A Deep Learning-Based Framework for Stock Analysis

Fausto Ricchiuti[1,†], Giancarlo Sperlí[1,*,†]

[1]University of Naples Federico II, via Claudio 21, Naples, 80131, Italy

Abstract

Artificial Intelligence (AI) has increasingly influenced financial markets, particularly in stock trading. However, existing AI-based forecasting methods often prioritize either sentiment analysis or technical indicators, neglecting a comprehensive approach that integrates multiple key factors. Additionally, traditional rule-based strategies frequently fail to account for broader market dynamics.

This paper is an extended abstract of a recent work in which we design a novel AI-driven advisory framework that leverages Long Short-Term Memory (LSTM) networks to enhance stock predictions by incorporating technical, contextual, and financial data. The model generates daily investment recommendations through an advanced Heuristic Stock Selection algorithm, refining decision-making based on forecasted trends. The framework was tested on 417 stocks and 67 cryptocurrencies over three years, demonstrating superior performance compared to existing models. Despite market downturns, the approach achieved a 41% profit in the stock market and a 39.38% return on cryptocurrency investments, showcasing its robustness across different financial environments.

Keywords

Stock Forecasting, Deep Learning, Stock Market, Multivariate Time Series

1. Introduction

The stock market operates as a platform where investors engage in the trading of company shares at agreed-upon prices [1]. As a fundamental pillar of the global economy, it offers the potential for financial gains, albeit with inherent risks that may surpass those of other investment avenues [2]. Consequently, forecasting stock market trends is a critical means to support practitioners and investors to mitigate investment risks, allowing stakeholders to make well-informed financial decisions.

The state-of-the-art methodologies can be categorized into economic and machine learning models. The former models are primarily focused on identifying linear dependencies within stock data [3]. However, the highly non-linear and volatile nature of financial markets presents considerable obstacles to their practical effectiveness [4]. In turn, machine learning models [5] have been designed to deal with these limitations [6].

However, machine learning performance is strongly influenced by both internal and external factors. The former mainly concerns the feature selection process, which is crucial for enhancing predictive accuracy [7], and the selection of an appropriate evaluation metric to assess model performance [3]. The latter is mainly due to the overwhelming volume of multimedia content published across different social media platforms [8]. Consequently, the inherent non-linearity and non-stationarity of stock market dynamics [7], coupled with the profound influence of external factors—such as public sentiment, expert analyses, and cybersecurity vulnerabilities [9].

In this paper, we design an Advisor Neural Network framework designed for daily investment decision-making, leveraging a Long Short-Term Memory (LSTM)-based Informative Stock Analysis model. The framework operates through a two-stage process: initially, it employs an LSTM-driven forecasting module that integrates technical indicators, contextual insights, and financial data to predict stock market movements and generate next-day investment recommendations. Additionally, seasonal patterns are incorporated into the model to enhance its ability to capture periodic fluctuations within stock market trends. To mitigate investment risks, we further propose a novel Heuristic Stock Selection

SEBD 2025: 33rd Symposium on Advanced Database Systems, June 16-19, 2025, Ischia, Italy

*Corresponding author.
†These authors contributed equally.
f.ricchiuti@studenti.unina.it (F. Ricchiuti); giancarlo.sperli@unina.it (G. Sperlí)

© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR-WS.org/Vol-4182/paper17.pdf

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073
published 2026-03-13

algorithm, which refines decision-making based on the deep learning module's output. The algorithm assigns a dynamic risk-reward score, where the positive score is inversely related to prediction errors, ensuring that stocks with lower forecast deviations are favored. Conversely, the negative score escalates with higher forecast inaccuracies, penalizing uncertain predictions and reducing the probability of financial losses. The proposed framework has been rigorously assessed using two datasets: the former encompasses over 400 NASDAQ-listed stocks, and the latter consists of 67 cryptocurrencies, both spanning a three-year period. To assess financial viability, we measured daily returns, ensuring the model's ability to navigate market risks effectively. Notably, stock transactions were executed at the market's opening and closing prices, with no adjustments for trading fees. Experimental findings indicate a notable capital appreciation exceeding 43%, even amid the bearish trend in NASDAQ during the analyzed quarter. Additionally, when applied to the cryptocurrency market, the framework yielded a 39.38% return on investment, further demonstrating its adaptability across different financial domains.

The structure of the paper is as follows: we revise the state-of-the-art approaches for stock forecasting in Section 2, while Section 3 presents our proposal integrating both financial and contextual data for providing stock suggestions. We evaluate the proposed framework on cryptocurrencies and stock markets domains, while Section 5 summarizes the main findings of our analysis.

2. Related Work

Stock market analysis has garnered significant attention from researchers and practitioners alike, driven by the need to understand the complex dynamics of stock price movements to minimize investment risks and enhance profitability. Recent studies, such as [9], have conducted extensive reviews of the literature, revealing a growing reliance on machine learning techniques for financial forecasting. Despite these advancements, the inherent non-linearity and volatility of stock markets continue to pose formidable challenges to predictive modeling [10].

To address these challenges, stock forecasting methodologies can be broadly categorized into model-based, statistical-based, and data-driven approaches. While model-based techniques [11, 12] struggle to accurately represent the intricate dependencies within financial data, statistical methods [13] often fall short due to the highly unpredictable nature of market trends. In contrast, data-driven approaches, particularly AI-driven models, have gained traction in financial research due to their ability to extract meaningful insights from vast and heterogeneous datasets [5].

Artificial Intelligence (AI) techniques have been widely applied to stock forecasting, with deep learning models offering promising results. Notably, Long Short-Term Memory (LSTM) networks have been demonstrated to outperform traditional econometric and machine learning models in time-series forecasting tasks [1]. However, several critical challenges remain, particularly regarding feature selection, where most studies rely on technical indicators and stock-specific data [14], often neglecting fundamental and contextual factors [1]. Alternative approaches, such as graph-based models [15], incorporate industry-centered relationships but still overlook the impact of real-time news and social media on stock price movements.

Another major challenge is the choice of evaluation metrics. Traditional measures like Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) are frequently used; however, risk-adjusted metrics and domain-specific indicators (e.g., return and volatility) provide a more accurate reflection of model performance [3]. Moreover, the unpredictable nature of financial markets, especially due to external shocks, limits the generalization capacity of AI models [16].

Recent research has explored integrating both textual and numerical data sources—such as news articles, social media sentiment, stock prices, and financial ratios—to improve prediction accuracy [17]. Notably, [18] found that news data is most effective for short-term forecasting (one-day horizon), whereas social media insights contribute valuable predictive power over multiple days. However, existing deep learning-based approaches tend to focus exclusively on sentiment analysis or technical indicators, without fully incorporating fundamental stock attributes or real-world contextual influences [19, 20].

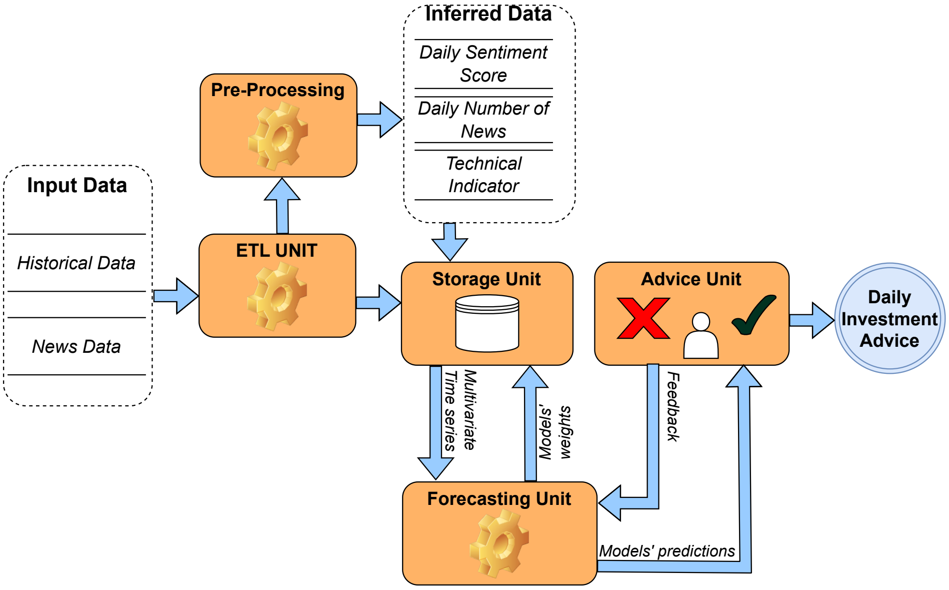


Figure 1: Overview of the proposed framework composed of three phases: i) *Data Ingestion*, ii) *Stock forecasting*, and iii) *Advice suggestions*.

Summarizing: i) we design a framework by combining historical financial data with news content more than focusing on a single modality [19]; ii) we design an heuristic selection algorithm using AI module output to identify stocks to invest on more than using handmade rules or technical indicators [21]; and iii) we integrate news content and financial information into the proposed framework to deal with unforeseen events.

3. Framework

Stock forecasting is a challenging task in financial analysis, involving the prediction of future stock prices as continuous numerical values [1]. The primary goal of enhancing stock forecasting is to reduce financial risks for institutional investors, such as government bodies, and to minimize potential losses in the broader financial market [22]. Hence, we design the proposed framework, whose architecture has been designed in Figure 1, to generate daily investment recommendations to assist investors in optimizing their financial strategies. The system is structured into three core components, guiding the process from data acquisition to stock selection: (i) Data Collection and Preprocessing, (ii) Predictive Stock Modeling, and (iii) Investment Recommendation Generation. Each module plays a crucial role in transforming raw financial and contextual data into actionable insights for informed decision-making.

3.1. Data Ingestion

This section provides a comprehensive analysis of the Data Ingestion module, detailing each stage from the initial data acquisition to the structured storage of processed financial and contextual information. While prior research [14] categorizes financial information into financial data and technical indicators, our approach extends this classification by incorporating a broader spectrum of contextual factors, thereby enhancing the robustness of stock market analysis. Specifically, we integrate the following key data sources: Financial Data, Technical Indicators, Seasonal Factors, and News Sentiment Data.

Historical financial data has been collected from Yahoo Finance[1]. We compute seasonality data in order to support forecasting module in learning the seasonality behavior of the stock behavior. News

[1]https://www.yahoo.com/author/yahoo-finance

data has been collected from the *EOD Historical Data*[2] platform, capturing the daily volume of news articles and the related sentiment scores for each security in terms of multivariate timeseries. We integrate five technical indicators (i.e., Awesome Oscillator, Relative Strength Index, Average True Range, Average Directional Movement Index, Aroon Indicator) in the proposed methodology, also removing the historical market data to reduce correlations between features.

3.2. Forecasting Unit

The Forecasting Unit serves as the computational core for designing, optimizing, and continuously refining the Neural Network architectures employed in predictive modeling. At this stage, we deploy and fine-tune Long Short-Term Memory (LSTM) networks, which have been empirically demonstrated to surpass conventional forecasting techniques in handling sequential dependencies and capturing temporal market fluctuations [2].

In this module, the relevant hyperparameter is the time window length, which determines the historical depth of stock price sequences used as input vectors for the LSTM-based forecasting model. This sliding window approach enables the neural network to learn latent temporal dependencies, facilitating the recognition of short-term trends, momentum shifts, and potential reversals within financial time-series data.

3.3. Advice Unit

The Advice Unit operates as a decision-making layer, leveraging the prediction scores generated by the Forecasting Unit to propose stocks that are expected to yield the highest returns, thereby offering actionable investment recommendations. Let S denote a set of securities, and this unit aims to recommend K stocks where the Daily Return is anticipated to outperform on the subsequent day. The Advice Unit can be conceptually divided into two primary functional components: i) Stocks Selection and ii) Advice Production.

During the first stage, a set of N forecasting models undergoes rigorous evaluation and iterative updates based on the data processed by the previous module. The output predictions are ranked, with the top-K highest predicted returns selected. This procedure not only identifies the most accurate models but also dynamically adapts to the current stock market conditions by selecting the models exhibiting the greatest predictive reliability for the forthcoming trading day. The models that are deemed most trustworthy based on today's input data are prioritized for future predictions.

In the last phase, the M models identified in the preceding step are employed to forecast the expected Daily Return for the next trading day. Using the outputs of these models, the predicted returns are again sorted in descending order. The K stocks with the highest predicted returns are then selected, forming the Daily Investment Advice, which serves as a recommendation for investment in the most promising stocks based on the generated forecasts. This multi-stage approach ensures that the selected stocks for the next trading day are backed by robust, contextually relevant predictions derived from the most reliable forecasting models.

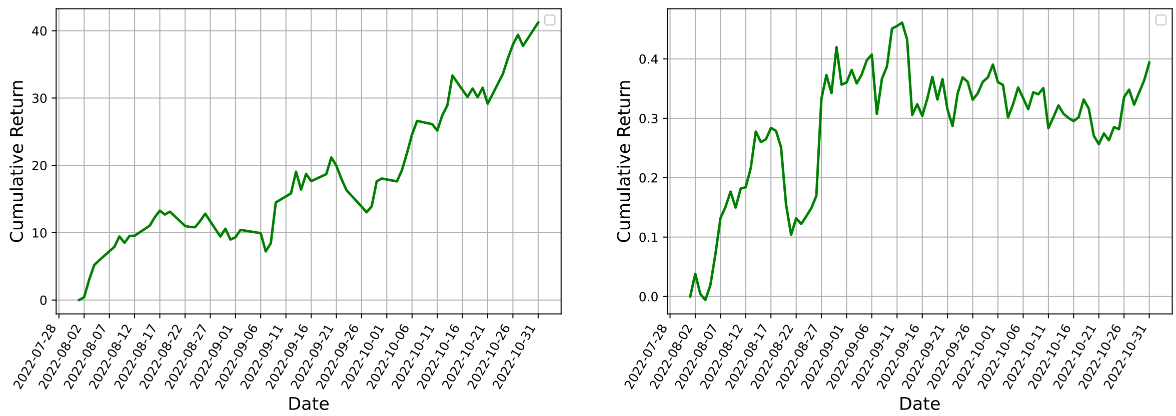
4. Experimental Analysis

The experimental investigation seeks to rigorously assess the efficacy of our proposal across two distinct financial domains: the stock market (497 stocks) and the cryptocurrency market (67 cryptocurrencies) over 3 3-month simulation period.

The third objective builds upon a subset of 417 stocks that have historical data available on the NASDAQ prior to October 2019. For each stock, we gathered data from 01/11/2019 to 31/10/2022 and trained an individual $LSTM$-based model for each.

Thus, the complete dataset comprises 690 entries per stock, with the following division: 480 and 140 samples are used for training and validation, while the remaining 70 entries compose the test set.

[2]https://eodhistoricaldata.com/



(a) Stock market.

(b) Cryptocurrencies.

Figure 2: Cumulative Percentage Economical Gain.

Similarly, an identical assessment was performed for the cryptocurrency market, based on a set of 67 cryptocurrencies, with data collection, training, and testing stages mirroring the stock market setup.

The experimental analysis was conducted on Google Colab (http://colab.research.google.com/), using the technology stack utilized for the analysis that is based primarily on Python 3.9, utilizing a variety of machine learning and deep learning libraries..

4.1. Results

This section presents the results of the simulation conducted on both the NASDAQ Stock and Cryptocurrency markets, focusing on the daily investment advice generated between 01/08/2022 and 31/10/2022 over 65 open market days. On each trading day, the framework evaluates the stocks, ranking them by evaluation score and selecting the top 50. These 50 stocks are then predicted for the next day's daily return, which is also sorted in descending order. The top 5 stocks with the highest expected returns are chosen as the daily investment advice.

A trading strategy was implemented to assess the efficiency of the proposal in generating advice for practitioners and investors. On a daily basis, the available capital was proportionally distributed across the five chosen stocks, purchasing at the market's opening price and selling at the closing price, with transaction costs excluded from the computation. The investment strategy maximized compound interest, calculating the overall economic gain based on the daily returns of the selected securities.

The framework achieved a 64.62% accuracy in predicting whether investments would result in a positive or negative return, with results distributed across 42 positive and 23 negative trading days.

In the next figures, we evaluate the performance of the proposed framework in terms of Cumulative Percentage Economic Gain. Specifically, daily Return reflects the growth of the available capital based on the returns from the previous day whilst Cumulative Percentage Economic Gain calculates the percentage change in capital compared to the initial value, providing a measure of the total gain up to day k

Figure 2a illustrates the progression of the Cumulative Percentage Economic Gain throughout the three-month simulation period. It is important to highlight that the framework delivers investment advice that results in a 41.21% increase in the initial capital. Furthermore, the same simulation methodology was applied to a set of 67 cryptocurrencies, with the simulation period extending over 92 trading days, compared to the 65 days of the stock market simulation, due to the cryptocurrency market being open every day. The analysis yielded an accuracy score of 59.78%, with positive and negative investment outcomes corresponding to 55 and 37 days, respectively. Figure 2b presents the Cumulative Percentage Economic Gain during the three-month cryptocurrency simulation period. The framework's advice led to a 39.38% economic gain from the initial capital.

We conduct a comparative analysis of the performance of our proposed method against several leading approaches on the test set. Initially, stock market behavior was predicted using an LSTM-based model, with its accuracy evaluated through common statistical metrics such as MAE, MSE, and RMSE to measure the divergence from the actual stock trends. Following this, the predictions from the model were incorporated into the stock selection strategy, which was then evaluated using the daily return metric—an essential indicator of investment performance.

The results presented in Table 1 illustrate the outcomes across both stock and cryptocurrency markets. Notably, our approach surpasses the baseline models by considering various feature types that provide a more comprehensive analysis of stock trends from multiple perspectives. Specifically, the integration of contextual data—such as news articles and sentiment analysis—alongside historical stock information significantly enhances the model's ability to predict future stock performance.

However, despite these improvements, Table 1 shows only a modest gain when compared to other baselines, particularly within the cryptocurrency market. This is likely due to the highly volatile and unpredictable nature of cryptocurrencies [23]. Unlike the stock market, where there are well-established reference points, predicting cryptocurrency price movements is especially challenging due to the lack of a regulated future market and the difficulty in identifying the key factors driving price fluctuations.

	Stock Market			Cryptocurrencies		
Approaches	**MAE**	**MAPE**	**RMSE**	**MAE**	**MAPE**	**RMSE**
[21]	3.294	0.023	4.305	4.112	0.029	5.375
[13]	3.228	0.023	4.219	4.030	0.028	5.268
[20]	3.146	0.022	4.113	3.991	0.028	5.217
[19]	2.950	0.021	3.856	3.870	0.027	5.059
[10]	3.490	0.025	4.562	4.172	0.029	5.454
Proposed	**2.728**	**0.019**	**3.566**	**3.749**	**0.026**	**4.900**

Table 1
Efficiency analysis of the proposal w.r.t. several state-of-the-art approaches.

In the final step, we benchmark our proposed advice strategy against the one presented by [24] across the entire test set. The results of this comparison, presented in Table 2, illustrate the performance on both the stock and cryptocurrency markets, with a focus on the cumulative return.

It is evident that our proposed approach surpasses the buy-and-hold strategy by 16.22% (45.05%) in the stock market and 16.47% (49.05%) in the cryptocurrency market. While the strategy outlined by [24] attempts to incorporate various features, it solely focuses on historical financial prices, neglecting external influences and a more comprehensive financial fundamental analysis.

5. Conclusion

This study presents an advanced investment advisory system built upon an $LSTM$-driven architecture, tailored for daily stock market recommendations with the objective of maximizing financial returns. The proposed framework integrates deep learning models to forecast stock movements by incorporating diverse data sources, such as historical market data, investor sentiment dynamics, and seasonal market patterns. Furthermore, a novel Heuristic Stock Selection mechanism refines the decision-making process by identifying securities exhibiting the highest predictability based on prior trading activity. To assess the efficiency of our proposal, a three-month simulation was conducted using over 400 NASDAQ stocks, yielding a 41.21% return on initial capital, despite a downturn in the market. It was also tested in the Cryptocurrency market, where it produced a 39.38% gain. Comparisons with state-of-the-art methods showed the proposed framework's superior performance in both markets. Additionally, the framework's low memory usage and fast training times were primarily attributed to the short three-day time window used during model training.

Future work will focus on expanding the dataset by including more markets and extending the time frame. Efforts will also be made to design more efficient trading strategies to further increase financial

	Cumulative Return	
	Stock	**Crypto**
Buy & hold strategy	28.31%	26.42%
[24]	35.46%	33.81%
Proposal	**41.21%**	**39.38%**

Table 2
Effectiveness comparison of the proposed advice strategy w.r.t. the buy & hold and the one designed by [24].

returns and explore methodologies for incorporating transaction costs into the investment recommendation framework. Additionally, exploring real-world relationships between financial products and their correlations will be a key area of development.

Funding

This work is supported by the Italian Ministry of University and Research (MUR) within the PRIN2022—ISALDI: Interpretable Stock Analysis Leveraging Deep multImodal models (CUP: E53D23008150006).

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] A. Thakkar, K. Chaudhari, A comprehensive survey on deep neural networks for stock market: The need, challenges, and future directions, Expert Systems with Applications 177 (2021) 114800. doi:https://doi.org/10.1016/j.eswa.2021.114800.

[2] K. Benidis, S. S. Rangapuram, V. Flunkert, Y. Wang, D. Maddix, C. Turkmen, J. Gasthaus, M. Bohlke-Schneider, D. Salinas, L. Stella, F.-X. Aubet, L. Callot, T. Januschowski, Deep Learning for Time Series Forecasting: Tutorial and Literature Survey, ACM Comput. Surv. 55 (2022). doi:10.1145/3533382.

[3] K. Olorunnimbe, H. Viktor, Deep learning in the stock market—a systematic survey of practice, backtesting, and applications, Artificial Intelligence Review (2022) 1–53. doi:https://doi.org/10.1007/s10462-022-10226-0.

[4] F. D. Paiva, R. T. N. Cardoso, G. P. Hanaoka, W. M. Duarte, Decision-making for financial trading: A fusion approach of machine learning and portfolio selection, Expert Systems with Applications 115 (2019) 635–655. doi:https://doi.org/10.1016/j.eswa.2018.08.003.

[5] M. M. Kumbure, C. Lohrmann, P. Luukka, J. Porras, Machine learning techniques and data for stock market forecasting: A literature review, Expert Systems with Applications 197 (2022) 116659. doi:https://doi.org/10.1016/j.eswa.2022.116659.

[6] Y. Yang, J. Guo, S. Sun, Y. Li, Forecasting crude oil price with a new hybrid approach and multi-source data, Engineering Applications of Artificial Intelligence 101 (2021) 104217. doi:https://doi.org/10.1016/j.engappai.2021.104217.

[7] H. H. Htun, M. Biehl, N. Petkov, Survey of feature selection and extraction techniques for stock market prediction, Financial Innovation 9 (2023) 26. doi:https://doi.org/10.1186/s40854-022-00441-7.

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

[9] T. Kehinde, F. T. Chan, S. Chung, Scientometric review and analysis of recent approaches to stock market forecasting: Two decades survey, Expert Systems with Applications 213 (2023) 119299. doi:https://doi.org/10.1016/j.eswa.2022.119299.

[10] J. Wang, Q. Cui, X. Sun, M. He, Asian stock markets closing index forecast based on secondary decomposition, multi-factor analysis and attention-based LSTM model, Engineering Applications of Artificial Intelligence 113 (2022) 104908. doi:https://doi.org/10.1016/j.engappai.2022.104908.

[11] C.-H. Cheng, J.-H. Yang, Fuzzy time-series model based on rough set rule induction for forecasting stock price, Neurocomputing 302 (2018) 33–45. doi:https://doi.org/10.1016/j.neucom.2018.04.014.

[12] L.-Y. Wei, A hybrid ANFIS model based on empirical mode decomposition for stock time series forecasting, Applied Soft Computing 42 (2016) 368–376. doi:https://doi.org/10.1016/j.asoc.2016.01.027.

[13] F. Hafiz, J. Broekaert, D. La Torre, A. Swain, Co-evolution of neural architectures and features for stock market forecasting: A multi-objective decision perspective, Decision Support Systems (2023) 114015. doi:https://doi.org/10.1016/j.dss.2023.114015.

[14] W. Chen, M. Jiang, W.-G. Zhang, Z. Chen, A novel graph convolutional feature based convolutional neural network for stock trend prediction, Information Sciences 556 (2021) 67–94. doi:https://doi.org/10.1016/j.ins.2020.12.068.

[15] W.-C. Huang, C.-T. Chen, C. Lee, F.-H. Kuo, S.-H. Huang, Attentive gated graph sequence neural network-based time-series information fusion for financial trading, Information Fusion 91 (2023) 261–276. doi:https://doi.org/10.1016/j.inffus.2022.10.006.

[16] F. Ronaghi, M. Salimibeni, F. Naderkhani, A. Mohammadi, COVID19-HPSMP: COVID-19 adopted Hybrid and Parallel deep information fusion framework for stock price movement prediction, Expert Systems with Applications 187 (2022) 115879. doi:https://doi.org/10.1016/j.eswa.2021.115879.

[17] X. Li, P. Wu, W. Wang, Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong, Information Processing & Management 57 (2020) 102212. doi:https://doi.org/10.1016/j.ipm.2020.102212.

[18] H. Dong, J. Ren, B. Padmanabhan, J. V. Nickerson, How are social and mass media different in relation to the stock market? A study on topic coverage and predictive value, Information & Management 59 (2022) 103588. doi:https://doi.org/10.1016/j.im.2021.103588.

[19] M.-Y. Chen, C.-H. Liao, R.-P. Hsieh, Modeling public mood and emotion: Stock market trend prediction with anticipatory computing approach, Computers in Human Behavior 101 (2019) 402–408. doi:https://doi.org/10.1016/j.chb.2019.03.021.

[20] S. Anbaee Farimani, M. Vafaei Jahan, A. Milani Fard, S. R. K. Tabbakh, Investigating the informativeness of technical indicators and news sentiment in financial market price prediction, Knowledge-Based Systems 247 (2022) 108742. doi:https://doi.org/10.1016/j.knosys.2022.108742.

[21] S. Banik, N. Sharma, M. Mangla, S. N. Mohanty, S. S., LSTM based decision support system for swing trading in stock market, Knowledge-Based Systems 239 (2022) 107994. doi:https://doi.org/10.1016/j.knosys.2021.107994.

[22] A. F. Kamara, E. Chen, Z. Pan, An ensemble of a boosted hybrid of deep learning models and technical analysis for forecasting stock prices, Information Sciences 594 (2022) 1–19. doi:https://doi.org/10.1016/j.ins.2022.02.015.

[23] H. Wang, X. Wang, S. Yin, H. Ji, The asymmetric contagion effect between stock market and cryptocurrency market, Finance Research Letters 46 (2022) 102345. doi:https://doi.org/10.1016/j.frl.2021.102345.

[24] P. Ghosh, A. Neufeld, J. K. Sahoo, Forecasting directional movements of stock prices for intraday trading using LSTM and random forests, Finance Research Letters 46 (2022) 102280. doi:https://doi.org/10.1016/j.frl.2021.102280.