original feature.

Jing et al. (2021) developed a novel hybrid model for predicting stock prices that incorporates both a sentiment analysis model and deep learning techniques. The authors utilised a combination of the Convolutional Neural Network model and Neural Network to analyse technical indicators from the stock market. The proposed hybrid model demonstrated superior performance, with a MAPE (Mean absolute percentage error) of 0.0449, which was notably lower than that of conventional models evaluated using the same dataset.

Analysis of machine learning on the stock price prediction

Some researchers explored the connection between social media content and news with stock market prices, aiming to determine the feasibility of using machine learning algorithms to predict stock market prices while considering the volatility of social, economic, and political news (Khan et al., 2020; Zhang et al., 2021; Mehta et al., 2021). Selecting the most influential details feature collection is a common and sometimes indispensable process when using machine learning algorithms for prediction with datasets. However, the presence of bots (Fan et al., 2020), spam tweets (Koukaras et al., 2022), and fake news (Fong et al., 2021; Aljabri et al., 2023; Domenico et al., 2021) on social media platforms like Twitter is a concern that was addressed through spam reduction and feature selection on the datasets used.

Maqbool et al. (2023) employed machine learning techniques to make forecasts about stock prices by combining sentiment scores of financial news with historical stock price data. The integration of these two sources of information allowed for more accurate predictions. Table 1 shows the selected stock markets for the study by Khan et al. (2020), along with their respective tweets and news counts. The researchers collected the stock market, social media, and news data for a two-year period from July 1, 2016, to June 30, 2018, along with S&P 500 index price data.

Table 1 - The stock markets chosen for this study

No.	Stock market	Ticker symbol	Country/stock exchange	Tweets count	News count
1	Karachi Stock Exchange	KSE	Pakistan	34	0
2	London Stock Exchange	LSE	United Kingdom	2535	53
3	New York Stock Exchange	NYSE	United States	12,538	0
4	HP Inc.	HPQ	NYSE	27,432	554
5	International Business Machines Corporation	IBM	NYSE	364,601	1700
6	Microsoft Corporation	MSFT	NASDAQ	168,901	3316
7	Oracle Corporation	ORCL	NYSE	51,328	799
8	Red Hat, Inc.	RHT	NYSE	18,120	212
9	Twitter, Inc.	TWTR	NYSE	380,472	2367
10	Motorola Solutions, Inc.	MSI	NYSE	6444	284
11	Nokia Corporation	NOK	NYSE	23,441	301

The research study included selecting and comparing twelve machine learning classifiers to assess their predictive performance in identifying future stock market trends.

Table 2 - The performance of the classification algorithms on the HPQ testing dataset

Classes	Metrics	Algorithms											
		GNB	MNB	SVM	LR	MLP	KNN	CART	LDA	AB	GBM	RF	ET
Positive	Precision (%)	66.00	62.00	67.00	69.00	66.00	77.00	72.00	72.00	65.00	83.00	81.00	75.00
	Recall (%)	54.00	100.00	96.00	82.00	84.00	67.00	77.00	86.00	98.00	85.00	89.00	73.00
	F-measure (%)	60.00	77.00	79.00	75.00	74.00	72.00	75.00	79.00	78.00	84.00	85.00	74.00
Neutral	Precision (%)	0.00	NA	NA	NA	NA	00.00	NA	NA	NA	NA	NA	NA
	Recall (%)	0.00	NA	NA	NA	NA	0.00	NA	NA	NA	NA	NA	NA
	F-measure (%)	0.00	NA	NA	NA	NA	0.00	NA	NA	NA	NA	NA	NA
Negative	Precision (%)	47.00	100.00	76 .00	57.00	53.00	57.00	59.00	68.00	82.00	75.00	79.00	58.00
	Recall (%)	54.00	2.00	23.00	40.00	30.00	65.00	53.00	47.00	16.00	72.00	67.00	61.00
	F-measure (%)	50.00	3.00	35.00	47.00	38.00	61.00	56.00	56.00	26.00	73.00	72.00	60.00

The study conducted stock market prediction experiments using social media, financial news, and both. The chosen algorithms for predicting stock market trends were trained and tested using tenfold Cross-Validation for ten days. Diagrams in Figures 1, 2, and 3 show the accuracies of algorithms prediction.

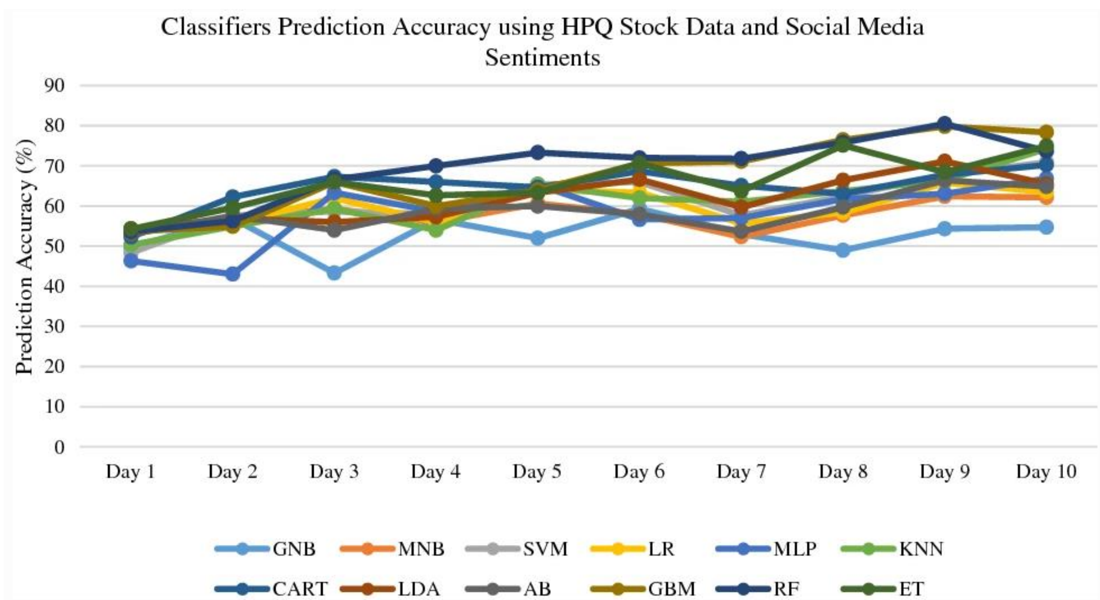

Figure 1 - Accuracy of the models based on the social median and stock information

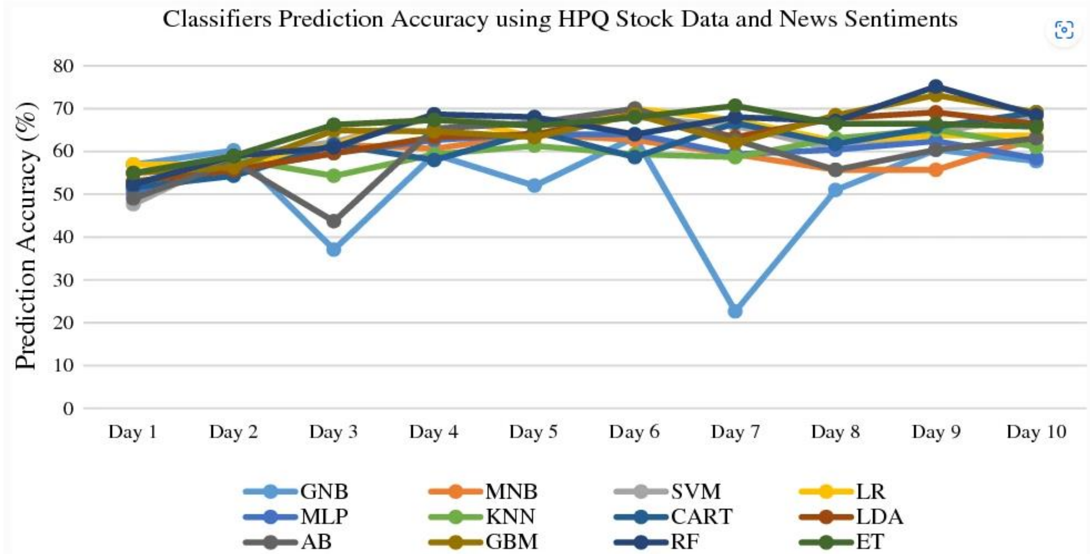

Figure 2 - Accuracy of the models based on the news and stock information

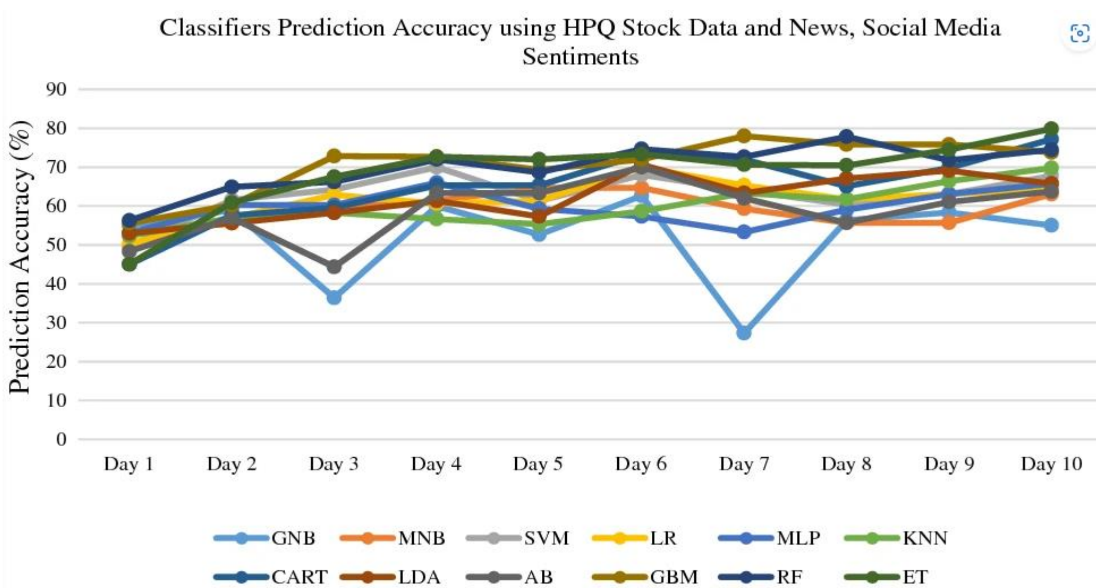

Figure 3 - Accuracy of the model for all data

Using financial and social media sentiments decreased the algorithm's highest accuracy, but accuracy improved after day three. The study demonstrated the potential impact of including social and economic data on accuracy.

In another study, Bansal et al. (2022) followed a methodology involving data collection from twelve companies using a dataset spanning from 2015 to 2021, training five and evaluating using performance metrics such as SMAPE, R2-Value, and RMSE (Chicco et al., 2021). Testing was done on eight trading days. The equations that were used to calculate are as below.

$$SMAPE = \frac{1}{n}\sum \frac{|forecast\ value\ - actual\ value|}{\frac{(|actual\ value| + |forecast\ value|)}{2}}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_{pred,i} - Y_i\right)^2}{n}}$$

$$R^2 = 1 - \frac{unexplained\ variation}{total}$$

Table 3 shows the SMAPE results obtained for the models tested on the dataset. A lower SMAPE value is better. The LSTM algorithm performed best with a SMAPE of 1.59.

Table 3 - SMAPE for all models

Parameter	**SMAPE (Symmetric Mean Absolute Percentage Error)**				
Algorithm	**K-Nearest Neighbors**	**Linear Regression**	**Support Vector Regression**	**Decision Tree Regression**	**Long ShortTerm Memory**
Adani Ports	10.99	9.18	9.51	10.68	1.65
Asian Paints	11.63	9.35	8.42	9.78	1.67
Axis Bank	16.67	10.37	5.62	8.48	1.88
HDFC	20.46	11.00	6.22	12.56	2.19
Hindustan Unilever	10.95	9.62	4.88	8.82	1.38
ICICI Bank	14.45	8.92	7.02	7.37	2.31
Kotak Bank	12.44	10.51	5.81	10.26	1.43
Maruti	15.92	11.13	3.09	13.92	1.32
NTPC	13.59	12.39	3.08	9.60	1.13
Tata Steel	16.08	13.10	5.75	8.06	1.75
TCS	15.84	9.95	4.41	10.05	1.40
Titan	12.90	3.50	3.33	11.39	1.06
Average	14.32	9.91	5.59	10.08	1.59

Table 4 shows the R2 results obtained for the models. The ideal R2 value is close to 1. The LSTM algorithm had the best performance with an R2 value of -0.11.

Table 4 - R2 for all models

Parameter	R^2 (R squared)				
Algorithm	**K-Nearest Neighbors**	**Linear Regression**	**Support Vector Regression**	**Decision Tree Regression**	**Long ShortTerm Memory**
Adani Ports	-0.22	-6.67	-1.76	-3.25	-0.90
Asian Paints	-1.66	-5.57	-2.21	-2.75	-0.45
Axis Bank	-5.32	-3.43	-1.05	-4.18	0.59
HDFC	-1.56	-2.47	-2.77	-3.04	-0.62
Hindustan Unilever	-2.03	-1.07	-0.35	-2.39	0.27
ICICI Bank	-2.59	-1.59	-2.72	-2.23	0.45
Kotak Bank	-3.78	-2.31	-3.38	-3.90	-0.01
Maruti	-1.01	-0.65	-2.61	-0.73	-0.99
NTPC	-6.90	-2.31	-1.11	-0.84	-0.02
Tata Steel	-2/16	0.48	-1/13	-2.01	0.80
TCS	-0.83	-1.18	-0.51	-0.92	-0.84
Titan	-0.21	-0.57	-0.12	-1.72	0.31
Average	-2.42	-2.27	-1.69	-2.33	-0.11

The ideal RMSE value is zero. Out of the five algorithmic models tested, the LSTM algorithm had the best performance with an RMSE of 22.55. The data is shown in Table 5.

Table 5 - RMSE for all models.

Parameter	RMSE (Root Mean Square Error)				
Algorithm	K-Nearest Neighbors	Linear Regression	Support Vector Regression	Decision Tree Regression	Long ShortTerm Memory
Adani Ports	29.78	43.76	37.94	43.25	16.22
Asian Paints	51.78	37.68	80.44	40.12	14.31
Axis Bank	47.41	60.03	66.23	48.65	15.77
HDFC	66.16	63.37	54.02	58.99	35.71
Hindustan Unilever	40.01	45.11	36.89	62.11	40.05
ICICI Bank	50.20	49.34	38.90	49.70	16.09
Kotak Bank	49.91	50.01	41.55	52.91	34.82
Maruti	84.72	73.4	21.70	63.22	12.94
NTPC	63.65	25.19	15.28	41.61	10.60
Tata Steel	70.17	50.26	54.18	71.07	22.80
TCS	67.13	55.68	58.74	44.14	30.56
Titan	56.36	60.39	50.49	24.46	20.83
Average	56.44	51.20	46.36	50.01	22.55

Bansal et al. (2022) proposed that LSTM outperformed the other algorithms with the least errors in SMAPE, R2, and RMSE.

Discussion

The research study involving semantic data reveals that the Random Forest classifier is the machine learning algorithm with the highest accuracy, and incorporating social and financial news has further enhanced the prediction accuracy of the algorithms (Khan et al., 2020). However, the study could have been improved by collecting semantic data from more social media platforms in addition to Twitter. Platforms such as Facebook, LinkedIn, and Instagram are also potential sources of information that can impact financial economies. Furthermore, financial news headlines were solely obtained from Business Insider, limiting the information scope. The study could have included more sources of financial news headlines to provide a more comprehensive view.

The second review focused on a research study conducted by Bansal et al. (2022). It was found that external factors, such as semantic data, were not considered in predicting stock market prices, which could impact the accuracy of the results. Including semantic data could produce more realistic outcomes in real-life situations.

Both studies conducted experiments on stock market data from a single country and continent. To increase the credibility of such research studies, applying the same machine learning prediction procedures to stock market data from different countries and continents would be beneficial.

Conclusion

To ensure credibility, a proposed solution to a problem must be universally accepted and effective in various situations. The stock market is a crucial aspect of every economy worldwide. Thus, it is essential to increase positive returns on investments for investors globally. This literature review aims to emphasise the significance of adding credibility to stock price data using semantic data from various social media platforms and financial news sources. The study advocates for the inclusion of data from sources such as Facebook, LinkedIn, and Instagram to enhance the accuracy of stock price predictions. Relying on Twitter data alone is insufficient due to the increasing presence of bots on the platform. By incorporating more data sources, the research findings in Khan et al.'s (2020) work will be more credible.

References

Aljabri, M., Zagrouba, R., Shaahid, A., Alnasser, F., Saleh, A. and Alomari, D.M. (2023). Machine learning-based social media bot detection: a comprehensive literature review. *Social Network Analysis and Mining*, 13(1). Available at: https://doi.org/10.1007/s13278-022-01020-5.

Aminimehr, A., Raoofi, A., Aminimehr, A. and Aminimehr, A. (2022). A Comprehensive Study of Market Prediction from Efficient Market Hypothesis up to Late Intelligent Market Prediction Approaches. *Computational Economics*, 60(2), pp.781–815. Available at: https://doi.org/10.1007/s10614-022-10283-1.

Bansal, M., Goyal, A. and Choudhary, A. (2022). Stock Market Prediction with High Accuracy using Machine Learning Techniques. *Procedia Computer Science*, 215, pp.247–265. Available at: https://doi.org/10.1016/j.procs.2022.12.028.

Chicco, D., Warrens, M.J. and Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, [online] 7(5), p.e623. Available at: https://doi.org/10.7717/peerj-cs.623.

Daul, S., Jaisson, T. and Nagy, A. (2022). Performance attribution of machine learning methods for stock returns prediction. *The Journal of Finance and Data Science*, 8, pp.86–104. Available at: https://doi.org/10.1016/j.jfds.2022.04.002.

Dhafer, A.H., Mat Nor, F., Alkawsi, G., Al-Othmani, A.Z., Ridzwan Shah, N., Alshanbari, H.M., Bin Khairi, K.F. and Baashar, Y. (2022). Empirical Analysis for Stock Price Prediction Using NARX Model with Exogenous Technical Indicators. *Computational Intelligence and Neuroscience*, 2022, pp.1–13. Available at: https://doi.org/10.1155/2022/9208640.

Domenico, G.D., Sit, J., Ishizaka, A. and Nunan, D. (2021). Fake news, social media and marketing: A systematic review. *Journal of Business Research*, 124(1), pp.329–341. Available at: https://doi.org/10.1016/j.jbusres.2020.11.037.

Fama, E. (1970). Efficient capital markets: A review of the theory. *The Journal of Financial Economics,* 49, 283-306. Available at: https://doi.org/10.1111/j.1540-6261.1970.tb00518.x

Fama, E. F. (1965). The behaviour of stock market prices. *The journal of Business*, 38, 34-105. Available at: http://www.e-m-h.org/Fama65.pdf [Accessed 1 Apr. 2022].

Fan, R., Talavera, O. and Tran, V. (2019). Social media bots and stock markets. *European Financial Management*. Available at: https://doi.org/10.1111/eufm.12245.

Fischer, T. and Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), pp.654–669. Available at: https://doi.org/10.1016/j.ejor.2017.11.054.

Fong, B. (2021). Analysing the behavioural finance impact of ‘fake news’ phenomena on financial markets: a representative agent model and empirical validation. *Financial Innovation*, 7(1). Available at: https://doi.org/10.1186/s40854-021-00271-z.

Han, Y., Liu, Y., Zhou, G. and Zhu, Y. (2021). Technical Analysis in the Stock Market: A Review. [online] papers.ssrn.com. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3850494 [Accessed 1 Apr. 2022].

Htun, H.H., Biehl, M. and Petkov, N. (2023). Survey of feature selection and extraction techniques for stock market prediction. *Financial Innovation*, 9(1). Available at: https://doi.org/10.1186/s40854-022-00441-7.

Javed Awan, M., Shafry Mohd Rahim, M., Nobanee, H., Munawar, A., Yasin, A. and Mohd Zain Azlanmz, A. (2021). Social Media and Stock Market Prediction: A Big Data Approach. *Computers, Materials & Continua*, 67(2), pp.2569–2583. Available at: https://doi.org/10.32604/cmc.2021.014253.

Jing, N., Wu, Z. and Wang, H. (2021) A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction. *Expert Systems with Applications*, 178, p.115019. Available at: https://doi.org/10.1016/j.eswa.2021.115019

Khan, W., Ghazanfar, M.A., Azam, M.A., Karami, A., Alyoubi, K.H. and Alfakeeh, A.S. (2020). Stock market prediction using machine learning classifiers and social media, news. *Journal of Ambient Intelligence and Humanized Computing*. Available at: https://doi.org/10.1007/s12652-020-01839-w.

Koukaras, P., Nousi, C. and Tjortjis, C. (2022). Stock Market Prediction Using Microblogging Sentiment Analysis and Machine Learning. *Telecom*, 3(2), pp.358–378. Available at: https://doi.org/10.3390/telecom3020019.

Leigh, W., Purvis, R., & Ragusa, J.M. (2002). Forecasting the nyse composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: a case study in romantic decision support. *Decision Support Systems*, 32, 361-377. Available at: https://doi.org/10.1016/S0167-9236(01)00121-X

Lohrmann, C. and Luukka, P. (2019). Classification of intraday S&P500 returns with a Random Forest. *International Journal of Forecasting*, [online] 35(1), pp.390–407. Available at: https://doi.org/10.1016/j.ijforecast.2018.08.004.

Lu, W., Li, J., Wang, J. and Qin, L. (2020). A CNN-BiLSTM-AM method for stock price prediction. *Neural Computing and Applications,* 33(10), pp.4741–4753. Available at: https://doi.org/10.1007/s00521-020-05532-z.

Maqbool, J., Aggarwal, P., Kaur, R., Mittal, A. and Ganaie, I.A. (2023). Stock Prediction by Integrating Sentiment Scores of Financial News and MLP-Regressor: A Machine Learning

Approach. *Procedia Computer Science*, 218, pp.1067–1078. Available at: https://doi.org/10.1016/j.procs.2023.01.086.

Mehta, P., Pandya, S. and Kotecha, K. (2021). Harvesting social media sentiment analysis to enhance stock market prediction using deep learning. *PeerJ Computer Science*, 7, p.e476. Available at: https://doi.org/10.7717/peerj-cs.476.

Nti, I.K., Adekoya, A.F. and Weyori, B.A. (2019). A systematic review of fundamental and technical analysis of stock market predictions. *Artificial Intelligence Review*, 53, pp.3007–3057. Available at: https://doi.org/10.1007/s10462-019-09754-z.

Yun, K.K., Yoon, S.W. and Won, D. (2021) Prediction of stock price direction using a hybrid GA-XGBoost algorithm with a three-stage feature engineering process. *Expert Systems with Applications,* 186, p.115716. Available at: https://doi.org/10.1016/j.eswa.2021.115716

Zhang, W., Gong, X., Wang, C. and Ye, X. (2021). Predicting stock market volatility based on textual sentiment: A non-linear analysis. *Journal of Forecasting*, 40(8). Available at: https://doi.org/10.1002/for.2777.