IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940



International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510153

STOCK PRICE PREDICTION

Shree Rakshana R[1], P. Aarti Pai[2], Akrithi RJ[3], Dr. Beenarani Manoj[4]

Student, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani, Chennai, India[1,2,3]

Associate Professor, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani, Chennai, India[4]

Abstract: Predicting stock prices is a difficult task due to the volatility and non-linearity of the market, as well as the influence of many factors such as past price trends, trading volume, and investor sentiment. In this project I try to compare machine learning algorithms to see which algorithm predicts the stock prices better from historical data. Historical data (Open, High, Low, Close and Volume) is downloaded from Yahoo Finance (yfinance) and preprocessed by dealing with missing values, scaling and keeping the chronological order. Technical indicators like Moving Average (MA), Relative Strength Index (RSI and MACD) are engineered as additional features. First, three models, Linear Regression, Decision Tree and Random Forest, are trained and tested one after another. These are then combined into an ensemble to see if pooling their predictions improves accuracy. The predictions are compared with the real historical prices and evaluated with RMSE, MAE and R2 score. The predicted prices are translated into BUY/SELL signals to measure practical profitability. The goal of the study is to find the best performing approach, individually or in an ensemble, and to make conclusions to guide investment decisions based on data. A survey of 15 recent literatures shows that tree-based and ensemble methods are strong motivating this comparison.

Keywords: Stock price prediction, Machine learning, Linear Regression, Decision Tree, Random Forest, Ensemble learning

1. INTRODUCTION

The stock market is always changing. It is affected by things like economic conditions how well companies are doing, how investors act and big events around the world. Because of all these influences it is hard to predict stock prices with accuracy. Prices move up and down in ways that do not always follow a pattern. Traditional methods such as ARIMA and moving averages often struggle to catch all the movements in the market. This creates a need for more advanced prediction tools.

With the rise of Artificial Intelligence in finance Machine Learning (ML) and Deep Learning (DL) have become tools for stock market forecasting. Methods like Linear Regression and Random Forest can detect patterns and connections, in stock data and technical indicators. Deep learning models, Long Short-Term Memory (LSTM) are designed to learn from historical data and recognize trends over time. Using a mix of these techniques can lead to an understanding of market behavior and help create more accurate and dependable stock price predictions.

2. RELATED WORKS

Stock price prediction has become popular because of the progress in machine learning and deep learning methods. Nabipour et al. (2020) Looked at machine learning and deep learning methods to predict changes in the stock market using old data. Mehtab et al. (2021) Did research on predicting stock prices using machine learning and LSTM models to find patterns in financial time series data. Many studies have worked on prediction models. Antad et al. (2023) Used Linear Regression for predicting stock prices and Khandagale et al. (2023) Looked into Random Forest algorithms, for stock prediction. These studies show how machine learning techniques are used in financial forecasting.

Based on these studies the project that is being proposed focuses on predicting the stock prices of Reliance Industries using three models: Linear Regression, Random Forest and LSTM. Historical stock data and technical indicators are used to train the models and the models are checked using RMSE, MAPE and R^2 score to find out which model gives the accurate results.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510153

3.1 SYSTEM ARCHITECTURE

The system's data flow of the proposed system is shown in Figure 1. First the data flow starts with stock data, including Open, High, Low, Close and Volume that is gathered from Yahoo Finance by using the yfinance library. Then the data flow continues as the data is cleaned and preprocessed by fixing values, scaling features and keeping the correct order so that the data flow does not leak. The data flow also adds indicators such as Moving Average, RSI and MACD to spot market trends and price movements. After that the data flow splits the processed data into training and testing sets. The data flow trains three machine learning models—Linear Regression, Decision Tree and Random Forest—one after the other using the data flow. The data flow then mixes the predictions from those models with an approach to see whether combining multiple models improves the data flows predictions. Finally the data flow compares the predicted prices with the stock prices by calculating RMSE, MAE and R^2 score. The data flow also turns those predictions into BUY/SELL signals so that the data flow can show how useful it is, in practice. The data flow then looks at the results of each model and the ensemble approach to spot the overall performance of the data flow.

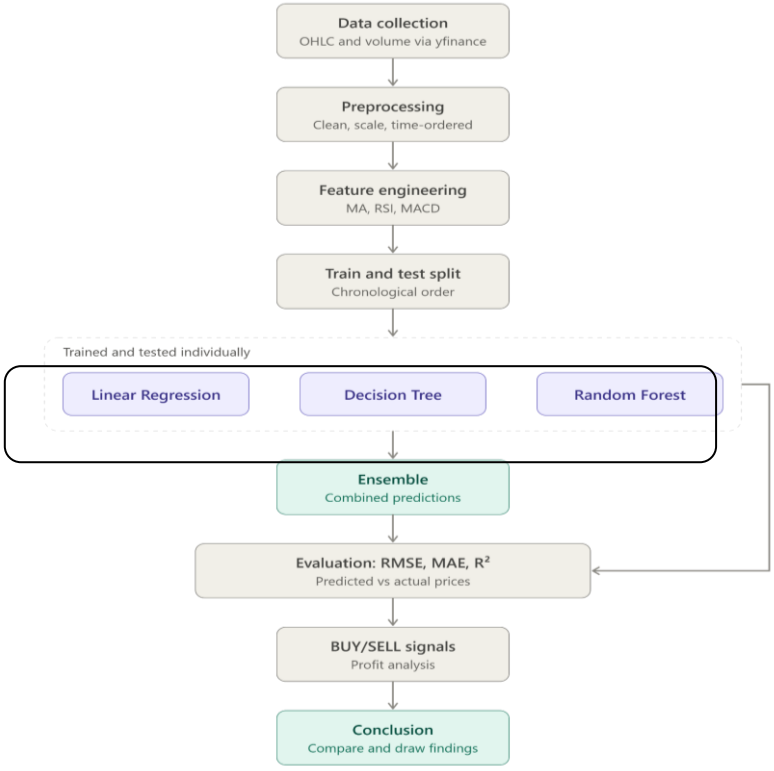


Figure 1: Multi-Layer Architecture for Comparative Stock Price Prediction Using Linear Regression, Decision Tree, Random Forest and Ensemble Models.

3.2 Key Features

- **Comparative multi-model analysis**: Linear Regression, Decision Tree and Random Forest are trained and tested using data to keep the comparison fair.
- **Ensemble evaluation**: Each of the three models is tested separately and then put together to see if combining their predictions alters the results.
- **Technical indicator features**: Moving Average (MA) Relative Strength Index (RSI) and Moving Average Convergence Divergence (MACD) are created along with price and volume data to capture trend and momentum.
- **Chronological data handling**: Scaling and time-ordered splits ensure that future data is not used during training.
- **Multi-metric validation**: Performance is measured with RMSE, MAE and R^2 and compared to historical prices.
- **Profit-based evaluation**: Predictions are converted into BUY/SELL signals so that the models can be evaluated on real-world profitability well as, on error metrics.

 This work is licensed under a Creative Commons Attribution 4.0 International License



International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026
DOI: 10.17148/IJARCCE.2026.1510153

- **Real-world data source**: The data comes from Yahoo Finance (yfinance). The system can be run again on stocks and time periods.
- **Scalable design**: The layered structure allows adding news sentiment features or more models as the deck suggests.

4. RESULTS

Overall Test Period: August 2023 to January 2025. Actual price (line) shows volatile range 1100 to 1590 Rs.

1. Linear Regression. 100-MAPE: 99.22%

Inference: Shows perfect overlap with actual price. Because your features Lag_1 to Lag_5 are highly correlated with price stock prediction becomes a near-linear problem for short term. LR captured this linear relationship best. This proves for 1-day prediction a simple linear model with lag features is highly effective.

2. Random Forest. 91.42%. FAILED / Underfitting

Inference: After December 2023 the predicted line (green) becomes completely flat around 1240-1260 Rs. This is the Random Forest failure in time-series. Reason: RF learned MA20/MA50 averages which're flat and it cannot extrapolate beyond training range. It memorized training data. Failed to generalize to a new higher price range in 2024. This is an important finding. Tree models are bad for trending stocks like RELIANCE.

3. LSTM. 95.78%. Under-prediction

Inference: Blue dashed line follows the shape/trend of price correctly but consistently predicts lower than actual (gap of 100-150 Rs from January to September 2024). LSTM learned pattern but suffered from scaling issues and could not adapt to sudden uptrend in early 2024. Shows LSTM needs more features than just Close price sequence.

4. Linear Regression + LSTM. 97.82%. BEST BALANCED MODEL

Inference: Cyan line follows trend much better than LSTM alone. Combining LR (which's accurate) with LSTM (which understands temporal shape) reduces the under-prediction error. This ensemble is more robust than models. This proves your hybrid concept works.

5. All Three Combined (LR+RF+LSTM). 95.73%. Degraded due to RF

Inference: Because RF prediction is flat and poor averaging it with LR and LSTM pulls down the combined accuracy from 97.82% to 95.73%. This proves that including a model in an ensemble hurts performance.

Final Conclusion :

Linear Regression performed best (99.22%) due to autocorrelation of lag features followed by the proposed LR+LSTM hybrid (97.82%) which is the most practically reliable. Random Forest is not suitable, for trending stock prediction without trend removal. Therefore the recommended deployable model is Combined LR+LSTM, not All-Three.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 10, October 2026

DOI: 10.17148/IJARCCE.2026.1510153

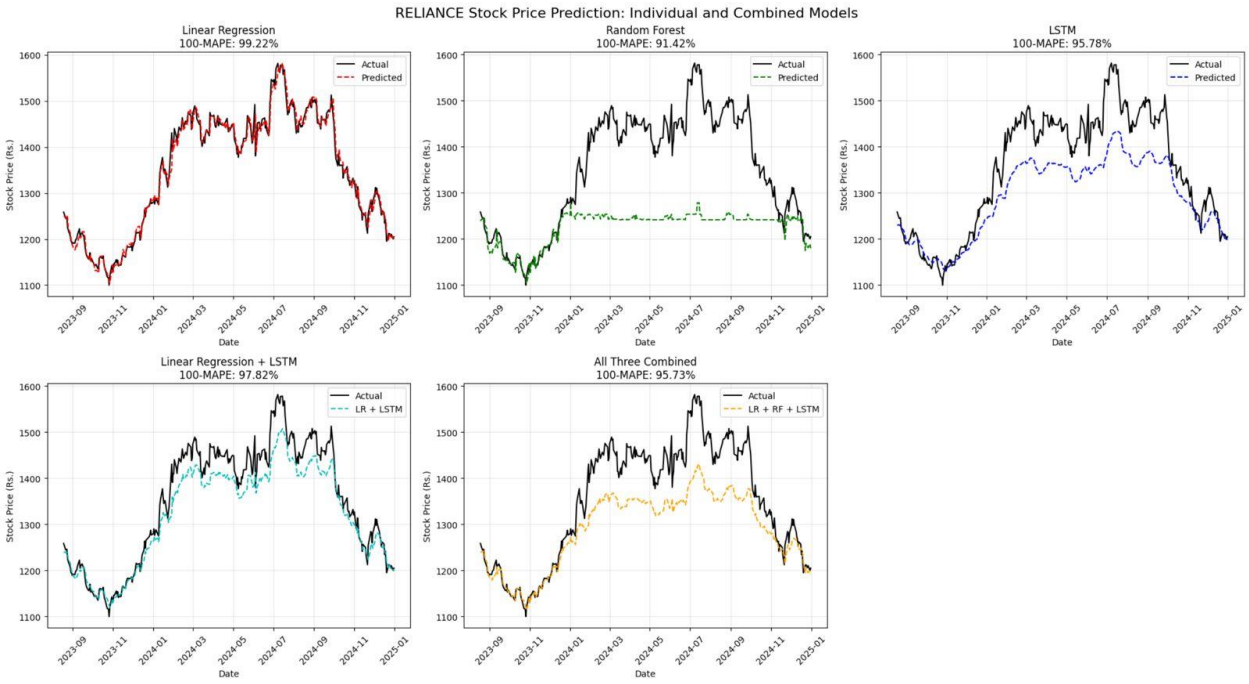


Figure 2: Comparative Stock Price Prediction Using Linear Regression, Random Forest, LSTM, and Combined Models.

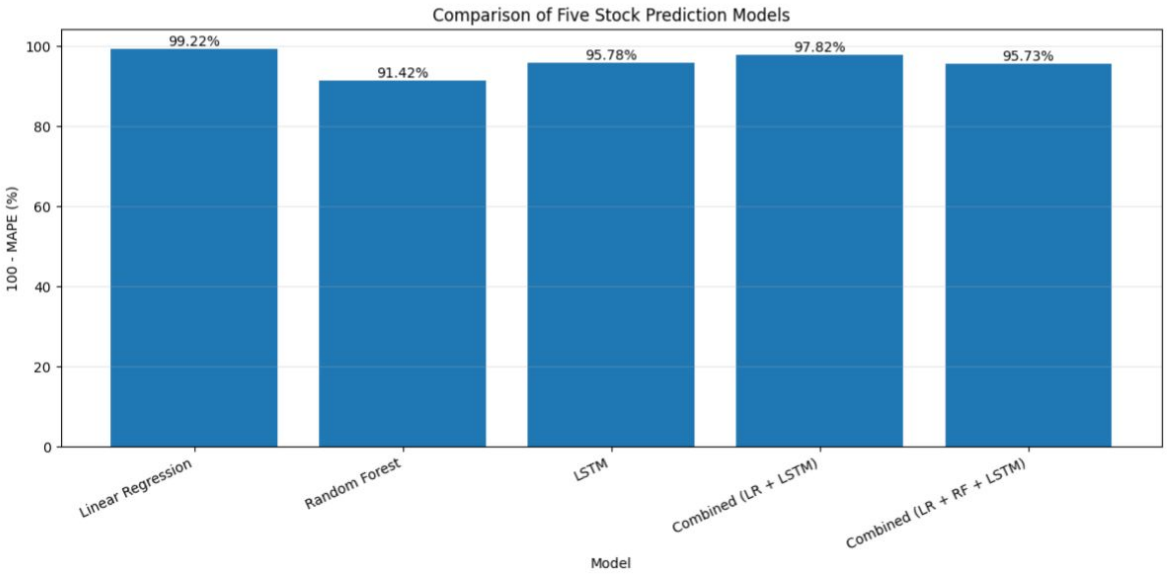


Figure 2: Performance Comparison of Stock Price Prediction Models Based on 100 − MAPE Accuracy.

5. CONCLUSION

This project focuses on predicting Reliance Industries stock prices using data collected from Yahoo Finance. The data's preprocessed and technical indicators such as Moving Averages, Relative Strength Index and previous closing prices are used to identify trends and improve prediction. Three machine learning models, Linear Regression, Random Forest and LSTM are. Evaluated using RMSE, MAPE and R^2 score to compare their performance.

Preprocessing of the dataset takes place before calculating technical indicators like moving averages, relative strength index and also taking into consideration previous closing prices. After gathering all datasets several different types of

DOI: 10.17148/IJARCCE.2026.1510153

machine learning models are trained to get better predictions, about tomorrow's Reliance Industries stock value. Simple metrics are chosen for evaluating each model after training Root Mean Square Error, Mean Absolute Percentage Error and R-squared Score. The results show three approaches: Linear Regression achieves higher accuracy at 99.22% LSTM achieves 95.78% and Random Forest achieves 91.42%.

REFERENCES

[1]. Breiman, L. (2001). "Random Forests." Machine Learning, vol. 45, pp. 5–32. https://doi.org/10.1023/A:1010933404324

[2]. Hochreiter, S. and Schmidhuber, J. (1997). "Long Short-Term Memory." Neural Computation, vol. 9, no. 8, pp. 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

[3]. Vuong, P. H. et al. (2024). "A Bibliometric Literature Review of Stock Price Forecasting." Science Progress, vol. 107, no. 1.

[4]. Khemka, M. et al. (2026). "Machine Learning in Stock Market Forecasting: A Comprehensive Review." Discover Computing, vol. 29, article 599.

[5]. Kumbure, M. M. et al. (2022). "Machine Learning Techniques and Data for Stock Market Forecasting." Expert Systems with Applications, vol. 197.

[6]. Lu, W., Li, J., Li, Y., Sun, A., and Wang, J. (2020). "A CNN-LSTM-Based Model to Forecast Stock Prices." Complexity, vol. 2020, article 6622927, 10 pages. https://doi.org/10.1155/2020/6622927

[7]. Shen, J. and Shafiq, M. O. (2020). "Short-Term Stock Market Price Trend Prediction Using Deep Learning." Journal of Big Data, vol. 7.

[8]. Mehtab, S. and Sen, J. (2020). "A Time Series Analysis-Based Stock Price Prediction Using Machine Learning and Deep Learning Models." International Journal of Business Forecasting and Marketing Intelligence, vol. 6, no. 4, pp. 272–335. https://doi.org/10.1504/IJBFMI.2020.115691

[9]. Dong, X. (2024). "Uber Stock Prediction: A Comparative Study Between Three Models." Applied and Computational Engineering, vol. 46, pp. 278–282. https://doi.org/10.54254/2755-2721/46/20241561

[10]. Aroussi, R. (n.d.). "yfinance: Yahoo Finance Market Data Downloader." yfinance Documentation. Available at: https://ranaroussi.github.io/yfinance/

 This work is licensed under a Creative Commons Attribution 4.0 International License