

A Systematic Review on Social Media Sentiment in Stock Price Prediction: Data Sources, Machine Learning Models, and the Evidence on Effectiveness

Naufal Andila Fauzan¹ and Muhammad Rifqi Arviansyah²

¹Digital Business Program, Faculty of Economic and Business, Universitas Padjadjaran, Jawa Barat, Indonesia

²Department of Management and Business, Faculty of Economics and Business, Universitas Padjadjaran, Bandung 40132, Indonesia.

*Corresponding author: naufal21002@mail.unpad.ac.id

ABSTRACT

As the rapid spread of information through social media has increased and sentiment-driven market movements have become more frequent, it's necessary to understand how social media influences stock price prediction. This paper presents a systematic literature review (SLR) of recent studies (2020–2024) focusing on stock price prediction that incorporates social media sentiment. The review follows the PRISMA framework, involving the selection and analysis of 22 peer-reviewed articles based on predefined criteria. The review examines the types of social media platforms used, the sources of stock price data, the machine learning models applied, and the limited evidence available on the effectiveness of incorporating social media sentiment into stock price prediction. The findings reveal that X (formerly Twitter) is the most commonly used social media platform for obtaining data, and historical stock prices were primarily retrieved from Yahoo Finance. Moreover, LSTM is the most utilized model across studies, followed by RNN and other deep learning variants. However, only single research directly compares prediction performance with and without the sentiment data, revealing a 2.2% improvement. This review summarizes current trends, highlights methodological limitations, and proposes directions for future studies, including the integration of transformer-based models and more intensive evaluation frameworks for sentiment impact.

Keywords: Systematic Literature Review (SLR), Social Media Sentiment, Sentiment Analysis, Stock Price Prediction.

1. INTRODUCTION

Stock markets play a crucial role in global economic advancement by enabling companies to raise capital and investors to acquire ownership in the hope of generating financial returns (Ashtiani & Raahemi, 2023; R. Kumar, 2014; Mishura, 2016). To explain stock market behavior, two main theories have been proposed. According to EMH theory, the efficient market price is fully mirrored by standard and available information that is so general that the suggestion is not testable, and the term fully mirrored needs to be specified first (Fama, 1970). Furthermore, this theory posits that market contributors are logical and reasoned economic entities that act on their self-interest and participate in decisions based on how outstanding they are, exchanging advantages and costs that are statistically accurate (Lo, 2005). Meanwhile, the second theory, which became part of behavioural finance, set out to confront this theory. AMH, or adaptive market hypothesis, explains that, unlike in EMH, where the market contributor is perfectly logical and reasoned, errors can occur, and these mistakes serve as learning points that can be improved upon in the future. As faulty happened, this theory suggests that a stock might not always trade according to its asset value (Lo, 2005).

There are many theories attempting to explain how a stock price moves; one of the most popular is the random walk theory. This theory posits that stock prices follow a random pattern; therefore, they cannot be predicted (Chitenderu, 2014). However, in recent years, many researchers have acknowledged that the financial market is predictable. Nevertheless, recent models and approaches still fail to satisfy the perfect stock market prediction, which remains unresolved. These circumstances have presented a challenge for researchers to develop new methods for predicting the stock market (Ashtiani & Raahemi, 2023).

To predict the stock market, researchers employ two approaches: technical modeling and fundamental modeling. Technical modeling identifies statistical patterns by observing phenomena that directly affect stock prices (Gerencsér et al., 2009). His approach utilizes the changes and variations in the past stock price to create a prediction of its future value. They use data that contains past historical stock prices and preprocess it, calculating specific indicators such as the mean average convergence divergence and the directional movement index, which are used as indicators. Then, they apply these indicators to the predictive models (Saud & Shakya, 2024).

The fundamental modelling approach involves assessing the fair value of the stock based on information about the business's present and future earnings, including the company's financial ratios, such as the earnings multiplier model, price-to-sales ratio, and price-to-book ratio (Wafi et al., 2015). Another essential factor in fundamental modelling is the implementation of data extraction techniques using neural language processing (NLP). This process involves ensuring that information from a text is fully extracted and interpreted appropriately based on its context, which makes it challenging. The text sources range from news and social media to financial reports (Ashtiani & Raahemi, 2023).

Text data from social media has been proven to be a suitable data source for stock price prediction. This statement has been confirmed by numerous studies that have demonstrated its capability to forecast stock market prices (Coyne et al., 2017; Derakhshan & Beigy, 2019). To transform meaningful insights from textual social media data into predictions of a stock's price, we extract sentiment using text transformation and representation methods that convert text into numeric variables as input for machine learning to create the prediction. Moreover, advancements in NLP have led to more advanced and efficient methods that enhance the performance of models (Ashtiani & Raahemi, 2023).

The process of identifying and classifying opinions in the form of text to understand the author's stance on specific issues and categorizing their opinion as positive, negative, or neutral is the definition of sentiment analysis (Tanaltay et al., 2021). Researchers commonly employ two approaches when performing sentiment analysis: the machine-learning approach and the lexicon-based approach. The machine learning approach focuses on classifying a text based on specified linguistic features. In contrast, the lexicon-based method utilizes a list of opinion words used to describe expressions or feelings (Hauthal et al., 2020). Sentiment analysis has been applied in various ways, including the finance market, where it is used to predict stock prices. For instance, the research (Tanaltay et al., 2021) proposes involving Twitter data as a variable indicator for stock price prediction.

While several literature reviews exist on stock market prediction, few focus specifically on the integration of social media sentiment and its measurable effectiveness. Past SLRs have

either explored broader applications of NLP in finances (Khadjeh Nassirtoussi et al., 2014; Pinto et al., 2021; Soni et al., 2022; Usmani & Shamsi, 2021), (Phillips et al., 2017) or reviewed sentiment-based forecasting using traditional news sources (Ashtiani & Raahemi, 2023).

This review paper aims to address this gap by analyzing recently published studies from 2020 to 2024 that incorporate social media sentiment into stock price prediction models, contributing fresh insights. In particular, we examine which social media platforms and stock price data sources are most commonly utilized, which machine learning models are most frequently applied, and what evidence exists regarding the effectiveness of incorporating social media sentiment into stock price prediction models. By focusing on stock price prediction that is driven by social media sentiment, this review provides summaries of current trends in data sources and ML models and a critical analysis of the methodological gaps, including direction for future research.

Overall, the paper's structure is as follows: Section 2 outlines the research methodology, including database selection, search strategy, and screening criteria, all based on the PRISMA framework. Section 3 presents the key findings related to each research question. Section 4 discusses the implications and identifies research gaps, while Section 5 concludes the review.

2. METHODS

The SLR, or systematic literature review, is defined as a process of analyzing, assessing, and understanding the available studies that are relevant to a specific field of interest, research question, or targeted topic (Kitchenham, 2004). This review was conducted using the PRISMA framework to provide accurate and transparent research (Sarkis-Onofre et al., 2021). Following the framework, the review process will consist of four main steps: defining the research questions, conducting a literature search, screening the search results to retain only the relevant studies, and extracting and analyzing the data.

2.1 Research Questions

Several factors influence the performance of the predictor model. These factors include the social media platform being used, the source of stock price data, the machine learning models employed, and a comparison of the effectiveness of prediction utilizing social media versus not using it. Based on these factors, we define the research questions, including the most common social media used for stock price prediction, and the primary stock exchange sources that provide the stock price data for the prediction, the machine learning model that the researchers mainly used to make the stock price prediction using social media sentiment analysis, and stock price prediction effectiveness when incorporating social media sentiment compared to not incorporating it.

2.2 Research Strategy

This review search process is based on the guidelines released by PRISMA, which follow three steps: identifying relevant papers, screening papers, and the inclusion stage. Figure 1 visualizes the selection process.

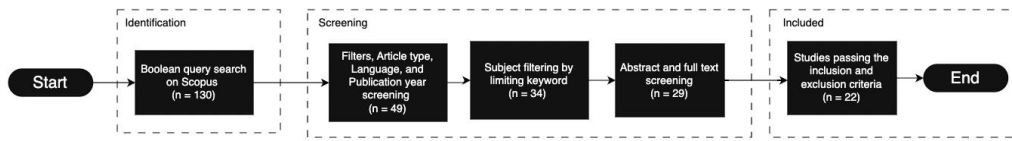


Figure 1. The PRISMA Diagram for The Study Selection Process.

During the identifying process, the review utilized research papers that are Scopus-indexed as Scopus stands as the largest global database for abstract and citation of scientific literature (Schotten et al., 2017) and has proven to be more comprehensive and data valid compared to other publishing, such as World of Science (WoS) (Pranckutė, 2021). The study gathered was also published between 2020 and 2024, and the queries were performed using keywords such as social media, sentiment analysis, and stock price prediction. To combine the keywords, we implement the AND logical operator. The search is restricted to titles, abstracts, and keywords.

2.3 Inclusion and Exclusion Criteria

Table 1. Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
The paper was published between the year 2020 to 2025	The study does not include specific keywords such as sentiment analysis, forecasting, social media, stock price, stock price prediction
The paper was written in English	The paper was not written in English
The article type research article	The research does not focus on the use of social media sentiment analysis to predict stock price

During the screening phase, to ensure that the studies gathered are relevant to our objective, inclusion and exclusion criteria are appropriately implemented. The rule is that papers that contain as little as one exclusion criterion must be excluded. Table 1 outlines the criteria for inclusion and exclusion that are applied.

Considering the criteria, we examine the title and abstract to identify relevant articles. If these points remain unclear, we will review the full text for further clarification. The results show that 22 articles meet the criteria and proceed to the next stage of analysis.

2.4 Data Analysis

During this process, the key findings from the primary studies that met the inclusion criteria were analyzed. Then, two types of data were extracted from each article. The first section provides information about the author, title, publication year, the journal in which the article was published, and a link to the article. The second one is the extraction of the key finding that is relevant to the research question, such as the social media platform used, stock

price source, the machine learning model being used, and the effectiveness compared to not incorporating social media sentiment, which is considered a relevant and crucial element to discuss in SLR.

2.5 Research Content

This review consists of three parts: data sources (including social media platforms and stock exchange sources), machine learning models, and a comparison of effectiveness. In the first part, we identify and categorize the data sources used for social media sentiment analysis and historical stock prices. The second part reviews various machine learning models applied during the sentiment analysis process for stock price prediction. The final section compares the accuracy of sentiment-based prediction with that of non-sentiment-based prediction. Additionally, in the conclusion, we highlight future exploration possibilities in the times to come. The data sources, machine learning models, and effectiveness comparison are reviewed in the subsequent sections.

3. RESULTS AND DISCUSSIONS

3.1 The Most Common Social Media and Main Stock Exchange Data Source

Social media is an internet-based platform that enables the creation and exchange of user-generated content, grounded in the fundamental ideology and Web 2.0 technology (Kaplan & Haenlein, 2010). The term “social media” was first used by a Tokyo-based online media company called Matisse in 1944. It was the early age of internet commercialization when social media started to show its growth. As time passed, the number of social media users increased significantly, making it one of the most important aspects of the modern internet. Now, social media users utilize the platform for various activities, such as social purposes, romantic activity, business matters, job searching, and even building a professional network (Aichner et al., 2021).

The majority of reviewed papers, or 65% of them, utilized X (formerly Twitter) to gather the text data being used for social media sentiment (Abdelfattah et al., 2024; Asgarian et al., 2023; Ho & Huang, 2021; Jin et al., 2023; S. K. Kumar et al., 2022; Ray et al., 2024; Tanaltay et al., 2021). X is a microblogging site that allows users to write and share their thoughts within the 140-character limit (Golbeck & Klavans, 2015). X data has been used numerous times for sentiment analysis purposes, as it has a character limit and has been utilized globally, representing opinions from diverse individuals across various demographics, including age, race, culture, and gender (Sailunaz & Alhajj, 2019). To retrieve X's user tweets based on the topic we want to explore in depth, an Application Programming Interface (API) is utilized. For example, X has been used to gather data about Turkish soccer clubs during a specific time (Tanaltay et al., 2021), collect Amazon company-related tweets (Peivandizadeh et al., 2024), and retrieve tweets that mention certain brands, such as Apple and Tesla, from 2021 to 2022 (Albahli & Nazir, 2024).

Moreover, the reviewed papers also utilize another social media platform as a source of data for sentiment analysis. These social media including Stocktwits (Abdelfattah et al., 2024; J. X. Liu et al., 2023), Weibo (Jiang et al., 2020), PTT Taiwan (J. Y. Huang et al., 2023), LinkedIn (Abdelfattah et al., 2024), Yahoo Message Board (Asgarian et al., 2023), Facebook

(Abdelfattah et al., 2024), Guba East Money (Zhou et al., 2021), and Dongfang Fortune Stock Bar (Q. Liu et al., 2023).

Aside from the social media sentiment data gathered from various platforms, historical data on the stock price is also needed to understand the stock's movement. To retrieve the historical data of the stock, this review found that Yahoo Finance (Abdelfattah et al., 2024; Asgarian et al., 2023; Ho & Huang, 2021; Jin et al., 2023; S. K. Kumar et al., 2022; Ray et al., 2024; Tanaltay et al., 2021), followed by NSE of India (Agrawal et al., 2024; Ahmed et al., 2022; Peivandizadeh et al., 2024), and Google Finance (Albahli et al., 2022; Albahli & Nazir, 2024) became most used stock exchange platform to do the job. These platforms are chosen as they have comprehensive data and are easy to access.

Moreover, the reviewed paper also use other stock exchange platform that put more attention towards the regional market such as Taiwan Stock Market (J. Y. Huang et al., 2023), Shanghai Stock Exchange (J. X. Liu et al., 2023), Chinese Stock Market (Zhou et al., 2021), cnYES (J. X. Liu et al., 2023), Guba (X. Huang & Song, 2021), and Xueqiu (X. Huang & Song, 2021). These regional stock exchanges are typically chosen to analyze a country's specific economic conditions, allowing for an understanding of local investor behavior and unique trading patterns on those markets.

3.2 The Most Common Machine Learning Model Used

Natural language processing, widely known as NLP, is an activity that involves a computer processing text or speech (Geman & Johnson, 2015). NLP, also referred to as computational linguistics, encompasses a range of learning processes and computational models used to address practical issues in the automatic analysis and understanding of human language, including both speech and text (Huynh-The et al., 2023). Sentiment analysis is an example of NLP implementation (Fang & Zhan, 2015).

This systematic literature review identified the machine learning models that are being applied during sentiment analysis as part of the stock price prediction process. The machine learning model that is frequently being used is the Long Short-Term Memory (LSTM) network, which was identified in 7 reviewed papers (Abdelfattah et al., 2024; Agrawal et al., 2024; J. Y. Huang et al., 2023; X. Huang & Song, 2021; Li & Zhang, 2023; Peivandizadeh et al., 2024; Swathi et al., 2022). This reflects the recent trend, as LSTM has been frequently applied to text classification activities due to its ability to comprehend long-term dependencies, commonly referred to as feedback connections (Singh et al., 2022).

Recurrent Neural Networks (RNNs) were the second most frequently used approach identified in three reviewed papers (Ahmed et al., 2022; Albahli & Nazir, 2024; Guo & Xie, 2024). RNNs were particularly popular in the context of text classification and language understanding. This model is shaped by the hidden state, which facilitates feedback loops during data processing while simultaneously producing the final output. The ability of sequential data processing that RNNs possess is the reason why many researchers prefer this model for handling complex data processing activities (Das et al., 2023).

The review also identified some other models that were being used less frequently. These models include Naïve Bayes (Tanaltay et al., 2021), Logistic Regression (Tanaltay et al., 2021), Extreme Learning Machine (ELM) (Albahli et al., 2022), as well as more advanced

such as Convolutional Neural Networks (CNNs) (Ho & Huang, 2021), Generative Adversarial Networks (GANs) (Asgarian et al., 2023), Multivariate Bayesian Structural Time Series (MBSTS) (Ray et al., 2024), Multi-Scale Temporal Convolutional Networks (M-TCN) (Ray et al., 2024), and finBERT (J. X. Liu et al., 2023).

3.3 Evidence on the Effectiveness of Incorporating Social Media Sentiment into Stock Price Prediction.

Researchers have found that combining historical stock prices with information gathered from social media can enhance the model's ability to create predictions, specifically in the case of stock prices (Nguyen et al., 2015). However, among the 22 reviewed papers, only one study (J. Y. Huang et al., 2023) directly held a comparative analysis between stock price prediction with and without the incorporation of social media sentiment.

The study by J. Y. Huang et al. (2023) finds that stock price prediction incorporating social media sentiment has a positive effect, leading to an enhancement in accuracy compared to one that does not integrate social media sentiment. It shows that stock price predictions that incorporate social media sentiment achieve 62.2% accuracy, a 2.2% increase compared to 60% accuracy without incorporating social media sentiment. This finding offers a glimpse of, yet limited, proof that supports the effectiveness of incorporating sentiment.

However, it is essential to keep in mind that this finding is based solely on one study, which suggests a significant gap in the literature. Many studies have assumed the effectiveness of social media sentiment when it was incorporated into the model, rather than empirically testing the assumption at its baseline. Future studies should examine this approach by implementing a controlled experiment to gain a better understanding of the actual value of social media sentiment in stock prediction tasks.

4. CONCLUSION

This systematic literature review was conducted to gain insight into various studies on the topic of using social media sentiment in stock price prediction. This review contributes to three main things. First, we will attempt to showcase which social media platform is most used for sentiment analysis and which stock exchange is the most frequently used as a data source for stock price projections. The researcher used data from various social media platforms as input for the sentiment analysis. X, formerly known as Twitter, has become the most widely used social media platform for researchers to gain insight into user sentiment. This is tied to the fact that data from X originated from global users, produced by various individuals with diverse backgrounds. X also proved to be a good sentiment due to its content availability and accessibility. Subsequently, historical data on a stock's past prices is also needed to forecast its future price. This review found that most researchers use Yahoo Finance to retrieve historical stock prices. Yahoo Finance was mainly used because it offers easy access and comprehensive data results.

Second, we identify the most commonly used machine learning models in sentiment analysis as part of the stock price prediction process. It has been discovered that the trend shows researchers mostly use Long Short-Term Memory (LSTM) for the sentiment analysis process. LSTM is frequently applied in text classification activities, given its ability to comprehend long-term dependencies, often referred to as feedback connections. However, the

researchers also employed other machine learning models, such as Recurrent Neural Networks and Naïve Bayes, among others.

Third, we show evidence on whether incorporating social media sentiment enhances the effectiveness of stock price prediction. Several factors influence the stock price, including the state of the economy, the company's financial condition, and others. While many researchers have claimed that sentiment from social media is a suitable aspect for forecasting stock prices, we found that most of the reviewed studies did not empirically test this claim. Only one study conducted a direct comparison of a model that incorporates corporate sentiment integration. The study found that the model accuracy increases by 2.2%. While this finding may seem to support the potential value of sentiment, it is insufficient to conclude the claim based on a single study. Therefore, rather than interpreting the finding as evidence of effectiveness, it highlights a notable gap that should be explored in future studies. Hence, future recommendations include implementing a direct evaluation of sentiment-influenced models against traditional baselines using standardized datasets to visualize the effect of social media sentiment on stock price prediction.

REFERENCES

- 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). <https://doi.org/10.3390/jtaer19010007>
- Agrawal, S., Kumar, N., Rathee, G., Kerrache, C. A., Calafate, C. T., & Bilal, M. (2024). Improving stock market prediction accuracy using sentiment and technical analysis. *Electronic Commerce Research*. <https://doi.org/10.1007/s10660-024-09874-x>
- Ahmed, D., Neema, R., Viswanadha, N., & Selvanambi, R. (2022). Analysis and Prediction of Healthcare Sector Stock Price Using Machine Learning Techniques: Healthcare Stock Analysis. *International Journal of Information System Modeling and Design*, 13(9). <https://doi.org/10.4018/IJISMD.303131>
- Aichner, T., Grünfelder, M., Maurer, O., & Jegeni, D. (2021). Twenty-Five Years of Social Media: A Review of Social Media Applications and Definitions from 1994 to 2019. In *Cyberpsychology, Behavior, and Social Networking* (Vol. 24, Issue 4). <https://doi.org/10.1089/cyber.2020.0134>
- Albahli, S., Irtaza, A., Nazir, T., Mehmood, A., Alkhalifah, A., & Albattah, W. (2022). A Machine Learning Method for Prediction of Stock Market Using Real-Time Twitter Data. *Electronics (Switzerland)*, 11(20). <https://doi.org/10.3390/electronics11203414>
- Albahli, S., & Nazir, T. (2024). Opinion mining for stock trend prediction using deep learning. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-024-19793-6>
- Asgarian, S., Ghasemi, R., & Momtazi, S. (2023). Generative adversarial network for sentiment-based stock prediction. *Concurrency and Computation: Practice and Experience*, 35(2). <https://doi.org/10.1002/cpe.7467>

- Ashtiani, M. N., & Raahemi, B. (2023). News-based intelligent prediction of financial markets using text mining and machine learning: A systematic literature review. *Expert Systems with Applications*, 217, 119509. <https://doi.org/10.1016/J.ESWA.2023.119509>
- Chitenderu, T. T. (2014). The Random Walk Theory And Stock Prices: Evidence From Johannesburg Stock Exchange. *International Business & Economics Research Journal*, 13(6).
- Coyne, S., Madiraju, P., & Coelho, J. (2017). Forecasting Stock Prices Using Social Media Analysis. *2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech)*, 1031–1038. <https://doi.org/10.1109/DASC-PICom-DataCom-CyberSciTec.2017.169>
- Das, S., Tariq, A., Santos, T., Kantareddy, S. S., & Banerjee, I. (2023). Recurrent Neural Networks (RNNs): Architectures, Training Tricks, and Introduction to Influential Research. In *Neuromethods* (Vol. 197). https://doi.org/10.1007/978-1-0716-3195-9_4
- Derakhshan, A., & Beigy, H. (2019). Sentiment analysis on stock social media for stock price movement prediction. *Engineering Applications of Artificial Intelligence*, 85, 569–578. <https://doi.org/10.1016/J.ENGAPPAI.2019.07.002>
- Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2). <https://doi.org/10.2307/2325486>
- Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data*, 2(1). <https://doi.org/10.1186/s40537-015-0015-2>
- Geman, S., & Johnson, M. (2015). Probabilistic Grammars and their Applications. In *International Encyclopedia of the Social & Behavioral Sciences: Second Edition*. <https://doi.org/10.1016/B978-0-08-097086-8.42161-6>
- Gerencsér, L., Torma, B., & Orlovits, Z. (2009). Fundamental Modelling of Financial Markets. *ERCIM News*, 2009.
- Golbeck, J., & Klavans, J. L. (2015). Introduction to Social Media Investigation: A Hands-on Approach. In *Introduction to Social Media Investigation: A Hands-on Approach*. <https://doi.org/10.1016/C2014-0-01104-5>
- Guo, K., & Xie, H. (2024). Deep learning in finance assessing twitter sentiment impact and prediction on stocks. *PeerJ Computer Science*, 10, e2018. <https://doi.org/10.7717/peerj-cs.2018>
- Hauthal, E., Burghardt, D., Fish, C., & Griffin, A. L. (2020). Sentiment Analysis. *International Encyclopedia of Human Geography, Second Edition*, 169–177. <https://doi.org/10.1016/B978-0-08-102295-5.10593-1>
- Ho, T. T., & Huang, Y. (2021). Stock price movement prediction using sentiment analysis and candlestick chart representation. *Sensors*, 21(23). <https://doi.org/10.3390/s21237957>

- 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). <https://doi.org/10.1007/s44196-023-00276-9>
- Huang, X., & Song, H. (2021). Investor Sentiment Combined with Multisource Information to Predict Stock Prices: An Analysis of China's A-Share Market. *Scientific Programming*, 2021. <https://doi.org/10.1155/2021/9094032>
- Huynh-The, T., Pham, Q. V., Pham, X. Q., Nguyen, T. T., Han, Z., & Kim, D. S. (2023). Artificial intelligence for the metaverse: A survey. *Engineering Applications of Artificial Intelligence*, 117, 105581. <https://doi.org/10.1016/J.ENGAPPAI.2022.105581>
- Jiang, H., Wang, X., & Zhu, M. (2020). Exploring propagation factors of social media moods for stock prices prediction. *Web Intelligence*, 18(3). <https://doi.org/10.3233/WEB-200441>
- Jin, D., Hu, Y., Chen, J., & Xia, M. (2023). Stock Price Trends Prediction Based on the Classical Models with Key Information Fusion of Ontologies. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 22(5). <https://doi.org/10.1145/3592599>
- Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, 53(1). <https://doi.org/10.1016/j.bushor.2009.09.003>
- Khadjeh Nassirtoussi, A., Aghabozorgi, S., Ying Wah, T., & Ngo, D. C. L. (2014). Text mining for market prediction: A systematic review. *Expert Systems with Applications*, 41(16), 7653–7670. <https://doi.org/10.1016/J.ESWA.2014.06.009>
- Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele University, UK and National ICT Australia*, 33. <https://doi.org/10.1.1.122.3308>
- Kumar, R. (2014). Stock Markets, Derivatives Markets, and Foreign Exchange Markets. *Strategies of Banks and Other Financial Institutions*, 125–164. <https://doi.org/10.1016/B978-0-12-416997-5.00005-1>
- Kumar, S. K., Akeji, A. A. A. R., Mithun, T., Ambika, M., Jabasheela, L., Walia, R., & Sakthi, U. (2022). Stock Price Prediction Using Optimal Network Based Twitter Sentiment Analysis. *Intelligent Automation and Soft Computing*, 33(2). <https://doi.org/10.32604/iasc.2022.024311>
- Li, M., & Zhang, Y. (2023). Integrating Social Media Data and Historical Stock Prices for Predictive Analysis: A Reinforcement Learning Approach. *International Journal of Advanced Computer Science and Applications*, 14(12). <https://doi.org/10.14569/IJACSA.2023.0141203>
- 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. <https://doi.org/10.7717/peerj-cs.1403>

- Liu, Q., Lee, W. S., Huang, M., & Wu, Q. (2023). Synergy between stock prices and investor sentiment in social media. *Borsa Istanbul Review*, 23(1). <https://doi.org/10.1016/j.bir.2022.09.006>
- Lo, A. W. (2005). Reconciling Efficient Market with Behavioral Finance: The ADAPTIVE MARKETS HYPOTHESIS. *The Journal of Investment Consulting*, 7(2).
- Mishura, Y. (2016). Introduction. *Finance Mathematics*, xi–xiv. <https://doi.org/10.1016/B978-1-78548-046-1.50006-5>
- Nguyen, T. H., Shirai, K., & Velcin, J. (2015). Sentiment analysis on social media for stock movement prediction. *Expert Systems with Applications*, 42(24), 9603–9611. <https://doi.org/10.1016/J.ESWA.2015.07.052>
- Peivandizadeh, A., Hatami, S., Nakhjavani, A., Khoshsim, L., Reza Chalak Qazani, M., Haleem, M., & Alizadehsani, R. (2024). Stock Market Prediction With Transductive Long Short-Term Memory and Social Media Sentiment Analysis. *IEEE Access*, 12, 87110–87130. <https://doi.org/10.1109/ACCESS.2024.3399548>
- Phillips, L., Dowling, C., Shaffer, K., Hodas, N., & Volkova, S. (2017). *Using Social Media to Predict the Future: A Systematic Literature Review*. <https://doi.org/10.48550/arXiv.1706.06134>
- Pinto, N., Da Silva Figueiredo, L., & Garcia, A. C. (2021). Automatic Prediction of Stock Market Behavior Based on Time Series, Text Mining and Sentiment Analysis: A Systematic Review. *Proceedings of the 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2021*, 1203–1208. <https://doi.org/10.1109/CSCWD49262.2021.9437732>
- Pranckutė, R. (2021). Web of Science (WoS) and Scopus: the titans of bibliographic information in today's academic world. In *Publications* (Vol. 9, Issue 1). <https://doi.org/10.3390/publications9010012>
- Ray, P., Ganguli, B., & Chakrabarti, A. (2024). Multivariate Bayesian Time-Series Model with Multi-temporal Convolution Network for Forecasting Stock Market During COVID-19 Pandemic. *International Journal of Computational Intelligence Systems*, 17(1), 170. <https://doi.org/10.1007/s44196-024-00525-5>
- Sailunaz, K., & Alhajj, R. (2019). Emotion and sentiment analysis from Twitter text. *Journal of Computational Science*, 36. <https://doi.org/10.1016/j.jocs.2019.05.009>
- Sarkis-Onofre, R., Catalá-López, F., Aromataris, E., & Lockwood, C. (2021). How to properly use the PRISMA Statement. In *Systematic Reviews* (Vol. 10, Issue 1). <https://doi.org/10.1186/s13643-021-01671-z>
- Saud, A. S., & Shakya, S. (2024). Technical indicator empowered intelligent strategies to predict stock trading signals. *Journal of Open Innovation: Technology, Market, and Complexity*, 10(4), 100398. <https://doi.org/https://doi.org/10.1016/j.joitmc.2024.100398>
- Schotten, M., El Aisati, M., Meester, W. J. N., Steiginga, S., & Ross, C. A. (2017). A brief history of Scopus: The world's largest abstract and citation database of scientific

- literature. In *Research Analytics: Boosting University Productivity and Competitiveness through Scientometrics*. <https://doi.org/10.1201/9781315155890>
- Singh, A., Dargar, S. K., Gupta, A., Kumar, A., Srivastava, A. K., Srivastava, M., Kumar Tiwari, P., & Ullah, M. A. (2022). Evolving Long Short-Term Memory Network-Based Text Classification. *Computational Intelligence and Neuroscience*, 2022. <https://doi.org/10.1155/2022/4725639>
- Soni, P., Tewari, Y., & Krishnan, D. (2022). Machine Learning Approaches in Stock Price Prediction: A Systematic Review. *Journal of Physics: Conference Series*, 2161(1), 012065. <https://doi.org/10.1088/1742-6596/2161/1/012065>
- Swathi, T., Kasiviswanath, N., & Rao, A. A. (2022). An optimal deep learning-based LSTM for stock price prediction using twitter sentiment analysis. *Applied Intelligence*, 52(12). <https://doi.org/10.1007/s10489-022-03175-2>
- Tanaltay, A., Langroudi, A. S., Akhavan-Tabatabaei, R., & Kasap, N. (2021). Can Social Media Predict Soccer Clubs' Stock Prices? The Case of Turkish Teams and Twitter. *SAGE Open*, 11(2). <https://doi.org/10.1177/21582440211004153>
- Usmani, S., & Shamsi, J. A. (2021). News sensitive stock market prediction: Literature review and suggestions. *PeerJ Computer Science*, 7, 1–36. <https://doi.org/10.7717/PEERJ-CS.490/SUPP-7>
- Wafi, Ahmed. S., Hassan, H., & Mabrouk, A. (2015). Fundamental Analysis Models in Financial Markets – Review Study. *Procedia Economics and Finance*, 30. [https://doi.org/10.1016/s2212-5671\(15\)01344-1](https://doi.org/10.1016/s2212-5671(15)01344-1)
- Zhou, Z., Gao, M., Xiao, H., Wang, R., & Liu, W. (2021). Big data and portfolio optimization: A novel approach integrating DEA with multiple data sources. *Omega (United Kingdom)*, 104. <https://doi.org/10.1016/j.omega.2021.102479>