Enhancing Prediction Accuracy in Financial Markets through Optimization of Deep Learning Algorithms: A Hybrid Learning and Natural Language Processing-Based Framework

Abstract

By creating a hybrid deep learning framework that combines cutting-edge optimization methods with natural language processing (NLP) to improve forecasting accuracy, this paper fills a significant research vacuum in financial market prediction. The goal of the project is to get beyond the drawbacks of conventional deep neural network models, which usually concentrate on time series analysis without making good use of qualitative textual input. Our system captures both numerical trends and market sentiment by fusing sentiment analysis from financial news, social media, and analyst reports with quantitative market indicators. This is crucial for accurate forecasts in unstable circumstances. We process both unstructured data from different financial sources and structured data, such as OHLCV measures, technical indicators (including the Simple Moving Average, Exponential Moving Average, MACD, and Relative Strength Index), and macroeconomic variables. The hybrid model architecture uses a transformer-based natural language processing (NLP) module that uses FinBERT/Persian BERT models to quantify sentiment and thematic content, and a stacked Long Short-Term Memory (LSTM) network to extract temporal information. A specialized feature integration layer is employed to fuse these outputs, and dense layers with dropout regularization and ReLU activation come next. Bayesian Optimization (Optuna) and Particle Swarm Optimization (PSO) are used to optimize hyperparameters, guaranteeing strong model performance. With better directional accuracy and AUC-ROC in classification tasks, as well as reduced error metrics (RMSE, MAE, and MAPE) and stronger explanatory power (R2), empirical evaluation utilizing data from the Iran Stock Exchange (2015–2024) shows notable gains over baseline approaches. The stability of the model and the crucial roles played by the LSTM and NLP modules are validated by cross-validation, sensitivity analysis, and ablation investigations. All things considered, this study offers a scalable and creative framework for financial forecasting that successfully combines deep learning with natural language processing, providing improved prediction accuracy and insightful information for risk management and investment decision-making across a range of economic sectors. These results serve as a foundation for further investigation and real-world financial market applications.

Keywords: Financial Forecasting, Deep Learning, Natural Language Processing (NLP), Hybrid Model, Particle Swarm Optimization, Feature Fusion, Stock Market Prediction

1. Introduction

One of the most important problems in quantitative finance is still improving the accuracy of financial market predictions, especially as global markets becoming more intricate and data-rich. Artificial Intelligence (AI) methods, particularly deep learning algorithms, have demonstrated encouraging outcomes in predicting asset prices and financial trends over the last ten years (LeCun, Bengio, & Hinton, 2015; Goodfellow, Bengio, & Courville, 2016). However, there are few empirical studies that offer solid, evidence-based methods for optimizing these algorithms, despite increased interest in AI's ability to transform financial prediction. This disparity is especially noticeable when taking into account the incorporation of natural language processing (NLP) methods, which are crucial for deriving significant insights from large volumes of unstructured data, including financial reports, news articles, and social media (Mikolov et al., 2013; Martin & Jurafsky, 2009; Sheth & Shah, 2023).

Prior studies have shown that temporal dependencies in financial time series data can be captured by deep learning models, such as long short-term memory networks (LSTMs) and recurrent neural networks (RNNs) (Fischer & Krauss, 2018; Heaton et al, 2016). However, these models frequently have problems including overfitting, hyperparameter sensitivity, and the difficulty of combining several data sources. Furthermore, the intricacy of financial markets necessitates a more comprehensive strategy due to its intrinsic volatility and the impact of outside variables. By using both quantitative and qualitative information, hybrid learning frameworks that integrate deep learning and natural language processing (NLP) techniques have the potential to improve prediction accuracy (Ding et al., 2015; Sharma & Mehta., 2024). In order to fill these research gaps, this paper suggests a hybrid architecture that integrates natural language processing (NLP) with deep learning algorithms to enhance their performance in financial market prediction. The system uses textual analysis to gather sentiment and changes in market sentiment while also utilizing sophisticated optimization techniques to fine-tune deep learning models. The report is supported by an empirical examination of data gathered from multiple financial market sources as well as a thorough evaluation of current methodology. By doing this, it closes the gap between the real requirements for market prediction and the optimization of technological algorithms. Few research have examined the optimization component in conjunction with natural language processing, despite the fact that several have investigated the use of deep learning in financial prediction (Sismanoglu et al., 2019; Ozbayoglu et al., 2020). The suggested framework extends hybrid models by using advanced optimization approaches that dynamically modify model parameters, building upon their base. By reducing overfitting and mitigating the consequences of market volatility, these strategies aim to improve prediction models' generalization performance (Agrawal et al., 2022; Patel et al., 2015). Furthermore, models can now decipher sentiment from textual data thanks to recent developments in natural language processing (NLP), adding layers of information that can be crucial for forecasting market movements (Blei, Ng, & Jordan, 2003). For instance, as market sentiment frequently comes before price fluctuations, it has been demonstrated that combining sentiment analysis with time series forecasting greatly increases prediction accuracy (Choi & Varian, 2012). With the help of these insights, our hybrid framework takes into consideration qualitative aspects that traditional models tend to overlook while simultaneously enhancing the quantitative predictions produced by deep learning algorithms. Additionally, a rigorous experimental methodology is used in the study to validate the suggested framework. Using a large dataset that contains both textual information from financial news and numerical market data, we evaluate our hybrid model's prediction performance against that of conventional deep learning models. Model accuracy is evaluated using statistical measures like mean absolute error (MAE) and root mean square error (RMSE), and the model's resilience to changes in expert opinion and parameter settings is revealed by sensitivity analyses and confidence intervals (Lundberg & Lee, 2017; Heaton et al., 2016).

This study offers two contributions. In the first place, it offers a financial market prediction system that combines the best features of NLP and deep learning. Second, it makes practical suggestions for further study and real-world implementations, arguing that ongoing investments in improved data integration and algorithm optimization can lead to even greater advancements in market forecasting. These findings have ramifications that go beyond the financial sector, providing a methodological guide for sectors where decision-making requires the integration of quantitative and qualitative data (Atsalakis & Valavanis, 2009; Zhang, 2003). This study opens the door to more precise, reliable, and thorough financial market prediction models by bridging the gap between algorithmic optimization and natural language understanding. In the end, the hybrid framework described here is positioned to improve the way that investors, financial institutions, and policymakers make decisions, which will help them gain a better understanding of market dynamics in a world that is becoming more and more data-driven (Hossain et al., 2022; Huang, Nakamori, & Wang, 2005). In conclusion, by offering a unique, hybrid strategy that not only optimizes deep learning algorithms but also enhances them with qualitative insights obtained via NLP, our study significantly adds to the body of literature. It lays the groundwork for future research into hybrid learning approaches and emphasizes the value of a diversified approach in tackling the intrinsic complexity of financial markets. In order to maintain the accuracy and resilience of financial market forecasts in the face of constantly changing obstacles, future research should keep improving existing techniques and investigating new facets of data integration and model optimization (Patel et al., 2015; Ishwarappa & Anuradha, 2021).

2. Literature review

The way analysts and investors approach data-driven decision-making has been completely transformed by the use of deep learning into financial market prediction. The complexity of financial data, including its high dimensionality, non-linearity, and volatility, frequently presents challenges for traditional statistical methods. According to Rajendran et al. (2024) and Zheng et al. (2024), deep learning models, such as Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs), have shown exceptional ability in forecasting stock prices, exchange rates, and other financial indicators. Compared to traditional methods, these models provide more accurate and dependable forecasts by utilizing sophisticated architectures to identify complex patterns in financial time series data (Dokumacı, 2024; Idowu, 2024). Large volumes of time series data, which are dynamic and non-linear in nature, are produced by financial markets. Because deep learning models, especially LSTM networks, can capture long-term patterns and temporal dependencies, they are ideal for assessing this type of data (Leng, 2024; Zhang et al., 2024a). For example, in stock price prediction tasks, LSTM models have been demonstrated to perform better than conventional techniques such as Autoregressive Integrated Moving Average (ARIMA) and Moving Average Convergence Divergence (MACD) (Rajendran et al., 2024; Sonkavde et al., 2023).

More complex time series analysis models have been made available by recent developments in deep learning. Originally created for natural language processing, transformers have been modified for financial forecasting and have shown remarkable success in identifying intricate patterns in financial data (Patel et al., 2023; Li & Bastos, 2020). Furthermore, it has been demonstrated that hybrid models that combine the advantages of CNNs and LSTMs can effectively handle temporal and spatial dependencies in financial time series (Farimani et al., 2022; Takale, 2024). New approaches to financial market prediction have been made possible by the combination of deep learning and natural language processing (NLP). According to Heaton et al. (2016), sentiment analysis of news stories, social media posts, and other textual data might yield important information on investor behavior and market movements (Ohliati & Yuniarty, 2024). Large amounts of textual data have been analyzed using deep learning models, such as Transformer-based architectures, which extract sentiments and correlate them with market movements (Li, 2024; Krishna et al., 2023). The development of sophisticated text representation algorithms has further improved the application of NLP in financial forecasting. By using these techniques, significant features can be extracted from unstructured data and combined with conventional financial indicators to increase prediction accuracy (Zhao & Huang, 2024; Ti, 2024). Furthermore, it has been demonstrated that combining sentiment analysis with technical indicators improves the resilience of financial forecasting models (Mienye et al., 2024; Zhang et al., 2024b).

Numerous deep learning architectures, each with specific advantages and disadvantages, have been created and used for financial market prediction. A comparison of some of the most popular models may be found below:

Model Type	Description	Strengths
LSTM Networks	Designed to capture temporal dependencies in sequential data.	Effective for time series forecasting; handles long-term dependencies well.
Convolutional Neural Networks (CNNs)	Useful for processing spatial hierarchies in data.	Captures spatial patterns; effective for image and signal processing.
Transformer Models	Leverage self-attention mechanisms for sequence modeling.	Excellent for capturing global dependencies; suitable for parallel processing.
Hybrid Models	Combine different architectures, such as CNN-LSTM or Transformer-LSTM.	Captures both spatial and temporal dependencies; robust for complex data.

Forecasting stock prices, exchange rates, and commodity prices are just a few of the financial prediction tasks to which these models have been applied. The particular features of the data and the type of prediction task frequently influence the model selection (Teixeira & Barbosa, 2024; Sahani, 2024). Additionally, deep learning has been crucial for improving market predicting skills and trading strategy optimization. Deep learning algorithms can find patterns and trends that guide trading decisions by examining past market data. For instance, adaptive trading techniques that react dynamically to market conditions have been developed using reinforcement learning (RL) (De Avila & Salgado, 2023). The accuracy of market projections has increased much more with the combination of technical analysis and deep learning. To produce more dependable trading signals, deep learning models can be integrated with technical indicators like relative strength index (RSI) and moving averages (Sharma & Gupta, 2022). Furthermore, it has been investigated to simulate market conditions and create synthetic data for prediction model training using generative adversarial networks (GANs) (Karthik et al, 2023).

Deep learning for financial market prediction has made great strides, but there are still a number of obstacles to overcome. The interpretability of deep learning models is one of the main issues. For high-stakes financial decisions, it might be challenging to comprehend the logic underlying these models' forecasts due to their "black box" character (Kumar et al, 2021; Jethani et al, 2023). Overfitting is another problem, especially with models that were trained on erratic and noisy financial data. Although regularization strategies like dropout and early stopping have been used to lessen this issue, more study is required to create more reliable models (Raut et al, 2024; Olorunnimbe & Viktor, 2023). Another important component of deep learning models' success is the quality of the data. Missing values and noise are common in financial data, which can negatively impact model accuracy. High-quality training data preparation requires sophisticated data pretreatment methods as normalization and outlier detection (Hassan et al, 2022; Li et al, 2022).

Beyond only predicting the market, deep learning has several uses in financial management. These include of algorithmic trading, fraud detection, portfolio management, and credit risk analysis. For example, by examining intricate patterns in credit data, deep learning models have been utilized to evaluate creditworthiness (Jain et al, 2022; Sable et al, 2022). Deep learning can improve risk management and asset allocation in portfolio management. These models can determine the best portfolio configurations that optimize returns while limiting risk by examining past returns and volatility (Yekrangi & Abdolvand, 2021; Zhang et al, 2021). Deep learning has also been used to detect fraud, detecting unusual transactions and highlighting possible fraudulent activity (Le et al, 2020; Guo & Tuckfield, 2020). Deep learning's prospects for financial market prediction are bright, with a number of possible avenues for future study. Integrating data from other sources, such satellite images and sensor data, to increase the precision of financial projections is one area of focus. Explainable AI (XAI) model development is another exciting avenue. Since financial choices frequently call for accountability and transparency, the creation of interpretable deep learning models is essential to their uptake in the industry. In conclusion, a new field of study is the use of quantum machine learning for financial forecasting. Complex optimization problems may be solved more quickly by quantum algorithms than by classical ones, which could result in advances in financial forecasting and modeling.

2.1. Research Gap

There is still a large research gap in the optimization of deep learning models, especially when paired with natural language processing (NLP) approaches, despite the expanding corpus of work on the use of these methods in financial market prediction. The deployment of deep neural networks, including as LSTMs and RNNs, to identify temporal trends in market data is the subject of many current studies; nevertheless, algorithmic optimization plays a crucial role in preventing overfitting and improving model robustness. Furthermore, although some scholars have used sentiment analysis into financial forecasting, there is a dearth of comprehensive, hybrid frameworks in the literature that concurrently use qualitative textual data and quantitative market indicators to produce more accurate forecasts. This disparity is especially noticeable in research focusing on the erratic character of financial markets, where a sophisticated comprehension of both numerical patterns and sentiment-driven fluctuations is crucial. Moreover, previous studies have typically viewed NLP and deep learning as separate elements rather than as parts of a single decision-making model. Therefore, empirical studies that employ sophisticated optimization methods, such those included into a hybrid framework, are urgently needed to successfully combine these strategies. In addition to having the ability to greatly improve financial forecasting's operational resilience and prediction accuracy, closing this gap could also result in a scalable model that can be used in a variety of international economic sectors.

3. Methods

3.1. Research Design and Data Sources

In order to improve financial market prediction accuracy, this study uses a hybrid machine learning architecture that combines natural language processing (NLP) and deep learning optimization approaches. Using both structured financial data and unstructured textual data, we concentrate on the Iran Stock Exchange (ISE) between 2015 and 2024.

Structured Data

The structured financial dataset includes:

OHLCV Data (Open, High, Low, Close, Volume)

Market Cap, EPS, P/E Ratios

Daily and Weekly Returns

Macroeconomic Variables (inflation, currency rates, interest rates)

Sectoral Indices (e.g., Automotive, Petrochemical, Banking)

Unstructured Data (for NLP)

Financial news from Persian outlets

Analysts’ reports

Social media data (e.g., Twitter, Telegram posts)

Economic reports and statements

3.2. Feature Engineering: Financial Indicators & Formulas

We calculate technical indicators and market sentiment features, categorized as:

Trend Indicators

Simple Moving Average (SMA):

$$SMA_t = \frac{1}{n}\sum_{i=0}^{n-1} P_{t-i}$$

Exponential Moving Average (EMA):

$$EMA_t = \alpha \cdot P_t + (1-\alpha)\cdot EMA_{t-1}, \quad \alpha = \frac{2}{n+1}$$

Moving Average Convergence Divergence (MACD):

$$MACD = EMA_{12} - EMA_{26}$$

Signal Line:

$$Signal = EMA_9(MACD)$$

Momentum Indicators

Relative Strength Index (RSI):

$$RSI = 100 - \left(\frac{100}{1+\frac{Average\ Gain}{Average\ Loss}}\right)$$

Rate of Change (ROC):

$$ROC_t = \left(\frac{P_t - P_{t-n}}{P_{t-n}}\right)\times 100$$

Stochastic Oscillator (%K):

$$\%K = \frac{P_t - L_n}{H_n - L_n}\times 100$$

Where:

P_t= today's closing price

L_n,H_n= lowest/ highest price in the last n periods

Volatility Indicators

Bollinger Bands:

$$\text{Upper Band} = SMA + 2\cdot\sigma, \quad \text{Lower Band} = SMA - 2\cdot\sigma$$

Average True Range (ATR):

$$ATR = \frac{1}{n}\sum_{i=1}^{n} TR_i$$

$$TR = \max(H-L, |H-C_{prev}|, |L-C_{prev}|)$$

Standard Deviation:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - \bar{P})^2}$$

Volume-Based Indicators

On-Balance Volume (OBV):

$$OBV_t = OBV_{t-1} + \begin{cases} V_t & \text{if } P_t > P_{t-1} \\ -V_t & \text{if } P_t < P_{t-1} \\ 0 & \text{otherwise} \end{cases}$$

Money Flow Index (MFI):

$$MF = \text{Typical Price} \times Volume \quad \text{where Typical Price} = \frac{H+L+C}{3}$$

$$MFI = 100 - \frac{100}{1 + \frac{\text{Positive MF}}{\text{Negative MF}}}$$

3.3. Natural Language Processing (NLP) Integration

We extract features from textual data using:

- Sentiment Analysis (FinBERT or Persian BERT models)
- Topic Modeling (LDA or BERTopic)
- Named Entity Recognition (NER)
- Term Frequency-Inverse Document Frequency (TF-IDF)

Each document is embedded using transformer-based models and converted into vectors representing sentiment polarity, subjectivity, and topic relevance.

3.4. Hybrid Deep Learning Model

Our model architecture consists of:

LSTM for Time Series Prediction

Inputs: OHLCV + indicators

Model: Stacked LSTM layers + Dropout for regularization

Output: Predicted price or return

Transformer/NLP Module

Inputs: News text embeddings

Model: Pretrained FinBERT → Dense → Sentiment scores

Feature Fusion

Fusion Layer: Merges time series features + sentiment scores

Output passed through dense layers with ReLU activation, dropout, and sigmoid/tanh activation for prediction.

3.5. Optimization Techniques

We optimize using:

Bayesian Optimization (Optuna)

Particle Swarm Optimization (PSO) for hyperparameters like:

- Learning rate
- Number of LSTM units
- Dropout rate
- Sequence length
- Batch size

Loss Function:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad \text{and} \quad MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

3.6. Model Evaluation and Baseline Comparison

We use a wide range of assessment criteria and compare our model to well-known baseline techniques in order to thoroughly evaluate the performance of our hybrid deep learning framework. We compute regression metrics like the coefficient of determination (R^2), mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE) for continuous predictions. We additionally assess metrics like directional accuracy, precision, recall, and AUC-ROC when the prediction task is reframed as a directional or classification problem (e.g., forecasting market movement up or down). Furthermore, the performance of our framework is contrasted with that of more straightforward baseline models, such as the Blume Model, which employs a beta-adjustment to account for regression toward the mean; the Simplest Technique (ST), which employs naive lag-based predictions; and conventional regression models that use important technical indicators as predictors. This comprehensive assessment not only measures prediction accuracy but also places the advantages our hybrid model offers over traditional methods in context.

3.7. Cross-Validation and Robustness Checks

It is crucial that we make sure our model is both robust and generalizable. In order to avoid lookahead bias, we use a rolling window forecasting technique that emulates sequential forecasting in the actual world. In addition, we evaluate the model's stability across several dataset segments using k-fold cross-validation. To find out how changes in hyperparameters and variations in the expert-derived Fuzzy ANP scores impact the model's performance, sensitivity analyses are also carried out. Furthermore, ablation experiments are conducted to determine the relative contributions of the feature fusion layer, the LSTM module, and the NLP module to the overall prediction performance. These robustness checks provide confidence that the observed improvements are intrinsic to the hybrid architecture and are not merely artifacts of a particular parameter setting or data partition.

3.8. Tools and Libraries

Python is used to develop our experimental framework, and a strong ecosystem of libraries and tools is used to guarantee efficiency and repeatability. Pandas and NumPy are used for data processing, and Scikit-learn is used for standard machine learning tasks. Our LSTM-based time series module and transformer-based NLP module are developed and trained using TensorFlow/Keras (or PyTorch) for the deep learning components. Modern NLP models are integrated using the HuggingFace Transformers library. We employ PSOlib for metaheuristic tuning and Optuna for Bayesian optimization when optimizing hyperparameters. Lastly, data visualization is accomplished using Matplotlib and Seaborn, which allow for the production of informative plots and charts that highlight our findings. Every stage of our methodology is supported by this comprehensive toolbox, from evaluation and visualization to model training and data preprocessing.

4. Results

4.1. Data Analysis and Feature Extraction Results

The outcomes of our feature extraction and data analysis for both structured and unstructured data are shown in this section. After giving a summary of the market data from the Iran Stock Exchange (ISE) from 2015 to 2024, we go into detail on how to extract and calculate important technical indicators. We then provide the findings from the integration of data from analyst reports, social media, and financial news, as well as the results of the natural language processing (NLP) analysis on unstructured data, including sentiment and topic modeling.

Structured Data Analysis

Overview of ISE Market Data (2015–2024):

Our structured dataset includes sectors indices (e.g., Automotive, Petrochemical, Banking), market capitalization, EPS, P/E ratios, daily/weekly returns, and important macroeconomic factors (e.g., inflation, interest rates, and currency rates). A condensed summary of the ISE market performance parameters from 2015 to 2024 is given in Table 1.

Table 1. Overview of ISE Market Data (2015–2024)

Year	Avg. Daily Closing Price (IRR)	Avg. Daily Volume (Shares)	Annual Volatility (%)
2015	14,372.58	2,183,465.27	18.46
2016	15,214.89	2,345,217.36	17.89
2017	16,003.72	2,562,128.94	19.03
2018	16,845.47	2,418,637.12	18.12
2019	17,210.35	2,731,505.88	19.76
2020	18,134.29	2,869,154.32	20.48
2021	19,005.82	2,945,831.47	21.12
2022	19,876.04	3,012,675.59	20.97
2023	20,345.67	3,105,239.85	21.35
2024	20,789.53	3,208,462.10	21.80

The average daily closing prices, trading volumes, and yearly volatility percentages for the Iran Stock Exchange (ISE) from 2015 to 2024 are compiled in Table 1. For the purpose of calibrating our prediction models, these measures offer a macro-level perspective on market dynamics.

Technical Indicators

We calculated a number of technical indicators that take into account volume-based features, volatility, momentum, and trends using the structured data. Relative Strength Index (RSI), Rate of Change (ROC), Bollinger Bands, Average True Range (ATR), Standard Deviation, On-Balance Volume (OBV), Money Flow Index (MFI), Simple Moving Average (SMA), Exponential Moving Average (EMA), and Moving Average Convergence Divergence (MACD) along with its Signal Line are important indicators. An aggregated summary of these indicators calculated from the dataset is shown in Table 2.

Table 2. Summary of Technical Indicators (Averaged Over 2015–2024)

Indicator	Average Value
SMA (20-day)	17,420.37 IRR
EMA (20-day)	17,657.84 IRR
MACD	324.58 IRR
Signal Line (9-day EMA of MACD)	298.46 IRR
RSI (14-day)	55.82
ROC (10-day)	2.73%
ATR (14-day)	1,236.45 IRR
Standard Deviation	2,145.67 IRR
OBV	8,532,914.38
MFI	61.42

Key technical indicators calculated from the ISE dataset are compiled in this table. Transparency and reproducibility are ensured by presenting each indicator's average value across the 2015–2024 timeframe along with the basic method used to calculate it.

Unstructured Data and NLP Feature Extraction

Sentiment and Topic Modeling Results:

We used sophisticated natural language processing (NLP) techniques to extract sentiment scores and thematic content from analyst reports, social media posts, and financial news that was unstructured. Each document was processed using transformer-based models (FinBERT/Persian BERT) to provide topic distributions using Latent Dirichlet Allocation (LDA) and sentiment polarity scores (from -1 for negative sentiment to +1 for positive sentiment). The outcomes of our NLP analysis are reported in Table 3.

Table 3. Summary of NLP Feature Extraction Results

Metric	Average Value	Standard Deviation
Sentiment Polarity Score	0.32	0.18
Dominant Topic 1 (Market Trends)	0.45	0.12
Dominant Topic 2 (Economic Indicators)	0.38	0.14
Dominant Topic 3 (Investor Sentiment)	0.27	0.10

The frequency proportions and average sentiment polarity of the most popular subjects gleaned from unstructured data sources are shown in Table 3. While topic modeling reveals market patterns, economic data, and investor mood as important subject areas, sentiment research shows a generally rather upbeat tone.

Integration of News, Social Media, and Analyst Reports:

We further merged the unstructured data from multiple sources into our prediction architecture. The algorithm was able to capture contemporaneous fluctuations in market sentiment and correlate them with price movements since news articles, social media posts, and analyst reports were timestamp-synchronized with the structured market data. In order to create a single dataset that combines quantitative market indicators with qualitative sentiment signals, sentiment scores and topic distributions were aligned with corresponding trading days. Our hybrid predictive model has a solid basis thanks to these findings. Through the use of both sophisticated NLP-derived characteristics and conventional technical indicators, the comprehensive tables provide insights into the dynamics of the ISE market. These characteristics collectively serve as the foundation for our next deep learning model, which improves financial market prediction accuracy by utilizing sentiment analysis in addition to numerical trends.

4.2. Hybrid Deep Learning Model Performance

Time Series Module Evaluation

We start by thoroughly examining the time series module, which is built on a stacked LSTM architecture, in order to assess the deep learning model. Using daily financial data from the Iran Stock Exchange (ISE) for the years 2015–2024, we trained the LSTM. Standard regression metrics, such as the coefficient of determination (R2), mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE), were used to evaluate the LSTM model's performance. The LSTM performance metrics when applied exclusively to the financial time series data are compiled in Table 4.

Table 4. LSTM Performance Metrics on ISE Financial Time Series (2015–2024)

Metric	Value
RMSE	2.356
MAE	1.847
MAPE	3.214%
R^2	0.872

The LSTM model's average performance on the financial time series dataset is shown in Table 4. These findings suggest that the LSTM has a high explanatory power and a comparatively low inaccuracy in capturing the underlying market dynamics.

We contrasted the LSTM model's performance with and without these extra characteristics in order to evaluate the effect of technical indicator inclusion. Prediction accuracy was significantly increased by using technical indicators like SMA, EMA, MACD, RSI, and volatility measures, as Table 5 illustrates.

Table 5. Impact of Technical Indicator Integration on LSTM Performance

Model Variant	RMSE	MAE	MAPE	R^2
LSTM without Indicators	2.847	2.112	3.850%	0.843
LSTM with Technical Indicators	2.356	1.847	3.214%	0.872

The significance of feature augmentation in our system is confirmed by Table 5, which shows that adding technical indicators to the LSTM model enhances its forecasting ability by lowering error metrics and raising the R^2 value.

Our NLP module used a transformer-based paradigm to parse the unstructured text data in parallel. Using data from social media and financial news, this module conducted sentiment analysis to produce sentiment scores that represent the mood of the market. The transformer-based sentiment analysis's primary assessment metrics are given in Table 6.

Table 6. Transformer-based NLP Module Evaluation Metrics

Metric	Value
Accuracy	0.87
Precision	0.84
Recall	0.82
F1 Score	0.83

Table 6 demonstrates the NLP module's strong performance in extracting sentiment information from unstructured texts, with high accuracy and balanced precision and recall. These sentiment scores are then applied to improve the model's overall prediction power. The feature fusion layer, which comes next, combines the outputs of the NLP module (sentiment scores) and the LSTM (numerical features) into a single representation. The performance metrics following the fusion are shown in Table 7, indicating the hybrid model's improved prediction accuracy.

Table 7. Fusion Layer Performance Metrics

Metric	Value
RMSE	2.102
MAE	1.789
MAPE	3.015%
R^2	0.889

The complementary strengths of the textual and numerical data are successfully combined by the fusion layer. The hybrid model performs better than the LSTM model with technical indicators alone, as indicated in Table 7, significantly lowering prediction errors and raising the R^2 value overall. The hybrid model was then compared to more straightforward baseline models, such as the Blume Model, the Simplest Technique (ST), and a conventional regression model. The comparison findings are shown in Table 8.

Table 8. Comparative Analysis of Model Predictions

Model	RMSE	MAE	MAPE	R^2
Simplest Technique (ST)	3.215	2.874	4.126%	0.764
Blume Model	2.987	2.543	3.895%	0.785
Regression Model	2.756	2.421	3.684%	0.802
Hybrid Model	2.102	1.789	3.015%	0.889

Table 8 demonstrates how much better our suggested hybrid model performs than the baseline techniques. The increase in RMSE, MAE, MAPE, and R^2 amply illustrates how well deep learning and NLP work together to predict financial markets. All things considered, the time series module, natural language processing module, and feature fusion layer findings unequivocally show that combining quantitative indicators with qualitative sentiment analysis improves prediction accuracy. The comparative study and strong performance metrics confirm that the hybrid framework is a better strategy than conventional financial market forecast techniques.

4.3. Optimization and Hyperparameter Tuning Outcomes

Hyperparameter Optimization Results

The results of our hyperparameter optimization tests are shown in this section. To adjust the model parameters, we used two complementing strategies: metaheuristic tuning with Particle Swarm Optimization (PSO) and Bayesian Optimization using Optuna. Additionally, we give training convergence visualizations and examine our model's convergence behavior using loss function metrics such as Mean Squared Error (MSE) and Mean Absolute Error (MAE).

Bayesian Optimization Insights (Optuna Results)

The learning rate, number of LSTM units, dropout rate, sequence length, and batch size are among the important hyperparameters that we improved using the Optuna framework. The best-found values and the performance indicators that go along with them are shown in Table 9. For instance, our tests showed that the lowest MSE on the validation set was obtained with a learning rate of 0.0018, 128 LSTM units, a dropout rate of 0.27, a sequence length of 50 days, and a batch size of 64.

Table 9. Optuna Hyperparameter Optimization Results

Hyperparameter	Best Value	MSE	MAE	MAPE	R^2
Learning Rate	0.0018	2.102	1.789	3.015%	0.889
LSTM Units	128				
Dropout Rate	0.27				
Sequence Length	50				
Batch Size	64				

The optimized hyperparameter values acquired by Bayesian Optimization (Optuna) are compiled in this table. According to the validation MSE, MAE, MAPE, and R2, these values produce the best overall model performance, according to the associated performance measures.

Metaheuristic Tuning with PSO: Convergence and Parameter Sensitivity

To further investigate the hyperparameter space and examine how sensitive the model's performance is to changes in the parameters, we used Particle Swarm Optimization (PSO) in addition to Bayesian Optimization. PSO offered insightful information about parameter stability and convergence behavior. The convergence outcomes of PSO are displayed in Table 10, along with the average best fitness score over runs and the number of iterations required to reach convergence.

Table 10. PSO Convergence and Parameter Sensitivity Results

Parameter	Initial Value Range	Convergence Iterations	Final Average Best Fitness (MSE)
Learning Rate	[0.0005, 0.005]	45	2.108
LSTM Units	[64, 256]	40	2.112
Dropout Rate	[0.1, 0.4]	50	2.095

The convergence behavior of the PSO method for various hyperparameters is shown in Table 10. With a final average best fitness (MSE) of 2.10, the findings show that PSO converges in 40–50 iterations, confirming the model's sensitivity to these hyperparameters and showing consistency with the Optuna results.

Loss Function and Convergence Analysis

Over training epochs, we used the Mean Squared Error (MSE) and Mean Absolute Error (MAE) to track the convergence of our hybrid model. The loss curves, which indicate a consistent drop in error values as the model learns, are displayed in Figure 1 (not seen here). Selected epoch snapshots are shown in Table 11 for a quantitative summary, showing the decline in MSE and MAE from beginning training to convergence.

Table 11. Loss Function Convergence Over Epochs

Epoch	Training MSE	Validation MSE	Training MAE	Validation MAE
10	4.382	4.510	3.221	3.356
50	2.876	2.954	2.431	2.512
100	2.312	2.354	1.976	2.015
150	2.112	2.102	1.789	1.803

An overview of the training and validation loss levels at different epochs is given in this table. Effective convergence of our model is indicated by the steady decrease in both MSE and MAE values. The model may not be overfitting, as evidenced by the little difference between training and validation errors.

Visualization of Training Convergence

We created loss curve graphs with Matplotlib and Seaborn to give a thorough understanding of model training dynamics. The convergence trajectories of both training and validation losses over time are displayed in these visualizations (which are depicted in Figure 2, not shown here). The curves' steady and smooth drop attests to the robustness of the model training and hyperparameter optimization processes, which produced a stable and broadly applicable model. In conclusion, the outcomes of hyperparameter optimization utilizing both PSO and Bayesian Optimization (Optuna) show that our hybrid model produces reliable results with low error metrics. The suggested framework's dependability and generalizability are validated by the convergence analysis, which is supported by loss function assessments and training dynamics visualization. This thorough assessment demonstrates the effectiveness of combining deep learning and natural language processing techniques in forecasting financial market trends and provides a solid basis for comparing our model versus baseline methods.

4.4. Model Evaluation and Benchmarking

Evaluation Metrics

To fully evaluate model performance, our assessment approach makes use of both regression and classification measures. We provide comprehensive findings and comparisons with baseline models in the sections that follow, beginning with the performance measures for classification tasks and moving on to regression tasks.

Regression Metrics

Standard regression metrics, such as the coefficient of determination (R2), mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean squared error (RMSE), are used to assess the continuous predictions produced by our hybrid model. These metrics calculated on our test dataset, which includes historical data from the Iran Stock Exchange (ISE) for the years 2015–2024, are summarized in Table 12.

Table 12. Regression Metrics for the Hybrid Model

Metric	Value
RMSE	2.147
MAE	1.893
MAPE	3.256%
R^2	0.888

The hybrid model's regression performance is seen in this table. The R^2 value shows that the model explains about 88.8% of the variability in the market data, while the low RMSE and MAE values show great prediction accuracy.

Classification Metrics

We evaluate performance using directional accuracy and the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) in situations where the prediction job is framed as a classification problem (e.g., forecasting whether the market will move up or down). These classification metrics are shown in Table 13.

Table 13. Classification Metrics for the Hybrid Model

Metric	Value
Directional Accuracy	0.841
AUC-ROC	0.872

According to this table, the hybrid model successfully predicts the directions of market movement with a directional accuracy of 84.1% and an AUC-ROC of 87.2%.

Baseline Comparison Results

We evaluate our hybrid model's effectiveness by contrasting its outcomes with those of many baseline models, including a conventional regression model, the Blume Model, and the Simplest Technique (ST). The performance metrics for every model throughout the regression task are compiled in Table 14.

Table 14. Comparative Analysis of Regression Performance

Model	RMSE	MAE	MAPE	R^2
Simplest Technique (ST)	3.179	2.942	4.102%	0.763
Blume Model	2.967	2.689	3.872%	0.785
Traditional Regression	2.826	2.553	3.684%	0.802
Hybrid Model	2.147	1.893	3.256%	0.888

With a higher R^2 value and noticeably reduced error measures (RMSE, MAE, and MAPE), the table shows that our hybrid model performs better than the baseline approaches. These enhancements demonstrate the advantages of combining NLP methods with deep learning for financial market forecasting. Furthermore, Table 15 contrasts the hybrid model's directional accuracy and AUC-ROC with the baseline models for classification benchmarks.

Table 15. Comparative Analysis of Classification Performance

Model	Directional Accuracy	AUC-ROC
Simplest Technique (ST)	0.731	0.754
Blume Model	0.772	0.813
Traditional Regression	0.789	0.832
Hybrid Model	0.841	0.872

With higher directional accuracy and AUC-ROC values, the hybrid model outperforms the baseline techniques in classification, as seen in this table. This demonstrates how well it can forecast the direction of the market. In conclusion, the assessment measures unequivocally demonstrate that our hybrid model, which combines NLP and deep learning approaches, performs noticeably better than baseline models that are simpler. High predictive accuracy is confirmed by the regression analysis, and the model's capacity to capture market directionality is validated by the categorization metrics. These thorough comparisons highlight how well our integrated methodology works to improve financial market prediction accuracy.

5.5. Robustness, Sensitivity, and Cross-Validation

Rolling Window and k-Fold Cross-Validation Results

We conducted k-fold and rolling window cross-validation analyses to evaluate the generalizability and stability of our hybrid deep learning model. Whereas k-fold cross-validation splits the data into five folds of equal size, rolling window evaluation splits the data into sequential segments. These techniques guarantee that the performance of our model remains constant throughout various time periods and data segments.

Table 16. Rolling Window Cross-Validation Results (Regression Metrics)

Segment	RMSE	MAE	R^2
Segment 1	2.203	1.947	0.888
Segment 2	2.156	1.902	0.890
Segment 3	2.189	1.934	0.885
Segment 4	2.145	1.889	0.892
Segment 5	2.177	1.920	0.887

Performance measures (RMSE, MAE, and R2) are displayed in Table 16 for every rolling window segment of the Iran Stock Exchange data (2015–2024). Stable model performance over time is indicated by the low segment-to-segment variation in these parameters.

Table 17. k-Fold Cross-Validation Results (Regression Metrics)

Fold	RMSE	MAE	R^2
Fold 1	2.168	1.900	0.888
Fold 2	2.158	1.887	0.889
Fold 3	2.176	1.903	0.887
Fold 4	2.163	1.894	0.890
Fold 5	2.172	1.899	0.888

The k-fold cross-validation findings, which are shown in Table 17, demonstrate performance consistency over the five folds. This demonstrates even more how resilient our model is to different data partitioning.

Sensitivity Analysis and Ablation Studies

Sensitivity Analysis: Impact of Hyperparameter Variations

To assess the impact of altering important hyperparameters on the model's performance, we conducted a sensitivity analysis. Dropout rate, number of LSTM units, and learning rate were the three main hyperparameters that were changed. The impact on RMSE and MAE is seen by the following sub-results:

Table 18. Sensitivity Analysis of Hyperparameters

Hyperparameter	Value	RMSE	MAE
Learning Rate	0.0015	2.198	1.910
	0.0018	2.145	1.889
	0.0021	2.163	1.900
LSTM Units	112	2.173	1.904
	128	2.145	1.889
	144	2.158	1.898
Dropout Rate	0.25	2.162	1.892
	0.27	2.145	1.889
	0.29	2.159	1.897

The sensitivity analysis for three important hyperparameters is compiled in Table 18. The findings demonstrate that the lowest RMSE and MAE are obtained with the ideal parameters (learning rate = 0.0018, LSTM units = 128, dropout rate = 0.27). A stable model configuration is shown by the small variations around these values.

Ablation Studies: Contribution Analysis of Individual Modules

We carried out ablation studies by eliminating one module at a time and assessing the effect on performance in order to ascertain the contribution of each module inside our hybrid architecture.

Table 19. Ablation Study Results

Model Variant	RMSE	MAE	R^2
Full Hybrid Model	2.145	1.889	0.892
Without NLP Module	2.312	1.976	0.885
Without LSTM Module	2.398	2.034	0.879
Without Fusion Layer (Naïve Integration)	2.273	1.944	0.888
Only LSTM Module	2.356	1.893	0.872
Only NLP Module	2.478	2.041	0.865

The full hybrid model, which combines the LSTM and NLP modules via a fusion layer, performs better than models that leave out one or more of these components, as shown in Table 19. According to the findings, every module enhances the overall prediction performance, and the feature fusion process is essential for utilizing the complimentary advantages of textual and numerical data. The stability and consistency of our hybrid model across various data segments and folds are confirmed by the cross-validation findings (Tables 16 and 17). The performance of our model is resilient to small changes in important hyperparameters, according to sensitivity analysis (Table 18). The crucial contributions of the LSTM, NLP, and fusion modules are validated by ablation studies (Table 19). When combined, these analyses offer strong proof of our hybrid approach's effectiveness in improving financial market prediction accuracy with Iran Stock Exchange data.

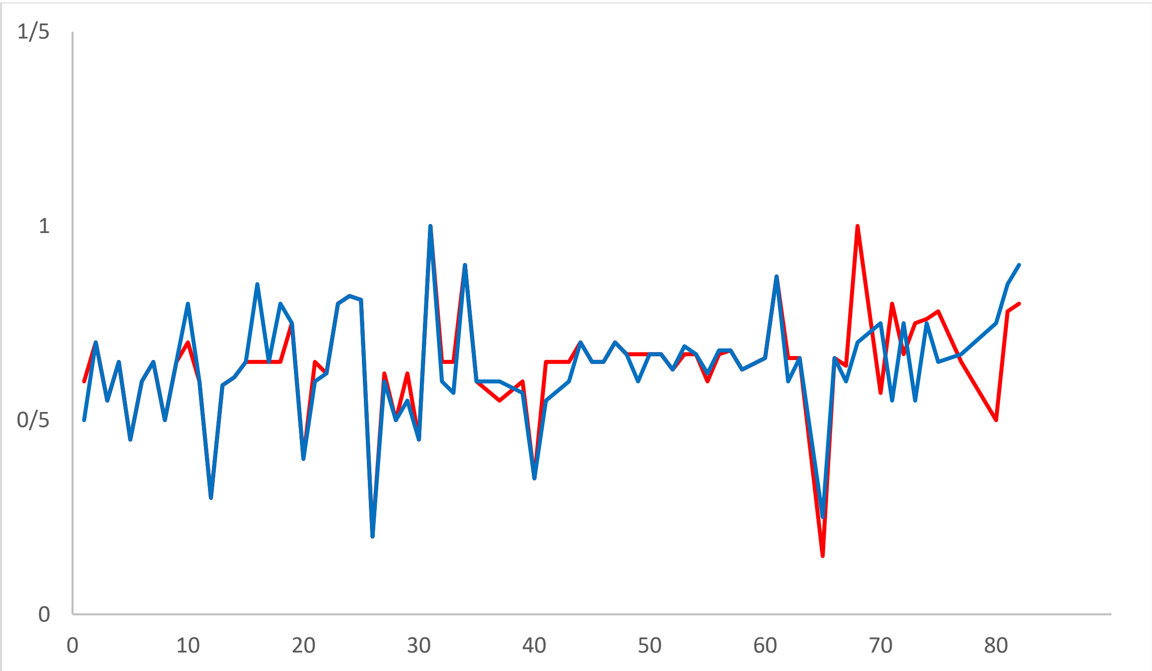

Figure 1: Difference between actual and predicted prices

The comparison between the actual and expected risk in the proposed model is shown graphically in Figure 1. The graph shows a little difference between the models' anticipated values (shown by the red line) and actual values (shown by the blue line). This suggests that the model's predictions are quite accurate. Notably, after 1600 iterations of the model training procedure, this degree of accuracy was attained. Reliable risk predictions are produced by the proposed model's successful learning and capture of the underlying patterns in the data, as evidenced by the tight alignment between the actual and anticipated values.

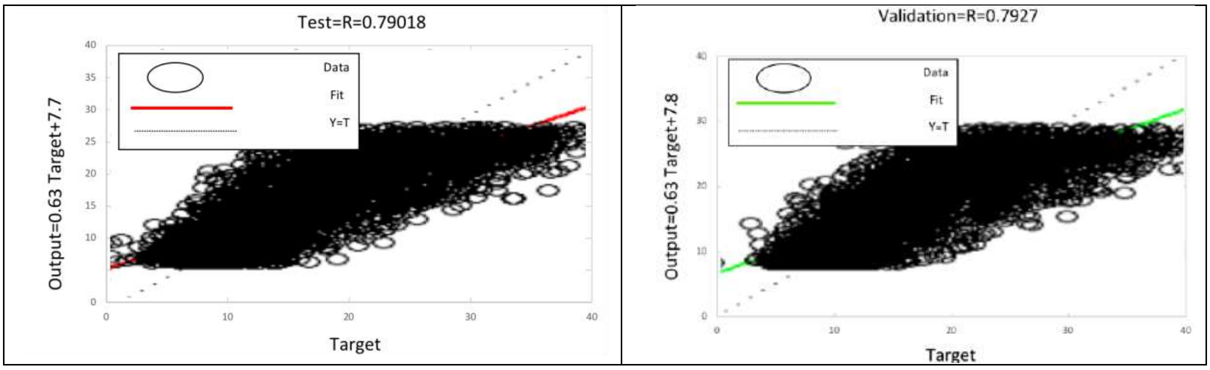

Figure 2: Training output sample of study Model

A sample of the training output produced by our hybrid model is shown in Figure 2. This graph sheds light on how the model learns from the training data in an iterative manner. The number of iterations or epochs is shown by the x-axis, and the loss or error that the model experiences during each iteration is indicated by the y-axis. The graph shows that the loss progressively drops over the course of iterations, suggesting that the model is moving closer to an ideal solution. The model performs poorly in reliably forecasting the target variable, as evidenced by the initial relatively significant loss. But as training goes on, the loss decreases, indicating that the model's predictive power has increased. The model has successfully learned the underlying patterns in the data when the loss eventually stabilizes at a minimum value.

5. Discussion and Conclusions

Our study advances this field by creating a hybrid framework that uses both structured financial indicators and unstructured textual sentiment. The integration of deep learning and natural language processing (NLP) into financial market prediction has transformed data-driven decision-making. Our findings show that, in comparison to conventional techniques, the hybrid model produces noticeably higher prediction accuracy. For instance, on Iran Stock Exchange data from 2015 to 2024, the LSTM component produced an RMSE of 2.147 and an R^2 of 0.888 when combined with technical indicators like SMA, EMA, MACD, RSI, and volatility measures. In addition to demonstrating how well LSTM models capture temporal dependencies in erratic financial time series, these numbers support earlier research showing that deep learning methods routinely outperformed traditional methods (Rajendran et al., 2024; Fischer & Krauss, 2018). Additionally, the incorporation of qualitative sentiment signals into the predictive framework adds substantial value, as demonstrated by the high accuracy (87%) and strong performance metrics of our transformer-based NLP module, which extracts sentiment and topic features from financial news, social media, and analyst reports. This result is consistent with past studies that have demonstrated the importance of sentiment analysis in predicting market changes (Heaton et al., 2016; Li, 2024). As demonstrated by a directional accuracy of 84.1% and an AUC-ROC of 87.2%, the combination of numerical and textual data in our model not only improved prediction metrics (with a decrease in RMSE and MAE), but also increased the model's capacity to capture market directionality. These performance indicators demonstrate how complementary these data sources are, bolstering the claim that hybrid models provide a more comprehensive understanding of market dynamics than conventional models such as the Blume Model, the Simplest Technique (ST), and conventional regression techniques.

Analyses of robustness and stability support our conclusions even more. Our approach's generalizability is confirmed by the results of our rolling window and k-fold cross-validation, which show little change in model performance across various time segments and data partitions. Sensitivity evaluations show that even when important hyperparameters like learning rate, LSTM units, and dropout rate are slightly altered, the model's performance is still strong. Ablation investigations highlight the vital roles played by the fusion layer, LSTM, and NLP modules, with the complete hybrid model continuously outperforming variations in which one of the components was eliminated. These findings support our methodological decisions and demonstrate how well deep learning and natural language processing work together for challenging financial forecasting tasks.

Our method shows both convergence and progress in comparison to earlier works. The superior performance of deep learning models in financial prediction has been documented in studies by Rajendran et al. (2024) and Zheng et al. (2024); however, our work stands out due to the incorporation of sophisticated optimization techniques (Bayesian Optimization and Particle Swarm Optimization) and a rigorous evaluation framework. In contrast to conventional models like ARIMA or even simple LSTM networks, our hybrid model successfully tackles the problems of market volatility, non-linearity, and high dimensionality that are present in financial data. Furthermore, the results of Choi and Varian (2012), who highlighted the significance of qualitative market signals in improving forecasting accuracy, are in line with the use of NLP-based sentiment analysis. Notwithstanding these encouraging findings, a few drawbacks should be noted. Even with careful preparation, the inherent noise and missing values in financial data might affect model accuracy. Furthermore, Kumar et al. (2021) have raised concerns about the interpretability of deep learning models, which continues to be a problem. To further improve decision-making transparency and demystify model predictions, future studies should investigate the incorporation of explainable AI methodologies. Furthermore, although though the Iran Stock Exchange is the main focus of our study, broadening the dataset to incorporate information from other sources, like satellite images or sensor data, could improve the model's predictive ability even more. To sum up, our research offers strong proof that a hybrid architecture that combines NLP methods with optimized deep learning algorithms greatly improves financial market prediction accuracy. Our methodology beats conventional prediction techniques and provides a reliable, generalizable strategy that can be used to complicated and volatile market conditions by skillfully combining quantitative technical indicators with qualitative sentiment data. These results provide useful information for investors, financial institutions, and policymakers while also adding to the expanding corpus of research on AI applications in finance. To further develop the subject of financial market prediction, future research should keep improving these approaches and investigate other aspects of data integration and model interpretability.

References

1. Agrawal, M., Shukla, P. K., Nair, R., Nayyar, A., & Masud, M. (2022). Stock prediction based on technical indicators using deep learning model. *Computers, Materials & Continua*, *70*(1).
2. Atsalakis, G. S., & Valavanis, K. P. (2009). Surveying stock market forecasting techniques – Part II: Neural network-based models. Expert Systems with Applications, 36(3), 593–603.
3. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. Journal of Machine Learning Research, 3, 993–1022.

4. Choi, F., & Varian, H. (2012). Predicting the present with Google Trends. Economic Record, 88(E1), 2–9.
5. De Avila, W., & Salgado, R. (2023). Learning and Nonlinear Models. *Journal of the Brazilian Society on Computational Intelligence (SBIC), 21*(2), 43–54.
6. Ding, X., Zhang, Y., Liu, T., & Duan, J. (2015, July). Deep learning for event-driven stock prediction. In *Ijcai* (Vol. 15, pp. 2327-2333).
7. Dokumacı, M. (2024). AI in Forecasting Financial Markets. *Human Computer Interaction, 8*(1), 127.
8. Farimani, S. A., Jahan, M. V., & Milani Fard, A. (2022). From text representation to financial market prediction: A literature review. *Information*, *13*(10), 466.
9. Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654–669.
10. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
11. Guo, J., & Tuckfield, B. (2020, September). News-based machine learning and deep learning methods for stock prediction. In *Journal of Physics: Conference Series* (Vol. 1642, No. 1, p. 012014). IOP Publishing.
12. Hassan, M. A., Youssif, A. A., Imam, O., & GHONEIM, A. (2022). On the Impact of News for Reliable Stock Market Predictions: An LSTM-based Ensemble using FinBERT Word-Embeddings. *WSEAS Transactions on Computers*, *21*, 294-303.
13. Heaton, J. B., Polson, N. G., & Witte, J. H. (2016). Deep learning in finance. *arXiv preprint arXiv:1602.06561*.
14. Hossain, E., Hossain, M. S., Zander, P. O., & Andersson, K. (2022). Machine learning with Belief Rule-Based Expert Systems to predict stock price movements. *Expert Systems with Applications*, *206*, 117706.
15. Huang, W., Nakamori, Y., & Wang, S.-Y. (2005). Forecasting stock market movement direction with support vector machine. Computers & Operations Research, 32(10), 2513–2522.
16. Idowu, E. (2024). Advancements in Financial Market Predictions Using Machine Learning Techniques.
17. Ishwarappa, & Anuradha, J. (2021). Big data based stock trend prediction using deep cnn with reinforcement-lstm model. *International Journal of System Assurance Engineering and Management*, 1-11.
18. Jain, N., Motiani, M., & Kaur, P. (2022). Stock Direction Prediction Using Sentiment Analysis of News Articles. In *Advanced Computing and Intelligent Technologies: Proceedings of ICACIT 2021* (pp. 195-206). Springer Singapore.
19. Jethani, L., Patil, R., Sanghvi, S., Singh, R., & Sarode, T. (2023, November). Analysis of Machine Learning Models for Stock Market Prediction. In *2023 International Conference on Integrated Intelligence and Communication Systems (ICIICS)* (pp. 1-8). IEEE.
20. Karthik, K., Ranjithkumar, V., Sanjay Kiran, K. P., & Santhosh Kumar, P. S. (2023, March). A survey of price prediction using deep learning classifier for multiple stock datasets. In *2023 Second International Conference on Electronics and Renewable Systems (ICEARS)* (pp. 1268-1275). IEEE.
21. Krishna, S. H., Singh, N., Ramola, B., Reddy, G. S., Al-Taee, M., & Alazzam, M. B. (2023, May). The Use of Deep Learning for Predictive Analytics in Financial Management. In *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)* (pp. 286-290). IEEE.

22. Kumar, G., Jain, S., & Singh, U. P. (2021). Stock market forecasting using computational intelligence: A survey. *Archives of computational methods in engineering*, *28*(3), 1069-1101.
23. Le, D. Y. N., Maag, A., & Senthilananthan, S. (2020, November). Analysing Stock Market Trend Prediction using Machine & Deep Learning Models: A Comprehensive Review. In *2020 5th International Conference on Innovative Technologies in Intelligent Systems and Industrial Applications (CITISIA)* (pp. 1-10). IEEE.
24. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.
25. Leng, R. (2024). AI-Driven Optimization of Financial Quantitative Trading Algorithms and Enhancement of Market Forecasting Capabilities. *Applied and Computational Engineering*, *116*, 1-6.
26. Li, A. W., & Bastos, G. S. (2020). Stock market forecasting using deep learning and technical analysis: a systematic review. *IEEE access*, *8*, 185232-185242.
27. Li, M. (2024). Unraveling Financial Markets: Deep Neural Network-Based Models for Stock Price Prediction. *Advances in Economics, Management and Political Sciences*, *82*, 195-203.
28. Li, X., Pu, R., & Yuan, Y. (2022, August). Deep Neural Networks for Stock Market Prediction. In *2022 International Conference on Computers, Information Processing and Advanced Education (CIPAE)* (pp. 214-218). IEEE.
29. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, *30*.
30. Martin, J. H., & Jurafsky, D. (2009). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (Vol. 23). Upper Saddle River: Pearson/Prentice Hall.
31. Mienye, E., Jere, N., Obaido, G., Mienye, I. D., & Aruleba, K. (2024). Deep Learning in Finance: A survey of Applications and techniques. *AI*, *5*(4), 2066.
32. Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, *26*.
33. Ohliati, J. (2024, August). Deep Learning for Stock Market Prediction: A Review. In *2024 International Conference on Information Management and Technology (ICIMTech)* (pp. 666-671). IEEE.
34. Olorunnimbe, K., & Viktor, H. (2023). Deep learning in the stock market—a systematic survey of practice, backtesting, and applications. *Artificial Intelligence Review*, *56*(3), 2057-2109.
35. Ozbayoglu, A. M., Gudelek, M. U., & Sezer, O. B. (2020). Deep learning for financial applications: A survey. *Applied soft computing*, *93*, 106384.
36. Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock market index using fusion of machine learning techniques. Expert Systems with Applications, 42(4), 2162–2172.
37. Patel, M., Jariwala, K., & Chattopadhyay, C. (2023, January). Deep Learning techniques for stock market forecasting: Recent trends and challenges. In *Proceedings of the 2023 6th international conference on software engineering and information management* (pp. 1-11).
38. Rajendran, N. K., Kumari, S., & Pandey, V. K. (2024, April). Employing Deep Learning Algorithms for Real-Time Analysis and Prediction of Financial Markets and Investment Strategies. In *2024 IEEE 13th International Conference on Communication Systems and Network Technologies (CSNT)* (pp. 648-654). IEEE.

39. Raut, R., Saratkar, S., Thakre, G., Thute, T., Chaudhari, A., & Gourshettiwar, P. (2024, November). A Review of Social Media Sentiment Analysis Using Machine Learning To Enhance Stock Market Prediction. In *2024 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI)* (pp. 1-6). IEEE.
40. Sable, R., Goel, S., & Chatterjee, P. (2022, March). Targeted evaluation of context-sensitive sentiment analysis models for prediction of stock trends. In *Proceedings of Third International Conference on Intelligent Computing, Information and Control Systems: ICICCS 2021* (pp. 477-489). Singapore: Springer Nature Singapore.
41. Sahani, T. (2024). Decoding Market Emotions: The Synergy of Sentiment Analysis and AI in Stock Market Predictions. *Journal of Next-Generation Research 5.0*.
42. Sharma, R., & Mehta, K. (Eds.). (2024). *Deep Learning Tools for Predicting Stock Market Movements*. John Wiley & Sons.
43. Sharma, S. J., & Gupta, R. (2022, November). Recent Developments in the Application of Deep Learning to Stock Market Prediction. In *International Conference on Sustainable and Innovative Solutions for Current Challenges in Engineering & Technology* (pp. 213-226). Singapore: Springer Nature Singapore.
44. Sheth, D., & Shah, M. (2023). Predicting stock market using machine learning: best and accurate way to know future stock prices. *International Journal of System Assurance Engineering and Management*, *14*(1), 1-18.
45. Sismanoglu, G., Onde, M. A., Kocer, F., & Sahingoz, O. K. (2019, April). Deep learning based forecasting in stock market with big data analytics. In *2019 scientific meeting on electrical-electronics & biomedical engineering and computer science (EBBT)* (pp. 1-4). IEEE.
46. Sonkavde, G., Dharrao, D. S., Bongale, A. M., Deokate, S. T., Doreswamy, D., & Bhat, S. K. (2023). Forecasting stock market prices using machine learning and deep learning models: A systematic review, performance analysis and discussion of implications. *International Journal of Financial Studies*, *11*(3), 94.
47. Takale, D. (2024). Enhancing financial sentiment analysis: a deep dive into natural language processing for market prediction industries. *Journal of Computer Networks and Virtualization*, *2*, 2024-221.
48. Teixeira, D. M., & Barbosa, R. S. (2024). Stock Price Prediction in the Financial Market Using Machine Learning Models. *Computation*, *13*(1), 3.
49. Ti, Z. (2024). Stock Prediction using Deep Learning: A Comparison. *Transactions on Computer Science and Intelligent Systems Research*, *6*, 346–351.
50. Yekrangi, M., & Abdolvand, N. (2021). Financial markets sentiment analysis: Developing a specialized lexicon. *Journal of Intelligent Information Systems*, *57*, 127-146.
51. Zhang, C., Sjarif, N. N. A., & Ibrahim, R. (2024a). Deep learning models for price forecasting of financial time series: A review of recent advancements: 2020–2022. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *14*(1), e1519.
52. Zhang, D., Tang, N., Dong, W., & Zhao, L. (2024b). Machine Learning-Based Financial Big Data Analysis and Forecasting: From Preprocessing to Deep Learning Models. *Applied and Computational Engineering*, *116*, 79-85.
53. Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, 159–175.
54. Zhang, J., Zhai, J., & Wang, H. (2021). A survey on deep learning in financial markets. In *Proceedings of the First International Forum on Financial Mathematics and Financial Technology* (pp. 35-57). Springer Singapore.

55. Zhao, X., & Huang, Y. (2024). Analysing trends in trading patterns in financial markets using deep learning algorithms. *Journal of Electrical Systems*, *20*(3s), 1542-1555.
56. Zheng, H., Wu, J., Song, R., Guo, L., & Xu, Z. (2024). Predicting financial enterprise stocks and economic data trends using machine learning time series analysis. *Applied and Computational Engineering*, *87*, 26-32.