Stock Price Prediction Using Python

Nripesh Narayan Vatsa[1], Onkar Jagtap[2], Tanishq Yeshwante[3], Pranay Hasabe[4], Bhoomi Narhire[5]
[1]*Lecturer, Dept. of AIML Diploma, ARMIET, Maharashtra, India*
[2,3,4]*Student, Dept. of AIML Diploma, ARMIET, Maharashtra, India*

Abstract: Stock market prediction is still difficult despite developments because of its intrinsic volatility and outside influences like geopolitical and economic policies. A number of variables, such as company performance, macroeconomic indicators, market sentiment, and world events, affect stock prices. Social media and news sentiment analysis are now crucial parts of contemporary stock prediction models. But issues like overfitting, noisy data, and market volatility continue to be major roadblocks. Because financial markets are unpredictable, no model can guarantee perfect accuracy, even with advances in predictive accuracy. By creating hybrid models, applying reinforcement learning, and leveraging different data sources, future research seeks to improve accuracy. Even though AI-powered stock forecasting is still developing, investors need to use risk management techniques in conjunction with technology breakthroughs to make the best choices. This study examines a number of stock price prediction techniques, such as deep learning methods (like Long Short-Term Memory (LSTM) networks), machine learning algorithms (like Random Forest and Support Vector Machines), and statistical models (like ARIMA). To increase accuracy, these models make use of sentiment analysis, trading volumes, historical stock prices, and macroeconomic variables

Keywords:- Stock price prediction, Machine learning, Pandas, Matplotlib

1.INTRODUCTION

Predicting the stock market has consistently been an important field of study in financial analytics.

Conventional approaches depend on past data and technical indicators to assess possible price changes. This initiative utilizes contemporary computational methods, incorporating Python libraries, to improve predictive accuracy to 81%. Through the use of the MACD indicator, the model seeks to examine stock trends and offer investors a structured method for market analysis.

Forecasting stock prices is a crucial element of financial market evaluation, significantly influencing investment choices. Investors, traders, and financial institutions depend on forecasting stock prices to enhance profits and reduce risks. Nonetheless, forecasting stock prices is a complicated endeavour because of the ever-changing and unstable characteristics of financial markets. Stock prices are affected by multiple elements, such as company performance, economic circumstances, international events, interest rates, and investor feelings. Conventional forecasting are commonly employed; however, they frequently do not account for the nonlinear and stochastic characteristics of stock fluctuations[1].

As artificial intelligence (AI), machine learning (ML), and deep learning (DL) continue to evolve, more advanced models have been created to improve the accuracy of stock price predictions. Methods like Long Short-Term Memory (LSTM) networks, Random Forest, and ARIMA models examine past stock price information and external influences to forecast upcoming trends. Incorporating sentiment analysis from economic indicators, financial news, and social media enhances prediction precision. This research examines different models for predicting stock prices, assesses their effectiveness, and underscores upcoming trends in financial forecasting

1.1Problem statement:

The stock market is highly unpredictable, driven by numerous external factors such as economic policies, political events, corporate earnings, and investor behavior. Traditional forecasting techniques rely on historical data and simple statistical models, but they often fail to account for market uncertainties and complex interactions between different influencing factors. Investors and traders need more reliable and efficient predictive models to make informed decisions and minimize financial losses.

1.2Main purpose

The main objective of this project is to develop and execute an AI-based stock price prediction model that improves forecasting precision and offers important insights for investors. Utilizing machine learning and deep learning methods, the model will examine past stock prices, technical indicators, and external elements like news sentiment and macroeconomic trends.

The objective of this project is to accomplish the following:

- Create and evaluate various predictive models (such as ARIMA, Random Forest, LSTM) to identify the best method for forecasting stock prices.
- Improve prediction precision by adding extra elements like trading volume, news sentiment, and economic metrics.
- Assess model performance employing statistical measures like Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R^2 score.
- Develop an intuitive graphic representation of projected stock prices to aid investors in making decisions.

By tackling the difficulties of predicting stock prices, this project will aid in the creation of more effective AI-driven financial models.

1.3 Objective of project

The stock price prediction project has significant potential for future advancements and improvements. One key area is real-time data streaming, which would enable the model to process live stock market data, making predictions more accurate andtimely. The integration of reinforcement learning can enhance decision-making by continuously adapting to market fluctuations. Additionally, incorporating more advanced deep learning architectures, such as Transformers and attention-based models, can improve the model's ability to capture complex market trends.

Another promising direction is multimodal data fusion, where financial reports, news articles, and social media sentiment are combined with technical indicators to enhance prediction accuracy. The project can also be extended to support automated trading strategies, where AI-driven algorithms make buy/sell decisions based on predictions. Moreover, by leveraging cloud computing and scalable deployment solutions, the system can handle large-scale data processing for multiple stock exchanges worldwide. Finally, integrating explainable AI (XAI) techniques will help users understand the reasoning behind predictions, increasing trust in AI-driven financial forecasts. These enhancements will further improve the reliability, efficiency, and usability of stock price prediction systems for traders and investors.

2. LITERATURE REVIEW

1. MujieSui et al. (2024):- Ensemble learning improves prediction accuracy compared to standalone models. Not peer-reviewed, ensemble models require more computation.

2. Nur Aliah Syahmina Mohd Nasiruddin et al. (2024): -Synthetic data improves model performance, SMOTE-ANN achieved 93% accuracyImbalanced dataset affects performance.

3. Xiaojing Fan et al. (2024):-xLSTM outperforms standard LSTM for long-term stock price predictionNot peer-reviewed, requires further validation.

4. Sonali Anted et al. (2023):-Linear Regression is effective for trend prediction but not for volatile markets,Cannot capture complex patterns.

5.Payal Soni, YogyaTewari, Deepa Krishnan (2022):-ML models outperform traditional statistical models; ensemble learning improves accuracy,Lack of real-time adaptability, dataset dependency.

6.Waist Khan et al. (2022):-Social media and news influence stock market predictions; ensemble methods improve accuracy, Accuracy varies based on data source reliability.

7.G. Bala Krishna et al. (2022):-LSTM models provide better stock price predictions compared to other ML algorithms,Requires large historical data for better accuracy.

8.Sidra Mehtab, Jaydip Sen, Abhishek Dutta (2021):-LSTM-based models with walk-forward validation improved prediction accuracy,Limited to NIFTY 50 index, requires hyperparameter tuning.

9.Zixin Hu, Yuqi Zhao, Matloob Khushi (2021):-Found that hybrid models (LSTM + DNN) outperform standalone models,High computational cost, requires extensive data preprocessing.

10.I Nyoman Sirayan et al. (2025):-GRU achieves lowest RMSE, Minmax Scaler improves performance

by 22.57%, Computationally expensive, scaling method choice affects accuracy.

3. PROJECT SCOPE

This project aims to develop a machine learning-based model for predicting stock prices by analysing historical data, market trends, and external factors such as news sentiment and macroeconomic indicators. Stock price data will be collected from reliable financial sources, and preprocessing techniques will be applied to handle missing values, remove inconsistencies, and standardize datasets. By leveraging past trends and key financial indicators, the model will generate insights that help investors anticipate future price movements with greater accuracy.

To enhance prediction accuracy, the project will integrate both machine learning and deep learning techniques. Traditional machine learning algorithms such as Linear Regression, Support Vector Machine (SVM), Random Forest, and XGBoost will be used for predictive analysis. Additionally, advanced deep learning models like Long Short-Term Memory (LSTM), Recurrent Neural Networks (RNN), and Transformer architectures will be implemented to improve time-series forecasting, allowing the model to capture complex patterns and long-term dependencies in stock price fluctuations.

A key aspect of the project is the development of a user-friendly web-based application that enables traders and investors to input stock details and obtain real-time forecasts. The model's performance will be assessed using evaluation metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) to ensure accurate and reliable predictions. By providing an interactive platform with intuitive visualizations, the project aims to assist users in making informed investment decisions and managing financial risks effectively.

Despite technological advancements, stock price prediction remains a complex challenge due to unforeseen market fluctuations caused by economic downturns, political events, and investor sentiment shifts. To further improve prediction accuracy, future enhancements may include real-time data streaming, integration of financial sentiment analysis, and reinforcement learning models for continuous learning and adaptation. Ultimately, this project seeks to establish a robust and dependable forecasting system that benefits investors, financial analysts, and researchers in their decision-making process

4.METHODOLOGY

The approach for predicting stock prices takes a systematic route, beginning with gathering data from trustworthy platforms like Yahoo Finance, Google Finance, and Alpha Vantage. The dataset comprises essential financial metrics including opening and closing prices, trade volume, and technical metrics such as moving averages and the Relative Strength Index (RSI). Furthermore, external elements like news sentiment and macroeconomic indicators are included to improve prediction accuracy. After gathering the data, preprocessing is conducted to address missing values, standardize numerical data, and create new features that may enhance the model's predictive performance. The dataset is subsequently divided into training and testing sets, guaranteeing that the model is trained on past data and assessed with unfamiliar data.

For selecting models, different methods are evaluated, including conventional statistical models like ARIMA for forecasting time series and GARCH for analysing volatility. Models such as Random Forest, Support Vector Machines (SVM), and XGBoost are utilized to identify intricate relationships in stock data. Deep learning models, especially Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), and Gated Recurrent Units (GRUs), are utilized to examine sequential data and identify long-term dependencies in stock price fluctuations. Hybrid models that integrate LSTM with CNN or XGBoost are investigated to enhance forecasting precision. The models go through training and refinement, during which hyperparameter adjustment is carried out using methods such as Grid Search and Random Search. The loss function, commonly Mean Squared Error (MSE) or Mean Absolute Error (MAE), is reduced by employing optimizers like Adam.

4.2 Purpose system

The proposed system aims to develop an advanced stock price prediction model using machine learning and deep learning techniques. This system will analyze historical stock price data, technical

indicators, and external factors such as financial news sentiment to improve forecasting accuracy. By leveraging artificial intelligence (AI), the model will provide investors and traders with reliable insights for decision-making.

The system will follow a structured workflow, starting with data collection from sources like Yahoo Finance and Google Finance. It will preprocess the data by handling missing values, normalizing price

4.2 System Architecture

The diagram represents a stock price prediction workflow using an LSSVR model. It involves fetching company data, training the model, predicting future stock prices, visualizing results, and saving the model.

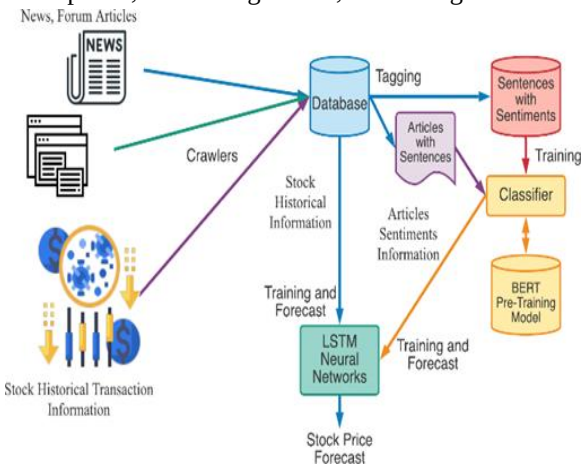
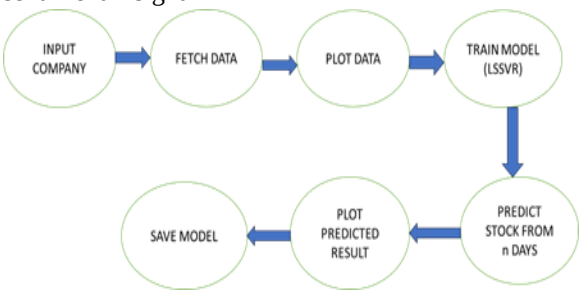


Fig.4.2 System Architecture

5.DETAILSOF DESIGN, WORKING AND PROCESS

5.1 Data flow diagram

The diagram illustrates a stock price prediction system using LSTM neural networks and sentiment analysis. It gathers stock data and news articles, processes sentiments using a BERT-based classifier, and forecasts stock prices based on historical trends and sentiment insight



5.2 Unified modelling language

The diagram depicts a stock price prediction process where moving averages (20, 50, and 200 days) are computed to identify trends. These trends are then used in Prediction Rule Ensembles and a DNN model to generate final predictions.

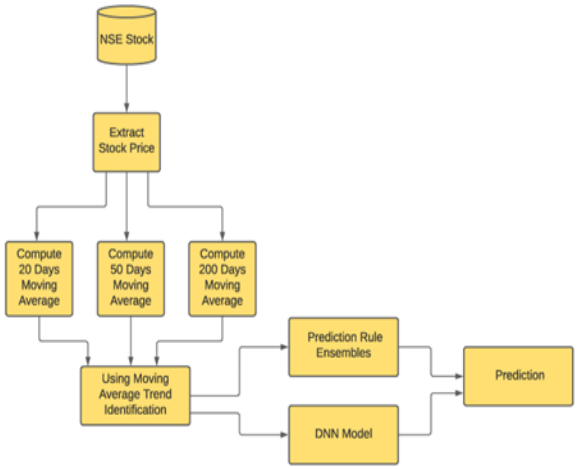


Fig 5.2.1 Class Diagram

5.3 Working of Project

The stock price forecasting project utilizes an organized method that combines data gathering, preprocessing, model training, assessment, and deployment. The process starts by gathering historical stock data from sources such as Yahoo Finance, Google Finance, and Alpha Vantage. This information encompasses Open, High, Low, Close (OHLC) prices, trading volume, and technical indicators like Moving Averages, RSI, and MACD. Furthermore, APIs from Twitter, Bloomberg, and Reuters are utilized to collect market news and social media sentiment, improving prediction precision. After the data is gathered, it goes through preprocessing, which addresses missing values, applies feature scaling, and creates new features such as volatility indices and sentiment scores. The sanitized dataset is subsequently divided into training (80%) and testing (20%) sets for assessing the model's performance.

For predicting stock prices, different models are utilized, such as conventional statistical models like ARIMA and GARCH, machine learning models like Random Forest and SVM, along with deep learning models including LSTM, CNN, and GRU. Hybrid models, like LSTM integrated with CNN or XGBoost, are also investigated for enhanced

6. RESULTS AND APPLICATIONS

The given chart illustrates the stock price movement of AAPL over a specific period, with buy and sell signals overlaid. The blue line represents the stock's closing price, fluctuating between 239 and 241 USD. Green triangles indicate buy signals, suggesting an expected upward trend, while red triangles indicate sell signals, predicting a possible decline. These signals are likely generated using a technical analysis model, possibly incorporating indicators such as moving averages or momentum-based algorithms. This visualization helps traders and investors make informed decisions by identifying optimal entry and exit points in the market.

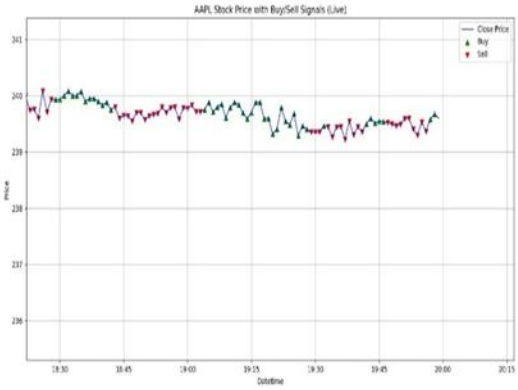


Fig. no. 6.1 stock graph (Buy-sell)

7CONCLUSION AND FUTURE SCOPE

7.1Conclusion

The project on stock price prediction showcases the efficiency of machine learning and deep learning models in anticipating stock market movements. Utilizing past stock data, technical metrics, and sentiment evaluation, the initiative improves forecasting precision, assisting investors in making knowledgeable choices. Different models such as ARIMA, Random Forest, SVM, LSTM, CNN, and GRU were evaluated, with deep learning methods like LSTM and hybrid models demonstrating greater effectiveness in identifying intricate market patterns. The model's effectiveness was assessed with metrics like MSE, RMSE, and MAPE, guaranteeing both reliability and precision. Ultimately, the launch of a web-based application enables users to obtain real-time stock price forecasts, rendering the system functional and scalable. Although it has its strengths, the model's accuracy could be enhanced by integrating real-time data streaming, sophisticated feature selection, and reinforcement learning methods. In summary, this initiative delivers a useful resource for traders and investors, presenting data-informed perspectives to manoeuvre through the stock market more effectively

7.2 Future Scope

The potential for stock price prediction is significant, with improvements in artificial intelligence (AI), deep learning, and quantum computing anticipated to boost prediction precision. The combination of real-time data analysis and high-frequency trading algorithms will facilitate quicker and more accurate decision-making. Additionally, integrating different data sources such as social media sentiment analysis, news trends, and worldwide economic indicators can enhance model effectiveness. Blockchain technology can additionally aid by providing transparent and secure financial information for examination. In general, forthcoming advancements will result in stronger, more flexible, and smarter stock market prediction systems, lowering investment risks and enhancing returns

REFERENCE

[1] Hu, Z., Zhao, Y., & Khushi, M. (2021). A Survey of Forex and Stock Price Prediction Using Deep Learning. Applied System Innovation.

[2] Mehtab, S., Sen, J., & Dutta, A. (2021). Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models. SV-LNCS.

[3] Soni, P., Tewari, Y., & Krishnan, D. (2022). Machine Learning Approaches in Stock Price Prediction: A Systematic Review. Journal of Physics: Conference Series.

[4] Antad, S. et al. (2023). Stock Price Prediction Website Using Linear Regression. ITM Web of Conferences.

[5] Sui, M. et al. (2024). An Ensemble Approach to Stock Price Prediction. Preprints.org.

[6] Nayak, S., Dash, R., & Majhi, B. (2020). Deep learning techniques for stock price prediction. Journal of Forecasting.

[7] Chakraborty, C., & Joseph, A. (2017). Machine learning at central banks. Bank of England.

[8] Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. Journal of Finance.
[9] Brownlee, J. (2018). Introduction to Time Series Forecasting with Python. Machine Learning Mastery.
[10] Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting.