

# Artificial intelligence-based stock market price prediction, a review

Vishwas H N<sup>1</sup>, Usha B A<sup>2</sup>, Sangeetha K N<sup>3</sup>

<sup>1</sup>BMS Institute of Technology and Management, Visvesvaraya Technological University, Belagavi, Karnataka, 590018, India

<sup>2</sup>BMS Institute of Technology and Management, Bengaluru, Karnataka 590119, India

<sup>3</sup>JSS Academy of Technical Education, Bengaluru, Karnataka, 560060, India

<sup>1</sup>Corresponding author

**E-mail:** <sup>1</sup>vishwashn@gmail.com, <sup>2</sup>ushaajay1@gmail.com, <sup>3</sup>sangeethakn@jssateb.ac.in

Received 27 February 2025; accepted 13 August 2025; published online 12 November 2025  
DOI <https://doi.org/10.21595/mme.2025.24861>



Copyright © 2025 Vishwas H N, et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Abstract.** The stock market, a cornerstone of the global financial system, is characterized by its dynamic and volatile nature, which makes accurate price-trend prediction challenging. However, traditional statistical models often fail to capture this complexity. Recent advancements in Artificial Intelligence (AI), particularly Machine Learning (ML) and Deep Learning (DL), have transformed stock market forecasting by using diverse datasets and algorithms. This review examines recent studies on AI methodologies for stock market price trend prediction models by analyzing architectures, datasets, performance metrics, and limitations, with a focus on hybrid models, sentiment analysis, and dataset diversity. Hybrid approaches, including the Multi-Model Generative Adversarial Network Hybrid Prediction Algorithm (MMGAN-HPA), K-means long short-term memory (LSTM), and LSTM autoregressive output (LSTM-ARO), improve predictive accuracy by combining statistical methods with deep learning. Sentiment analysis models such as Stock Senti WordNet (SSWN) and Hybrid Quantum Neural Network (HQNN) integrate social media sentiment to capture market dynamics. Real-time frameworks that use stream processing show promise for high-frequency trading applications. This review addresses key challenges including data noise, nonstationarity, overfitting risks, and black-box model interpretability. Solutions include GAN-based synthetic data generation, transformer-based architectures such as SpectralGPT, and optimization techniques for computational efficiency. This review provides a taxonomy of AI-based approaches, while identifying gaps for future research. These findings highlight the potential of AI in financial forecasting while emphasizing the need for interdisciplinary collaboration to address its limitations in data quality, methodology, interpretability, and ethics.

**Keywords:** artificial intelligence (AI), autoregressive integrated moving average artificial neural network (ARIMA-ANN) (time-series forecasting), generative adversarial network (GAN), K-means long short-term memory (LSTM), stock senti wordnet (SSWN), hybrid quantum neural network (HQNN), spectral GPT.

## 1. Introduction

The stock market, a critical component of the global financial system, reflects economic health, and serves as a platform for investment and wealth creation. The accurate prediction of stock market price trends is essential for investors, financial institutions, and policymakers to optimize portfolio management, mitigate risks, and make informed decisions. However, the nonlinear, dynamic, and stochastic nature of financial markets poses significant challenges to traditional forecasting methods, such as Autoregressive Integrated Moving Average Artificial Neural networks (ARIMA) and GARCH [1, 2]. These models often struggle to capture the complex dependencies and rapid fluctuations inherent to financial data.

Recent advancements in Artificial Intelligence (AI) have revolutionized stock market predictions by introducing advanced computational models capable of analyzing vast datasets and

uncovering intricate patterns. Machine Learning (ML) and Deep Learning (DL) techniques have demonstrated superior performance in handling nonlinearity and large-scale data compared to traditional statistical approaches [3, 4]. Models such as Long Short-Term Memory (LSTM) networks capture temporal dependencies in sequential data [5], whereas hybrid frameworks such as K-means LSTM enhance the predictive accuracy by integrating clustering with temporal modeling [6].

In addition, Generative Adversarial Networks (GANs) address challenges such as data sparsity by generating synthetic datasets to improve model robustness against volatility [7]. Sentiment analysis models such as Stock Senti WordNet (SSWN) leverage social media data to incorporate public sentiment into predictive frameworks, further enhancing short-term forecasting accuracy [8]. Explainable AI frameworks such as Hybrid Quantum Neural Networks (HQNN) integrate sentiment analysis with quantum computing principles to improve interpretability while maintaining high accuracy [9].

This review aims to provide a comprehensive analysis of AI-based methodologies for stock market price trend prediction by evaluating existing models, datasets, and performance metrics and identifying research gaps and future opportunities.

The remainder of this paper is organized as follows. Section 2 highlights the significance of the topic and Section 3 describes the problem statement. Section 4 outlines the objectives and Section 5 provides a comprehensive literature review. Section 6 details the taxonomy employed and Section 7 discusses the datasets used. Section 8 presents a comparative analysis of various algorithms, and Section 9 identifies gaps and challenges. Section 10 explores future research directions and Section 11 concludes the paper.

## 2. Significance of the topic

The study of stock market price trend prediction holds immense importance because of its profound impact on investment strategies, risk management, and economic decision making. The integration of AI with financial analytics has transformed the traditional forecasting methods. The significance of this study is analyzed through its economic, technological, and practical implications.

1) Economic Impact: Stock market predictions influence investment decisions at the individual and institutional levels, shaping the global financial ecosystem.

- Investment Decisions: Accurate predictions empower investors to make informed decisions and enhance wealth creation. Institutional investors use predictive models to optimize portfolio allocation.

- Economic Stability: Stock trends indicate economic performance. A bullish market signals investor confidence and stimulates economic growth through corporate investment [10].

- Global Financial Impact: The adoption of AI in financial forecasting is growing rapidly. The financial services sector invested \$35 billion in AI projects in 2023, with the global AI in finance market projected to grow at a CAGR of 30.6 %, reaching \$190 billion [10] by 2030. Even marginal improvements in the prediction accuracy can result in significant financial gains.

2) Technological Advancements: AI integration has transformed stock market prediction beyond traditional forecasting models.

- Revolutionizing Traditional Models: Traditional models, such as ARIMA and GARCH, cannot handle nonlinear relationships and large datasets effectively. AI models, including LSTM networks, transformer architectures, and hybrid approaches, capture complex financial patterns and temporal dependencies in the financial data.

- Enhanced Predictive Accuracy: AI models show superior accuracy compared with traditional methods. Techniques such as hierarchical frequency decomposition combined with deep learning have shown significant improvements in capturing complex financial patterns and improving prediction accuracy [11]. LSTMs achieve 72 % accuracy in price movement predictions, whereas CNN-based models achieve 90 % accuracy [10] for trend identification.

– Real-Time Analysis: AI-based forecasting enables faster and more accurate real-time decision making, which is critical in high-frequency trading. Transformer-based models reduce prediction latency by 57 % [10] and enhance trading efficiency.

3) Risk Management: Accurate stock market predictions help mitigate risks in volatile financial markets.

– Volatility Mitigation: Financial markets are inherently volatile; as demonstrated during the COVID-19 pandemic, the CBOE Volatility Index spiked by 43.2 % in 2020, highlighting significant market uncertainty. AI-driven models can help to anticipate market movements and implement hedging strategies to reduce losses.

– Portfolio optimization: AI models help fund managers dynamically adjust portfolios based on predicted market conditions. Hybrid models such as SeroFAM with Genetic Algorithms improve portfolio returns by optimizing trading strategies. This approach reduces exposure to high-risk assets during downturns while capitalizing on growth opportunities.

– Crisis Prevention: Early detection of potential market crashes through predictive analytics helps policymakers take preemptive measures to stabilize markets and protect investments.

4) Data utilization: AI-driven stock market prediction leverages historical and real-time data to enhance the accuracy.

– Large-Scale Historical Data: AI models use historical datasets to identify the trends and patterns often overlooked by traditional models.

– Technical Indicators: The inclusion of technical indicators such as moving averages, RSI, ADX, Bollinger Bands, CMO, and MACD in predictive models enhances their ability to analyze market behavior and improve forecasting accuracy.

– Social Media Sentiment Analysis: AI models, such as Stock Senti WordNet, analyze social media sentiment and incorporate sentiment data to improve the prediction accuracy by 18% for volatile stocks.

Multisource Data Integration: Hybrid GAN-LSTM or Transformer architectures combine diverse data sources, such as historical prices, macroeconomic indicators, and real-time news sentiment for accurate forecasts.

The significance of stock market price trend prediction lies in its ability to influence investment decisions, mitigate risks, and leverage technological advancements to revolutionize financial forecasting. AI models provide unparalleled predictive accuracy and real-time decision-making capabilities by utilizing large-scale historical data, technical indicators, and social media sentiments. As global adoption of AI in finance grows, this field continues to drive innovation, reshape financial markets, and contribute to economic prosperity. However, challenges such as data noise, ethical considerations, and model scalability must be addressed to ensure equitable and transparent application of AI in financial forecasting.

### 3. Problem statement

The problems in recent research analytics of stock market forecasting are defined and determined in this section.

1) Algorithm and hyperparameter selection: Choosing the right algorithm and tuning its hyperparameters are complex and often require domain expertise and extensive experimentation.

2) Black Box Nature of Models: Many deep learning models operate as “black boxes,” making it difficult for stakeholders to interpret how predictions are made. This lack of transparency reduces the trust in AI-driven systems.

3) Overfitting and generalization: Traditional and some ML models tend to overfit historical data and perform well on training sets but poorly on unseen or future datasets.

4) Feature Selection Complexity: Identifying the most predictive features (e.g., historical prices, volume, sentiment, and economic indicators) is challenging because of the dynamic and nonlinear nature of financial markets. Irrelevant or redundant features can degrade the model performance.

5) Regime Shifts and Adaptability: Financial markets are highly dynamic and undergo frequent regime shifts owing to economic crises, policy changes, or global events. Many existing models have struggled to adapt to these changes.

6) Data Quality Issues: Stock market datasets are often incomplete, noisy, or inconsistent. However, ML/DL models require high-quality and clean data for reliable prediction.

7) Nonstationarity of Financial Data: Stock prices exhibit changes in the mean and variance over time, making it difficult to model them using static assumptions.

8) Market volatility and nonlinearity: Stock prices are influenced by a wide range of factors, including macroeconomic, microeconomic, and geopolitical factors, which are often nonlinear, interdependent, and multifaceted patterns and trends, making modeling and prediction difficult.

9) Ethical and Regulatory Concerns: The use of AI in finance raises ethical concerns regarding fairness, accountability, and potential misuse (e.g., market manipulation). Regulatory frameworks that address these concerns have evolved.

#### 4. Literature review

This section provides a comprehensive review of the recent advancements in artificial intelligence (AI)-based methodologies for predicting stock market price trends. This discussion highlights the key approaches to address forecasting challenges, including machine learning (ML) and deep learning (DL) models and their applications. By examining state-of-the-art techniques, this review highlights their strengths, limitations, and contributions to improving prediction accuracy and decision making in dynamic financial markets.

Polamuri et al. [3] implemented a “Multi-Model Generative Adversarial Network Hybrid Prediction Algorithm (MMGAN-HPA)” to effectively forecast stock prices. This method fine-tunes the hyperparameters using Bayesian optimization with reinforcement learning. The proposed method combines Generative Adversarial Networks (GANs) with traditional machine learning models to address market volatility and data noise. The GAN-related approach, with its generator and discriminator, effectively handles time-series stock data to improve the predictions. GANs generate synthetic data for training, whereas hybridization with ARIMA and LSTM improves predictive accuracy. MMGAN-HPA outperformed the standalone models and traditional GANs by reducing the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). This method excels at handling nonlinear market patterns and provides robust predictions. Although this study demonstrates the scalability of the model across different stock markets, its limitations include reliance on hyperparameter tuning, and increased computational requirements. The authors conclude that MMGAN-HPA has significant potential for accurate stock market forecasting.

Bandhu et al. [12] introduced a real-time stock price prediction framework that integrates stream processing with deep learning to handle high-velocity financial data. The model uses Apache Kafka for stream processing and combines LSTM and GRU to capture the temporal dependencies of the stock prices. This study focuses on improving prediction accuracy while ensuring low-latency processing for real-time applications. A comparative analysis of the traditional models demonstrated superior performance in terms of the MAE and RMSE. The framework was tested on real-time stock data streams, thereby demonstrating its scalability to market conditions. This study emphasizes the importance of preprocessing techniques to enhance the model performance. The ability of the framework to fetch complex dependencies leads to an increased stock price forecasting accuracy. Despite its advantages, the model is limited by the vanishing gradient problem, which affects prediction accuracy. However, the computational complexity and infrastructure requirements are drawbacks. The study concludes that this approach is effective for real-time financial decision-making.

Gandhudi et al. [9] implemented an Explainable Hybrid quantum neural network (HQNN) framework to study the effects of social media sentiments on stock market price predictions. This method uses a hybrid genetic-based quantum neural network for the stock price analysis and data

fusion of transformers. The combination of LSTM neural networks and quantum computing principles helps identify complex patterns in the data for better predictions. FinBERT is used to classify tweets as positive, negative, or neutral sentiments in the HQNN model. The results show that tweet sentiment improves the stock price prediction accuracy by 18 %, especially for volatile stocks. SHAP values were used to interpret the model predictions and identify the key influencing factors. The HQNN outperforms classical neural networks in terms of speed and accuracy for large datasets. However, challenges remain in terms of quantum computing hardware scalability and real-world application. This method uses the average sentiment scores of tweets for the data combinations. The study concludes that HQNNs are promising for integrating social media data into financial forecasting, while maintaining the transparency of the model.

Gülmez [13] introduced an improved deep Long Short-Term Memory (LSTM) network using the Artificial Rabbits Optimization Algorithm (ARO) model (LSTM-ARO). The ARO algorithm optimizes key LSTM parameters, such as the learning rate, layers, and neurons, to improve prediction accuracy and reduce overfitting. The hidden layers evaluate the data, and the output is calculated using the current inputs. The hidden state and cell input help forecast stock prices, with ARO using two steps for less error forecasting than traditional models. The built-in LSTM parameters help the ARO find the best architecture for prediction. The model demonstrated superior performance on multiple stock datasets compared to traditional LSTM and other optimization models, such as particle swarm optimization (PSO) and Genetic Algorithms (GA), achieving a lower Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE). The study shows ARO's robustness of ARO in navigating complex optimization landscapes and handling nonlinear patterns in stock data. However, the deep LSTM model with optimization is computationally expensive, and ARO exhibits slow convergence and an inefficient search. This hybrid approach is promising for advanced financial forecasting applications.

Chandola et al. [5] introduced a framework for predicting stock price directional movements using deep-learning models. The hybrid model combines Word2Vec for feature extraction and LSTM for temporal dependency. This study uses historical stock price data and technical indicators such as the RSI, MACD, and moving averages. The model achieves high accuracy in directional forecasting, outperforming traditional machine learning models such as Support Vector Machine (SVM) and Random Forests. The authors addressed class imbalance by using SMOTE. Limitations include computational intensity and potential overfitting, with multisource and multifrequency factors complicating neural-network management. This framework is effective for directional stock-price predictions.

Sangeetha et al. [14] presented an Evaluated Linear Regression Based Machine Learning (ELR-ML) algorithm for forecasting stock market trends. The method uses Ordinary Least Squares optimization to minimize prediction variability. This study incorporates historical price data and technical indicators such as moving averages and Bollinger Bands. ELR achieved competitive results compared to Decision Trees and Support Vector Machines, particularly in short-term predictions. The simplicity and interpretability of the model render it suitable for applications that require limited computational resources. The performance metrics showed consistent results across the datasets, although their linear nature limited their ability to capture the complex relationships. The study concludes that the ELR is effective for quick financial forecasting, especially for small-scale investors.

Albahli et al. [8] implemented Stock Senti WordNet (SSWN), which adds sentiment analysis to stocks to help traders make decisions. This framework uses real-time Twitter data and Natural Language Processing (NLP) to analyze sentiment from tweets, categorizing them as positive, negative, or neutral, generating sentiment scores, and integrating them with historical stock data to enhance prediction accuracy. These sentiment scores quantify market feelings about specific stocks, and can be used for news stories and earnings reports. The SSWN utilizes the Random Forest and SVM algorithms, and the results show that Twitter sentiment improves the model's predictive capability. Sentiment research can help minimize risk by avoiding stocks through dropping sentiments and diversifying assets based on market sentiment. The authors emphasized

preprocessing steps to improve the data quality. These limitations include data scope, sentiment accuracy, context sensitivity, and integration challenges, particularly because of the dynamic nature of social media. Despite these limitations, the model shows a strong potential for short-term stock prediction, concluding that real-time social media sentiment analysis with machine learning provides valuable financial market insights.

Zhang and Chen [4] introduced a two-stage stock price prediction model combining Variational Mode Decomposition (VMD) with ensemble machine learning techniques. In the first stage, VMD decomposes historical stock price data into multiple intrinsic mode functions (IMFs), isolating frequency components and reducing noise. Support Vector Machine Regression (SVR) was used to find a function that minimises prediction errors within a tolerance margin while predicting the end values of the input features. SVR uses kernel functions to handle the nonlinear relationships between features and objective values. In the second stage, the IMFs serve as inputs to an ensemble learning model comprising Gradient Boosting Decision Trees (GBDT), random forests (RF), and Extreme Gradient Boosting (XGBoost) for predictions. An Extreme Learning Machine (ELM) generates excellent generalizations by improving the base-learning speed. The authors showed that this approach improves the prediction accuracy and stability compared to single-model methods, evaluated using the Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE). This study highlights VMD's importance of VMD in addressing nonstationary and noisy stock market data. The ensemble model captured both the linear and nonlinear patterns. However, SVR faces challenges in hyperparameter tuning, computational complexity, and kernel selection, whereas ELM has limited flexibility and the risk of overfitting. The complexity of the model may limit its scalability. The results confirm that the proposed method outperforms baseline models, offering a robust framework for stock price forecasting.

Diqi et al. [7] implemented the StockGAN framework, a Generative Adversarial Networks (GAN) method, to provide the best accuracy in stock market forecasting. The model uses GAN's ability to learn complex data distributions through a generator to simulate stock price patterns and a discriminator to distinguish between actual and generated data. The generators and discriminators were trained using argumentative learning methods to improve the model and to defend against overfitting. Historical stock prices from major markets were used as inputs for preprocessing to improve data quality. The authors compared StockGAN with traditional models such as ARIMA and LSTM and demonstrated its superior accuracy in capturing nonlinear relationships in financial data. Evaluation metrics, including Mean Squared Error (MSE) and R-squared ( $R^2$ ), confirmed its effectiveness. GAN offer advantages such as generating realistic data, enhancing training with synthetic data, and modeling complex distributions. Although GANs require careful hyperparameter tuning to avoid mode collapse, this study demonstrates the model's robustness against market volatility. The model has limitations including high computational demands and training instability. The study concludes that StockGAN's dynamic training process and its effectiveness in unsupervised learning make it a powerful tool for machine learning and financial forecasting in volatile markets.

Leon Lai Xiang Yeo a et al. [15] introduced dynamic portfolio rebalancing with lag-optimized trading indicators using SeroFAM and genetic algorithms. This study explores portfolio rebalancing by integrating Self-Regulating Fuzzy Associative Memory (SeroFAM) with genetic algorithms (GAs) to optimize trading indicators. The SeroFAM model adaptively learns and adjusts to market dynamics by addressing the lag in the traditional trading indicators. Genetic algorithms were used to fine-tune the indicator parameters for optimal performance under diverse market conditions. The methodology tested on financial datasets demonstrates significant improvements in portfolio returns compared to static rebalancing strategies. Key metrics such as the Sharpe Ratio and Cumulative Returns highlight the effectiveness of this hybrid approach in managing risk while maximizing profitability. The authors emphasize the interpretability of SeroFAM for market trends and decision making. However, they noted the computational challenges of GA optimization and scalability for larger portfolios. This study contributes to the development of adaptive AI-based portfolio-management strategies.

T. Swathi et al. introduced Teaching and Learning Based Optimization model with Long Short-Term Memory (TLBO-LSTM), a hybrid model combining Teaching-Learning-Based Optimization (TLBO) and Long Short-Term Memory (LSTM) networks for stock price prediction using Twitter sentiment analysis. The model integrates sentiment scores from tweets with historical stock price data to improve the prediction accuracy. TLBO optimizes LSTM key hyperparameters, such as the learning rate and the number of layers and hidden units, by leveraging TLBO's global optimization capabilities, eliminating manual tuning, and improving generalization for nonlinear and nonstationary financial datasets. The model processes historical stock prices, technical indicators, and sentiment scores through the teaching phase (global search for optimal hyperparameters) and the learning phase (local refinement). The optimized LSTM captured the temporal dependencies in the sequential data. The LSTM network was optimized using a metaheuristic algorithm Adam optimizer to fine-tune the hyperparameters, reduce the training time, and improve computational efficiency. The optimizer is well suited for large datasets, reducing the training time and enhancing the convergence rate, while improving the computational efficiency. Past gradients have been incorporated to achieve these improvements. The experimental results showed superior performance over ARIMA and standalone LSTMs with higher accuracy and lower error rates. However, TLBO-LSTM has limitations including computational complexity, parameter sensitivity, overfitting risk, and challenges with noisy tweets. This study demonstrates the importance of social media sentiment in market movements, and the potential of combining deep learning with alternative data sources for financial forecasting. A Generative Adversarial Network with Long Short-Term Memory (GAN-LSTM) combines GANs with LSTMs to address data imbalance by generating synthetic datasets for training to enhance the robustness against noisy data, thereby offering complementary solutions for stock market prediction, whereas TLBO-LSTM focuses on hyperparameter optimization.

Chen et al. [6] implemented a K-means-LSTM hybrid model to forecast the stock prices of Chinese commercial banks. The approach begins with K-means clustering, which groups stocks with similar price patterns, thereby reducing data complexity and improving the model focus. These clustered datasets were fed into a Long Short-Term Memory (LSTM) network, which captured the temporal dependencies and nonlinear trends in stock prices. This model can be used for static and dynamic stock price forecasting of a wide range of equities with different tendencies. By clustering stocks with similar price movements, tailored forecasting using cluster-specific models improves prediction accuracy by training the LSTM on more homogeneous data. By concentrating on pertinent patterns inside each cluster, this method enables LSTM models to accurately represent intricate temporal relationships in the stock values. The hybrid model was benchmarked against standalone LSTM and traditional machine learning methods, showing improved prediction accuracy measured using the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The authors highlighted the model's scalability for large datasets but noted challenges in selecting optimal K-means parameters, determining the number of clusters, and potential LSTM overfitting. The study concludes that the K-means-LSTM hybrid method is promising for stock price forecasting, particularly in sector-specific scenarios, such as banking.

Yanli Zhao Guang Yang et al. [16] introduced hybrid Self-Attention Deep Long Short-Term Memory (SA-DLSTM) to predict stock-price movements. SA-DLSTM is a deep learning-based framework that combines multiple models to improve the prediction accuracy. The framework integrates Convolutional Neural Networks (CNNs) for feature extraction, Long Short-Term Memory (LSTM) networks for temporal dependencies, and neural networks for final predictions. The model uses historical stock data and external factors such as market indices and economic indicators to create a comprehensive dataset. After applying self-attention to the input sequences, the attention weights highlight the relevance of each time step, enabling focus on the key features. The weighted representations were processed through LSTM layers to capture temporal relationships and patterns in stock prices. The integration of self-attention and deep LSTM layers improved the capacity of the model to analyze and predict intricate time-based stock price trends. The results show that the framework outperforms standalone models, such as CNN, LSTM, and

traditional approaches, in terms of accuracy and F1-score. This framework can handle complex nonlinear relationships in stock data, and is robust to noise. Its modularity allows additional data sources and components. However, it faces challenges in terms of computational complexity when combining self-attention with deep LSTM networks, and requires large datasets to avoid overfitting.

Hong et al. [10] introduced SpectralGPT, a Spectral Remote Sensing Foundation Model, a transformer-based model following generative pretraining principles for spectral remote sensing data that aligns with AI-based stock market prediction. SpectralGPT's ability to capture sequential patterns, perform multi-target reconstruction, and handle high-dimensional data addresses financial forecasting challenges, including the integration of multiple variables (e.g., historical prices, technical indicators, and external factors). Its pretraining on large datasets parallels financial AI models that leverage historical market data to enhance the predictive accuracy. The processing of spatial, spectral, and temporal data using SpectralGPT suggests its adaptability to financial datasets. The attention mechanisms of the model can focus on key market indicators, similar to the prioritization of spectral features in remote sensing tasks. Spectral decomposition reduces market volatility noise and generative scenario simulation aids risk management. Its ability to handle diverse inputs demonstrates cross-domain applicability. UrbanSAM, designed for urban segmentation, offers insights into financial forecasting through multiscale contextual analysis and cross-attention mechanisms, enabling the alignment of diverse data sources while preserving temporal-spatial fidelity. Its spatiotemporal correlations analyze market influences and its dynamic adaptation suits nonstationary financial time series and risk assessments. UrbanSAM integrates real-time data for short-term market predictions and enables generalization across heterogeneous datasets. Although neither model directly addresses stock market prediction, its methodologies for handling noise, complex data, and sequential dependencies have inspired advancements in financial forecasting.

Etaf Alshawarbeh et al. [17] introduced a hybrid Autoregressive Integrated Moving Average Artificial Neural Network (ARIMA -ANN) for analyzing and predicting high frequency datasets, with a focus on financial time series. The Auto-Regressive Integrated Moving Average (ARIMA) model handles the linear components of the data, whereas an Artificial Neural Network (ANN) addresses nonlinear relationships, thereby creating a complementary framework. This hybrid model overcomes the limitations of the standalone ARIMA, such as its inability to capture complex nonlinear patterns, and the standalone ANN, which struggles to handle seasonality and trends. The methodology was tested on high-frequency financial datasets and achieved significant improvements in accuracy metrics, such as the Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE), compared to the individual models. This study highlights the importance of preprocessing steps, such as stationarity checks and data normalization, for hybrid model performance. Although the results were promising, the authors noted challenges in terms of the computational efficiency and the need for careful hyperparameter tuning. This study highlights the potential of hybrid statistical AI models for precise forecasting of high-frequency financial environments.

Tian and Pei [18] implemented the Echo State Network - Sparrow Search Algorithm (SSA-ESN) to predict stock market prices. This method uses Echo State Networks (ESNs) optimized by the Sparrow Search Algorithm (SSA) for hyperparameter optimization. The ESN is a reservoir computing model known for efficiently processing time-series data, owing to its lightweight architecture and dynamic memory capabilities. Traditional ESNs often have suboptimal hyperparameter settings, such as reservoir size and spectral radius, which are addressed by SSA. SSA avoids local optima and searches for large spaces, thereby enabling an effective ESN parameter optimization. This combination helps identify trends in stock prices, with SSA improving the data quality for ESN modeling. The authors showed that the SSA-optimized ESN outperformed standard ESNs and other models (ARIMA and LSTM) in terms of prediction accuracy using metrics such as the Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE). The model demonstrated the ability to capture nonlinear patterns while



maintaining the computational efficiency. Despite its advantages, this method faces challenges in terms of the computational complexity and parameter tuning. This study demonstrated the potential of combining evolutionary algorithms with reservoir computing for financial forecasting.

Lu and Xu [19] introduced a Time-series Recurrent Neural Network (TRNN) as a novel architecture to improve the efficiency and accuracy of stock price predictions. The TRNN leverages Long Short-Term Memory (LSTM) units to capture short-term fluctuations and long-term dependencies in financial time-series data. The model uncovers temporal patterns and adapts them to dynamic market conditions by integrating the recurrence mechanisms. Time-series decomposition was employed to handle nonstationary data by dividing them into trend, seasonal, and residual components for the analysis. Optimization techniques were incorporated to reduce the computational overhead while maintaining accuracy. The experimental results show that the TRNN outperforms traditional Recurrent Neural Networks (RNN), LSTMs, and GRUs using metrics such as the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). However, this model also has several limitations. Its high computational demands make it resource-intensive for real-time trading and less scalable for large datasets. TRNNs are prone to overfitting with poorly generalized training data and require careful regularization. Furthermore, hyperparameter tuning is time consuming, which may hinder its practical implementation. Despite these challenges, this study highlights TRNN's potential of TRNNs to advance deep learning models for financial forecasting.

## 5. Taxonomy

The taxonomy of artificial intelligence (AI)-based stock market price trend prediction models provides a structured framework for categorizing methodologies based on their algorithms, data dependencies, and application contexts. This section outlines the classification of these models into distinct categories, and highlights their strengths, limitations, and data dependencies.

Fig. 1 shows a structured taxonomy based on general advancements in AI in financial markets.

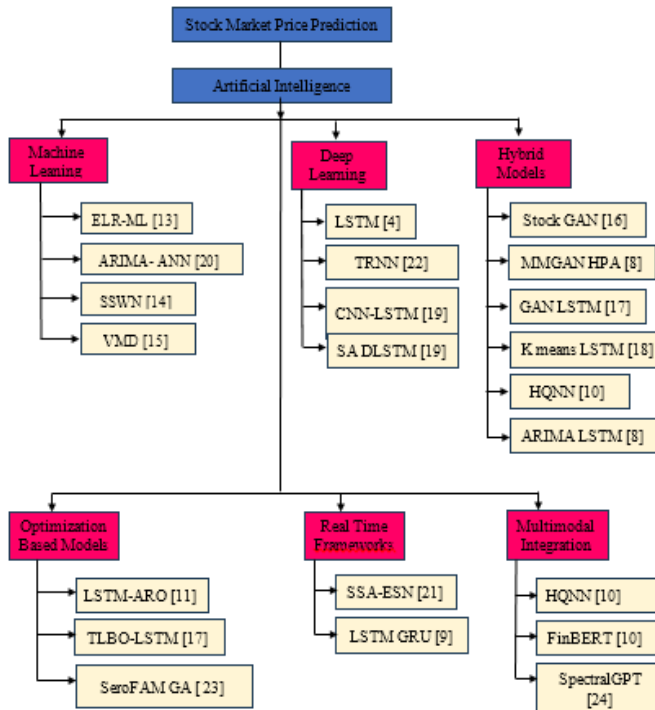


Fig. 1. Represents the taxonomy of stock market

## 5.1. Machine learning models

Machine learning (ML) models are fundamental to stock market prediction because of their simplicity and interpretability. These models rely on statistical techniques and structured datasets such as historical price data and technical indicators, and require extensive feature engineering to achieve effective predictions.

- Linear Regression-Based Models: enhance traditional regression through feature selection and parameter optimization.

Example 1: Evaluated Linear Regression (ELR-ML) [14] enhances traditional regression by incorporating feature selection and parameter optimization.

Example 2: ARIMA-ANN hybrid model integrates ARIMA for linear trends with ANN for nonlinear relationships [17].

Strengths: Simple, interpretable, and suitable for resource-limited applications.

Limitations: The limited ability to capture nonlinear relationships in stock data. Computational Efficiency and Hyperparameter Tuning.

Dataset Dependency: Structured data, such as OHLC prices and technical indicators (e.g., moving averages and Bollinger Bands).

Code Availability: The Python repository for ARIMA is implemented in “statsmodels” and ANN models can be built using TensorFlow or PyTorch.

- Support Vector Machines (SVM): effective for classification and regression.

Example 1: Stock Senti WordNet (SSWN) integrates SVM with sentiment analysis for short-term predictions [8].

Example 2: SVR combined with Variational Mode Decomposition (VMD) for denoising financial data [11].

Strengths: Handles nonlinear patterns effectively.

Limitations: Computationally intensive during hyperparameter tuning, context sensitivity.

Dataset Dependency: Sentiment data (e.g., Twitter and news) combined with historical price data.

Code Availability: Python code for SVM is implemented in scikit-learn, while sentiment analysis can be performed using NLTK or TextBlob.

## 5.2. Deep learning models

Deep learning (DL) models capture nonlinear dependencies and temporal patterns in financial data. These models require large-scale datasets, and are computationally intensive.

- Recurrent Neural Networks (RNNs).

Example 1: Long Short-Term Memory (LSTM) networks optimized with ARO for hyperparameters are widely used for sequential data because of their ability to model temporal dependencies [13].

Example 2: Time-series Recurrent Neural Network (TRNN) for capturing short- and long-term dependencies [19].

Strengths: Effective modeling of temporal dependencies in time-series forecasting [20].

Limitations: Risk of overfitting without proper regularization; computationally expensive.

Dataset Dependency: Large-scale historical price data with temporal features.

Example 3: Echo State Network with Sparrow Search Algorithm (SSA-ESN): Uses SSA for hyperparameter optimization in ESNs [18].

Strengths: Efficient for nonlinear time series data.

Limitations: Sensitive to initial conditions, parameter tuning challenges.

Dataset Dependency: historical price data with temporal features, domain-specific data, and optimized hyperparameters.

Code Availability: ESN and SSA libraries available in Python under the MIT license and their integration can be performed using pyESN or EsnTorch.

– Convolutional Neural Networks (CNNs): CNNs is effective for feature extraction from time series data.

Example 1: CNN-LSTM hybrid model for feature extraction and temporal modeling [21].

Example 2: CNN-based feature extraction for integrating technical indicators such as RSI and MACD [6].

Strengths: High accuracy in directional forecasting and robust feature extraction.

Limitations: Prone to overfitting; high computational costs.

Dataset Dependency: Historical prices and technical indicators.

Code availability: CNN-LSTM is available under the MIT license and has a well-documented repository.

– Self-Attention Mechanisms.

Example 1: Self-Attention Deep LSTM (SA-DLSTM) focuses on the key temporal features. CNNs have been combined for feature extraction, self-attention for sequence weighting, and long short-term memory (LSTM) for temporal patterns [16].

Example 2: Transformer-based models such as SpectralGPT have been adapted for high-dimensional financial data [10].

Strengths: Modular frameworks; handle noisy and nonlinear data effectively.

Limitations: Require large datasets; computationally intensive.

Code Availability: Transformer models are implemented using the Hugging Face library.

### 5.3. Hybrid models

Hybrid models combine the strengths of multiple methodologies to enhance the predictive accuracy, robustness, and generalizability. They integrate diverse datasets such as sentiment data, historical prices, and event-driven data.

– Generative Adversarial Networks (GANs).

Example 1: StockGAN generates synthetic datasets to address issues such as data sparsity, as highlighted in the dataset comparison section, and to improve the robustness in volatile markets [7].

Example 2: MMGAN-HPA combines GANs with ARIMA and LSTM to handle noisy and nonlinear data. Focus on synthetic data generation and hyperparameter tuning via Bayesian optimization and reinforcement learning [3].

Example 3: GAN-LSTM: Combines GANs for data augmentation and LSTMs for temporal dependency modeling. Improves robustness against noisy data.

Strengths: Superior performance in handling noisy datasets effective for data augmentation.

Limitations: Risk of mode collapse; computationally expensive, training instability.

Dataset Dependency: GAN-generated synthetic data and historical price data.

Code availability: The GANs implementations are available in TensorFlow/Keras and PyTorch.

– K-means-LSTM: Clustering-enhanced methods, such as K-means integrated with deep learning models, improve predictive accuracy by identifying similar patterns before applying temporal modeling techniques [22].

Example: Combines K-means clustering for preprocessing with LSTMs for tailored forecasting [6].

Strengths: Improves accuracy by clustering similar patterns before prediction.

Limitations: Challenges in determining optimal cluster numbers; risk of overfitting.

Dataset Dependency: Historical prices and clustered patterns.

Code availability: K-means is available in scikit-learn and LSTM is implemented in TensorFlow/Keras and PyTorch.

– Explainable Hybrid Quantum Neural Networks (HQNN).

Example: Integrate quantum computing principles with LSTM and FinBERT for explainability and sentiment analysis [9].

**Strengths:** This enhances interpretability using SHAP values, and improves prediction accuracy by incorporating investor sentiment as a key feature.

**Limitations:** Scalability challenges due to quantum hardware constraints.

**Dataset Dependency:** Sentiment data and structured price data.

**Code Availability:** HQNN is available under custom License.

– ARIMA-LSTM.

**Example:** ARIMA-LSTM combines traditional statistical methods with deep learning to address both linear and nonlinear dependencies in stock price trends [3].

**Strengths:** Captures both linear trends and complex patterns effectively.

**Limitations:** Computational complexity due to hybrid architecture.

**Dataset Dependency:** Historical prices

#### 5.4. Optimization-based models

Optimization-based models enhance traditional algorithms by fine-tuning parameters to achieve better accuracy and reduce overfitting.

– LSTM-ARO (Artificial Rabbits Optimization).

**Example:** ARO was used to optimize LSTM hyperparameters [13].

**Strengths:** It reduces overfitting and outperforms traditional optimization methods, such as PSO and GA.

**Limitations:** Slow convergence during optimization; high computational cost.

**Dataset Dependency:** Historical prices and parameter-optimized datasets.

**Code Availability:** LSTM is implemented in TensorFlow/Keras and PyTorch. The ARO library is available under MIT license.

– TLBO-LSTM (Teaching-Learning-Based Optimization).

**Example:** Combines TLBO with LSTMs for sentiment-integrated stock price prediction.

**Strengths:** It handles nonlinear trends effectively and reduces RMSE/MAPE values.

**Limitations:** High computational requirements; risk of overfitting without proper regulation.

**Dataset Dependency:** Sentiment data and historical prices.

**Code Availability:** No prebuilt libraries are available for TLBO-LSTM.

– SeroFAM with Genetic Algorithms (GA): (Genetic algorithm-based optimization).

**Example:** Uses Self-Regulating Fuzzy Associative Memory and GA for portfolio rebalancing [15].

**Strengths:** Improves Sharpe Ratio, interpretable for decision-making.

**Limitations:** Computational challenges with GA.

**Dataset Dependency:** Historical prices, technical indicators and Portfolio Performance Metrix

**Code availability:** SeroFAM GA is implemented in MATLAB and is available under a custom license.

#### 5.5. Real-time prediction frameworks

Real-time adaptability is critical for high-frequency trading applications, where immediate decision-making is required.

– Stream Processing Frameworks.

**Example:** Apache Kafka integrated with hybrid deep learning models like LSTM-GRU [12].

**Strengths:** Scalable architecture suitable for dynamic market conditions. Low latency.

**Limitations:** Infrastructure complexity; vanishing gradient problems in GRU/LSTMs.

**Dataset Dependency:** Real-time stock price streams.

**Code availability:** Apache Kafka has Python clients such as “confluent-kafka” for stream processing

– Echo State Networks Optimized by Sparrow Search Algorithm (SSA-ESN).

**Example:** Combines lightweight reservoir computing with SSA optimization [18].

Strengths: Efficient memory usage and fast computation.

Limitations: Sensitive to initial conditions and parameter tuning challenges.

Dataset Dependency: High-frequency real-time data.

## 5.6. Multimodal integration models

These models integrate structured numerical data with unstructured textual sources (e.g., news or social media sentiments) to improve predictive accuracy.

- Sentiment Analysis Integration.

Example: HQNN incorporates FinBERT sentiment analysis into quantum neural networks [9].

Strengths: Captures investor sentiment effectively.

Limitations: Scalability issues due to quantum computing infrastructure.

Dataset Dependency: Sentiment data and historical prices.

Code Availability: HQNN is available under custom License.

- Transformer-Based Models.

Example: SpectralGPT leverages transformer architectures to process high-dimensional multimodal data [10]. The Transformer models are implemented using the Hugging Face library.

Strengths: Handles diverse data sources; reduces noise.

Limitations: High computational costs; requires advanced preprocessing.

Dataset Dependency: Historical prices, technical indicators, tweets, and news articles.

Code Availability: Transformer models are implemented using the Hugging Face library

This taxonomy highlights the diversity of the AI methodologies employed to address the complexities of forecasting in financial markets. Each category addresses specific challenges such as noise reduction, real-time adaptability, and multimodal integration, while presenting the unique strengths and limitations of the methods. Machine learning models offer simplicity and interpretability but lack scalability for complex datasets. Deep learning architectures excel at capturing nonlinear patterns but face challenges related to their computational cost and explainability. Hybrid approaches effectively combine the strengths of different methodologies but require significant computational resources, whereas Optimization-based Models enhance robustness but require faster convergence techniques. Real-time frameworks ensure adaptability to dynamic markets but require robust infrastructure. Finally, multimodal integration models demonstrate the potential of combining diverse data sources to enhance predictive accuracy but require advanced preprocessing and scalability improvements. This taxonomy organizes methodologies into clear categories and highlights their strengths, limitations, and dependencies on the datasets. Future research should focus on refining these approaches by incorporating advanced optimization techniques, leveraging emerging technologies, improving explainability, and addressing ethical considerations in financial AI systems to improve the efficiency of financial forecasting systems.

## 6. Dataset comparison

Datasets are fundamental to the development of AI-based stock market price trend prediction models, significantly influencing their accuracy, robustness, and generalizability. This section provides a comparative analysis of datasets used in recent studies, focusing on their characteristics, applications, and associated challenges.

### 6.1. Types of datasets used in stock market prediction

The datasets employed in stock market predictions are broadly categorized, as listed in Table 1.

**Table 1.** Dataset types

Dataset type	Examples	Key features	Applications	Strengths	Challenges
Historical price data	S&P 500, NASDAQ, Dow Jones	It Includes Open, High, Low, Close (OHLC) prices and trading volumes	Time-series forecasting and trend analysis	It is widely available; essential for trend analysis	Nonstationarity, noise, and a lack of contextual information
Technical indicators	RSI, Bollinger Bands, CMO, ADX, ATR%, MACD, Moving Averages	Derived features summarizing price trends and momentum	Feature engineering for machine learning models	Summarizes key patterns; reduces noise	This May oversimplify complex relationships; dimensionality issues
Macroeconomic data	GDP growth, inflation rates, Interest rates	Provides broader economic context influencing market trends	Incorporating external economic factors into prediction models	Provides broader economic context influencing market trends	Integration with market-specific features and temporal alignment
Sentiment data	Twitter sentiment, news data	Textual data reflecting public or investor sentiment about specific stocks or markets	Short-term market sentiment analysis and prediction	Reflects real-time public/investor opinions.	Noisy and unstructured data; requires sophisticated preprocessing and NLP techniques
Event-driven data	Earnings reports, M&A data, general election results data	Focuses on specific events that can cause price fluctuations	Event-based prediction and anomaly detection	Highly relevant for short-term predictions	Sparse and irregular occurrences; high dependency on event labeling and timing
Synthetic data	GAN-generated data	Artificially generated data that mimics real-world stock price patterns	Augmenting datasets to overcome noise, sparsity, or imbalance	<ul style="list-style-type: none"> <li>– Addresses data sparsity or imbalance issues</li> <li>– Can improve model robustness during training</li> </ul>	Requires careful tuning of GANs to avoid mode collapse and maintain data realism
Multi-Temporary Stock Data (MTS)	Daily/weekly/monthly stock sequences (e.g., sequence modeling in deep learning architectures like RCA-BiLSTM models)	It captures multi-scale temporal dependencies (daily, weekly, monthly)	Sequence modeling in deep learning architectures like RCA-BiLSTM or DQN strategies for stock prediction	<ul style="list-style-type: none"> <li>– Captures both short-term volatility and long-term trends</li> <li>– It Enhances feature extraction through reverse cross-attention mechanisms</li> </ul>	<ul style="list-style-type: none"> <li>– High computational cost</li> <li>– Requires extensive preprocessing to effectively align temporal scales</li> </ul>

## 6.2. Comparative analysis of datasets

A comparative analysis of the data sources along with their key features, applications, and associated challenges is presented in Table 2.

**Table 2.** Comparative analysis of datasets

Author	Dataset source	Key features	Applications	Challenges
Subba Rao Polamuri et al. [8]	Historical prices with GAN-generated synthetic data	It Combines synthetic and real-world data for robust training; integrates GAN with ARIMA and LSTM	Predicts nonlinear and complex stock patterns using MMGAN-HPA	It relies on the quality of synthetic data and the computationally intensive training process
Kailash Chandra Bandhu et al. [12]	Real-time stock price streams (Apache Kafka)	Preprocessed high-frequency data with normalization and outlier handling	Real-time prediction for financial decision-making using LSTM and GRU	Computational complexity and vanishing gradient problems in LSTM/GRU models
Manoranjani Gandhudi et al. [9]	Twitter sentiment data + historical stock prices	Integrates sentiment analysis using FinBERT and SHAP values for explainable predictions	Incorporates social media sentiment into stock price prediction for volatile stocks	Context sensitivity of sentiment data; scalability issues with quantum computing hardware
Burak Gülmez [13]	Historical stock prices	Optimizes LSTM hyperparameters using Artificial Rabbits Optimization (ARO)	Captures long-term dependencies in time-series data for improved accuracy	High computational cost due to deep LSTM architecture and optimization process
Deeksha Chandola et al. [5]	Historical stock prices + technical indicators (RSI, MACD)	Addresses class imbalance using SMOTE; integrates CNN for feature extraction and LSTM for temporal modeling	Predicts directional movement of stock prices with high accuracy	Computationally intensive; potential for overfitting due to reliance on deep learning models
Saleh Albahli et al. [8]	Real-time Twitter data + historical stock prices	Uses NLP techniques to generate sentiment scores integrated with Random Forest and SVM	Improves predictive capability by incorporating market sentiment	Sentiment scope and accuracy; dynamic nature of social media introduces unpredictability
Jun Zhang & Xuedong Chen [4]	Historical stock prices	Variational Mode Decomposition (VMD) to decompose data into intrinsic mode functions, combined with ensemble learning models like Gradient Boosting Decision Trees (GBDT) and Random Forest	Improved prediction accuracy and stability by addressing nonstationary and noisy data	Computational complexity of ensemble models and challenges in hyperparameter tuning
Mohammad Digi et al. [7]	Historical stock prices	Utilizes StockGAN to generate synthetic data and train models for better prediction accuracy	Effective for handling market volatility and generalizing across datasets	GANs require careful tuning to avoid issues like mode collapse
Yufeng Chen et al. [6]	Historical stock prices (Chinese commercial bank stock)	Combines K-means clustering for grouping similar stock price patterns with LSTM for temporal dependency modeling	Sector-specific stock price prediction (e.g., banking sector)	Challenges in determining the optimal number of clusters and risk of overfitting in LSTM models

Yanli Zhao et al. [16]	Historical stock prices + external factors (market indices, economic indicators)	Combines CNN for feature extraction, LSTM for temporal dependencies, and self-attention mechanisms for relevance weighting	Robust prediction of intricate stock price movements with improved accuracy and F1-score	High computational complexity and resource demands for combining self-attention with deep LSTM networks
Danfeng Hong et al. [10]	Spectral remote sensing data (conceptual alignment with financial data)	Uses SpectralGPT, a transformer-based model, to handle sequential patterns, high-dimensional data, and diverse variables	Conceptually adaptable to financial data for integrating diverse data sources like historical prices, sentiment analysis, and technical indicators	Not originally designed for financial data, requiring significant adaptation for stock market prediction

A comparative analysis of datasets reveals that while historical price data remain the most commonly used, the integration of alternative data sources, such as sentiment analysis and event-driven data, has shown great potential for improving predictive accuracy. However, challenges such as data quality, integration complexity, and real-time processing persist, necessitating further advancements in data preparation and integration techniques to address these issues.

## 7. Comparative analysis

Artificial Intelligence (AI) techniques, including Machine Learning (ML) and Deep Learning (DL), have significantly advanced stock market price trend predictions. This section presents a comparative evaluation of various AI-based methodologies, focusing on their models, datasets, performance metrics, and limitations. The analysis highlights how these approaches address challenges such as market volatility, data noise, and real-time processing while identifying gaps for future research.

### 7.1. Hybrid models for enhanced prediction

Hybrid models combine multiple algorithms to leverage their strengths and improve prediction accuracy. These approaches often integrate Generative Adversarial Networks (GANs), Long Short-Term Memory (LSTM), and optimization techniques.

A comparative analysis of the hybrid methodologies, strengths, limitations, and performance criteria is presented in Table 3.

These hybrid approaches demonstrate the potential of combining traditional statistical models with deep-learning techniques to address challenges such as noise, volatility, and nonlinearity.

### 7.2. Real-time prediction frameworks

Real-time stock price prediction requires models that are capable of processing high-velocity data streams while maintaining low latency.

A comparative analysis of Real-Time Prediction methodologies, strengths, limitations, and performance criteria is presented in Table 4.

These frameworks emphasize the importance of preprocessing techniques, such as normalization and outlier removal, while leveraging tools, such as Apache Kafka, to handle real-time data effectively.



**Table 3.** Comparative analysis of hybrid models

Author	Methodology	Strengths	Limitations	Performance criteria
Subba Rao Polamuri et al. [3]	MMGAN-HPA: Combines GANs with ARIMA and LSTM; fine-tuned using Bayesian optimization and reinforcement learning	It Handles nonlinear market patterns; reduces Mean Absolute Error (MAE) and Root Mean Square Error (RMSE)	Training instability, reliance on hyperparameter tuning, high computational requirements	MAE, RMSE, Correlation
Burak Gülmez [13]	LSTM-ARO: Uses Artificial Rabbits Optimization (ARO) to optimize LSTM hyperparameters.	It has a Superior accuracy in capturing long-term dependencies; reduces overfitting	High computational cost; slow convergence of the ARO algorithm	MSE, RMSE
Jun Zhang & Xuedong Chen [4]	Two-stage model combining Variational Mode Decomposition (VMD) with ensemble machine learning techniques	It reduces noise in data; captures both linear and nonlinear patterns using ensemble models	Hyperparameter tuning complexity; computational demands of ensemble learning	- Mean Absolute Percentage Error (MAPE), (RMSE)
Mohammad Diqi et al. [7]	StockGAN: GAN-based model generating synthetic stock price data to improve training robustness	It is effective for volatile markets; generalizes well across datasets	Mode collapse in GANs and high computational requirements	MSE, R <sup>2</sup>
Leon Lai Xiang Yeo et al. [15]	SeroFAM with Genetic Algorithms	The MACD and RSI were combined to improve the accuracy of identifying trend reversals and overbought/oversold conditions	High dependency on historical data leads to lag and potential overfitting. SeroFAM is in MATLAB with custom license, not directly usable in Python	Accuracy
Deeksha Chandola et al. [5]	CNN-LSTM Hybrid Model	It combines CNN feature extraction with LSTM temporal modeling for high directional accuracy	It is computationally intensive and prone to overfitting without proper regularization	Accuracy

**Table 4.** Comparative analysis of real-time prediction frameworks

Author	Methodology	Strengths	Limitations	Performance criteria
Kailash Chandra Bandhu et al. [12]	Framework integrating Apache Kafka for stream processing with LSTM and GRU models	Real-time adaptability; superior performance in terms of MAE and RMSE	Computational complexity; vanishing gradient problem in deep learning models	MAE, RMSE
Manoranjana Gandhudi et al. [9]	Hybrid Quantum Neural Network (HQNN) with sentiment analysis for real-time predictions	Improves accuracy by incorporating social media sentiment; explainable using SHAP values	Scalability challenges with quantum computing hardware; reliance on average sentiment scores	SHAP

### 7.3. Sentiment analysis models

Sentiment analysis integrates alternative data sources, such as social media posts or financial news, to capture the influence of public opinion on market trends.

A comparative analysis of the Sentiment Analysis methodologies, strengths, limitations, and performance criteria is presented in Table 5.

**Table 5.** Comparative analysis of sentiment analysis models

Author	Methodology	Strengths	Limitations	Performance criteria
Manoranjana Gandhudi et al. [9]	Hybrid Quantum Neural Network (HQNN): Incorporates FinBERT sentiment analysis for Twitter data	It improves prediction accuracy for volatile stocks by up to 18 %	Limited contextual understanding of sentiment; scalability issues due to quantum computing constraints	Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared
Saleh Albahli et al. [8]	Stock Senti WordNet (SSWN): Combines Random Forest and SVM with real-time Twitter sentiment analysis	Enhances decision-making for short-term predictions; reduces risk by analyzing sentiment trends	Noisy and unstructured social media data; sentiment context sensitivity remains a challenge for accuracy improvements	MAPE, RMSE, R2

Sentiment-driven models enhance short-term predictions, but require sophisticated preprocessing pipelines to effectively handle noisy or unstructured input data.

### 7.4. Simple models for interpretability

Simple models prioritize interpretability over complexity, making them ideal for resource-limited applications and for small-scale investors.

A comparative analysis of the simple models, such as the linear regression-based ML and SeroFAM methodologies, strengths, limitations, and performance criteria, is presented in Table 6.

**Table 6.** Comparative analysis of simpler models

Author	Methodology	Strengths	Limitations	Performance criteria
J. Margaret Sangeetha & K. Joy Alfia [14]	Linear Regression-Based Machine Learning Approach for stock price forecasting.	Simple interpretable approach suitable for small-scale investors & institutions.	Limited ability to capture nonlinear relationships in stock datasets.	RMSE, MAE, MAPE, R <sup>2</sup>
Leon Lai Xiang Yeo et al. [15]	SeroFAM with Genetic Algorithms for dynamic portfolio rebalancing.	Improves portfolio returns; interpretable results for decision-making.	Computational challenges with genetic algorithm optimization; scalability issues for larger portfolios.	Historical Prices

This comparative analysis highlights the transformative potential of AI in stock market price trend prediction by using hybrid models, real-time frameworks, sentiment integration, and interpretable approaches. Although hybrid methods address nonlinearity and volatility, real-time frameworks enable a dynamic adaptability. Sentiment analysis enhances short-term forecasting

by incorporating alternative data sources. Simpler models remain valuable for resource-constrained applications but lack the sophistication needed for complex markets. Future research should focus on optimizing the computational efficiency, enhancing the model interpretability, and leveraging advanced architectures to effectively address existing challenges.

## 8. Gaps and challenges

The application of artificial intelligence (AI) to stock market price trend prediction has demonstrated significant advancements. However, several critical gaps and challenges limit their robustness, generalizability, and practical applicability. These challenges can be categorized into four main areas: data challenges, methodological limitations, interpretability issues, and ethical and regulatory concerns.

### 8.1. Data challenges

The noisy, dynamic, and nonstationary nature of financial data poses significant obstacles to the accurate prediction of stock market performance.

#### 1) Noisy and Nonstationary Data:

Financial time series are highly volatile and influenced by unpredictable events (e.g., geopolitical crises and economic policies), making it difficult for models to identify persistent patterns.

Example: Models such as MMGAN-HPA [3] attempt to mitigate noise using synthetic data, but face challenges in generalizing across markets.

Mitigation Strategies: Advanced preprocessing techniques (e.g., wavelet transforms and spectral decomposition) and hybrid models (e.g., ARIMA-LSTM) show promise in addressing these issues.

#### 2) Data Quality and Accessibility:

High frequency or alternative data sources (e.g., sentiment data from social media) are often expensive and difficult to obtain, thereby limiting research accessibility.

Example: Bandhu et al. [12] highlighted the importance of preprocessing pipelines for cleaning and normalizing financial datasets.

Mitigation Strategies: GAN-based models, such as StockGAN [7], can generate high-quality synthetic data to augment sparse datasets.

#### 3) Data Sparsity in Emerging Markets:

Emerging markets often lack historical data and exhibit incomplete datasets, limiting the ability of the models to generalize globally.

Example: Predictive models trained on U.S. markets often underperform when applied to emerging markets such as Southeast Asia [9].

Mitigation Strategies: Adaptive frameworks, such as SeroFAM [15], can be dynamically adjusted to diverse datasets.

### 8.2. Methodological limitations

AI models face inherent limitations in handling the complexities of financial markets.

#### 1) Overfitting to Historical Data:

Many models perform well in back testing but fail in real-world scenarios, especially during unprecedented events, such as the COVID-19 pandemic.

Example: Gülmez [13] highlights overfitting issues in LSTM-ARO models.

Mitigation Strategies: Regularization techniques, cross-validation, and continual learning frameworks can improve the robustness. When selecting an appropriate financial model, aligning model complexity with dataset characteristics to avoid overfitting and ensure generalizability [23].

#### 2) Limited Multimodal Integration:

Combining structured (e.g., historical prices) and unstructured (e.g., news sentiment) data remains a challenge.

Example: Hybrid models such as HQNN [9] integrate sentiment analysis, but struggle with scalability and data fusion.

Proposed Solutions: Frameworks such as SA-DLSTM [16], which incorporates self-attention mechanisms, show potential for multimodal data integration.

### 3) Real-Time Adaptation and Scalability:

High computational costs limit the deployment of advanced models such as transformers in real-time trading applications.

Example: Models such as SpectralGPT [15] demonstrate high accuracy, but are computationally intensive.

Mitigation Strategies: Lightweight architectures such as SSA-ESN [18] optimize the computational efficiency for real-time applications.

## 8.3. Interpretability and explainability

The black-box nature of many AI models hinders their adoption in financial markets, owing to a lack of transparency.

### 1) Black-Box Models:

Deep-learning models such as LSTMs and GANs provide limited insights into how predictions are made.

Example: Gandhudi et al. [9] used SHAP values to improve explainability but noted the limited adoption of such techniques.

Proposed Solutions: Developing explainable AI (XAI) frameworks tailored to financial markets is critical for enhancing stakeholder trust.

### 2) Causal Understanding:

Most models focus on correlation rather than causation, which leads to unreliable predictions during structural market changes.

Example: Models such as MMGAN-HPA [3] excel in identifying patterns, but fail to understand the underlying drivers of market movements.

Proposed Solutions: Integrating domain knowledge with AI models can improve causal reasoning and prediction reliability.

## 8.4. Ethical and regulatory concerns

The increasing adoption of AI in financial markets raises ethical and regulatory challenges.

### 1) Market Manipulation and Fairness:

AI prediction-driven algorithmic trading can amplify market volatility and create unfair advantages for large institutions.

Example: Flash crashes caused by AI-driven high-frequency trading algorithms are a growing concern.

Proposed Solutions: Regulatory frameworks and ethical guidelines for AI deployment are essential for ensuring market fairness.

### 2) Bias and Inequality:

Models trained on biased datasets may reinforce existing inequalities, disadvantaging certain market participants.

Example: Sentiment analysis models such as SSWN [8] may favor large-cap stocks because of data availability bias.

Mitigation Strategies: Ensuring diverse and unbiased datasets during training is crucial.

## 8.5. Emerging technologies and future directions

Emerging technologies, such as quantum computing and federated learning, offer promising solutions, but remain underexplored.

### 1) Quantum Computing:

Hybrid quantum neural networks (HQNNs) [9] demonstrate the potential to address computational inefficiencies, but face hardware limitations.

### 2) Federated Learning:

Decentralized training frameworks can improve data privacy and accessibility but require robust convergence strategies.

Addressing these gaps and challenges in AI-based stock market prediction requires interdisciplinary efforts that span finance, machine learning, data science, and ethics. Addressing these gaps and challenges is critical for advancing the AI-based stock market predictions. The identified issues – spanning data quality, methodological limitations, interpretability, and ethical concerns – highlight the need for innovative approaches to ensuring the robustness, scalability, and fairness of financial forecasting systems.

## 9. Future directions

Despite significant advancements in artificial intelligence (AI)-based stock market price trend predictions, critical gaps and challenges persist. These include data-related issues, such as noisy and nonstationary data, methodological limitations, limited interpretability, and ethical concerns. Addressing these challenges requires an interdisciplinary approach combining expertise in finance, machine learning, data science, and ethics. Key research directions for advancing the field toward more reliable, scalable, adaptable, explainable, and ethical AI-driven financial forecasting systems.

### 9.1. Enhancing generalization across market conditions

#### Challenges:

- Financial markets are highly dynamic with regime shifts caused by economic crises, geopolitical events, and regulatory changes.
- Models often overfit historical data, limiting their ability to generalize to unseen scenarios, particularly during sudden market disruptions.

#### Future Directions:

- Adaptive Learning: Develop models capable of incremental or continual learning to adapt dynamically to new market conditions without requiring complete retraining.
- Meta-Learning: Implementation of meta-learning frameworks to enable rapid adaptation to new datasets with minimal fine-tuning, improving performance in volatile environments.
- Hybrid Approaches: Traditional statistical models (e.g., ARIMA) are combined with machine-learning models (e.g., LSTMs and transformers) to capture both linear and nonlinear dependencies.

Example: Gülmez [13] demonstrated the potential of LSTM-ARO in long-term dependency modeling, but highlighted scalability challenges.

### 9.2. Real-time prediction and scalability

#### Challenges:

- High-frequency trading and financial decision making require real-time predictions; however, existing models often struggle with computational intensity and latency, particularly during periods of market volatility.

#### Future Directions:

- **Lightweight Architectures:** Explore lightweight models like Echo State Networks (ESNs) optimized with evolutionary algorithms (e.g., Sparrow Search Algorithm [18]) to handle real-time data efficiently.

- **Edge Computing:** AI models are deployed closer to data sources (e.g., stock exchanges) to reduce latency and improve prediction speeds.

- **Online Learning:** Develop algorithms, such as Online Gradient Descent, to continuously update model weights in real time as new data arrive.

Example: Bandhu et al. [12] proposed a hybrid framework that combined Apache Kafka for stream processing and deep learning for low-latency applications.

### 9.3. Handling noisy and nonstationary data

Challenges:

- Financial datasets are inherently noisy, nonstationary, and prone to anomalies from unexpected events (e.g., geopolitical crises and flash crashes).

- Emerging markets often lack sufficient historical data, which leads to data sparsity and reduced model generalizability.

Future Directions:

- **Noise Reduction:** Advanced preprocessing techniques, such as wavelet transforms and spectral decomposition, are used to filter out noise while preserving meaningful patterns. SpectralGPT [10] demonstrates potential for reducing noise in high-dimensional data.

- **Anomaly Detection:** Anomaly detection techniques are integrated to identify and exclude the outliers caused by market shocks.

- **Data Augmentation:** Leverage GAN-based models (e.g., StockGAN [7]) to generate synthetic datasets for training and to enhance model robustness.

- **Standardized Benchmarks:** Establish globally standardized datasets and evaluation metrics to ensure fair and consistent model comparison.

Example: Polamuri et al. [3] used GANs to address noise in financial datasets and improve prediction accuracy but noted challenges in hyperparameter tuning.

### 9.4. Integration of multimodal data

Challenges:

- Stock prices are influenced by both structured (e.g., historical prices and technical indicators) and unstructured (e.g., news sentiment and macroeconomic reports) data.

- Temporal misalignment between these data sources complicates predictive accuracy.

Future Directions:

- **Multimodal Fusion:** Develop advanced fusion techniques (e.g., early fusion and late fusion) to effectively integrate structured and unstructured data.

- **Graph Neural Networks (GNNs):** Use GNNs to model relationships between entities (e.g., stocks, sectors, macroeconomic factors) to capture dependencies within financial systems.

- **Advanced NLP Models:** Leverage domain-specific NLP tools such as FinBERT or Bloomberg GPT to analyze textual data and extract actionable insights.

Example: Zhao et al. [16] demonstrated the effectiveness of self-attention mechanisms in integrating multimodal data to improve the accuracy.

### 9.5. Explainable AI (XAI) frameworks

Challenges:

- The “black-box” nature of many AI models limits their adoption owing to a lack of transparency, which is critical for regulatory compliance and stakeholder trust.

- Most models focus on correlations rather than causation, which leads to unreliable

predictions during structural market changes.

Future Directions:

- Feature Importance Analysis: Use tools like SHAP (SHapley Additive exPlanations (SHAP) or Local Interpretable Model-Agnostic Explanations (LIME) to enhance transparency and interpretability and highlight the contributions of individual features.

- Attention Mechanisms: Integrate attention mechanisms to identify key drivers of stock price movements.

- Causal Inference: Combine causal reasoning with AI models to better understand market dynamics and drivers.

Example: Gandhudi et al. [9] used SHAP values to enhance the interpretability of hybrid quantum neural networks and to improve stakeholder trust.

## 9.6. Ethical and regulatory considerations

Challenges:

- AI-driven trading systems can exacerbate market volatility and systemic risks, raising ethical concerns regarding fairness and transparency.

- Models trained on biased datasets may perpetuate inequalities, disadvantaging certain market participants.

Future Directions:

- Ethical AI Guidelines: Collaborate with regulatory bodies to establish ethical standards for AI deployment in financial markets.

- Bias Mitigation: Diverse unbiased datasets were used during model training to reduce the prediction biases.

- Transparency Standards: Define industry-wide standards for explainability and accountability in AI-driven trading systems.

Example: Sentiment analysis models such as SSWN [8] demonstrate the importance of addressing biases to ensure fair predictions.

## 9.7. Leveraging emerging technologies

Challenges:

- Emerging technologies, such as quantum computing and federated learning, are underexplored in financial AI applications because of hardware and scalability limitations.

Future Directions:

- Quantum Computing: Develop hybrid quantum neural networks (HQNNs) to address computational inefficiencies and improve performance on high-dimensional data.

- Federated Learning: Implement decentralized training frameworks to enhance data privacy and accessibility, particularly for sensitive financial datasets.

Example: HQNNs [9] show promise for enhancing the computational efficiency, although hardware constraints remain a limitation.

Addressing the gaps and challenges in AI-based stock market prediction requires interdisciplinary efforts that span finance, machine learning, data science, and ethics. Future research should prioritize the following:

- Developing robust models to handle noisy, nonstationary data and adapting dynamically to changing markets.

- Improving interpretability to enhance stakeholder trust and regulatory compliance.

- Exploring emerging technologies, such as quantum computing and federated learning, to enhance computational efficiency and data privacy.

- Establishing standardized benchmarks to enable fair comparisons of predictive models.

## 10. Conclusions

The application of artificial intelligence (AI) to stock market price trend prediction has significantly advanced financial forecasting by addressing the complexities of volatile and nonlinear financial markets. This review synthesizes insights from recent advancements, showcasing the diversity of methodologies, datasets, and hybrid approaches that have emerged in this field and highlights the transformative potential of AI-driven models in enhancing predictive accuracy, robustness, and scalability.

Hybrid approaches, such as ARIMA-ANN [17] and K-means-LSTM [6], have demonstrated their ability to integrate statistical methods with deep learning techniques, effectively capturing both linear and nonlinear dependencies in the data. GAN-based models, such as MMGAN-HPA [3] and StockGAN [7], show promise in handling noisy datasets by generating high-quality synthetic data, whereas frameworks, such as LSTM-ARO [13], optimize temporal dependencies through advanced hyperparameter tuning. Sentiment analysis models such as Stock Senti WordNet (SSWN) [8] leverage real-time social media data to improve short-term predictions. Frameworks such as HQNN [9], which integrate FinBERT for sentiment analysis, have shown up to an 18 % improvement in prediction accuracy for volatile stocks, emphasizing the importance of incorporating diverse data sources, including historical prices, technical indicators, and social media sentiments. Emerging technologies, such as transformer-based architectures (e.g., SpectralGPT [10]), further enhance the ability to process high-dimensional financial data by capturing sequential patterns and reducing noise. Explainable AI (XAI) techniques, such as SHAP values integrated into HQNN frameworks [9] are paving the way for greater transparency and trust in AI-driven systems.

Despite these advancements, several challenges persist, including data-quality issues, overfitting risks, computational complexity, and ethical concerns. Addressing these gaps requires interdisciplinary collaboration between finance, machine learning, and ethics to develop adaptive learning frameworks, multimodal data integration techniques, and robust ethical guidelines.

In conclusion, AI-based stock market predictions have immense potential to transform financial forecasts. By balancing innovation with interpretability and fairness, future research could enable scalable, transparent, and ethical AI systems to meet the evolving demands of global financial markets.

## Acknowledgements

The authors have not disclosed any funding.

## Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Author contributions

Vishwas H N: conceptualization, data curation, formal analysis, investigation, methodology, writing the article. Dr. Usha B A: investigation, methodology, visualization, validation, supervision, resources. Sangeetha K N: investigation, methodology, validation.

## Conflict of interest

The authors declare that they have no conflict of interest.



## References

- [1] A. Almaafi, S. Bajaba, and F. Alnori, "Stock price prediction using ARIMA versus XGBoost models: the case of the largest telecommunication company in the Middle East," *International Journal of Information Technology*, Vol. 15, No. 4, pp. 1813–1818, May 2023, <https://doi.org/10.1007/s41870-023-01260-4>
- [2] F. Pokou, J. Sadefo Kamdem, and F. Benhmad, "Hybridization of ARIMA with learning models for forecasting of stock market time series," *Computational Economics*, Vol. 63, No. 4, pp. 1349–1399, Nov. 2023, <https://doi.org/10.1007/s10614-023-10499-9>
- [3] S. R. Polamuri, D. K. Srinivas, and D. A. Krishna Mohan, "Multi-model generative adversarial network hybrid prediction algorithm (MMGAN-HPA) for stock market prices prediction," *Journal of King Saud University – Computer and Information Sciences*, Vol. 34, No. 9, pp. 7433–7444, Oct. 2022, <https://doi.org/10.1016/j.jksuci.2021.07.001>
- [4] J. Zhang and X. Chen, "A two-stage model for stock price prediction based on variational mode decomposition and ensemble machine learning method," *Soft Computing*, Vol. 28, No. 3, pp. 2385–2408, Jun. 2023, <https://doi.org/10.1007/s00500-023-08441-0>
- [5] D. Chandola, A. Mehta, S. Singh, V. A. Tikkiwal, and H. Agrawal, "Forecasting directional movement of stock prices using deep learning," *Annals of Data Science*, Vol. 10, No. 5, pp. 1361–1378, Aug. 2022, <https://doi.org/10.1007/s40745-022-00432-6>
- [6] Y. Chen, J. Wu, and Z. Wu, "China's commercial bank stock price prediction using a novel K-means-LSTM hybrid approach," *Expert Systems with Applications*, Vol. 202, p. 117370, Sep. 2022, <https://doi.org/10.1016/j.eswa.2022.117370>
- [7] M. Digi, M. E. Hiswati, and A. S. Nur, "StockGAN: robust stock price prediction using GAN algorithm," *International Journal of Information Technology*, Vol. 14, No. 5, pp. 2309–2315, Apr. 2022, <https://doi.org/10.1007/s41870-022-00929-6>
- [8] S. Albahli, A. Irtaza, T. Nazir, A. Mehmood, A. Alkhalifah, and W. Albattah, "A machine learning method for prediction of stock market using real-time twitter data," *Electronics*, Vol. 11, No. 20, p. 3414, Oct. 2022, <https://doi.org/10.3390/electronics11203414>
- [9] M. Gandhudi, A. P. J. A., U. Fiore, and G. G. R., "Explainable hybrid quantum neural networks for analyzing the influence of tweets on stock price prediction," *Computers and Electrical Engineering*, Vol. 118, p. 109302, Aug. 2024, <https://doi.org/10.1016/j.compeleceng.2024.109302>
- [10] A. Pagliaro, "Artificial intelligence vs. efficient markets: a critical reassessment of predictive models in the big data era," *Electronics*, Vol. 14, No. 9, p. 1721, Apr. 2025, <https://doi.org/10.3390/electronics14091721>
- [11] Y. Li, L. Chen, C. Sun, G. Liu, C. Chen, and Y. Zhang, "Accurate stock price forecasting based on deep learning and hierarchical frequency decomposition," *IEEE Access*, Vol. 12, pp. 49878–49894, Jan. 2024, <https://doi.org/10.1109/access.2024.3384430>
- [12] K. C. Bandhu, R. Litoriya, A. Jain, A. V. Shukla, and S. Vaidya, "An improved technique for stock price prediction on real-time exploiting stream processing and deep learning," *Multimedia Tools and Applications*, Vol. 83, No. 19, pp. 57269–57289, Dec. 2023, <https://doi.org/10.1007/s11042-023-17130-x>
- [13] B. Gülmez, "Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm," *Expert Systems with Applications*, Vol. 227, p. 120346, Oct. 2023, <https://doi.org/10.1016/j.eswa.2023.120346>
- [14] J. M. Sangeetha and K. J. Alfia, "Financial stock market forecast using evaluated linear regression based machine learning technique," *Measurement: Sensors*, Vol. 31, p. 100950, Feb. 2024, <https://doi.org/10.1016/j.measen.2023.100950>
- [15] L. L. X. Yeo, Q. Cao, and C. Quek, "Dynamic portfolio rebalancing with lag-optimised trading indicators using SeroFAM and genetic algorithms," *Expert Systems with Applications*, Vol. 216, p. 119440, Apr. 2023, <https://doi.org/10.1016/j.eswa.2022.119440>
- [16] Y. Zhao and G. Yang, "Deep learning-based integrated framework for stock price movement prediction," *Applied Soft Computing*, Vol. 133, p. 109921, Jan. 2023, <https://doi.org/10.1016/j.asoc.2022.109921>
- [17] E. Alshawarbeh, A. T. Abdulrahman, and E. Hussam, "Statistical modeling of high frequency datasets using the ARIMA-ANN hybrid," *Mathematics*, Vol. 11, No. 22, p. 4594, Nov. 2023, <https://doi.org/10.3390/math11224594>

- [18] Z. Tian and L. Pei, "Stocks price prediction based on optimized echo state network by sparrow search algorithm," *International Journal of Dynamics and Control*, Vol. 12, No. 8, pp. 2764–2777, Mar. 2024, <https://doi.org/10.1007/s40435-024-01405-8>
- [19] M. Lu and X. Xu, "TRNN: An efficient time-series recurrent neural network for stock price prediction," *Information Sciences*, Vol. 657, p. 119951, Feb. 2024, <https://doi.org/10.1016/j.ins.2023.119951>
- [20] H. N. Bhandari, B. Rimal, N. R. Pokhrel, R. Rimal, K. R. Dahal, and R. K. C. Khatri, "Predicting stock market index using LSTM," *Machine Learning with Applications*, Vol. 9, p. 100320, Sep. 2022, <https://doi.org/10.1016/j.mlwa.2022.100320>
- [21] S. Rath, N. R. Das, and B. K. Pattanayak, "Stacked BI-LSTM and E-optimized CNN-a hybrid deep learning model for stock price prediction," *Optical Memory and Neural Networks*, Vol. 33, No. 2, pp. 102–120, Jul. 2024, <https://doi.org/10.3103/s1060992x24700024>
- [22] M. Li, Y. Zhu, Y. Shen, and M. Angelova, "Clustering-enhanced stock price prediction using deep learning," *World Wide Web*, Vol. 26, No. 1, pp. 207–232, Apr. 2022, <https://doi.org/10.1007/s11280-021-01003-0>
- [23] P. Banerjee and R. Nayak, "Recommendations on financial models for stock price prediction," *SN Computer Science*, Vol. 5, No. 1, p. 178, Jan. 2024, <https://doi.org/10.1007/s42979-023-02507-4>
- [24] D. Hong et al., "SpectralGPT: Spectral remote sensing foundation model," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 46, No. 8, pp. 5227–5244, Aug. 2024, <https://doi.org/10.1109/tpami.2024.3362475>



**Vishwas H N** received his B.E. degree in computer science and engineering from Kuvempu University in 2000 and an M.Tech degree in computer science and engineering from Visvesvaraya Technological University (VTU) in 2017. He is currently a research scholar pursuing his Ph.D. under VTU, specializing in artificial intelligence and machine learning. His research interests span artificial intelligence, machine learning, deep learning, and their real-world applications. He is particularly focused on leveraging AI techniques to solve complex challenges in domains such as stock market prediction, pattern recognition, data-driven decision-making, and automation. His work includes developing innovative algorithms for analyzing large-scale data, improving prediction accuracy, and addressing challenges like model interpretability and robustness in AI systems. He is deeply committed to advancing the field of AI by exploring its potential in areas like financial forecasting, intelligent systems, and adaptive learning models.



**Dr. Usha B A** is a dedicated motivational individual committed to maximizing learning opportunities in diverse academic settings using consistent and organized practices. Energetic and ambitious professional with 18 years of experience in teaching. She has completed her research in the area of information security. She obtained her Ph.D. from Visvesvaraya Technological University, Belagavi, Karnataka, India in the Year 2016. She has published more than 35 research papers in reputed international journal and conferences where some of them are Scopus indexed and also has good impact factors. Currently, she is working as Professor in the Department of Information Science and Engineering, BMS Institute of Technology and Management, Bengaluru. Her research interests include IoT, information security, cyber security, data mining, artificial intelligence and machine learning.



**Dr. Sangeetha K N** is presently working as Assistant Professor in JSS Academy of Technical Education, Bengaluru. She completed her bachelor's degree, master's degree and Doctor of Philosophy in Electronics and Communication Engineering from Visveswaraya Technological University, Belagavi. Her area of specialization includes Digital Communication, Image Processing, Information Security, Cryptography, and Python Programming. She has 13 years of teaching experience so far. She is an enthusiastic, committed individual who uses systematic, consistent methods to maximize learning opportunities in a variety of academic environments. She was selected as a reviewer for reputed international conference. Published papers in reputed International Journal and Conferences