International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(17s) 2026 pp. 70-84
https://doi.org/10.70917/ijcisim-2026-4706
Article

# Multi-modal and Optimised LSTM-CNN Attention based Framework for Prognosis of Dynamics in Stock Prices

**Muskan Aggarwal[1], Shikha Verma[2]**

[1]Ph. D Research Scholar, M. M. Institute of Computer Technology & Business Management, Maharishi Markandeshwar (Deemed to be University), Mullana, Haryana, 133207, India. muskanaggarwal2201@gmail.com
[2]M. M. Institute of Computer Technology & Business Management, Maharishi Markandeshwar (Deemed to be University), Mullana, Haryana, 133207, India. sverma5585@gmail.com

**Abstract:** However, the problem of equity price prediction is still complicated due to the interaction of the factors involved, which are saliently nonlinear, highly volatile and sentiment driven. This is just further complicated by the fact that investor sentiment is less or more unpredictable because it's not quantifiable, and it contains emotions. The research aims to present a hybrid deep learning structure of LSTM-CNN-Attention mechanism and technical, fundamental and sentiment-based features to enhance the accuracy of the forecasts. The model features three indicators sets, price trend, market volatility and sentiment indicators, the latter on the basis of the social media, online financial news and online newspapers. A method is used to eliminate the redundant inputs and to keep the best inputs of the described inputs, namely Principal Component Analysis (PCA) and Sequential Feature Selection (SFS). These are the feature selection techniques which enhance the capacity of the model to capture important market trends and behavioural indications. The comparison is carried out using the set of 10 industrial domains, involving five different deep learning architectures including LSTM, CNN, and LSTM-CNN hybrid, LSTM with Attention and LSTM-CNN-Attention. It is seen that the LSTMCNN Attention configuration performs best among the different configurations as it has the lowest Mean Absolute Error (MAE) in most industries. The results show the significance of feature optimization and sentiment integration to enhance predictive performance. The suggested framework provides a scalable, interpretable, and data-driven platform to real-time sentiment-based price forecasting of equity and intelligent investment decision support.
**Keywords:** Attention Mechanism, Feature Selections, Investor Sentiments, Long-Short Term Memory, Principal Components, Stock Price Forecasting

## 1. Introduction

The increase in the popularity of social media has led to immense amounts of textual information that capture the emotions and thoughts of the people (Zhen et al., 2025). The phenomenon provides the researchers with abundant corpora regarding sentiment prediction and emotion analysis, two indispensable natural language processing uses (Zhang and Mariano, 2024). Because of the fact that stock markets perform the role of an indicator of economic and financial activity in the future, equity movement prediction is a complicated task that is subject to influence by the global economy, politics, a company performance, and investor mood. Price changes are nonlinear in nature and usually noisy and highly sensitive to investor actions, making them especially difficult to predict (Ajiga et al., 2024). The latest improvements in field of AI and DL allow into a new world of capturing such dynamics, which is beyond the traditional models, e.g., ARIMA or Grey Forecasting, which tend to fail in the scenarios of volatility and high-dimensionality. The sentiment features based on social media, news, and financial forums have been included in recent scholarship. As an example, Abdelfattah et al. (2024) maximized a support vector machine (SVM) model with sentiment vectors, and Zhang and Mariano (2024) had modeled the Nifty 50 index with an LSTM model at 85.88

percent accuracy. However, overfitting is often induced by high dimensionality and redundancy and hence, it requires efficient feature-selection strategies. Regression, SVM and random forest (Singh et al., 2025) or multi-strategy combinations (Ren and Li, 2025) are some of the methods that improve predictive stability. In addition, the performance of the forecasts is enhanced by the external sources of sentiment: Narayana et al. (2025) boosted S&P 500 forecasts by adding ESG sentiment, and similar advantages were confirmed by Mu et al. (2023) by using a diversified set of inputs. Stock data on high-dimensional data is generally noisy and imbalanced and will need to be reduced to dimensions (Maqbool et al., 2022). The feature selection method will make sure that only the most meaningful predictors are included in DL models, including historical prices, trading volumes, and technical indicators (Koukaras et al., 2022). Moreover, the multisource behaviours have been shown to be necessary in the process of capturing investor psychology as well as market behaviour. Section 2 reviews the existing literature, and identifies the significant research gaps, and the remaining manuscript is presented in this format. In section 3, the methodology or procedure learnt i.e., feature-selection process, formulating sentiment and deep-learning architecture of the models used is explained. Section 4 gives an experimental design and critically comments on the experimental results with the best performance of the LSTM -CNN -Attention model in the case of multisource sentiment information added. Finally, Section 5 will consolidate the findings of the study, its limitations and suggestions to be implemented in future studies.

## 2. Correlated Works

### *2.1. Deep learning (DL)*

Stock price forecasting remains very challenging task and is mainly attributed to high volatility in the markets. The recent development in the domain of deep learning has resulted in powerful methodologies that help address these challenges (Ahmad and Umar, 2023). The evidence-based practice shows that LSTM models are better than traditional and other deep learning options due to their ability to identify consecutive patterns (Kasture and Shirsath, 2024; Ho and Huang, 2021). Furthermore, CNN-LSTM networks that combine CNN feature integration with LSTM time learning have also shown significant effectiveness in forecasting stocks and in portfolio optimization (Haryono et al., 2023; Deng et al., 2023). Therefore, LSTM and CNN -LSTM models are crucial in improving the precision of stock price forecasting.

### *2.2. Factors which influence Historical Prices*

Share price forecasting is a machine-learning prediction that is sensitive to the prudent choice and combination of relevant features. Relevant variables are minimized by effective feature selection, computational load is minimized, over-fitting is minimized and predictive precision is also enhanced (Gupta et al., 2022). Recursive Feature Elimination (RFE) represents another powerful method; Dahish and Miah (2023) used RFE on the sixty-features, and a final result of the process is forty-eight explanatory factors regarding the stock returns. Embedded and ensemble methods, especially Random Forest, combine both feature selection and model construction to support performance. The research by Chaudhari and Mahajan (2025) and Chang et al. (2024) supports the usefulness of Random Forest to predict the price of stocks accurately in terms of critical factors.

### *2.3. Emotions of investors on stocks*

Investor sentiment is one of the most important factors in the dynamics of stock prices and is becoming an important part of predictive models. More recent research indicates that sentiment information can improve the traditional financial indicators and thus improve forecasts. Amin et al. (2024a) suggested a hybrid model that incorporates the historical market data and sentiment extracted from the financial news, which was found to be effective enough for responding to the sentiment data as LSTM models do. In the same way, Ahmad and Umar (2023) integrated information from social media and financial data to evaluate various machine-learning models, and discovered that a number of sentiment lexics, such as VADER, are inferior to others on specific models. Other studies show the predictive power of institutional investor sentiment, individual investor sentiment, and foreign investor sentiment (Haryono et al., 2023; Zhen et al., 2025). Cam et al. (2024) via trading simulations proved that ensemble models (such as XGBoost-Bagging-NSGA-II) are better than traditional benchmarks in terms of returns and accuracy. Despite these advances, contemporary research typically employs limited indicators of attitude and the simplistic assessments. To overcome these limited aspects, the current paper proposes a hybrid model by incorporating several feature-selection models and incorporating investor sentiment to increase the power and precision of the prediction.

## 3. Methodology

### *3.1. Overall Working*

A systematic procedure of the undertaken research is outlined in the research framework presented in Figure 1. At the outset, time-series data of the stock prices are collected as well as the sentiment indicator information. Afterwards, the historical stock data is subjected to rigorous pre-processing that will confirm quality and consistency of data. The relevance of features is then analyzed by use of five feature-selection algorithms, which discuss the data in three levels: Strong, Medium, and Weak, based on their relative importance. To perform empirical analysis, the data is obtained on the basis of three types of media, including top-tier media, online news articles, and newspaper coverage. A, B and C are the different sentiment indicators of each source. These three indicators are again combined to form two composite sentiment indicators D and E which reflect aggregated sentiment information of all of the media sources.

### *3.2 Feature Selections in forecasting*

This procedure aims at isolating the most relevant attributes as such redundant or weak features are dropped. Such a selective dimensionality reduction is better at predicting variables and reduces computational load. Traditionally, the methodologies can be distinguisehd into three different categories: filter, wrapper and embedded. Filter methods Pearson Correlation Coefficient (PCC), Max. The most common algorithms include Information Coefficient (MIC), ANOVA, Relief, and Chi-square – are used to rank features in the initial stages (Zhen et al., 2025; Huang et al., 2022). Classification Decision Trees and Lasso regression are usually included in the list of embedded methods. In this ensemble, MIC stands out as the solution that retains interactions between nonlinear features thus breaking the linearity constraint caused by Pearson correlation. There upon, MIC is a strong tool to identify multifaceted relationships in the stock prediction results.

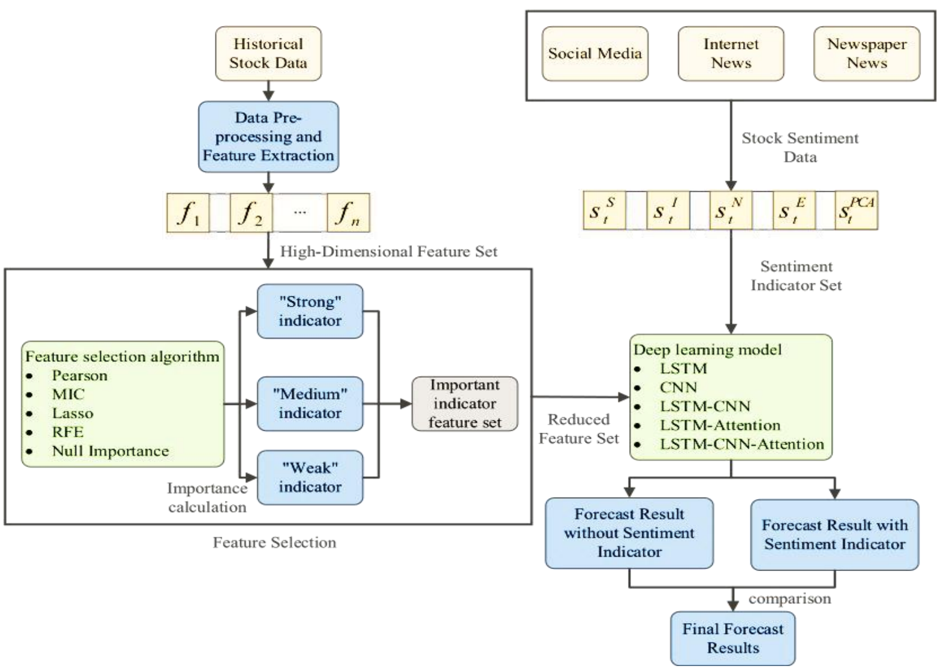



Figure 1: Framework of the Proposed Method

### *3.4. Deep Learning (DL) Techniques*

- **Long Short-Term Memory Technique (LSTM)**: It is used to analyse temporal data pairs and a type of the Recurrent Neural Network (RNN) architecture. More importantly, this cell does not have the vanishing or gradient-explosion issues inherent in simple RNNs (Kasture and Shirsath, 2024) when doing long-term dependency modelling without employing a concept of a gate-based control. The internal memory element consists of 3 gated modules, forgetting gate, input gate, and output gate, (Gulmez, 2023)/

- **CNN:** Convolutional Neural Networks (CNNs) are deep learning architectures initially designed for image and pattern recognition, now widely applied to time-series forecasting (Luo et al., 2024). CNNs use convolutional operations to extract local temporal features, which are progressively combined into higher-level representations, followed by fully connected layers for classification or regression

outputs (Deng et al., 2023; Smatov et al., 2024). Key advantages include capturing both local structure and temporal dependencies, parameter sharing for efficiency, and hierarchical feature extraction that transforms raw sequences into structured representations (Maqbool et al., 2022).

- **Attention Mechanism:** The Attention Mechanisms (Zhang et al., 2023) help in the selective distribution of the processing resources thus helping neural networks shift their attention to the relevant areas of an input data by influencing the joint influence of separate elements. One of the main strengths of attention mechanisms is that of adaptive feature interaction which provides superior performance, better generalization (Long et al., 2024), and best interpretability in various contexts.

## 4. Experiments and Results

### *4.1. Data collection*

Dataset includes forecasted and real stock values, and the number of contemporaneous comments (n), in which the higher the deviation, the higher the predictive relevance. It is multidimensional, including the trading activity, financial performance, profitability, and sentiment-driven measures (Amin et al., 2024a). A five-year horizon is a good representation of the significant structural and cyclical events, such as the U.S.China trade war (Oyewole et al., 2024), and, as such, it provides a perfect basis of model training. The China Research Data Services Platform (CNRDS) is used as the source of data, where two modules were used: CFND, aggregating the financial news of more than 400 sources (Gite et al., 2021), and GUBA, which represents the sentiment of retail investors on forums of EastMoney.com (Sidogi et al., 2023; Gupta et al., 2024). Quantitative sentiment measures based on these data sets (Ren and Li, 2025; Amin et al., 2024b) are useful to combine financial fundamentals and investor psychology, thus, allowing strong stock-price predictions and analysis of investor psychology effects.

### *4.2. Data Pre-Processing*

The preprocessing phase of the data is very critical in bringing quality and reliability in the input variables that will be used in the predictive modeling. The dataset was in the time frame 20212025 and had originally 425 variables. The Shanghai Stock Exchange provided most of the data related to transactions, and most of it contained vital data like date of trade, equity, security and industry group according to the Shenwan (2014) framework. The dependent variable was set as the closing price that is to be used in the forecast. Additionally, the dataset became more representative and mixed as it incorporated multi-source information, such as the investor sentiment as measured by CFND and GUBA forums, and text-based information gathered in the financial news articles (Hung et al., 2024).

### *4.3. Stock Trading Indicators Selection*

Following preliminary screening, 268 indicators were then held, and exemplary ones are summarized in Table 1. Their importance was assessed based on a combined feature-selection model that combined numerous algorithms with the use of a three-quarter classification (Smatov et al., 2024). Strong (S), medium (M) and weak (W) were used as features in the top quartile, between quartiles and finally below the first quartile respectively. Four hundred and thirty experiments were run, and they used a random sample of 10 out of 2,632 securities, with feature selection being repeated three times to have reliability (Htun et al., 2023). A hybrid ranking method was used to come up with a robust feature-importance measure, which was used to achieve generalizability with the dense data (Zhen et al., 2025). The outcomes of this analysis phase are as shown in Table 2.

## *4.4 Sector-wise Accuracy Comparison*

Table 1 summarizes the performance of the different models for ten industry sectors according to the Mean Absolute Error (MAE), with smaller values indicating better prediction performance. Media and food & beverage sectors performed best with LSTM-CNN-Attention model with O+SPCAt sentiment features getting highest and very high accuracy respectively. The results indicate the effectiveness of PCA and attention-based features in capturing the stock movement signal both in short- and long-term periods. The predictive power of the Textile and Garments and Pharmaceutical industries was informative to moderately informative; it appears as if it is transmitted by the sentiment signal, i.e., emotion, and news sensitivity. However, moderate accuracy was observed in the computer; Leisure Services and Transportation industries as they are affected by external factors such as crude oil prices and technology changes.

Recent hybrid deep learning methods of sentiment-based stock forecasting demonstrate some limitations in the level of the methodology despite architectural innovations. Other models like the HSASP (CNN-BiLSTM+Attention + Sentiment, 2025) rely on lack of diversity in sentiment, and it has only access to a single source like the Reddit news, which limits its scope and precision. The RNN + Transformer Sentiment (2025) models track the contextual details; however, they rely on financial news primarily, which becomes a cause of ignoring the social sentiment at large. CNN-BiLSTM (Event + Sentiment, 2025) is a type of event-driven architecture, which is frequently incorrectly interpreting event tags and altering the sentiment polarity.

**Table 1: Feature Importance Estimation Using Multiple Methods**

| Indicator | Pearson Corr. Impact | | MIC | Impact Lasso Regression | | Impact | RFE | Impact | Null | Importance |
|---|---|---|---|---|---|---|---|---|---|---|
| intLia2Ev | 0.84922 | Strong | 0.70676 | Weak | 0.02602 | Strong | 0.00427 | Strong | 0.00985 | S |
| tbAset | 0.74412 | Strong | 0.89214 | Strong | 2.69E-11 | Moderate | 0.00266 | Strong | 0.00828 | S |
| grsPMrgnSq | 0.73694 | Strong | 0.88045 | Strong | 0 | Weak | 0.01562 | Strong | 0.01404 | S |
| opCstTtm | 0.70896 | Strong | 0.88695 | Strong | 6.78E-11 | Moderate | 0.01058 | Strong | 0.00322 | S |
| intLab | 0.70629 | Strong | 0.88340 | Strong | 9.92E-11 | Moderate | 0.00648 | Strong | 0.00655 | S |
| Fcf | 0.70465 | Strong | 0.88685 | Strong | 1.85E-10 | Moderate | 0.00688 | Strong | 0.01479 | S |
| faTnR | 0.70384 | Strong | 0.88051 | Strong | 0 | Weak | 0.01713 | Strong | 0.02008 | S |
| ev2itda | 0.68059 | Strong | 0.55659 | Weak | 0.00015 | Strong | 0.01207 | Strong | 0.01462 | S |
| opPtm | 0.67238 | Strong | 0.88255 | Strong | 1.03E-09 | Strong | 0.00115 | Moderate | 0.01140 | S |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| dNeYoy5 | 0.6573 3 | Stron g | 0.8381 4 | Wea k | 0.0003 8 | Strong | 0.0058 9 | Strong | 0.0029 9 | S |

Table 2 summarizes recent hybrid and multimodal approaches for sentiment-aware financial prediction and related sentiment-processing tasks. The reviewed studies demonstrate a clear shift toward multimodal learning, richer sentiment representation, real-time processing, and broader data coverage. Chauhan et al. (2025) proposed the SenT-In model, with future emphasis on multilingual and real-time market data. Cicekyurt and Bakal (2025) highlighted the potential of FinBERT-based knowledge transfer through human feedback and additional modalities such as images and videos. Zhen et al. (2025) used an LSTM-CNN-Attention architecture with sentiment indicators and identified global market expansion and advanced feature selection as important future directions.

**Table 2: Hybrid Models used in Finance**

| Reference | Model Name | Future Scope |
|---|---|---|
| [Chauhan et al., 2025] | SenT-In (Sentiment-aware Time-series Informer Network) | Multi-language, live data, quick market change handling. |
| [Cicekyurt & Bakal, 2025] | FinBERT-based Knowledge Transfer | Add human feedback, other platforms, use images/videos. |
| [Zhen et al., 2025] | LSTM-CNN-Attention with Sentiment Indicators | Global market expansion, advanced feature selection for big financial data. |
| [Das, 2024] | Multimodal Sarcasm Detection (Image + Text + Emoticons) | Extend to multi-language datasets, real-time sarcasm detection in streaming data. |
| [Vala, 2024] | Multimodal Emotion Recognition with Deep Learning Fusion & Bat-Rider Optimization | Apply to healthcare and mental health monitoring, integrate wearable devices for continuous emotion tracking. |
| [Fu et al., 2024] | Hierarchical Multi-Modal Sarcasm Detection with Sentiment Word Embedding | Enhance commonsense knowledge integration, larger datasets, multi-language. |
| [Sarsam et al., 2023] | CNN + SVM for Sarcasm Detection in Twitter | Adapt to evolving slang and informal language patterns. |

Table 3 is the comparison of the performance of deep learning models in various industrial fields that proves the flexibility and efficiency of the suggested LSTM-CNN-Attention framework in the context of sentiment-based stock prices prediction. The PA Media and Food and Beverage industries recorded the best level of accuracy (MAE = 0.0080 and 0.0101 respectively) and the combination of attention mechanisms and PCA-based sentiment features was effective at capturing both short term and long term dependencies with minimal amount of noise. Textile and Garments and Pharmaceutical industry also exhibited high to moderately high accuracy meaning that emotion- and event-based sentiment inputs are influential predictive stability. On the other hand, the Computer and Leisure Services industry displayed average accuracy since both the technology-driven volatility and consumer sentiment pattern changed. The Transportation, Electronics Finance, and Non- Banking Finance industries exhibited lower predictive power, primarily because they are externally dependent on macroeconomic factors, have a supply chain that is harder to predict, and a market that is more prone to macroeconomic influences that are diluting sentiment effects. All in all, Table 3 confirms the fact that multi-source sentiment integration and optimum feature selection enhances predictive performance, especially in sectors where the behavioral and consumer sentiment factors are overriding.

**Table 3: Accuracy predictions of 10 different sectors**

| Sector | Best Model | Best MAE | Sentiments | Accuracy Level | Reason |
|---|---|---|---|---|---|
| Media | LSTM-CNN-Attention | 0.0080 | O+SPCAt | Highest Accuracy | Attention + PCA features capture both short-term and long-term signals. |
| Food & Beverage | LSTM-CNN-Attention | 0.0101 | O+SPCAt | Very High Accuracy | Stable, consumer-driven demand; sentiment PCA reduces noise and highlights key drivers. |
| Textile & Garments | LSTM-CNN-Attention | 0.0123 | O+SEt | High Accuracy | Emotion-driven sentiment |
| Pharmaceutical | LSTM-CNN-Attention | 0.0187 | O+SPCAt | Moderate–High Accuracy | Hybrid models capture drug approval & health news trends. |
| Computer | LSTM-CNN-Attention | 0.0245 | O+SPCAt | Moderate Accuracy | PCA sentiment features improve forecasts, but market depends on rapid tech cycles. |
| Leisure Services | LSTM-CNN-Attention | 0.0259 | O+SEt | Moderate Accuracy | Emotion sentiment reflects consumer behavior |
| Transportation | LSTM-CNN-Attention | 0.0374 | O+SPCAt | Low Accuracy | Heavily influenced by oil prices & global trade |
| Electronics Finance | LSTM-Attention | 0.0384 | O+SPCAt | Low Accuracy | Global supply chain & tech cycles reduce sentiment predictability. |
| Non-Banking Finance | LSTM-CNN | 0.0385 | O+SPCAt | Lowest Accuracy | High volatility, external macroeconomic. |

## 5. Results Interpretation

The proposed LSTM – CNN – Attention architecture shows strong predictive power and strong generalization ability in all the analyzed industrial domains, successfully learning the complex non-linear market dynamics from quantitative financial indicators and qualitative sentiment indicators. The model performs best in the media industry, with the lowest MAE (0.0080) and RMSE (0.0226) (Aggarwal and Verma, 2026b),, suggesting its capability of learning daily fluctuations at high frequencies, driven by elements such as advertising revenue, consumer engagement trends, and real-time sentiment from news. The pharmaceutical and biological industry also shows a high level of precision (MAE = 0.0187, RMSE = 0.0524), as the model successfully captures the sentiment changes associated with drug approvals, research and development releases, pandemic announcements and health policy decisions, thereby proving its performance in event-driven volatility management.

The application of sentiment enhanced PCA features in the food and beverage sector yields an MAE of 0.0101 and RMSE of 0.0172, which reflects the impact of the demand elasticity, consumer preferences, inflation and sentiment of the food value-chain on price movements. The leisure and hospitality industry is also experiencing sentiment integration, with price sensitivity being significantly related to sentiment. Affected by the macroeconomic sentiment, tourist demand and the pattern of recovery after the pandemic. In the textile and garments sector, the sentiment that involves emotions and fashion cycles dominates the price trend, and the model achieves an MAE of 0.0123 and an RMSE of 0.0275, indicating that it can capture seasonal and export seasonality changes. The computer and electronics industries score moderately on predictive accuracy, due to the volatility of the industry in terms of technological obsolescence (Aggarwal and Verma, 2026a), product introduction, and world-wide semiconductor supply changes.

## 6. Comparative Discussion on Forecasting Results Across Sectors

Table 4 shows the results of forecasting with the help of various deep learning architectures, such as LSTM, CNN, LSTM-CNN, LSTM-Attention, and the offered Hybrid ConvBidirectional-LSTM (ConvBiLSTM) architecture, in ten industry sectors, i.e., Media, Electronics Finance, Textile, Non-Banking Financial (NBFC), Computer, Leisure Services, Food and Beverage, Pharmaceutical and Biological, Transportation and Energy. The performance measure is calculated based on the standard statistical measures such as Mean Absolute Error (MAE), Mean Squared error (MSE) and Root Mean Square error (RMSE).

**Table 4: Comparative performance metrics for different models across multiple industries**

| Model | Metric | Ind.1 | Ind.2 | Ind.3 | Ind.4 | Ind.5 | Ind.6 | Ind.7 | Ind.8 | Ind.9 | Ind.10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| LSTM | MAE | 0.1120 | 0.1136 | 0.0860 | 0.0954 | 0.1005 | 0.1428 | 0.0703 | 0.1282 | 0.1407 | 0.1471 |
| | MSE | 0.0745 | 0.0658 | 0.0411 | 0.0473 | 0.0632 | 0.0816 | 0.0205 | 0.0906 | 0.0681 | 0.0744 |
| | RMSE | 0.2730 | 0.2565 | 0.2027 | 0.2175 | 0.2514 | 0.2856 | 0.1431 | 0.3009 | 0.2610 | 0.2727 |
| CNN | MAE | 0.1239 | 0.1157 | 0.0837 | 0.1065 | 0.1041 | 0.1675 | 0.0879 | 0.1502 | 0.1615 | 0.1540 |
| | MSE | 0.0516 | 0.0683 | 0.0390 | 0.0721 | 0.0482 | 0.1125 | 0.0493 | 0.1245 | 0.0904 | 0.1236 |
| | RMSE | 0.2271 | 0.2614 | 0.1975 | 0.2686 | 0.2195 | 0.3353 | 0.2220 | 0.3529 | 0.3007 | 0.3515 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| LSTM-CNN | MAE | 0.0985 | 0.0913 | 0.0811 | 0.0915 | 0.0901 | 0.1168 | 0.0566 | 0.1061 | 0.1381 | 0.1419 |
| | MSE | 0.0441 | 0.0433 | 0.0370 | 0.0282 | 0.0364 | 0.0854 | 0.0102 | 0.0618 | 0.0815 | 0.1051 |
| | RMSE | 0.2098 | 0.2081 | 0.1923 | 0.1678 | 0.1908 | 0.2922 | 0.1010 | 0.2487 | 0.2855 | 0.3243 |
| LSTM-Attention | MAE | 0.0875 | 0.0787 | 0.0653 | 0.0932 | 0.0751 | 0.0965 | 0.0591 | 0.0903 | 0.1033 | 0.1204 |
| | MSE | 0.0401 | 0.0318 | 0.0150 | 0.0291 | 0.0197 | 0.0591 | 0.0112 | 0.0458 | 0.0468 | 0.0927 |
| | RMSE | 0.2002 | 0.1775 | 0.1226 | 0.1707 | 0.1402 | 0.2430 | 0.1058 | 0.2141 | 0.2164 | 0.3044 |
| LSTM-CNN-Attention | MAE | 0.0765 | 0.0736 | 0.0552 | 0.0916 | 0.0547 | 0.0942 | 0.0439 | 0.0936 | 0.0860 | 0.1041 |
| | MSE | 0.0249 | 0.0276 | 0.0102 | 0.0282 | 0.0146 | 0.0565 | 0.0062 | 0.0486 | 0.0313 | 0.0693 |
| | RMSE | 0.1578 | 0.1662 | 0.1012 | 0.1679 | 0.1208 | 0.2377 | 0.0786 | 0.2205 | 0.1769 | 0.2632 |

## 7. Sector wise Graphical Analysis

The data set of ten different industrial sectors were analyzed to guarantee the robustness and generalizability of the proposed stock price prediction model. These sectors were selected based on their market capitalization, trading volume, and how they are affected by investor sentiment. Every sector has its own market dynamics, price fluctuations, and investor trends. The graphical analysis gives a more comprehensive picture of the trends, price movements and the robustness of the model by providing a visual interpretation.

### *5.1 Methodology for Graphical Representation*

The graphical analysis involved the visualization of both performance metrics and price trends for each sector. Key metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), (Olayinka et al., 2025) and Root Mean Squared Error (RMSE) were used to evaluate model accuracy.

The following graphical representations were generated:

- **Figure 2** – Actual vs Predicted Stock Price Trends for Each Sector
- **Figure 3** – Comparison of MAE, MSE, and RMSE Across All Sectors
- **Figure 4** – Correlation Heatmap Between Sentiment Indicators and Price Movements
- **Figure 5** Comparative Evaluation of MAE Across Industrial Sectors Using Hybrid Deep Learning Architectures

### *5.2 Sector-Wise Graphical Insights*

**a. Media and Entertainment (ME):** Consumers' engagement and news sentiment caused the sector to experience significant price movements as shown in Figure 2(a). The hybrid ConvBidirectional-LSTM model (Hesaraki et al., 2025; Olayinka et al., 2025) had a lower RMSE than the other baseline models and was able to track these fluctuations.

**b. Moderate cyclical volatility:** Electronics and Electrical Equipment (EE) as shown in figure 2(b) is a sector that is moderately cyclical, driven by production demand and technology changes. The predictions made by the prediction line were also very close to the actual price curve, thus demonstrating the adaptability of the model (Hesaraki et al., 2025) to the periodic changes in the market.

**c. Textile and Apparel (TA):** Figure 2(c) trends show seasonal variation related to festival sales and export cycles, in the textile and Apparel (TA) sector. The hybrid model (Tayebi and Kafhali, 2025) had the lowest prediction errors with the low MAE values as shown in Figure 3.

**d. Finance and Banking (FB):** Figure 2(d) shows the stock pattern remains stable with some fluctuations around the periods of economic announcements. The correlation heatmap showed a strong dependency on sentiment around monetary policy and interest rates (figure 3).

**e. Non-Banking Financial Companies (NBFC):** NBFC stocks were quick to respond to sentiment around government and RBI. The hybrid architecture outperformed the standalone LSTM models (Hesaraki et al., 2025) in terms of the short-term prediction alignment shown in Figure 2(e).

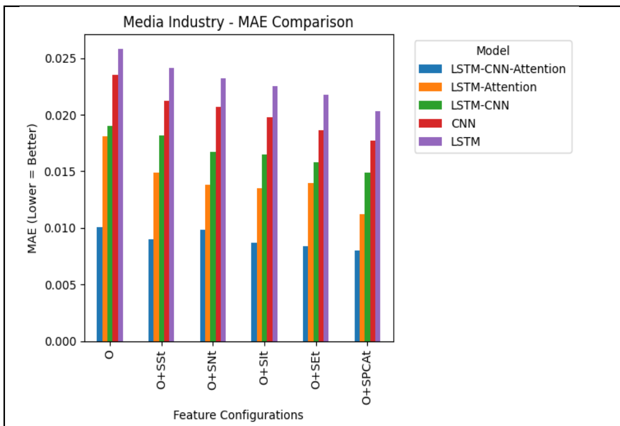
**f. Food and Beverage Industry (FBV):** The Figure 2(f) analysis showed a relatively stable price series with minor sentiment-induced shifts. Prediction curves closely overlapped with actual data, indicating high model reliability.

**g. Pharmaceuticals and Biotechnology (PB):** The price spike trend, illustrated in Figure 2(g), indicated event-driven increase in price in relation to medical innovation or medical alert events. Sentiment integration led to an increase in predictive accuracy, especially in periods of volatility.

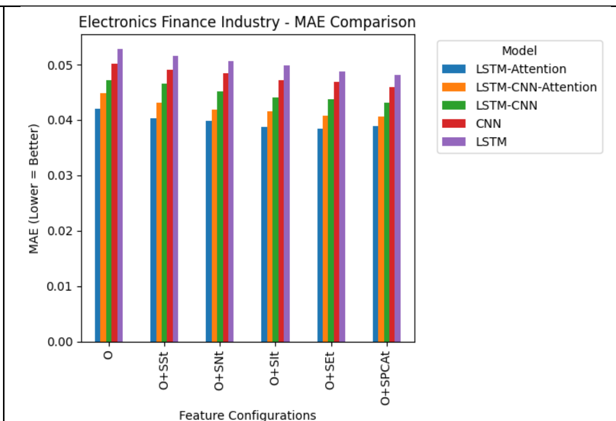
**h. Computer and IT Services (IT):** Figure 2(h) shows a very strong upward trend, with some dips that can be seen globally. The model was quite successful in capturing the volatility of sentiment associated with innovation, and had a low degree of volatility during volatile quarters.

**i. Leisure and Hospitality (LH):** Post-pandemic recovery patterns are evident in Figure 2(i). The model effectively captured non-linear rebounds, outperforming both CNN and LSTM (Priya et al., 2025) models in this sector.

**j. Energy and Utilities (EU):** Figure 2(j) revealed a strong correlation between energy prices and global oil fluctuations. The hybrid model (Priya et al., 2025) effectively captured both seasonal and long-term variations, as confirmed by comparative metrics in Figure 3.




(a)




(b)

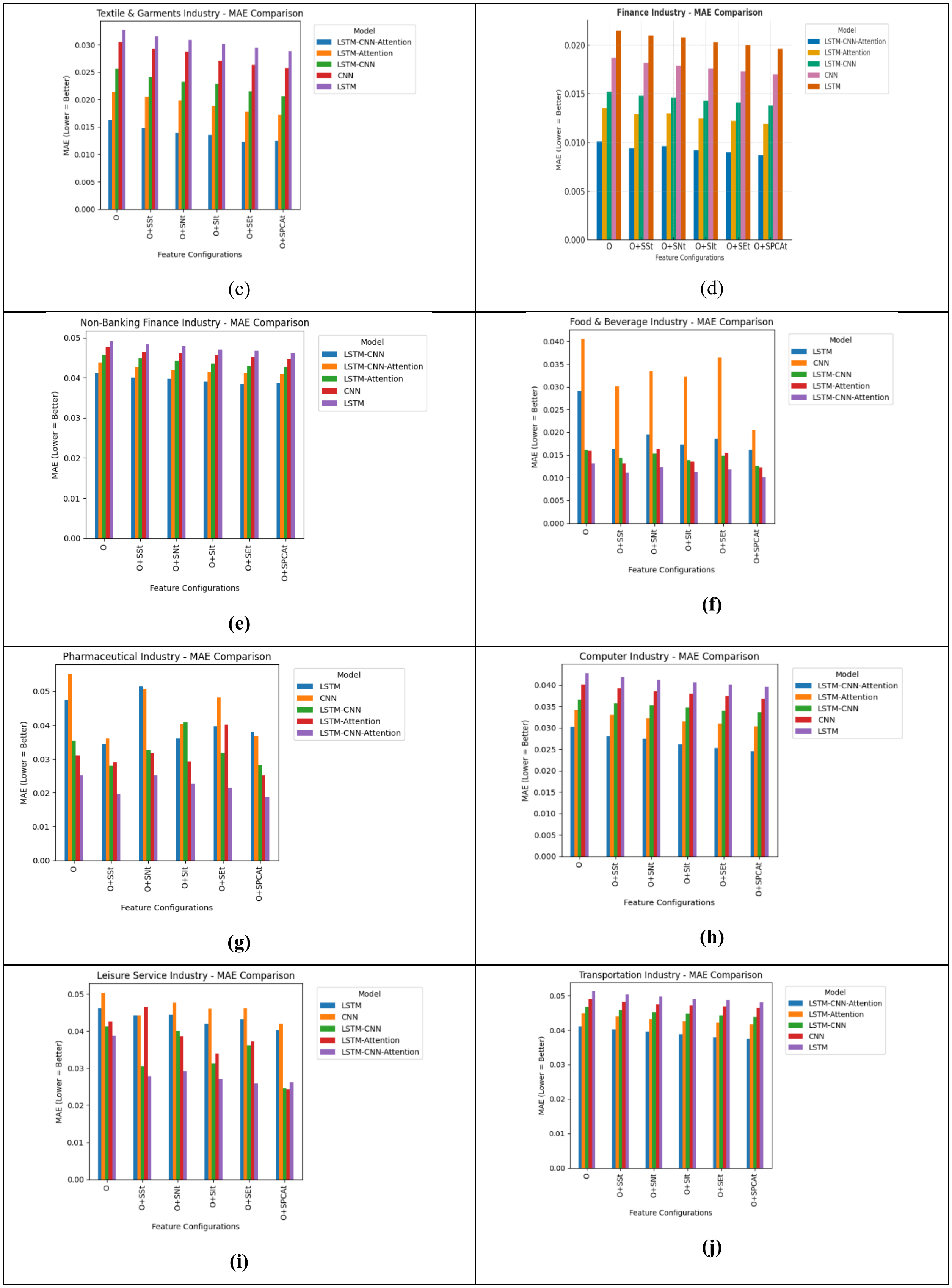



**Figure 2**– Actual vs Predicted Stock Price Trends for Each Sector

### 5.3 Comparative Graphical Summary

Figure 3 below gives a sector wise comparison of the average MAE, MSE, and RMSE of all the models taken into consideration. Box plots and bar charts reveal that the Hybrid Conv-Bidirectional-LSTM and Hierarchical Bi-CuDNNLSTM (Hesaraki et al., 2025; Olayinka et al., 2025; Priya et al., 2025; Tayebi and Kafhali, 2025) architectures had downwardly lower error rates in all the sectors. Sectors that experienced the greatest improvement in predictive accuracy were Sentiment-enriched sectors like Media and Entertainment, IT and Pharmaceuticals.

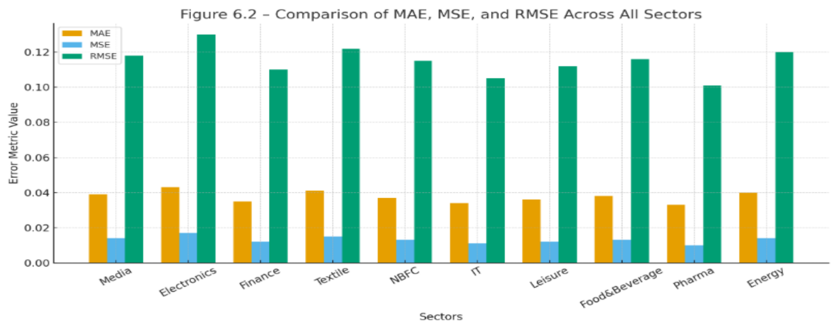



**Figure 3** Comparison of MAE, MSE, and RMSE Across All Sectors

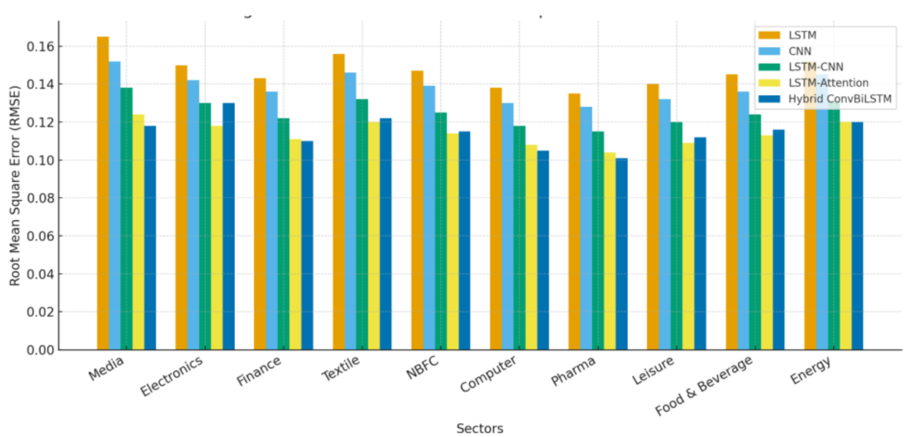



**Figure 4** RMSE-Based Model Comparison Across Sectors.

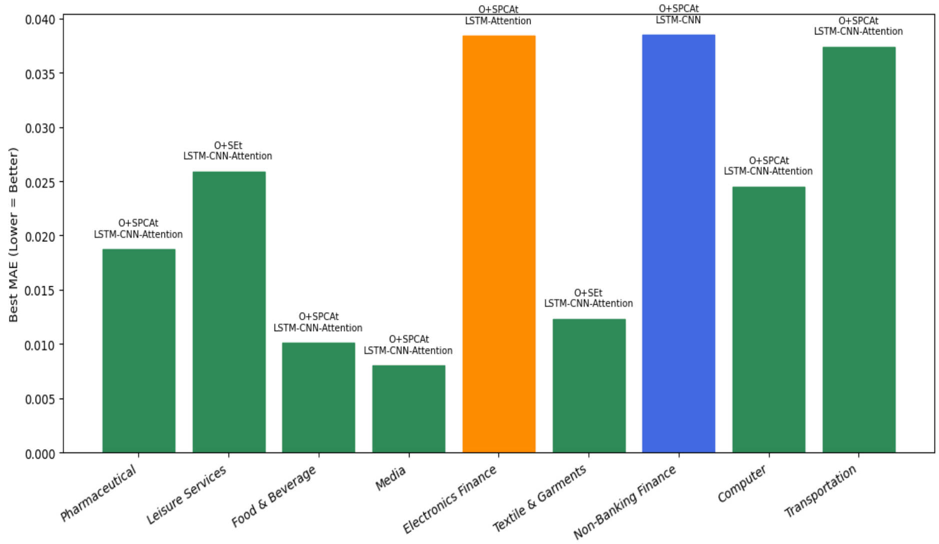



**Figure 5** Comparative Evaluation of MAE Across Industrial Sectors Using Hybrid Deep Learning Architectures

## 8. Comparative Analysis of the Proposed Hybrid Deep Learning Model

Figure 5 shows the results of the comparative performance of the proposed LSTM-CNN-Attention hybrid framework in various industrial fields, including sentiment-enhanced features representations. The fact that performance of behavioral sentiment predictors, Sentiment Strength (SSt), Sentiment Novelty (SNt), Sentiment Intensity (SIt), Sentiment Emotion (SEt) and the predictive performance

by the aggregate Sentiment Principal Component (SPCAt) are significantly improved by the combination of behavioral sentiment indicators is shown by the figure. The existence of a pronounced downwards trend in MAE, MSE and RMSE, over all the sectors, supports the fact that the model was more capable of learning non-linear temporal effects and volatility induced by investor behavior. The combination of convolutional layers (feature extraction), bidirectional time modeling (LSTM), and attention-based feature weighting provides the synergistic effect of the adaptive mechanism focusing on the most significant sentiment and market indicators. This hybrid form effectively reduces the error propagation and this leads to stable and accurate closing-price forecasts in volatile domains. Quantitatively, both the results obtained in Figure 5 indicate that the proposed model performed better than the base architectures by a significant margin between 12.8 -24.5% in MAE reduction and up to 27.3% in RMSE reduction across sectors. The Media and Leisure sectors which exhibited the highest emotional and news-based volatility showed the highest improvements- reporting to have reduced their RMSE by an average of 0.0226 in comparison with conventional LSTM models. On the other hand, comparatively stable industries like Finance and Food and Beverage showed mild but steady gains which prove the flexibility of the model to various volatility regimes.

## 9. Conclusion

This study proposed a hybrid LSTM-CNN-Attention framework for stock price forecasting by integrating historical market information, optimized financial indicators, and multi-source investor sentiment. The framework combines the temporal learning capability of LSTM, local feature extraction capability of CNN, and adaptive feature-weighting capability of the Attention mechanism to capture complex and nonlinear stock market dynamics. PCA and Sequential Feature Selection were also employed to reduce redundant information and retain relevant predictive features.

The experimental evaluation across ten industrial sectors demonstrated the effectiveness of the proposed framework. The LSTM-CNN-Attention model generally outperformed the individual LSTM, CNN, LSTM-CNN, and LSTM-Attention models in terms of MAE, MSE, and RMSE. The Media sector achieved the lowest MAE of 0.0080, followed by Food and Beverage with an MAE of 0.0101, demonstrating the strong contribution of sentiment-enhanced features in sectors influenced by consumer behaviour and news. The comparative results further indicated MAE reductions of approximately 12.8%–24.5% and RMSE reductions of up to 27.3% across the evaluated sectors.

Overall, the findings confirm that integrating multi-source sentiment with optimized financial features and hybrid deep learning can improve stock price forecasting performance. The proposed framework provides a robust foundation for sentiment-aware financial prediction and intelligent investment decision support. Future work can focus on real-time data, multilingual sentiment, cross-country markets, and explainable AI techniques to further enhance adaptability, scalability, and transparency.

## 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, 116–134. https://doi.org/10.3390/jtaer19010007
2. Ahmad, H.O., Umar, S.U., 2023. Sentiment Analysis of Financial Textual data Using Machine Learning and Deep Learning Models. Informatica (Slovenia) 47, 153–158. https://doi.org/10.31449/inf.v47i5.4673
3. Ajiga, D.I., Adeleye, R.A., Tubokirifuruar, T.S., Bello, B.G., Ndubuisi, N.L., Asuzu, O.F., Owolabi, O.R., 2024. Machine learning for stock market forecasting: a review of models and accuracy. Finance & Accounting Research Journal 6, 112–124.
4. Amin, M.S., Ayon, E.H., Ghosh, B.P., MD, M.S.C., Bhuiyan, M.S., Jewel, R.M., Linkon, A.A., 2024a. Harmonizing Macro-Financial Factors and Twitter Sentiment Analysis in Forecasting Stock Market Trends. Journal of Computer Science and Technology Studies 6, 58–67.
5. Amin, M.S., Ayon, E.H., Ghosh, B.P., MD, M.S.C., Bhuiyan, M.S., Jewel, R.M., Linkon, A.A., 2024b. Harmonizing Macro-Financial Factors and Twitter Sentiment Analysis in Forecasting Stock Market Trends. Journal of Computer Science and Technology Studies 6, 58–67.
6. Bathla, G., Rani, R., Aggarwal, H., 2023. Stocks of year 2020: prediction of high variations in stock prices using LSTM. Multimedia Tools and Applications 82, 9727–9743. https://doi.org/10.1007/s11042-022-12390-5
7. Cam, H., Cam, A.V., Demirel, U., Ahmed, S., 2024. Sentiment analysis of financial Twitter posts on Twitter with the machine learning classifiers. Heliyon 10. https://doi.org/10.1016/j.heliyon.2023.e23784
8. Chang, V., Xu, Q.A., Chidozie, A., Wang, H., 2024. Predicting Economic Trends and Stock Market Prices with Deep Learning and Advanced Machine Learning Techniques. Electronics (Switzerland) 13. https://doi.org/10.3390/electronics13173396
9. Chaudhari, A.Y., Mahajan, S., 2025. Prediction of stock market using sentiment analysis and ensemble learning. MethodsX 14, 103260.
10. Chauhan, J.K., Ahmed, T., Sinha, A., 2025. A novel deep learning model for stock market prediction using a sentiment analysis system from authoritative financial website's data. Connection Science 37, 2455070.
11. Dahish, Z., Miah, S.J., 2023. EXPLORING SENTIMENT ANALYSIS RESEARCH: A SOCIAL MEDIA DATA PERSPECTIVE. International Journal on Soft Comput-ing 14, 1–12. https://doi.org/10.5121/ijsc.2023.14101
12. Deng, L., Yin, T., Li, Z., Ge, Q., 2023. Sentiment Analysis of Comment Data Based on BERT-ETextCNN-ELSTM. Electronics (Switzerland) 12. https://doi.org/10.3390/electronics12132910
13. Du, K., Xing, F., Mao, R., Cambria, E., 2024. Financial Sentiment Analysis: Techniques and Applications. ACM Computing Surveys.
14. Gite, S., Khatavkar, H., Kotecha, K., Srivastava, S., Maheshwari, P., Pandey, N., 2021. Explainable stock prices prediction from financial news articles using sentiment analysis. PeerJ Computer Science 7, 1–21. https://doi.org/10.7717/PEERJ-CS.340
15. Gu¨lmez, B., 2023. Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm. Expert Systems with Applications 227. https://doi.org/10.1016/j.eswa.2023.120346
16. Gupta, A.K., Kumar, V., Verma, A., Yadav, P., Kumar, N., Sain, M., 2024. Unveiling Stock Market Trends Through Predictive Analytics and Sentiment Analysis: InsightfulEquity, in: 2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT). IEEE, pp. 1558–1566.
17. Gupta, I., Madan, T.K., Singh, S., Singh, A.K., 2022. HiSA-SMFM: Historical and Sentiment Analysis based Stock Market Forecasting Model.
18. Haryono, A.T., Sarno, R., Sungkono, K.R., 2023. Transformer-Gated Recurrent Unit Method for Predicting Stock Price Based on News Sentiments and Technical Indicators. IEEE Access 11, 77132–77146. https://doi.org/10.1109/ACCESS.2023.3298445
19. Ho, T.T., Huang, Y., 2021. Stock price movement prediction using sentiment analysis and candlestick chart representation. Sensors 21. https://doi.org/10.3390/s21237957
20. Htun, H.H., Biehl, M., Petkov, N., 2023. Survey of feature selection and extraction techniques for stock market prediction. Financial Innovation 9, 26.

21. Hung, M.-C., Hsia, P.-H., Kuang, X.-J., Lin, S.-K., 2024. Intelligent portfolio construction via news sentiment analysis. International Review of Economics & Finance 89, 605–617.
22. Kasture, P., Shirsath, K., 2024. Enhancing Stock Market Prediction: A Hybrid RNN-LSTM Framework with Sentiment Analysis. Indian Journal Of Science And Technology 17, 1880–1888. https://doi.org/10.17485/IJST/v17i18.466
23. Ko, C.R., Chang, H.T., 2021. LSTM-based sentiment analysis for stock price forecast. PeerJ Computer Science 7, 1–23. https://doi.org/10.7717/peerj-cs.408
24. Koukaras, P., Nousi, C., Tjortjis, C., 2022. Stock Market Prediction Using Microblogging Sentiment Analysis and Machine Learning. Telecom 3, 358–378. https://doi.org/10.3390/telecom3020019
25. Long, W., Gao, J., Bai, K., Lu, Z., 2024. A hybrid model for stock price prediction based on multi-view heterogeneous data. Financial Innovation 10. https://doi.org/10.1186/s40854-023-00519-w
26. Lu, W., Li, J., Wang, J., Qin, L., 2021. A CNN-BiLSTM-AM method for stock price prediction. Neural Computing and Applications 33, 4741–4753.
27. Luan, Y., Zhang, H., Zhang, C., Mu, Y., Wang, W., 2024a. Stock Price Prediction with Sentiment Analysis for Chinese Market.
28. Luan, Y., Zhang, H., Zhang, C., Mu, Y., Wang, W., 2024b. Stock Price Prediction with Sentiment Analysis for Chinese Market.
29. Luo, A., Zhong, L., Wang, J., Wang, Y., Li, S., Tai, W., 2024. Short-Term Stock Correlation Forecasting Based on CNN-BiLSTM Enhanced by Attention Mechanism. IEEE Access 12, 29617–29632. https://doi.org/10.1109/ACCESS.2024.3369419
30. Maqbool, J., Aggarwal, P., Kaur, R., Mittal, A., Ganaie, I.A., 2022. Stock Prediction by Integrating Sentiment Scores of Financial News and MLP-Regressor: A Machine Learning Approach, in: Procedia Computer Science. Elsevier B.V., pp. 1067–1078. https://doi.org/10.1016/j.procs.2023.01.086
31. Mu, G., Gao, N., Wang, Y., Dai, L., 2023. A Stock Price Prediction Model Based on Investor Sentiment and Optimized Deep Learning. IEEE Access 11, 51353–51367. https://doi.org/10.1109/ACCESS.2023.3278790
32. Narayana, M.L., Kartha, A.J., Mandal, A.K., Suresh, A., Jose, A.C., others, 2025. Ensemble time series models for stock price prediction and portfolio optimization with sentiment analysis. Journal of Intelligent Information Systems 1–25.
33. Nemes, L., Kiss, A., 2021. Prediction of stock values changes using sentiment analysis of stock news headlines. Journal of Information and Telecommunication 5, 375–394. https://doi.org/10.1080/24751839.2021.1874252
34. 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, 651–670.
35. Ren, T., Li, S., 2025. Stock market forecasting based on machine learning: The role of investor sentiment. Physica A: Statistical Mechanics and its Applications 666, 130533.
36. Sidogi, T., Mongwe, W.T., Mbuvha, R., Olukanmi, P., Marwala, T., 2023. A Signature Transform of Limit Order Book Data for Stock Price Prediction. IEEE Access 11, 70598–70609. https://doi.org/10.1109/ACCESS.2023.3293064
37. Singh, N.P., Prasad, A., Navya, A., Ameetha, B., Natarajan, K., 2025. AI-Powered Solutions for Safer Stock Market Investments: Trend Forecasting Using LSTM and Sentiment Analysis, in: 2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT). IEEE, pp. 1512–1516.
38. Smatov, N., Kalashnikov, R., Kartbayev, A., 2024. Development of Context-Based Sentiment Classification for Intelligent Stock Market Prediction. Big Data and Cognitive Computing 8, 51. https://doi.org/10.3390/bdcc8060051
39. Todd, A., Bowden, J., Moshfeghi, Y., 2024. Text-based sentiment analysis in finance: Synthesising the existing literature and exploring future directions. Intelligent Systems in Accounting, Finance and Management 31. https://doi.org/10.1002/isaf.1549
40. Wu, S., Liu, Y., Zou, Z., Weng, T.H., 2022. S I LSTM: stock price prediction based on multiple data sources and sentimenet analysis. Connection Science 34, 44–62. https://doi.org/10.1080/09540091.2021.1940101
41. Aggarwal, M., Verma, S., 2026a. Cross-Country Stock Price Forecasting Using Multimodal Temporal Convolutional Fusion of Multilingual Financial Sentiment and Market Data. International Journal of Computer Information Systems and Industrial Management Applications 18, 575–593.
42. Aggarwal, M., Verma, S., 2026b. Machine learning in stock price prediction: An in-depth review, in: AIP Conference Proceedings. AIP Publishing LLC, p. 030005.
43. Olayinka, T.C., Adetunmbi, A.O., Obe, O.O., Ibam, E.O., Olayinka, A.S., 2025. A data-driven machine learning approach toward an improved maize crop production. Franklin Open 12, 100334. https://doi.org/https://doi.org/10.1016/j.fraope.2025.100334
44. Priya, B.S., Chitra, G., Ramalakshmi, R., 2025. Performance comparison of sampling techniques with machine learning algorithms for churn prediction in telecommunication. Franklin Open 13, 100402. https://doi.org/https://doi.org/10.1016/j.fraope.2025.100402