REAL-TIME STOCK MARKET PRICE PREDICTION USING MACHINE LEARNING

Dr. N. Kiranmai[1],*, P. Rakshitha[2], T. Abhinay[2], T. Shiva[2], N. Rishitha[2]
[1]Associate Professor, Department of CSE (DS), TKR College of Engineering & Technology, Meerpet, Telangana 500097
[2]B.Tech (Scholar), Department of CSE (DS), TKR College of Engineering & Technology, Meerpet, Telangana 500097
Correspondence: nkiranmai@tkrcet.com

ABSTRACT
After the COVID-19 ended, the global economy gradually recovered. Due to the nonlinearity, complexity, and high noise of financial time series, stock price prediction has become one of the most challenging tasks in the stock market. we propose a real-time stock market price prediction based on Long Short-Term Memory (LSTM) to simultaneously improve the fitting and accuracy of stock price prediction. Key algorithms like LSTM (Long Short-Term Memory), ARIMA (Auto Regressive Integrated Moving Average), and Random Forest will be explored for their effectiveness in forecasting stock trends. Data preprocessing and feature engineering techniques are applied to improve model reliability and prediction accuracy. The developed system integrates an interactive dashboard that visualizes both historical and predicted stock trends, enabling users to monitor market movements and gain insights into potential price behavior. By combining machine learning models with real-time data acquisition and visualization, the proposed approach provides a practical tool for supporting data-driven financial analysis and decision-making.

Keywords: LSTM, Price Prediction, ARIMA, Random Forest, Machine learning, stock market.

1. INTRODUCTION

The stock market plays a vital role in the global economy, serving as a platform where investors trade financial securities and influence economic growth. Stock prices are highly dynamic and fluctuate continuously in response to various factors such as company performance, investor sentiment, global events, and macroeconomic indicators. Predicting these price movements has long been a challenge due to the market's inherent volatility and complexity. Traditional statistical models often fall short in capturing the nonlinear and rapidly changing patterns of stock market data, leading to the need for more advanced approaches [1].

One of the key difficulties in stock market analysis lies in its real-time nature, where prices update within milliseconds. Investors and traders require fast, accurate, and reliable predictions to make informed decisions and maximize returns while minimizing risks. However, the vast volume of streaming financial data, combined with unpredictable external influences, makes real-time forecasting extremely challenging [2]. This complexity calls for systems that can learn from massive datasets, identify hidden patterns, and adapt quickly to new market conditions.As Data Science students, we worked together to design and implement the complete machine learning pipeline. We collected and preprocessed both historical and live data, created financial features such as moving averages, and applied models like Random Forest and LSTM.

Machine learning (ML) has emerged as a powerful tool for tackling such challenges in financial forecasting. Unlike traditional models, ML algorithms are capable of handling high-dimensional data, capturing nonlinear relationships, and improving accuracy through continuous learning. Techniques such as regression models, support vector machines, recurrent neural networks (RNNs), and long short-term memory (LSTM) networks are increasingly used to predict stock prices by learning from historical and real-time data streams [3]. The adaptability and scalability of ML make it highly suitable for real-time applications in the stock market.

Real-time stock price prediction using ML has significant implications for both individual investors and large financial institutions. For day traders, it can enhance decision-making and profitability by offering short-term insights, while institutional investors can leverage it for portfolio optimization and risk management. Furthermore, automated trading systems, powered by ML-driven predictions, have the potential to execute trades at optimal times with minimal human intervention. As financial markets continue to evolve with increasing speed and complexity, the importance of real-time predictive models becomes more critical.

This study focuses on exploring machine learning techniques to predict stock market prices in real time, leveraging historical and live data to enhance forecasting accuracy. The objective is to design a system that can process high-frequency financial data, extract meaningful patterns, and provide predictions with minimal latency [4]. By evaluating different ML models and comparing their performance, the research aims to contribute to the growing field of financial analytics and demonstrate how intelligent algorithms can support informed investment strategies in fast-paced markets. This complexity calls for systems that can learn from massive datasets, identify hidden patterns, and adapt quickly to new market conditions.

2. LITERATURE SURVEY

Research in stock market prediction has been a longstanding area of interest, combining finance, statistics, and computational methods [5]. Traditional approaches relied heavily on econometric models such as Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) [6-7]. These models aimed to capture linear relationships and volatility patterns within historical stock prices. While useful for trend analysis, such methods struggled with the highly nonlinear and chaotic nature of financial markets, limiting their effectiveness for real-time prediction tasks.

With the growth of machine learning, more advanced algorithms such as Support Vector Machines (SVM), Random Forests, and k- Nearest Neighbors (k-NN) were introduced for stock price forecasting [8]. These approaches demonstrated improved predictive power by leveraging complex patterns in historical and technical indicator data. However, many of these models were still static, requiring retraining when new data became available [9]. As a result, their application in real-time environments was constrained, particularly in high-frequency trading scenarios where rapid adaptability is crucial [10].

In recent years, deep learning methods have significantly influenced stock market prediction research. Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRU) have been extensively employed to capture temporal dependencies in sequential financial data. Convolutional Neural Networks (CNN) have also been adapted to identify spatial correlations among multivariate time series [11-14].

Another important direction in related work involves the integration of alternative data sources, such as social media sentiment, news articles, and macroeconomic indicators [15]. Studies have demonstrated that incorporating sentiment analysis using natural language processing (NLP) techniques can enhance the robustness of prediction models by accounting for market psychology and investor behaviour [16-18].

3. PROPOSED METHODOLOGY

3.1 Data Collection and Preprocessing

The initial stage of the proposed system focuses on acquiring both historical and real-time stock market data from reliable financial data sources such as Yahoo Finance or Alpha Vantage. The collected dataset includes essential attributes such as Date, Open, High, Low, Close, and Volume (OHLCV), which represent the trading behavior of stocks over a period of time. Since raw financial datasets may contain incomplete records, noise, or inconsistencies, appropriate preprocessing techniques are applied to clean and organize the data before it is used for model training and analysis

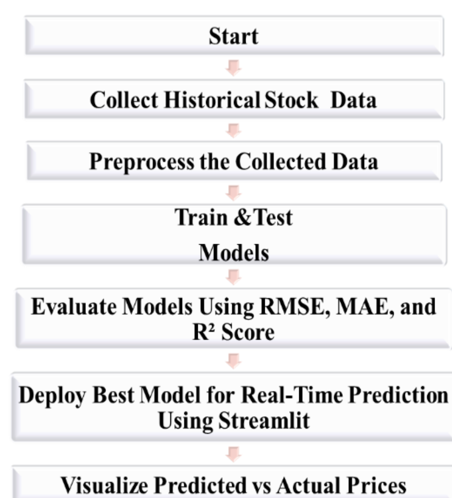

Fig 3.1 Process Flow of Real-Time Stock Market Prediction.

3.2 Feature Engineering

Feature Feature engineering is an essential step in enhancing the performance of machine learning models. In this project, additional financial indicators are derived from historical stock market data to better capture market behavior such as trends, momentum, and price volatility. Indicators including Moving Average (MA), Relative Strength Index (RSI), and Moving Average Convergence Divergence (MACD) are calculated from the dataset to provide meaningful information that can assist the model in learning important market patterns. These indicators help the model understand market conditions more effectively by highlighting price movements and trading signals. By incorporating these engineered features, the prediction model can capture complex relationships in stock data and improve forecasting accuracy.

3.3 Model Training and Prediction

Model training involves building and optimizing an LSTM (Long Short-Term Memory) model to learn patterns from historical stock market data. The processed dataset, including technical indicators and lag features, is split into training, validation, and testing sets. The model is trained to predict future opening and closing prices using Mean Squared Error (MSE) as the loss function and the Adam optimizer for efficient learning. Techniques like dropout layers and early stopping are applied to improve accuracy and prevent overfitting. Once trained, the model is saved for later use.

3.4 Real-Time Data Acquisition and Preprocessing

In the prediction phase, live stock data is fetched through APIs, pre processed in the same way as the training data, and passed into the saved model. The system the predicts the next opening and closing prices in real time. These predictions are continuously updated and displayed through a dashboard, enabling users to monitor market trends and make informed decisions quickly and efficiently. This real-time data integration allows the system to respond quickly to market fluctuations and update predictions accordingly.

3.5 Model Deployment and Dashboard Integration

After the training and evaluation stages are completed, the prediction model is deployed within a web-based platform that allows users to interact with the system. The front-end interface is developed using Streamlit, which provides a simple and interactive dashboard environment. Through this interface, users can select specific stocks, analyze historical price data, and observe predicted price movements in real time. The dashboard presents both actual and predicted stock prices using dynamic charts and visualizations for better understanding. Users can easily compare past market trends with predicted outcomes to evaluate the model's performance. The visual interface makes complex prediction results easier to interpret for both technical and non-technical users. This interactive dashboard enhances the usability of the system and allows users to monitor stock market behavior efficiently.

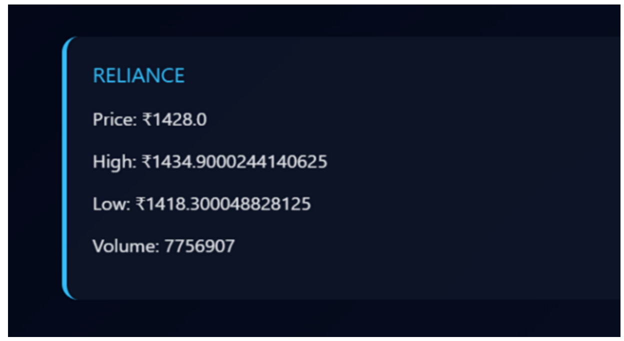

Fig 3.2 Real-Time Stock Market Dashboard Interface

3.6 Performance Evaluation

To determine the effectiveness of the prediction system, several evaluation metrics are applied. These metrics help assess how closely the predicted stock prices match the actual market values. In this project, commonly used performance indicators such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the R^2 Score are used to evaluate the model. By comparing the predicted results with the real stock prices, these metrics provide a clear understanding of the model's accuracy and reliability. A lower error value indicates better prediction performance and stronger model effectiveness.

3.7 Result Visualization and Output

The final stage of the methodology focuses on presenting the prediction results in a clear and understandable visual format. The predicted stock prices are compared with the actual stock prices and displayed using graphical charts. These visualizations enable users to observe how closely the predicted values follow the actual market trends. By comparing both values, users can evaluate the performance of the prediction model and gain insights into stock price movements. The system generates interactive charts that illustrate historical data along with predicted prices, making the results easier to analyze and interpret.

Fig 3.3 Comparison of Predicted and Actual Stock Price

4. ARCHITECTURE

The architecture of the proposed system is designed to enable real-time stock market price prediction using machine learning techniques. The system combines several functional components such as data collection, preprocessing, model training, prediction generation, and visualization through an interactive dashboard. These components work together to process stock market information and generate accurate predictions that can assist users in understanding market behavior.

4.1 Data acqusition

The data acquisition module represents the first stage of the system architecture. Its primary function is to gather both historical and real-time stock market data from reliable financial APIs such as Yahoo Finance or Alpha Vantage. The collected dataset includes essential attributes such as opening price, closing price, highest price, lowest price, trading volume, and timestamp information. These parameters describe stock market activity over time and provide the necessary input for training and evaluating machine learning models.

4.2 Data preprocessing

The collected stock market data often contains missing values, noise, and inconsistencies that can affect the accuracy of machine learning models. Therefore, preprocessing techniques are applied to clean and prepare the dataset. Preprocessing steps include handling missing values, removing anomalies, normalizing data values, and converting the dataset into a structured time-series format. Feature scaling methods such as Min-Max normalization are used to ensure that all input variables remain within the same range.

4.3 Machine Learning Model Processing

After completing the preprocessing stage, the prepared dataset is provided to machine learning algorithms for training and prediction. In this project, the Long Short-Term Memory (LSTM) network is selected as the main prediction model because of its ability to learn sequential patterns present in time-series data. The LSTM model analyzes previous stock price
values to identify hidden relationships and long-term dependencies within the dataset. Based on these learned patterns, the model predicts the upcoming closing price of the stock. In addition to the LSTM model, other algorithms such as Random Forest and ARIMA are implemented to compare their performance and evaluate the effectiveness of different prediction approaches.

4.4 Prediction Generation

Once the machine learning model has been successfully trained, the system begins generating predictions using real-time market data. Stock price data collected from financial APIs is provided as input to the trained model in order to estimate the next price value. The prediction process continues as new market data becomes available, allowing the system to update forecasts dynamically. This capability enables investors and analysts to observe potential future price movements and better understand possible market trends.

4.5 Visualization and User Interface

The final stage of the architecture focuses on presenting the prediction results through a user-friendly interface. An interactive dashboard is developed using Streamlit to allow users to monitor stock data and prediction outcomes easily. Dashboard displays both historical and predicted stock prices using graphical charts and visual representations. These visual tools help users analyze market behavior and compare predicted values with actual stock prices.

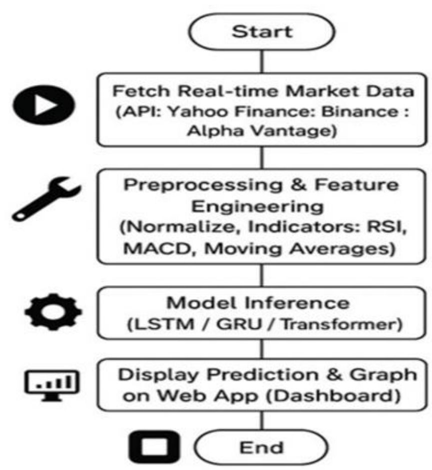

Fig 4.1Architecture of Real-Time Stock Market Price Prediction

5. RESULT

The proposed system was evaluated using historical stock market data obtained from financial APIs. The dataset included key attributes such as Open, Close, High, Low, and Volume, which represent trading activity across different time periods. After collecting the data, preprocessing operations such as normalization, handling missing values, and converting the dataset into a time-series format were applied before training the machine learning model. These preprocessing steps helped ensure that the dataset was well-structured, consistent, and suitable for accurate prediction.

The system also provides graphical visualization of historical stock price movements. These visual representations allow

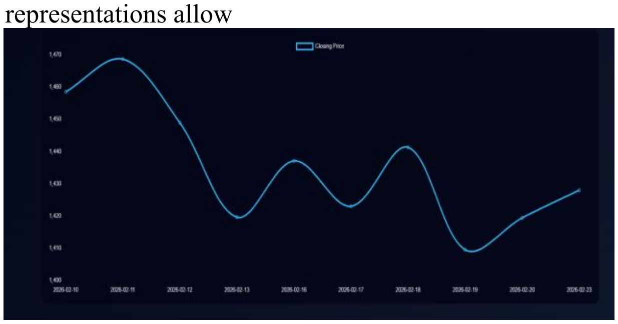

Fig 5.1 Stock Price Trend Visualization

The prediction module generates real-time stock price forecasts using the trained LSTM model. The model analyzes previous stock price values as input and predicts the next closing price. As new market data becomes available, the system continuously updates the prediction results to maintain real-time forecasting capability.

Fig 5.2 Real-Time Stock Price Prediction Output

This This real-time prediction feature enables users to monitor potential future price movements and analyze stock market trends dynamically. The predicted results are displayed through the system's interactive interface, allowing investors and analysts to easily interpret the prediction outcomes. Based on the most recent market data, the system generates updated predictions and displays them on the dashboard for quick analysis.

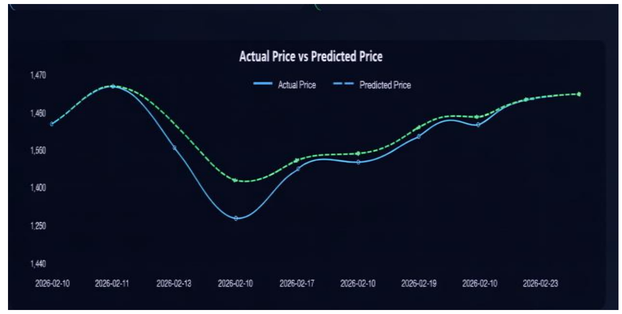

Fig.5.3. Different security vulnerability

To assess the performance of the prediction model, the system compares predicted stock prices with actual market values. The comparison is visualized through charts where both predicted and real values are plotted together. When the predicted line closely follows the actual stock price trend, it indicates that the model has successfully learned the underlying patterns within the data. Performance metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) are used to evaluate prediction accuracy. The results indicate that the LSTM model effectively captures time-series dependencies and produces reliable short-term stock price predictions.

6. CONCLUSION & FUTURE SCOPE

This The proposed project introduces a real-time stock market price prediction system developed using machine learning techniques. The system collects historical stock market data and processes it using preprocessing methods such as data cleaning, normalization, and time-series transformation. Machine learning algorithms including Long Short-Term Memory (LSTM), ARIMA, and Random Forest are utilized to analyze historical stock patterns and generate predictions. These models help uncover relationships within financial time-series data and contribute to improving prediction accuracy.

Among the implemented models, the LSTM algorithm produced better results for time-series forecasting tasks. Its ability to capture sequential relationships and long-term dependencies in stock market data enables the model to generate reliable short-term price predictions. By learning patterns from historical market behavior, the model can identify underlying trends that influence stock price movements and use this information to forecast future values.

The developed system also integrates real-time stock market data through financial APIs, enabling the model to update predictions continuously as new data becomes available. The prediction outcomes are presented through an interactive visualization dashboard that displays both historical and predicted stock prices using graphical charts. This visual interface allows users to easily interpret market trends and evaluate the accuracy of the prediction model.

Although the proposed system provides effective prediction results, stock markets remain highly dynamic and are influenced by various external factors such as economic conditions, global events, and investor sentiment. Future improvements may include incorporating additional data sources such as financial news, social media sentiment, and macroeconomic indicators to further enhance prediction performance. In addition, the system can be extended by implementing advanced deep learning techniques such as GRU networks, transformer-based architectures, or hybrid machine learning models to improve forecasting capability.

REFERENCES

[1] X. R. Hou, K. Wang, C. Zhong, Z. Wei, "ST-Trader: A Spatial-Temporal Deep Neural Network for Modeling Stock Market Movement," *IEEE/CAA Journal of Automatica Sinica*, Vol. 8, No. 5, pp. 1015-1024, May 2021.
Suman, B., & Jadhav, P. P. (2023). Advancements in routing algorithm techniques for wireless sensor networks. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3).

[2] Suman, B., & Jadhav, P. P. (2023). Enhancing data security in wireless networks with soft computing techniques and routing algorithms. International Journal of Applied Engineering & Technology, 5(4).

[3] Srinivas, B. S., Krishna, V., Sathish, K., Naresh, K., & Banala, R. (2024). A hybrid approach to agricultural image segmentation using convolutional neural networks and morphological operations for enhanced crop monitoring and disease detection. Frontiers in Health Informatics.

[4] Geetha, L. S., El-Ebiary, Y. A. B., Srinivasa Rao, B., Rautrao, R. R., Mastan Rao, T. S., Venkata Naga Ramesh, J., & Al-Omari, O. (2025). Challenges and solutions in agile software development: A managerial perspective on implementation practices. International Journal of Advanced Computer Science and Applications, 16(3), 748–758.

[5] Balamurugan, K., Sudhakar, G., Xavier, K. F., Bharathiraja, N., & Kaur, G. (2025). Human-machine interaction in mechanical systems through sensor enabled wearable augmented reality interfaces. Measurement: Sensors, 39, 101880

[6] Abshalomu, Y., Jyothi, Y., Balamurugan, K., & Selvaraj, R. (2023). Effect of varied cashew nut ash reinforcement in aluminum matrix composite. Advances in Materials Science and Engineering, 2023(1), 3383777

[7] N, Bharathiraja, Minu, M. S., Vijay, R., Rajalakshmi, M., Vidyullatha, P., & Balamurugan, K. (2025). Development of Hybrid Explainable Artificial Intelligence With Swin Vision Transformer Intrusion Detection for Securing VANETs From Attacks. Transactions On Emerging Telecommunications Technologies, 36(10).

[8] Latchoumi, T. P., Parthiban, L., Raja, K., Balamurugan, K., & Parthiban, R. (2023). Secured smart manufacturing systems using blockchain technology for industry 4.0. In Integrating Blockchain and Artificial Intelligence for Industry 4.0 Innovations (pp. 281-294). Cham: Springer International Publishing

[9] Ananthajothi, K., Balamurugan, K., Divya, D., & Latchoumi, T. P. (2026). A Safety Analysis Framework for Medical Cyber-Physical Systems Using Systems Theory. Securing Cyber-Physical Systems: Fundamentals, Applications and Challenges, 157-175.

[10] Latchoumi, T. P., Parthiban, L., Balamurugan, K., Raja, K., Vijayaraj, J., & Parthiban, R. (2023). A framework for low energy application devices using blockchain-enabled IoT in WSNs. In Integrating Blockchain and Artificial Intelligence for Industry 4.0 Innovations (pp. 121-132). Cham: Springer International Publishing

[11] Balamurugan, K., Deepthi, T., Subramanian, A. K., Banerjee, A., Agarwal, D., Biswas, A., & Sinha, A. (2023). A study on the mechanical properties of rare earth-based aluminium composite. Journal of The Institution of Engineers (India): Series D, 104(1), 15-25

[12] Arunkarthikeyan, K., & Balamurugan, K. (2020). Studies on the effects of deep cryogenic treated WC–Co insert on turning of Al6063 using multi-objective optimization. SN applied Sciences, 2(12), 2103.

[13] Pavan, M. V., Balamurgan, K., & Balamurgan, P. (2021). Wear experiments on PLA-Cu composite filament printed in different FDM conditions. Turkish Journal of Computer and Mathematics Education, 12(9), 2245-2251

[14] Muthu, M. A. (n.d.). A hybrid deep CNN model for brain tumor image multi-classification. International Journal of Engineering Research and Science & Technology (IJERST).

[15] Muthu, M. A. (n.d.). Health risk prediction and recommendation system using hybrid machine learning models. International Journal of Engineering Research and Science & Technology (IJERST).

[16] Muthu, M. A. (2016). Performance analysis of cloud computing centers using M/G/m/m+r queuing systems. International Journal of Research in Engineering, Science and Technologies.

[17] Muthu, M. A. (n.d.). Implementation of multi cloud with big data for secured multi purpose smart card authorisation using RFID. International Journal.

[18] Krishna, V., Sumalatha, C., Raju, Y. D. S., & Mohan, K. V. M. (2022). Analysis of heart disease prediction using machine learning classification algorithms. Journal of Optoelectronics Laser.