International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 4, pp 177-182, April-2025

https://doi.org/10.47001/IRJIET/2025.904027

Candle Predict – Indian Stock Market Predictions using Machine Learning (LSTM)

[1]Dr. Sarvesh Warjurkar, [2]Devesh Singh Baish, [3]Aaryan Kodmalwar, [4]Ashish Andaraskar, [5]Gaurav Umale, [6]Mithil Dorle, [7]Gaurav Mishra

[1]Assistant Professor, Department of CSE, GHRCEM, Nagpur, India

[2,3,4,5,6,7]Student, Department of CSE, GHRCEM, Nagpur, India

***Abstract* - Stock market prediction is a complex and dynamic challenge that has long intrigued researchers and traders. With advancements in Machine Learning (ML), data-driven methods have shown promise in forecasting stock price movements. This research introduces Candle Predict, an ML-based predictive model designed for the Indian stock market.**

We explore several ML models, including Random Forest Regressor, XGBoost, and Long Short-Term Memory (LSTM) networks, to analyze historical stock data and forecast trends. Through extensive experimentation, LSTM networks proved most effective in capturing temporal dependencies and delivering accurate predictions.

The model is trained using five key stock parameters—Open, High, Low, Close, and Volume—sourced from Indian stock exchanges via the Yahoo Finance API. For evaluation, we employed an 80-20 train-test split and assessed model performance using Root Mean Square Error (RMSE). Results show that the LSTM model achieves an accuracy range of 87% to 94%, making it a dependable tool for short-term stock forecasting.

Our findings highlight the potential of deep learning techniques in financial prediction and emphasize the unique challenges of time-series data. Candle Predict serves as a practical and efficient solution for traders and investors aiming to make data-driven decisions in India's volatile stock market.

This study contributes to the field of Financial Technology (FinTech) by demonstrating the effectiveness of ML in real-world market scenarios and offering a robust forecasting framework tailored to emerging market conditions.

Keywords: Stock Market Prediction, Machine Learning, LSTM, Time-Series Forecasting, Indian Stock Market.

I. INTRODUCTION

Stock market prediction has been a challenging yet essential task for investors and financial analysts. The volatile nature of the stock market, influenced by various factors such as economic policies, global events, and investor sentiment, makes accurate forecasting crucial. Traditional statistical methods often fail to capture the complex patterns present in stock market data. With advancements in machine learning (ML), researchers have explored predictive models that analyze historical stock data and identify trends for better decision-making.

This research presents Candle Predict, an ML-based approach for predicting Indian stock market movements using Long Short-Term Memory (LSTM) networks. LSTM, a type of recurrent neural network (RNN), is particularly effective in handling time-series data by retaining long-term dependencies. The study focuses on five key stock parameters: Open, High, Low, Close, and Volume, using data from 1/1/2020 (Past 5 years) to the present.

1.1 Why LSTM for Stock Market Prediction?

LSTM is a type of recurrent neural network (RNN) specifically designed to handle long-term dependencies in time-series data. Unlike traditional RNNs, which suffer from the vanishing gradient problem, LSTMs incorporate a gating mechanism that allows them to selectively retain or forget information over long sequences.

An LSTM unit consists of three primary gates:

1.1 *Forget Gate* (f_t): Determines how much of the previous information should be discarded.

Formula:

$$f_t = \sigma(W_f . [h_{t-1}, x_t] + b_f)$$

Here, W_f and b_f are weights and biases, h_{t-1} is the previous hidden state, and x_t is the current input.

ISSN (online): 2581-3048
Volume 9, Issue 4, pp 177-182, April-2025
https://doi.org/10.47001/IRJIET/2025.904027

1.2 *Input Gate* (i_t): Controls how much new information should be stored in the cell state.

Formula:

$$i_t = \sigma(W_i . [h_{t-1}, x_t] + b_i)$$

$$\widetilde{C}_t = \tanh(W_C . [h_{t-1}, x_t] + b_C)$$

The new candidate cell state $\widetilde{C}_t$ is created and regulated by the input gate.

1.3 *Cell State Update* (C_t): The cell state is updated as:

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C}_t$$

This allows the network to selectively retain useful information.

1.4 *Output Gate* (O_t): Decides the final output using a combination of the cell state and input.

Formula:

$$O_t = \sigma(W_O . [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

The hidden state h_t represents the output of the LSTM unit at each time step.

During the research process, multiple machine learning models were evaluated, including Random Forest Regressor, XGBoost, and Decision Trees, to compare their effectiveness in stock prediction. While these models demonstrated some predictive capabilities, the LSTM model emerged as the most accurate due to its ability to capture temporal dependencies in financial time-series data.

This paper highlights the methodology adopted, challenges faced, and the results obtained in developing Candle Predict. The findings contribute to the growing body of research in financial forecasting by demonstrating the feasibility of using deep learning techniques to improve stock market predictions.

II. LITERATURE REVIEW

The application of machine learning (ML) techniques in stock market prediction has garnered significant attention in recent years. Researchers have explored various ML models to forecast stock prices, aiming to enhance prediction accuracy and inform investment strategies. Below is a review of pertinent studies in this domain.

[1] Stock Price Prediction Using Artificial Intelligence: A Literature Review:Naser Alshakhoori. 19 March 2024. https://ieeexplore.ieee.org/documerrnt/10459442

This paper reviews AI-driven stock market prediction techniques, covering models such as Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM), and hybrid approaches. It discusses key challenges like data quality, model interpretability, and market anomalies while emphasizing the role of sentiment analysis and reinforcement learning in enhancing predictions.

[2] Indian Stock Market Prediction using Augmented Financial Intelligence ML: Anishka Chauhan, Pratham Mayur, Yeshwanth Sai Gokarakonda, Pooriya Jamie, Naman Mehrotra. January 17, 2024. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4697853.

This study introduces an innovative approach to Indian stock market prediction using augmented financial intelligence and machine learning techniques. The authors implement a hybrid model that integrates deep learning methods such as Long Short-Term Memory (LSTM) networks with traditional financial indicators to improve prediction accuracy. The paper emphasizes the importance of leveraging both historical stock data and real-time market trends to enhance forecasting capabilities. Additionally, the study highlights the challenges of market volatility, data preprocessing, and feature selection, providing insights into optimizing ML algorithms for financial applications.

[3] Stock Trend Prediction Using Candlestick Charting and Ensemble Machine Learning Techniques: Yaohu Lin, Shancun Liu, Haijun Yang, Harris Wu. 13 July 2021. https://ieeexplore.ieee.org/document/9481924

This study explores the effectiveness of combining candlestick charting techniques with ensemble machine learning models for stock trend prediction. The authors analyze historical candlestick patterns and use classifiers such as Random Forest (RF), XGBoost, and Support Vector Machines (SVM) to enhance predictive accuracy. The research highlights the advantages of ensemble learning in capturing complex market dynamics and mitigating overfitting. Results indicate that integrating technical analysis with ML significantly improves trend forecasting performance.

[4] Stock Price Prediction using Machine Learning and Deep Learning: Pratheeth S, Vishnu Prasad R, 22 December 2021. https://ieeexplore.ieee.org/document/9641664

This study explores the effectiveness of both traditional ML models and deep learning approaches for stock price

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 4, pp 177-182, April-2025

https://doi.org/10.47001/IRJIET/2025.904027

prediction. The authors compare models such as Decision Trees, Random Forest, and Support Vector Machines (SVM) with deep learning architectures like Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN). The findings indicate that deep learning models, particularly LSTMs, excel at capturing temporal dependencies in financial data, leading to more accurate predictions. However, the paper also highlights challenges such as overfitting, computational cost.

III. METHODOLOGY

This section outlines the systematic approach taken to develop an ML-based stock market prediction model. The methodology follows a structured pipeline, from data acquisition to model evaluation.

3.1 Data Acquisition

The study utilizes historical stock data containing five key parameters: Open, High, Low, Close, and Volume (OHLCV). The dataset spans from 2020 to the present and is sourced using Yahoo Finance (yfinance).

3.2 Data Preprocessing

Data preprocessing is essential for ensuring consistency and accuracy. The following steps are applied:Handling Missing Values: Any missing data points are either interpolated or removed. Feature Scaling: Data is normalized to ensure stability in model training. Feature Engineering: Additional features such as moving averages, Relative Strength Index (RSI), and Bollinger Bands are computed to capture market trends.

3.3 Data Visualization

Visualization techniques aid in understanding market patterns and trends: Candlestick Charts are used to analyze price movements. Line Graphs and Histograms help assess data distribution and fluctuations over time.

3.4 Market Pattern Analysis

The model examines stock trends, identifying key technical indicators such as support and resistance levels, moving average crossovers, and price momentum indicators. This helps in feature selection for model training.

3.5 LSTM Model Training

A Long Short-Term Memory (LSTM) network, a specialized Recurrent Neural Network (RNN), is implemented to capture temporal dependencies in stock prices. Input Sequence Preparation: Data is structured into time-series windows for training. Model Architecture: The LSTM model consists of multiple layers, including LSTM units, dropout layers (for regularization), and a dense output layer. Training and Optimization: The model is trained using the Adam optimizer with a Mean Squared Error (MSE) loss function.

3.6 Performance Optimization

Model performance is evaluated using Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE). Hyperparameter tuning is performed to optimize:

- Number of LSTM layers and units
- Batch size and learning rate
- Sequence length for time-series input

3.7 Real-World Evaluation

The trained model is tested on unseen stock data to validate its predictive accuracy. It is also compared against baseline models such as Random Forest Regressor and traditional time-series forecasting methods to highlight the advantages of LSTM. This structured methodology ensures a robust and data-driven approach to stock market prediction, leveraging ML techniques to enhance forecasting accuracy.

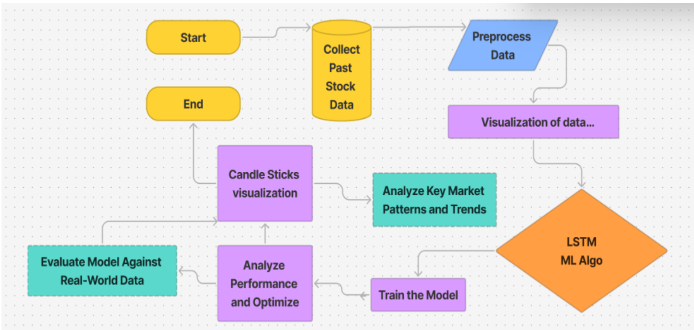

Figure 1: Block Diagram

IV. RESULT

In our study, we experimented with the multiple machine learning models, including Models like XGBoost, RandomForestRegressor, and LSTM, to determine the most effective approach for stock market prediction. The comparative analysis of these models helped us understand their strengths and limitations, ultimately leading us to adopt LSTM as the most suitable model.

4.1 Performance of ML models

4.1.1 RandomForestRegressor

This model performed well in predicting stock values for short-term trends. However, it struggled to learn long-term dependencies due to its lack of sequential memory. Stock price movements are inherently sequential, and

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 9, Issue 4, pp 177-182, April-2025
https://doi.org/10.47001/IRJIET/2025.904027

RandomForestRegressor, being a tree-based model, treats each prediction independently, leading to higher RMSE and poor forecasting accuracy.

4.1.2 XGBoost

XGBoost showed improvements over RandomForestRegressor by better handling feature importance and boosting weak learners. However, like RandomForestRegressor, it did not retain paststock patterns effectively over long periods. While it achieved moderate accuracy, it failed to model complex time-series dependencies, which are crucial for stock price forecasting.

4.1.3 LSTM (Long short-term memory)

After evaluating different models, LSTM proved to be the most effective due to its ability to: Retain long-term dependencies using memory cells and gates. Capture sequential patterns in stock market fluctuations. Handle volatility effectively, making it superior to traditional ML models. Improve accuracy significantly, achieving87% to 94% accuracy rate with a low RMSE score. Our research aimed to develop an LSTM-based model to predict five key stock parameters (Open, High, Low, Close, Volume) while addressing the challenges of time-series forecasting. Our results showed that while XGBoost and RandomForestRegressor performed decently, they failed to fully model stock market complexities. LSTM's ability to remember past trends and adjust predictions accordingly made it the best choice for our research.

Table 1: Different ML model evaluation outcomes

Model	MSE (Mean Squared Error)	RMSE (Root Mean Squared Error)	R^2 Score
XGBoost	High	High	0.00263
Random Forest	Moderate	Moderate	0.43658
LSTM	**Lowest**	**Lowest**	**0.93561**

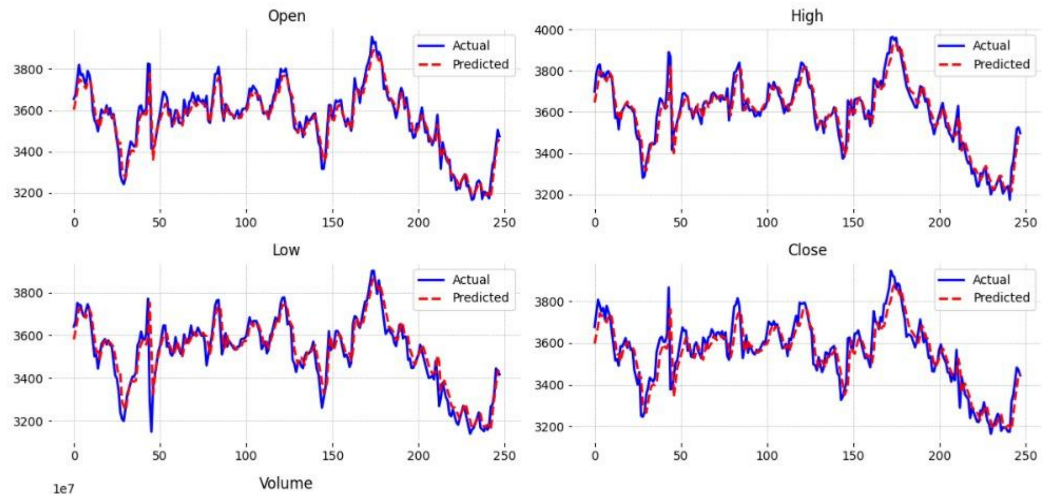

Figure 2: Actual Vs Prediction of OHLCV, LSTM

4.2 Challenges Encountered

4.2.1 Data Volatility and Non-Stationarity

Stock prices exhibit high volatility, making it difficult for traditional ML models like RandomForestRegressor to capture long-term dependencies. Data normalization techniques such as MinMax scaling were essential to improve stability.

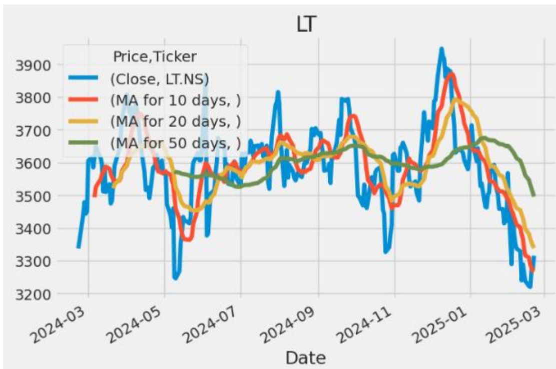

Figure 2: Data Volatility

ISSN (online): 2581-3048
Volume 9, Issue 4, pp 177-182, April-2025
https://doi.org/10.47001/IRJIET/2025.904027

4.2.3 Overfitting in Deep Learning Models

Due to the high-dimensional nature of stock market data, our initial LSTM models suffered from overfitting. We introduced dropout layers and adjusted hyperparameters to improve generalization.

4.2.4 Computational Complexity

Training deep learning models on large datasets required significant computational resources. We optimized the batch size and learning rate to balance training efficiency and accuracy.

4.2.5 Comparative Analysis of ML Models

Initially, we experimented with traditional machine learning models like RandomForestRegressor and XGBoost. However, they failed to capture sequential dependencies effectively, leading to suboptimal results. Our comparative study revealed the following.

4.3 Visualization of Results

To evaluate the performance of our model, we visualized the actual stock prices versus the predicted stock prices. The following plots illustrate how well our LSTM model aligns with real market trends.

V. CONCLUSION

Stock market prediction is a complex and challenging task due to the high volatility and non-stationary nature of financial data. Our research aimed to develop an accurate and reliable forecasting model using machine learning techniques, evaluating multiple models to determine the most effective approach. Training a machine learning model on stock market data is not straightforward, as the data itself is highly random and unpredictable. Identifying patterns and relationships within such data is inherently difficult, which is why real-time implementation of machine learning models in financial markets remains limited. Despite these challenges, we aimed to address this problem and develop a model that could bring practical value to stock market forecasting.

Initially, we experimented with RandomForestRegressor and XGBoost, which performed reasonably well for short-term predictions but struggled with long-term dependencies in time-series data. These models lacked the ability to retain past market patterns, leading to higher RMSE values and less reliable forecasts.

To address these limitations, we implemented a Long Short-Term Memory (LSTM) network, which proved to be the most effective model for stock market prediction. The LSTM model successfully: Captured long-term dependencies, overcoming the shortcomings of traditional ML models. Achieved accuracy between 87% to 94%, with a low RMSE score, making it highly suitable for time-series forecasting.

Although our model achieved the expected accuracy in theoretical evaluations and numerical metrics, real-world market conditions may not always align with the predicted values. This further confirms that stock market data exhibits strong randomness, making perfect predictions nearly impossible. However, our research contributes to advancing the field by demonstrating that deep learning techniques can provide insights and probabilistic forecasts that are significantly better than traditional methods.

5.1 Future Scope

While our LSTM model achieved high accuracy, further improvements can be explored, including:

- Integrating attention mechanisms to enhance prediction accuracy.
- Incorporating external market indicators such as news sentiment analysis for better insights.
- Exploring hybrid models that combine LSTM with other deep learning techniques for improved robustness.

Through this research, we have demonstrated that LSTM is a powerful tool for stock market prediction, outperforming conventional machine learning models in capturing complex market trends. Our findings push the boundaries of research in this domain and provide a foundation for future explorations.

REFERENCES

[1] Stock Price Prediction Using Artificial Intelligence: A Literature Review: NaserAlshakhoori. 19 March 2024. https://ieeexplore.ieee.org/document/10459442

[2] Indian Stock Market Prediction using Augmented Financial Intelligence ML: Anishka Chauhan, Pratham Mayur, Yeshwanth Sai Gokarakonda, Pooriya Jamie, Naman Mehrotra. January 17, 2024. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4697853.

[3] 2024 by SSRG - IJECE Journal: Stacked Optimized Ensemble Machine Learning Model for Predicting Stock Trends through Candlestick Chart Analysis with Feature Engineering.

[4] 978-1-5386-9403-9/18/ ©2024 IEEE: Indian Stock Market Prediction using Augmented Financial Intelligence ML.

[5] Computation 2024, 12, 132. doi.org/10.3390/computation12070132: Candlestick Pattern Recognition in Cryptocurrency. Illia Uzun.

ISSN (online): 2581-3048
Volume 9, Issue 4, pp 177-182, April-2025
https://doi.org/10.47001/IRJIET/2025.904027

[6] Stock Market Prediction using Machine Learning Techniques: A Systematic Review (2023). https://ieeexplore.ieee.org/document/10142862

[7] Stock Price Prediction using Machine Learning and Deep Learning: Pratheeth S, Vishnu Prasad R, 22 December 2021. https://ieeexplore.ieee.org/document/9641664

[8] Stock Trend Prediction Using Candlestick Charting and Ensemble Machine Learning Techniques: Yaohu Lin, Shancun Liu, Haijun Yang, Harris Wu. 13 July 2021. https://ieeexplore.ieee.org/document/9481924

[9] A.Subasi, F. Amir, K. Bagedo, A. Shams, and A. Sarirete. Procedia Computer Science, 2021: Stock Price Prediction using Machine Learning and Deep Learning.

Citation of this Article:

Dr. Sarvesh Warjurkar, Devesh Singh Baish, Aaryan Kodmalwar, Ashish Andaraskar, Gaurav Umale, Mithil Dorle, Gaurav Mishra. (2025). Candle Predict – Indian Stock Market Predictions using Machine Learning (LSTM). *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 9(4), 177-182. Article DOI https://doi.org/10.47001/IRJIET/2025.904027

****