



Social Media Sentiment for Predicting Stock Market Returns and Supporting Investment Decision Making

Anika, Dr. Neeraj Chauhan

¹Research scholar, Institute of management studies and research, Maharshi Dayanand University Rohtak
Email- anikachugh42@gmail.com

²Institute of management studies and research , Maharshi Dayanand University Rohtak
Email- neerajchauhan22@gmail.com

Article History

Received Date: 15/06/2026
Revised Date: 21/08/2026
Accepted Date: 12/09/2026
Published Date: 24/09/2026

Abstract

Social media is one of the key channels for real-time data that captures investor views, expectations and perceptions. This study was to see if social media sentiment could enhance the way we predict stock market returns and make investment decisions. This involved a quantitative longitudinal study with 726 observations made from January 2022 to December 2024. Hang Seng Technology returns, market returns, sentiment indicators, overnight sentiment, social media attention measures, exchange rate changes and interbank rate changes were analysed. Descriptive statistics and correlation analysis were carried out, and later machine learning models were created using linear regression, Random Forest and Gradient Boosting. Performance of models was evaluated based on mean absolute error, root mean squared error, coefficient of determination and the directional accuracy. The contemporaneous sentiment was moderately positively related to Hang Seng Technology returns while the overnight sentiment was weakly, but not significantly, related. There was no consistent improvement in performance when using sentiment-enhanced models over market-only models. Linear regression had the best directional accuracy, and Random Forest had a marginal decrease in prediction error once sentiment variables were introduced. However, feature-importance analysis revealed sentiment measures brought valuable information in addition to the overall market factors. Overall, the social media sentiment seemed more useful as a secondary market indicator, rather than as a standalone forecasting signal. It therefore has most value in combination with traditional financial data in a wider framework of investment decision support systems.

Keywords:

- Investment decision making;
- Machine learning;
- Social media sentiment;
- Stock market prediction;
- Stock returns

Introduction

In the era of digital communication, the dynamics of financial information generation, dissemination and interpretation have quickly changed for market players. Today, social media gives vast amounts of real-time data about the investors' opinions, expectations, uncertainty and reactions to financial events. This type of information may serve as a supplement to traditional financial metrics as market valuations are not only determined by fundamental and technical factors, but also by investors' perceptions and actions. Sentiment analysis of social media has thus become a valuable tool for mining information about the market from online conversations and aiding investment decisions (Hasselgren et al., 2022). Meanwhile, the use of artificial intelligence (AI) in the financial markets is growing, enhancing the prediction, portfolio evaluation, and investment decision-making processes (Colombo et al., 2023).

Investor sentiment is especially relevant because the process of investing can be far more than rational use and understanding of economic data. But trading behaviour can be influenced by behavioral biases, emotional reactions, and shifts in market expectations, which can lead to fluctuations in stock prices and returns. Investors' sentiment and behavioural biases have been shown to affect investment decisions in emerging financial markets, especially in uncertain times (Parveen et al., 2023). Xia & Madni (2024) suggest that behavioural factors are not insignificant in influencing investors' decision-making process from similar evidence from the Chinese stock market. The effects of these could be exacerbated in the case of social media, where opinion and market narratives can be spread quickly to large groups of investors. Thus, relation between information from social media and investment behaviour is gaining more significance (Sathya & Prabhavathi, 2024).

The development of natural language processing, machine learning and computational finance has opened the door to using unstructured information on the internet as quantifiable sentiment metrics. AI techniques can be used to detect positive, negative, and neutral financial opinions in social-media content and then translate them into quantitative signals that can be used in forecasting systems (Qian et al., 2022). Machine learning methods have also come to the forefront in stock market prediction due to their ability to capture complex and possibly nonlinear relationships between the elements of the financial world (Jain & Vanzara, 2023). There is also systematic evidence of the on-going use of AI in various stock prediction tasks, but with variations in model structure, input features and assessment methods impacting on predictive power (Lin & Marques, 2024).

Sentiment data with traditional market indicators could give more complete predictive data than just the historic prices. Machine-learning classifiers that use a social media and financial news combination have shown the potential of textual information for stock-market prediction (Khan et al., 2022). Likewise, social sentiment models combined with machine-learning approaches have been devised to forecast stock-market dynamics and pinpoint potential valuable links between social sentiment and financial shifts (Zaman et al., 2023). Moreover, AI-powered big data analytics can help in financial decision making by consolidating information from various sources and providing analytical insights from the complex data of the financial markets (Pillai, 2023). Analysis of social media data for forecasting price movements has been tried in various contexts, such as cryptocurrency markets, which shows the

general relevance of social media for financial intelligence systems (Yasir et al., 2023). Furthermore, supervised machine-learning methods have shown promise in the potential to assist financial decisions in stock-market scenarios (Sharma et al., 2023).

Although these developments are important, it remains to be seen whether sentiment in the social media can offer valuable extra information on stock returns and whether such predictions can be used to support investment decisions. Existing studies mainly aim at predicting the accuracy of forecasts, or sentiment classification, or investor behaviour. Combining sentiment indicators, market variables, predictive modelling and directional decision signals can offer a better picture of the usefulness of social-media information practically. Hence, in the present study, we focus on the predictive power of social-media sentiment with respect to stock-market returns and how it can help in data-driven investment decision making.

Research Objectives:

- To examine the relationship between social-media sentiment indicators and stock-market returns.
- To develop and evaluate machine-learning models for predicting stock-market returns using sentiment and financial-market variables.
- To assess whether sentiment-enhanced return predictions can support market-direction assessment and investment decision making.

2. Materials and Methods

2.1 Research Design

The research employed a quantitative type of longitudinal design to assess the potential of using social media sentiment to enhance the predictability of stock market returns and to aid in data-oriented investment decision making. The financial and sentiment observations were analysed daily in chronological order. The analytical model used included descriptive analysis, correlation analysis, and machine learning forecasting with the aim of keeping the temporal consistency of the predictors and the return under analysis.

2.2 Data Source and Study Variables

The data spanned the time from January 2022 to December 2024 and included 726 trading-day observations with 13 variables (Xu, 2026). The main market variables are the Hang Seng Index (HSI), the Hang Seng Technology Index (HST), the Hang Seng Index return (RHSI) and the Hang Seng Technology Index return (RHST). Other financial factors were RMB, RRGB, SHIBOR, and RSHIBOR. The sentiment index (STI), which was used to represent social media information, was calculated as an average of overnight sentiment (STI_on), attention volume (AV), and number of overnight posts (n_posts_on). RHST was considered as the outcome variable of choice.

2.3 Data Preprocessing and Feature Engineering

The date variable was put into chronological form and observations ranked by trading date. All numeric variables were checked for missing and invalid data. Lagged versions of selected market and sentiment variables were created to only use information that was available prior to the return to be predicted. Sentiment measured the previous day (RHST, RHSI, RRGB, RSHIBOR, STI) were used as predictive features, and sentiment measured on the same day (STI_on) was used separately as sentiment taken before the market opens.

2.4 Descriptive and Correlation Analysis

Descriptive statistics for principal financial and sentiment variables were obtained, such as mean, standard deviation, minimum and maximum values. Temporal patterns were evaluated by plotting the stock return and sentiment indicators during the time periods examined. A correlation analysis was conducted, to assess the strength and direction of relationship between sentiment measures, stock returns and financial variables. The correlation of overnight sentiment and market returns was considered.

2.5 Machine Learning Model Development

This was accomplished using three predictive approaches. Linear regression was used as the baseline model, and Random Forest and gradient boosting/XGBoost were used to model nonlinear relationships and interactions between predictors. Two modelling specifications were tested: first a baseline model with conventional financial variables; and second a sentiment-enhanced model with STI, STI_on, AV and n_posts_on. This comparison was applied to evaluate the added value in predicting using social media information.

2.6 Model Training, Validation, and Performance Evaluation

Due to the time-series nature of the data, the observations were not randomly shuffled but were split into a training and testing phase. The first observations were used for model development, and later observations were kept out of sample for out of sample evaluation. Mean absolute error, root mean squared error, coefficient of determination, and directional accuracy were used to evaluate the model performance. A lower MAE and RMSE value and a higher R2 and directional accuracy showed better prediction performance.

2.7 Feature Importance and Investment Decision Support

The differentiability of sentiment and financial predictors was investigated by analyzing feature importance from tree-based models. STI, STI_on, AV, and n_posts_on were looked at. Predicted returns were also transformed to directional signals with positive predicted returns implying expected upward movement and negative predicted returns implying expected downward movement. These signals were compared to the actual return directions to see how useful they can be for investment decision support.

2.8 Statistical Analysis and Software

In the cases where inferential analysis was performed, a p value of less than 0.05 was considered statistically significant. The analysis was designed to be implemented in Python, with standard libraries for data processing, statistical analysis, machine learning and visualization. Descriptive statistics, correlation estimates, model-performance measures, feature-importance results, and graphical comparisons of observed and predicted returns were employed to report results.

3. Results

3.1 Descriptive Characteristics of Sentiment and Market Variables

A total of 726 trading observations were analysed. The returns of the Hang Seng Technology Index (RHST) averaged -0.00037 per day and had a standard deviation of 0.02579 per day, which was quite significant. The mean value of the sentiment index (STI) was 0.3293 while overnight sentiment (STI_on) was on the other hand comparatively higher with 0.4683 . Social-media activity also differed significantly with an average of 392.82 (attention volume) and an average of 197.25 overnight posts. Table 1 shows descriptive data for the key variables in the study.

Table 1. Descriptive statistics of sentiment and financial-market variables

Variable	N	Mean	SD	Minimum	Median	Maximum
RHST	726	-0.00037	0.02579	-0.11685	-0.00097	0.20051
RHSI	726	-0.00023	0.01631	-0.06567	-0.00034	0.08693
STI	726	0.32931	0.10615	-0.20746	0.33458	0.71876
STI_on	725	0.46825	0.15027	-0.19802	0.48849	0.78882
AV	726	392.82	162.93	68.00	359.50	1876.00
n_posts_on	725	197.25	177.92	33.00	155.00	3299.00
RRMB	726	0.00111	0.01773	-0.10080	-0.00010	0.06810
RSHIBOR	726	-0.00112	0.01173	-0.08000	0.00000	0.10700

The temporal pattern of RHST demonstrated frequent positive and negative movements throughout the study period, including several episodes of relatively pronounced volatility (Figure 1). Overnight sentiment also fluctuated substantially over time, although its variation was generally smoother than that of daily market returns (Figure 2).

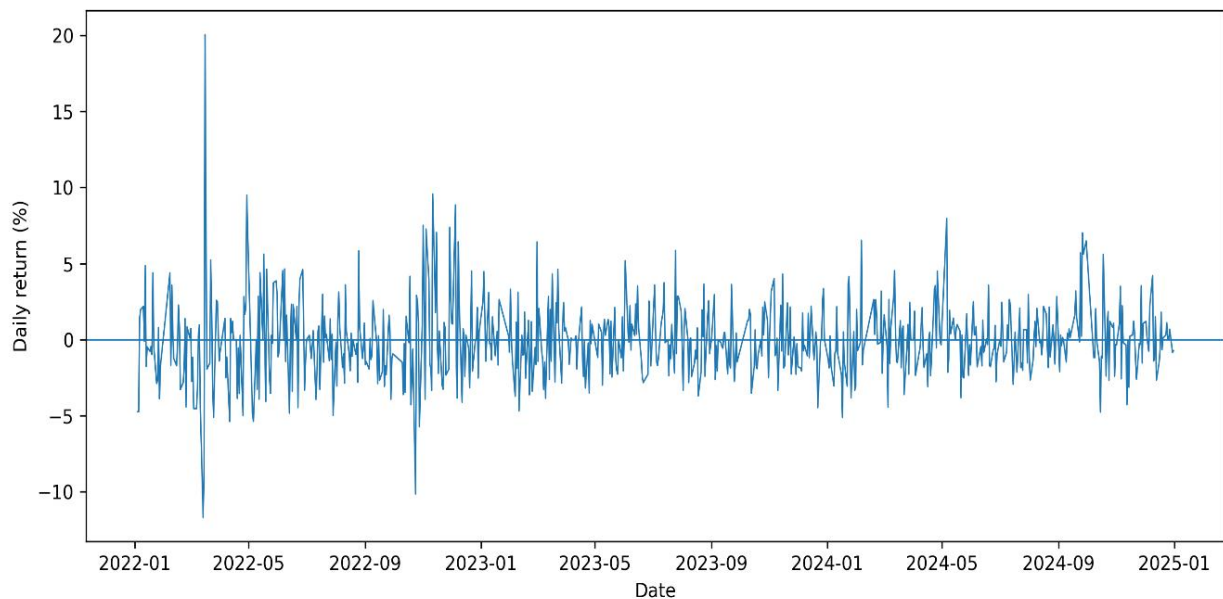


Figure 1. Temporal pattern of Hang Seng Technology Index daily returns

Daily RHST values demonstrate substantial short-term volatility throughout the observation period, with both positive and negative return episodes and several relatively large market movements.

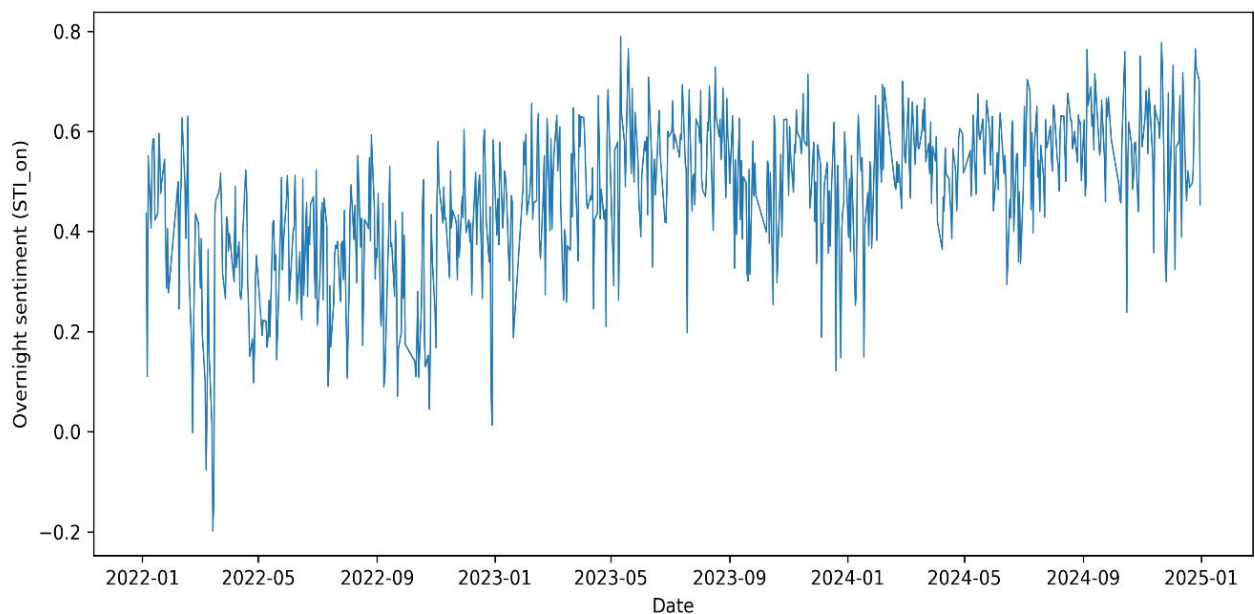


Figure 2. Temporal variation in overnight social-media sentiment

Overnight sentiment remained predominantly positive but varied across the study period, indicating changing social-media perceptions and providing a time-varying sentiment measure that could be incorporated into return-prediction models.

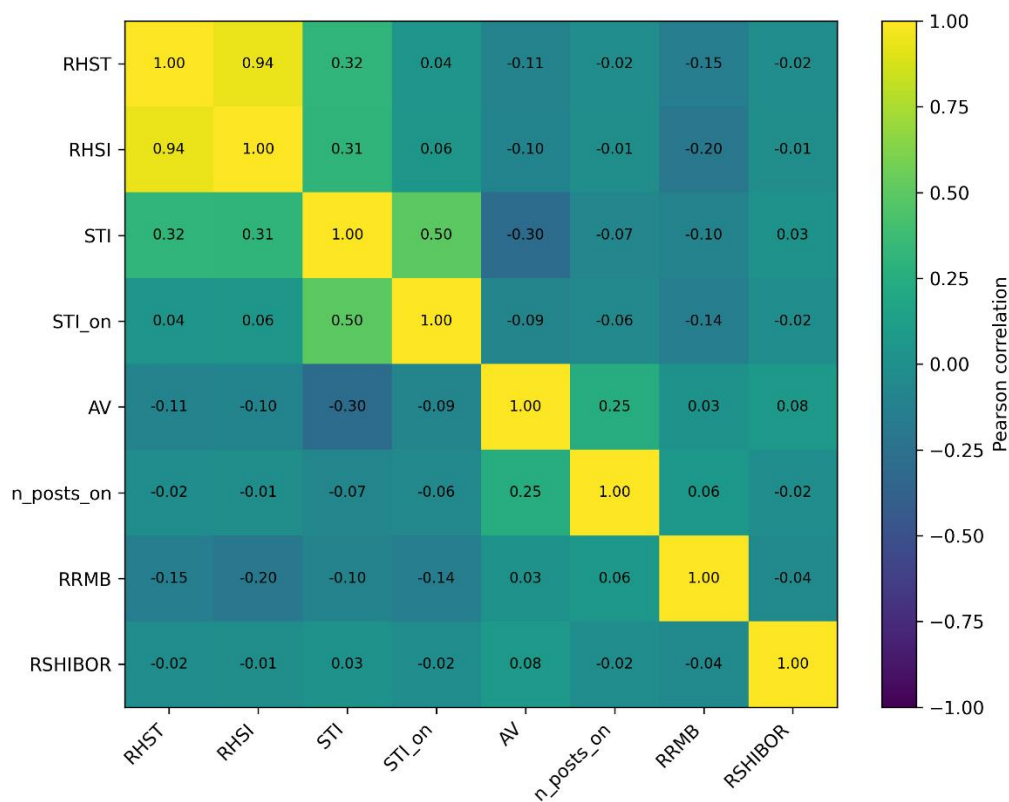
3.2 Association Between Social-Media Sentiment and Stock Returns

The correlation analysis found that RHST was highly positively correlated with the return of the broad market index (Hang Seng Index return (RHSI); $r = 0.938$, $p < 0.001$). There was a moderate positive relationship between STI (contemporaneous sentiment index) and RHST ($r = 0.319$, $p < 0.001$). Overnight sentiment was only weakly and not significantly associated with RHST ($r = 0.039$, $p = 0.290$), however. There was a weak negative correlation between attention volume and RHST ($r = -0.107$, $p = 0.004$) as well as between RRMB and RHST ($r = -0.147$, $p < 0.001$). The results were complete and are displayed in Table 2 and the overall correlation structure is depicted in Figure 3.

Table 2. Correlations of sentiment and financial variables with RHST

Variable	N	Pearson r	p-value	Interpretation
RHSI	726	0.938	<0.001	Strong positive
STI	726	0.319	<0.001	Moderate positive
STI_on	725	0.039	0.290	Very weak, NS
AV	726	-0.107	0.004	Weak negative
n_posts_on	725	-0.019	0.613	Very weak, NS
RRMB	726	-0.147	<0.001	Weak negative
RSHIBOR	726	-0.018	0.634	Very weak, NS

NS = not statistically significant.

**Figure 3. Correlation matrix of sentiment and financial-market variables**

The matrix shows the strongest relationship between RHST and RHSI, while contemporaneous sentiment demonstrates a moderate positive association with RHST. Overnight sentiment and overnight post volume show comparatively weak direct associations with returns.

3.3 Machine-Learning Prediction Performance

A total of 725 observations were obtained, after building lagged predictors and discarding observations that were not available to forecast. The first 580 observations were used for model development with the remaining 145 observations being used for chronological out-of-sample testing. The market-only and sentiment-enhanced specifications' performance is shown in Table 3. For the market only models, the lowest test RMSE was 0.02002 for linear regression and the highest value for the directional accuracy was 56.55% for linear regression. Sentiments related predictors did not lead to consistent improvement across models. For sentiment-enhanced linear regression, the RMSE value was 0.02007, R2 value was -0.056 and the directional accuracy was 54.48%. The sentiment enhanced Random Forest had an RMSE value of 0.02010 while Gradient Boosting had an RMSE value of 0.02034. The R² values

of all models were negative which means that the models accounted for little to none of the variation in the out-of-sample daily returns.

Table 3. Out-of-sample predictive performance of machine-learning models

Predictor specification	Model	MAE	RMSE	R ²	Directional accuracy
Market-only	Linear Regression	0.01461	0.02002	-0.051	56.55%
Market-only	Random Forest	0.01486	0.02029	-0.079	52.41%
Market-only	Gradient Boosting	0.01476	0.02015	-0.065	53.79%
Sentiment-enhanced	Linear Regression	0.01472	0.02007	-0.056	54.48%
Sentiment-enhanced	Random Forest	0.01479	0.02010	-0.059	52.41%
Sentiment-enhanced	Gradient Boosting	0.01487	0.02034	-0.084	51.72%

3.4 Incremental Contribution of Social-Media Sentiment

There were differences in the incremental contribution of sentiment across the models (Table 4). Although sentiment variables were added, there was a small improvement from Random Forest with a decrease of 0.50% for MAE and 0.93% for RMSE. It's R² changed by +0.020, but there was no change in its directional accuracy. On the other hand, sentiment inclusion slightly degraded the results of MAE and RMSE for linear regression and Gradient Boosting. The accuracy of direction decreased by about 2.07% for both models.

Table 4. Change in predictive performance after addition of sentiment variables

Model	MAE improvement	RMSE improvement	Change in R ²	Change in directional accuracy
Linear Regression	-0.74%	-0.25%	-0.005	-2.07 pp
Random Forest	+0.50%	+0.93%	+0.020	0.00 pp
Gradient Boosting	-0.71%	-0.93%	-0.020	-2.07 pp

Positive improvement values indicate lower prediction error after adding sentiment; pp = percentage points.

These findings indicate that although sentiment was associated with returns at the contemporaneous level, its incremental value for prospective out-of-sample forecasting was small and model-dependent.

3.5 Feature Importance of Sentiment-Enhanced Models

The forecasting ability of the overall models was modest, but sentiment variables were meaningful in the internal structure of the tree-based models. Lagged RHSI was the most important feature (0.219), lagged STI was the second most important feature (0.194) and overnight sentiment was the third most important feature (0.149) in Random Forest. Lagged STI (0.225), RRMB (0.173) and RHSI (0.161) were the most important predictors in Gradient Boosting. Table 5 shows the summary of the feature importance for the Random Forest specification and Figure 4 shows the visualization of the feature importance.

Table 5. Feature importance in the sentiment-enhanced Random Forest model

Predictor	Importance
Lagged RHSI	0.219
Lagged STI	0.194
STI_on	0.149
Lagged RRMB	0.113
Lagged RHST	0.090

Lagged RSHIBOR	0.089
n_posts_on	0.079
Lagged AV	0.067

Broader market returns were the strongest predictor but lagged social-media sentiment and overnight sentiment also occupied prominent positions, suggesting that sentiment contained information used by the model even though its addition produced only limited improvement in out-of-sample performance.

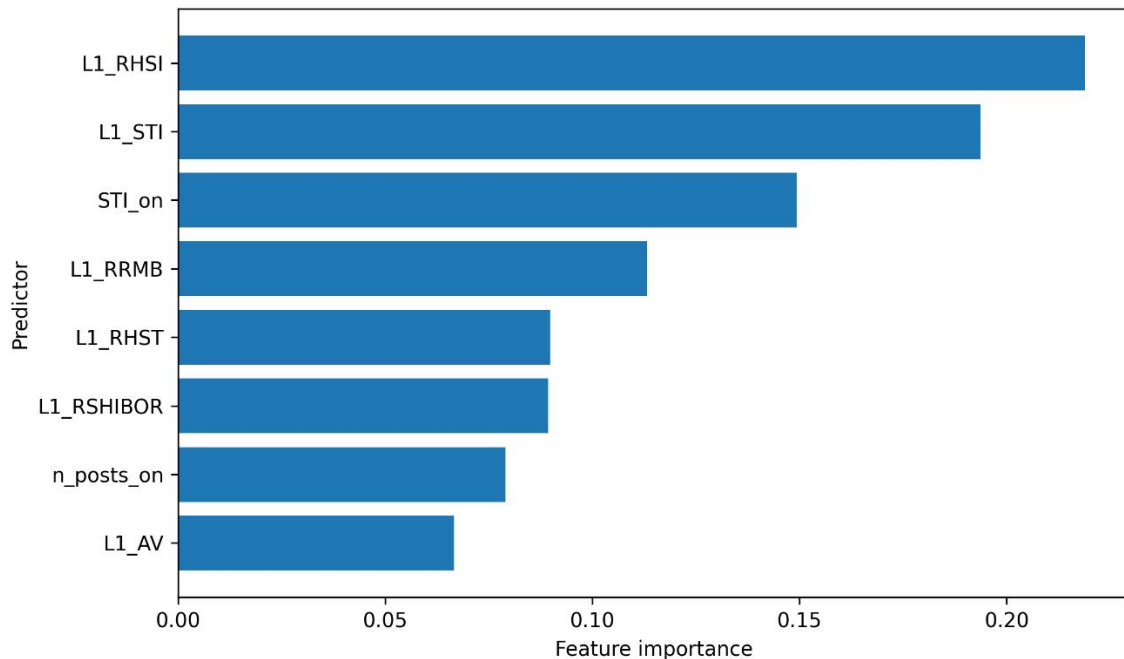


Figure 4. Feature importance in the sentiment-enhanced Random Forest model

3.6 Investment Decision-Support Performance

The sentiment enhanced linear regression model achieved a classification accuracy of 54.48% in the direction. The accuracy, recall and F1 score for the positive-return days were 53.41%, 65.28% and 58.75%, respectively. The model was found to correctly classify 47 positive-return days, 32 negative-return days and directionally misclassify 66 test observations of 145 total observations. These findings indicate that the model delivered only partial market-direction signals and thus could be regarded as a market-direction analytical decision-support signal, but not as a trading rule.

Overall, results indicate strong contemporaneous links between social-media sentiment and stock-market returns, but weaker support for the view that sentiment significantly enhances stock-market returns forecasts for the following periods, especially for STI. The results thus differentiate association and predictive usefulness, the latter of which is still restricted with the machine-learning specifications tested.

4. Discussion

The study investigated the possibility of using social media sentiment to predict stock market returns and aid in data-driven investment decisions. The results indicated that social media sentiment today had moderate and positive relationship with Hang Seng Technology returns, while sentiment yesterday was only weakly and insignificantly associated with returns today. That's relevant because it means that sentiment could be stronger than predictive of future market moves. Past studies have also shown that sentiment-based data can aid in predicting stock movements, especially when advanced sentiment processing methods are employed (Abdelfattah et al., 2024). Financial decision-making research also suggests that investor decisions can be impacted by the sentiment of social media, but the predictive value may differ, depending on the market and analytical context (Yoon & Oh, 2022).

The predictive modelling results revealed that sentiment-enhanced models did not always outperform the market-only models. Linear regression yielded the greatest directional accuracy of the methods examined, while there was

little reduction in prediction error with the addition of sentiment variables for Random Forest. However, these results suggest that sentiment information might provide some incremental value, but it probably will be small when compared to more commonly used market factors. Some research papers that have explored both technical and sentiment indicators have reported better forecasting results, indicating that sentiment might be strongly dependent on feature construction and model architecture (Agrawal et al., 2025). Awad et al. (2023) have also shown that more sophisticated methods, including deep reinforcement learning, can learn nonlinear models that provide good performance in complex financial scenarios, but not necessarily in every market environment.

The negative prospective effect of overnight sentiment in the current analysis contrasts with studies on other financial areas in which sentiment has been found to have a greater predictive effect. For instance, AI sentiment models have been effectively used in cryptocurrency to predict volatility and trends, as online discourse about cryptocurrencies is prolific (Srivastava et al., 2023). Likewise, tweet-based and explainable hybrid models have shown that the content of social media can impact stock price forecasting under certain conditions (Gandhudi et al., 2024). Twitter investor sentiment has also been successfully applied in the context of renewable-energy stock forecasting, contributing to the general concept that investor-provided information in the online sphere can include market signals (Herrera et al., 2022). Social-network sentiment has also been proven to help boost the accuracy of stock prediction in deep-learning models under certain contexts (Huang et al., 2023).

The results are further explained by the feature-importance results. Lagged broader market returns continued to be one of the most influential predictors, while other factors such as lagged sentiment and overnight sentiment were also relatively influential in the tree-based models. This agrees with previous findings that the technical indicators could be complemented by sentiment indicators, enhancing the informativeness of predictive systems (Lee et al., 2024). Similarly, explainable deep-learning technologies have highlighted the need to comprehend what variables impact predictions rather than simply focusing on overall accuracy (Muhammad et al., 2024). Other studies on applications of neural networks have also revealed that market context, input selection, and the validation strategy play significant roles in the predictive outcome of the model (Oyewole et al., 2024). In addition, LSTM models enhanced with social media have also shown the usefulness of sequential information when sentiment is temporally aligned with market outcomes (Peivandizadeh et al., 2024).

In terms of investment decisions, the models were shown to have only limited ability to differentiate between good days and bad days given the observed level of directionality (around 54%). This aligns with findings suggesting sentiment can serve as a better indicator of overall volatility or trends rather than for making precise daily return predictions (Liu et al., 2023). However, sentiment analysis along with other analysis methods have been suggested to be useful for investors as a decision support tool if used with caution (Shashidhar et al., 2023). Other studies have also pointed out the importance of sentiment analysis for investment decision support in highly dynamic markets, such as hybrid CNN-LSTM sentiment analysis (Roumeliotis et al., 2024) and multimodal sentiment analysis like blending technical indicators with social sentiment (Wang et al., 2023).

There are several caveats to be noted. The analysis was limited to a particular market situation and a rather short period of observation per day, and the investment decision was based on the direction of predicted returns, not on actual observable investor behaviour. Future research needs to include longer timeframes, other ways of building sentences, more frequent data and more sophisticated sequential or ensemble models. Further testing in other markets would also shed light on the generalizability of the relatively small incremental value of sentiment identified in this study.

5. Conclusion

In this study, the researchers have used the daily financial and sentiment-related information to assess the usefulness of social media sentiment in predicting stock market returns and investment decision making. Results demonstrated that the contemporaneous social media sentiment was positively correlated with Hang Seng Technology returns, suggesting that social media sentiment captures current market sentiment, to a certain degree. But overnight sentiment did not show strong association with lagged returns, indicating that sentiment was more of a contemporaneous indicator than a good stand-alone predictor. The machine-learning analysis also revealed that sentiment variables did not necessarily lead to better out-of-sample predictive power. Linear regression performed best in terms of directional accuracy of the models tested, and the forecasting error of Random Forest was marginally decreased after sentiment variables were added. However, feature-importance analysis showed that besides the general market and financial features lagged sentiment and overnight sentiment provided additional valuable information. The results indicate that social media sentiment can supplement traditional market metrics, while offering limited incremental predictive value, which is model specific. In an investment context, the models gave only limited assistance in the determination of the direction of the market, but not in making consistently accurate daily forecasts. As such, sentiment indicators should be used not as standalone trading signals, but as supplemental tools to help make decisions. Future research needs to consider longer time periods, more frequent observations, other

sentiment indicators, more sophisticated sequential models, and validation across markets to enhance robustness and applicability.

References

1. Abdelfattah, B. A., Darwish, S. M., & Elkaffas, S. M. (2024). Enhancing the prediction of stock market movement using neutrosophic-logic-based sentiment analysis. *Journal of theoretical and applied electronic commerce research*, 19(1), 116-134.
2. Yoon, J., & Oh, G. (2022). Investor herding behavior in social media sentiment. *Frontiers in Physics*, 10, 1023071.
3. Agrawal, S., Kumar, N., Rathee, G., Kerrache, C. A., Calafate, C. T., & Bilal, M. (2025). Improving stock market prediction accuracy using sentiment and technical analysis: S. Agrawal et al. *Electronic Commerce Research*, 25(5), 4103-4126.
4. Awad, A. L., Elkaffas, S. M., & Fakhr, M. W. (2023). Stock market prediction using deep reinforcement learning. *Applied System Innovation*, 6(6), 106.
5. Srivastava, S., Pant, M., & Gupta, V. (2023). Analysis and prediction of Indian stock market: a machine-learning approach. *International Journal of System Assurance Engineering and Management*, 14(4), 1567-1585.
6. Gandhudi, M., Alphonse, P. J. A., Fiore, U., & Gangadharan, G. R. (2024). Explainable hybrid quantum neural networks for analyzing the influence of tweets on stock price prediction. *Computers and Electrical Engineering*, 118, 109302.
7. Colombo, J. A., Akhter, T., Wanke, P., Azad, M. A. K., Tan, Y., Edalatpanah, S. A., & Antunes, J. (2023). Interplay of cryptocurrencies with financial and social media indicators: an entropy-weighted neural-MADM approach. *Journal of operational and strategic analytics*, 1(4), 160-172.
8. Hasselgren, B., Chrysoulas, C., Pitropakis, N., & Buchanan, W. J. (2022). Using social media & sentiment analysis to make investment decisions. *Future Internet*, 15(1), 5.
9. Herrera, G. P., Constantino, M., Su, J. J., & Naranpanawa, A. (2022). Renewable energy stocks forecast using Twitter investor sentiment and deep learning. *Energy Economics*, 114, 106285.
10. Huang, J. Y., Tung, C. L., & Lin, W. Z. (2023). Using social network sentiment analysis and genetic algorithm to improve the stock prediction accuracy of the deep learning-based approach. *International Journal of Computational Intelligence Systems*, 16(1), 93.
11. Jain, R., & Vanzara, R. (2023). Emerging trends in AI-based stock market prediction: A comprehensive and systematic review. *Engineering Proceedings*, 56(1), 254.
12. Khan, W., Ghazanfar, M. A., Azam, M. A., Karami, A., Alyoubi, K. H., & Alfakeeh, A. S. (2022). Stock market prediction using machine learning classifiers and social media, news: W. Khan et al. *Journal of Ambient Intelligence and Humanized Computing*, 13(7), 3433-3456.
13. Lee, H., Kim, J. H., & Jung, H. S. (2024). Deep-learning-based stock market prediction incorporating ESG sentiment and technical indicators. *Scientific Reports*, 14(1), 10262.
14. Lin, C. Y., & Marques, J. A. L. (2024). Stock market prediction using artificial intelligence: A systematic review of systematic reviews. *Social Sciences & Humanities Open*, 9, 100864.
15. Muhammad, D., Ahmed, I., Naveed, K., & Bendeche, M. (2024). An explainable deep learning approach for stock market trend prediction. *Heliyon*, 10(21).
16. Oyewole, A. T., Adeoye, O. B., Addy, W. A., Okoye, C. C., Ofodile, O. C., & Ugochukwu, C. E. (2024). Predicting stock market movements using neural networks: A review and application study. *Computer Science & IT Research Journal*, 5(3), 651-670.
17. Parveen, S., Satti, Z. W., Subhan, Q. A., Riaz, N., Baber, S. F., & Bashir, T. (2023). Examining investors' sentiments, behavioral biases and investment decisions during COVID-19 in the emerging stock market: a case of Pakistan stock market. *Journal of Economic and Administrative Sciences*, 39(3), 549-570.
18. Peivandizadeh, A., Hatami, S., Nakhjavani, A., Khoshshima, L., Qazani, M. R. C., Haleem, M., & Alizadehsani, R. (2024). Stock market prediction with transductive long short-term memory and social media sentiment analysis. *IEEE Access*, 12, 87110-87130.
19. Pillai, V. (2023). Integrating AI-driven techniques in big data analytics: Enhancing decision-making in financial markets. *International Journal of Engineering and Computer Science*, 12(07), 10-18535.
20. Qian, C., Mathur, N., Zakaria, N. H., Arora, R., Gupta, V., & Ali, M. (2022). Understanding public opinions on social media for financial sentiment analysis using AI-based techniques. *Information Processing & Management*, 59(6), 103098.
21. Liu, J. X., Leu, J. S., & Holst, S. (2023). Stock price movement prediction based on Stocktwits investor sentiment using FinBERT and ensemble SVM. *PeerJ Computer Science*, 9, e1403.
22. Sathya, N., & Prabhavathi, C. (2024). The influence of social media on investment decision-making: examining behavioral biases, risk perception, and mediation effects. *International Journal of System Assurance Engineering and Management*, 15(3), 957-963.
23. Sharma, N., Soni, M., Kumar, S., Kumar, R., Deb, N., & Shrivastava, A. (2023). Supervised machine learning method for ontology-based financial decisions in the stock market. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 22(5), 1-24.
24. Shashidhar, R., Aditya, V., Srihari, N., & Subhash, M. H. (2023, November). Empowering investors: Insights from sentiment analysis, FFT, and regression in Indian stock markets. In *2023 International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics (AIKIIIE)* (pp. 01-06). IEEE.

25. Roumeliotis, K. I., Tselikas, N. D., & Nasiopoulos, D. K. (2024). LLMs and NLP models in cryptocurrency sentiment analysis: A comparative classification study. *Big Data and Cognitive Computing*, 8(6), 63.
26. Wang, Z., Hu, Z., Li, F., Ho, S. B., & Cambria, E. (2023). Learning-based stock trending prediction by incorporating technical indicators and social media sentiment. *Cognitive Computation*, 15(3), 1092-1102.
27. Xia, Y., & Madni, G. R. (2024). Unleashing the behavioral factors affecting the decision making of Chinese investors in stock markets. *Plos one*, 19(2), e0298797.
28. Xu, Z. (2026). Replication Code for "Social Media Investor Sentiment and the Hang Seng Tech Index: Asymmetric Bidirectional Predictability and Regime Dependence" (Version v1) [Computer software]. Zenodo. <https://doi.org/10.5281/zenodo.21759882>.
29. Yasir, M., Attique, M., Latif, K., Chaudhary, G. M., Afzal, S., Ahmed, K., & Shahzad, F. (2023). Deep-learning-assisted business intelligence model for cryptocurrency forecasting using social media sentiment. *Journal of Enterprise Information Management*, 36(3), 718-733.
30. Zaman, N., Ghazanfar, M. A., Anwar, M., Lee, S. W., Qazi, N., Karimi, A., & Javed, A. (2023). Stock market prediction based on machine learning and social sentiment analysis.