Stock price prediction using ML Algorithm

Srinithi N
Department of Information Technology, Sri Krishna Adithya College of Arts and Science, Coimbatore, Tamil Nadu, India
srinithin285@gmail.com

Dr. S. Malathi
Department of Information Technology, Sri Krishna Adithya College of Arts and Science, Coimbatore, Tamil Nadu, India
malathis@skacas.ac.in

Abstract

Stock market prediction has long been a challenging task due to its dynamic, volatile, and non-linear nature. Traditional forecasting methods often fail to capture complex patterns and dependencies in financial data, leading to inaccurate predictions. With the advancement of machine learning (ML) algorithms, it has become possible to analyze large volumes of historical stock data, identify hidden trends, and generate more reliable forecasts. This study presents a machine learning-based approach for predicting stock prices using algorithms such as Linear Regression, Random Forest, and Long Short-Term Memory (LSTM) networks. The system leverages historical stock datasets, including features like opening price, closing price, volume, and market indicators, to train predictive models. Data preprocessing techniques such as normalization and feature selection are applied to improve accuracy and reduce noise. The proposed system evaluates the performance of different ML algorithms by comparing metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R^2 score. Results demonstrate that advanced models like LSTM outperform traditional approaches by effectively capturing temporal dependencies in stock price movements.

Introduction

The stock market plays a crucial role in the global economy, influencing investment decisions, corporate growth, and financial stability. However, predicting stock prices remains a highly complex task due to the volatile, dynamic, and non-linear nature of financial markets. Traditional statistical methods often fail to capture hidden patterns and dependencies within large datasets, leading to limited accuracy in forecasting. Investors and analysts therefore seek more advanced tools that can provide reliable insights for decision-making.

Recent advancements in machine learning (ML) have opened new possibilities for financial forecasting by leveraging algorithms capable of learning from historical data and identifying intricate relationships among market variables. ML models such as Linear Regression, Random Forest, and Long Short-Term Memory (LSTM) networks can process vast amounts of stock data, including opening and closing prices, trading volume, and market indicators, to generate predictive outcomes. These models not only improve accuracy but also adapt to changing market conditions, making them suitable for dynamic environments. Stock price prediction using machine and advanced algorithms for financial decision-making.

Literature Study

Early studies on stock price prediction relied heavily on statistical and econometric models such as Autoregressive Integrated Moving Average (ARIMA) and linear regression. While these methods provided basic forecasting capabilities, they struggled to capture the non-linear and highly volatile nature of financial markets. Their limited ability to process large datasets and complex dependencies often resulted in inaccurate predictions, highlighting the need for more advanced techniques.

With the rise of machine learning, researchers began applying algorithms such as Support Vector Machines (SVM), Random Forests, and Artificial Neural Networks (ANN) to financial forecasting. These models demonstrated improved accuracy by learning hidden patterns in historical data and incorporating multiple features like trading volume, market indicators, and sentiment analysis. Studies showed that ensemble methods and tree-based algorithms could outperform traditional statistical approaches in short-term prediction tasks.

Recent literature emphasizes the use of deep learning models, particularly Long Short-Term Memory (LSTM) networks, which excel at capturing temporal dependencies in sequential data. LSTM-based approaches have proven highly effective in modeling stock price movements and reducing prediction errors. Furthermore, integration with IoT and cloud-based platforms has enabled real-time data collection and analysis, making predictive systems more scalable and accessible to investors and analysts. These advancements underline the growing importance of intelligent, data-driven tools in financial decision-making. By integrating textual information with numerical stock data, these models capture both quantitative trends and qualitative market sentiment.

Significance of the Study

The significance of this study lies in its ability to improve the accuracy of stock price predictions by leveraging machine learning algorithms. Traditional forecasting methods often fall short in capturing the complex, non-linear patterns of financial markets. By applying models such as Random Forest and LSTM, this study demonstrates how intelligent systems can learn from historical data and generate more reliable predictions, helping investors make informed decisions.

This research empowers traders and financial analysts with a data-driven approach to investment strategy. Instead of relying solely on intuition or static models, the system uses real-time data and algorithmic analysis to identify trends and potential market movements. The integration of machine learning enhances decision-making by reducing human bias and improving responsiveness to market fluctuations.

Accurate stock prediction plays a vital role in managing financial risk. By forecasting potential price movements, investors can better assess entry and exit points, hedge against losses, and optimize portfolio performance. The study also contributes to broader financial analytics by offering insights into market behavior, volatility patterns, and the impact of external factors and safer investment practices.

.Proposed System

The proposed system is built around a machine learning framework that processes historical stock data to predict future prices. It includes modules for data collection, preprocessing, feature selection, model training, and prediction. The system uses datasets containing stock prices, volume, and market indicators, which are cleaned and normalized before being fed into the ML models.

Multiple machine learning algorithms are implemented and compared, including Linear Regression, Random Forest, and Long Short-Term Memory (LSTM) networks. Each model is trained using historical data and evaluated using performance metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R^2 score. The system selects the best-performing model to generate predictions based on its accuracy and ability to capture market trends.

The final model is deployed in a user-friendly interface that allows investors and analysts to input stock symbols and view predicted price trends. Visualization tools such as line charts and candlestick graphs are integrated to display historical and forecasted data. This makes the system practical for real-world use, offering a smart, accessible solution for financial forecasting.

The proposed system is designed with scalability in mind, allowing it to handle large datasets and adapt to different stock exchanges or financial instruments. Beyond price prediction, the framework can be extended to incorporate sentiment analysis from news articles and social media, providing a more holistic view of market behavior. This flexibility ensures that the system can evolve with emerging technologies and remain relevant for diverse financial forecasting applications.

Methodology

The methodology begins with collecting historical stock data, including features such as opening price, closing price, volume, and market indicators. This raw data undergoes preprocessing steps like normalization, handling missing values, and feature selection to ensure model efficiency. Once cleaned, the dataset is split into training and testing sets. Various machine learning algorithms—such as Linear Regression, Random Forest, and LSTM—are then trained on the data to learn patterns and relationships that influence stock price movements.

After training, each model is evaluated using performance metrics like Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R^2 score to determine its predictive accuracy. The best-performing model is selected for deployment in a user-friendly interface that allows users to input stock symbols and view forecasted trends. Visualization tools such as line graphs and candlestick charts are integrated to enhance interpretability and real-time updates to market changes.

Limitations

One major limitation of the system is its dependence on high-quality historical stock data. Incomplete, noisy, or outdated datasets can significantly reduce model accuracy and reliability. Moreover, sudden market events such as political instability or natural disasters are difficult to capture through past data alone, making predictions less effective during volatile periods.

Machine learning models may struggle to generalize across different stocks or market conditions. A model trained on one company's data might not perform well when applied to another due to differences in trading behavior and external influences. Additionally, complex models like LSTM are prone to overfitting, where they perform well on training data but poorly on unseen data, limiting their practical use.

Stock prices are influenced by a wide range of external factors such as news sentiment, global economic trends, and investor psychology—elements that are not always reflected in numerical data. Without integrating sentiment analysis or real-time news feeds, the system may miss critical signals that impact market movements, reducing its predictive power in fast-changing environments.

Conclusion

This study demonstrates the effectiveness of machine learning algorithms in predicting stock prices by analyzing historical data and identifying hidden patterns. By implementing models such as Linear Regression, Random Forest, and LSTM, the system provides a comparative analysis of their performance, highlighting the strengths of advanced deep learning techniques in capturing temporal dependencies. The project successfully shows how data-driven approaches can enhance.

The proposed system offers significant value to investors, traders, and financial analysts by providing a reliable tool for informed decision-making. Through visualization and user-friendly interfaces, the system translates complex predictions into actionable insights. This not only supports better investment strategies but also contributes to risk management by helping users anticipate market fluctuations and optimize portfolio performance.

While the system achieves promising results, it also opens avenues for further research and improvement. Integrating sentiment analysis, real-time data streams, and hybrid models could enhance prediction accuracy and adaptability in volatile markets. Overall, this project underscores the growing importance of machine learning in financial analytics and sets the foundation for more intelligent, scalable, and holistic forecasting systems in the future.

References

1. Prediction of Stock Price Using Machine Learning Techniques International Journal of Trend in Scientific Research and Development (IJTSRD).

2. Stock Market Prediction Using Machine Learning International Journal for Multidisciplinary Research (IJFMR).

3. Hybrid Approaches for Stock Market Forecasting Using Sentiment Analysis and ML Elsevier – Expert Systems with Applications.

4. Deep Learning for Stock Price Prediction: A Comparative Study IEEE Access. This study compares deep learning models like LSTM and CNN with traditional ML approaches in current dynamic market sectors.