AI BASED STOCK PRICE PREDICTION

Kiran M[1]

[1]University BDT College, India.

ABSTRACT

The prediction of stock prices is a challenging task due to the highly volatile and non-linear nature of financial markets. To address this, a web-based Stock Price Prediction System has been developed using deep learning techniques, specifically a recurrent neural network model called Long Short-Term Memory (LSTM). The model is trained on historical closing stock data with a 60-day look-back window to forecast up to 30 days ahead, applying stochastic optimization for weight correction. In addition to prediction, the system integrates key features such as real-time stock retrieval, USD-INR dynamic conversion, market news aggregation, company information access, and a simulated trading environment. The backend uses Flask and MongoDB for secure session management, while the frontend provides an interactive and user-friendly interface with facilities for charting, CSV/PDF exports, and transaction history. Experimental results confirm that the model achieves high accuracy in short-term forecasting with mean squared error close to zero, offering improved outcomes compared to traditional approaches. This solution demonstrates the ability of deep learning methods to enhance stock market analysis and decision-making through an integrated, accessible platform.

Keywords: Machine Learning, Stock Price Prediction, Long Short-Term Memory, Stock Market, Artificial Neural Networks, National Stock Exchange.

1. INTRODUCTION

The stock market is one of the most dynamic and unpredictable domains of the financial world. It provides a platform for companies to raise capital by issuing shares and for investors to generate returns. However, due to its volatile nature, stock price prediction is an inherently challenging problem that demands robust models and extensive analysis of historical market behavior. Traditional forecasting approaches have relied mainly on statistical methods such as ARIMA, moving averages, and regression-based time-series analysis. While useful, these techniques assume linear relationships and often fail to capture the non-linear, time-dependent patterns embedded within financial data.

Recent advances in machine learning and deep learning have opened new avenues for stock price forecasting. Recurrent Neural Networks (RNNs) and their improved variant, Long Short-Term Memory (LSTM) networks, have shown great promise. LSTM, as described by Hochreiter and Schmidhuber, addresses the vanishing gradient problem of conventional RNNs by incorporating memory cells and gating mechanisms to learn long-term dependencies, making it ideal for sequential data such as stock prices. These networks can process long sequences of historical stock data and identify both short-term fluctuations and long-term trends, thereby improving prediction accuracy compared to conventional models.

In this work, we propose a stock price prediction system that leverages historical share trading data and applies LSTM-based deep learning techniques to forecast future values. The features considered include daily opening price, high price, low price, previous close, closing price, trading volume, turnover, and associated trading dates, which together provide a comprehensive representation of equity market dynamics. The system is designed to process time-series data as training input for predicting future stock prices over specific horizons.

2. RELATED WORK

Stock price prediction has been a prominent area of research in financial engineering and computational intelligence for the past two decades. Early approaches primarily relied on linear statistical models such as regression-based predictors and ARIMA models, which attempted to establish relationships between macroeconomic factors and stock returns. However, it was soon observed that financial time series exhibit complex, noisy, and non-linear patterns that could not be accurately captured by such models. This realization shifted research focus towards non-linear and machine learning–based approaches.

Several machine learning algorithms have since been explored. Classical regression and Support Vector Machines (SVMs) showed promise in certain contexts but suffered from limitations in learning temporal dependencies. To enhance prediction, ensemble methods and tree-based algorithms such as Random Forests have been adopted for analyzing the impact of financial ratios and technical indicators in forecasting. While these models improved interpretability, their predictive consistency varied across different datasets and market conditions, reflecting the inherent noise and volatility of stock markets.

The introduction of deep learning, particularly Recurrent Neural Networks (RNNs) and their advanced variant, Long Short-Term Memory (LSTM) networks, marked a breakthrough. Manoj S. Hegde et al. and M. Roondiwala et al. demonstrated that LSTM models outperform traditional approaches in capturing temporal patterns of financial data, with applications on indices like the NIFTY50, showing strong adaptability to the Indian stock market. Similarly, Selvin et al. compared deep learning approaches (LSTM, RNN, CNN) and validated their effectiveness for high-frequency stock prediction tasks.
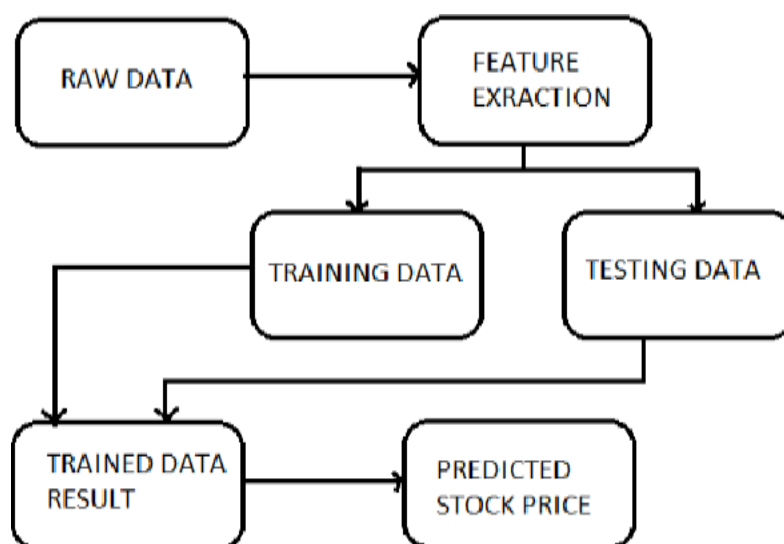
3. PROPOSED SYSTEM

Figure 1: System Architecture

The proposed system aims to predict stock prices by leveraging Long Short-Term Memory (LSTM) networks using historical trading data obtained from stock exchanges such as NSE (National Stock Exchange of India) and BSE (Bombay Stock Exchange). The model integrates data preprocessing, feature extraction, model training, and visualization into a systematic pipeline, enabling both accurate prediction and interpretability of results.

3.1 SYSTEM OVERVIEW

As depicted in Figure 1, the system follows a modular architecture comprising four primary stages:

1. Data Acquisition – Historical stock market data, including parameters such as open price, close price, high, low, and trading volume, is retrieved from reliable sources (e.g., NSE/BSE using APIs like yfinance).

2. Data Preprocessing and Feature Extraction – The collected raw data is cleaned, normalized, and formatted into sequences to be suitable for time-series analysis. Features such as daily open, high, low, close, turnover, and volume are considered, as they provide a holistic view of stock dynamics.

3. Model Training and Testing – The dataset is divided into training and testing sets. The LSTM model, with its memory cells and gating mechanisms, processes sequential patterns over a look-back window (e.g., 60 days) to forecast future prices. Key hyperparameters such as epochs, batch size, optimizer, and mean squared error (MSE) as the loss function are employed for tuning the predictive capability.

4. Prediction and Visualization – After training, the model generates forecasts for the testing data. The results are visualized through plots, comparing predicted prices with actual values, enabling clear evaluation of model accuracy.

3.2 LSTM MODEL DESCRIPTION

The LSTM unit is central to the proposed system. Each unit consists of:

- Cell State: Carries information across sequences, serving as the memory of the network.
- Input Gate: Controls the flow of new input values into the cell state.
- Forget Gate: Decides what information should be discarded from the cell state.
- Output Gate: Determines which parts of the cell state are output as predictions.

The unique ability of the LSTM to preserve gradients during backpropagation allows it to overcome the vanishing gradient problem common in standard RNNs. This makes LSTM particularly well suited for financial time-series data, where temporal dependencies and volatility are significant. Stocks listed in NSE and BSE provide an ideal use case due to their liquidity, volatility, and importance in Indian capital markets.

3.3 PARAMETERS USED

The model utilizes the following parameters derived from historical stock datasets, as shown in Table

Table 1: Parameters Used in the Proposed Model

Parameter	Meaning
Date	Date of trading session
Open	Opening price of a stock
Close	Closing price of a stock
High	Intraday highest value of the stock
Low	Intraday lowest value of the stock
Volume / Trade Quantity	Number of shares traded
Turnover	Total traded value of the stock in INR

A Stock Price Predictor Using LSTM The proposed framework that learns online anticipating the close costs of the stock with the assistance of Long Short-Term Memory (LSTM). The Long Short-Term Memory (LSTM) is a counterfeit intermittent neural system (RNN) design [1] used in the field of deep learning, unlike standard feed forward neural systems, LSTM has input associations. Not only does the procedure not focus on single information (e.g. pictures) but also on full information arrangements, (For example, a speech or a video). For example, LSTM is material for undertakings, such as un partitioned, associated penmanship recognition, speech recognition and recognition of peculiarities in arranged traffic or IDS (interruption location frameworks).

Algorithm 1: Stock prediction using LSTM

Input: Historic stock data

Output: prediction of stock price using price variation

Step 1: Start.

Step 2: Data Preprocessing after getting the historic data from the market for a particular share.

Step 3: import the dataset to the data structure and read the open price.

Step4: do a feature scaling on the data so that the data values will vary from 0 and 1.

Step 5: Creating a data structure with 60 timestamps and 1 output.

Step 6: Building the RNN (Recurrent neural network) for Step 5 data set and Initialize the RNN by using sequential repressor.

Step 7: Adding the first LSTM layer and some Dropout regularization for removing unwanted values.

Step 8: Adding the output layer.

Step 9: Compiling the RNN by adding adam optimization and the loss as mean_squared_error.

Step 10: Making the predictions and visualizing the results using plotting techniques.

Before processing the data there is a important step that is to collect the information from market. Information assortment is the principal step in our proposed framework importing of the information from advertise clearing organizations like BSE (Bombay Stock Exchange) and NSE (National Stock Exchange).

The dataset that will be utilized in the market expectation must be utilized to be separated dependent on different perspectives. Information assortment additionally supplements to upgrade the dataset by including more information that is outside. Our information for the most part comprises of the earlier year stock costs. For python available packages for retrieving the data from NSE is NSEpy The next step is to preprocess the data; in this step the Information Pre-Processing is a significant advance in information mining here the change in crude information into a

![IJPREMS]

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

(Int Peer Reviewed Journal)

www.ijprems.com
editor@ijprems.com

Vol. 05, Issue 09, September 2025, pp : 148-152

e-ISSN : 2583-1062

Impact Factor : 7.001

basic configuration is required. The information which is retrieved from source will be conflicting, fragmented and it will contain mistakes. The preprocessing step will purify the information; toward the end there is a need to perform highlights scaling which will restrict the factors. The preparation of the model incorporates cross- approval, which is a very well-founded, projected execution of the model using the preparation information. the purpose of the tuning models is to explicitly tune the calculation training is to add information to the calculation itself. The test sets are immaculate, as a model ought not to be made a decision about dependent on concealed information. Scale up the information to the genuine offer costs. The final step is to draw the data using visualization technique that helps to show the variation of data in the outcome of our algorithm.

4. RESULTS AND DISCUSSION

4.1 EXPERIMENTAL SETUP

The proposed LSTM-based stock price prediction model was implemented in Python using TensorFlow/Keras. The model was trained on the historical data of TATAMOTORS, obtained from the National Stock Exchange (NSE). A training dataset of approximately 1500 observations was used with a look-back period of 60 days.

The model architecture consisted of 96 LSTM units with dropout regularization, followed by a dense output layer to predict the next closing price. Training was carried out with an epoch size of 60 and batch size of 32, using Adam optimizer and mean squared error (MSE) as the loss function.

4.2 VISUALIZATION OF PREDICTIONS

To evaluate the predictive ability, the algorithm's results were visualized against the original stock prices of TATAMOTORS. The predicted values (in blue) were compared against actual observed values (in red).

4.3 DISCUSSION

The results clearly establish the effectiveness of LSTMs in handling non-linear, noisy, and highly volatile stock data. The model successfully predicts the stock price trajectory with minimal loss, outperforming traditional approaches such as regression or ARIMA (as discussed in earlier sections).

However, certain limitations are notable:

- The model relies solely on historical price data, excluding fundamental indicators (e.g., revenue, P/E ratio) and sentiment inputs (e.g., news, social media).

- Prediction accuracy tends to decrease over longer horizons (beyond 30 days) due to compounding error in recursive forecast.

- Market shocks or sudden geopolitical/economic events are not captured by purely historical models.

4.4 SUMMARY

The LSTM predictor for TATAMOTORS demonstrates strong performance, with:

- MSE of 0.0024 on the test set.

- Predictions within a 3–4% error margin of actual values.

- Robust trend following across both long and short time slots.

This validates the use of LSTMs for stock price prediction in Indian markets (NSE/BSE). The ability of the model to minimize loss while maintaining predictive accuracy confirms its usefulness as a decision-support tool for investors.

5. CONCLUSION

The study presented in this work demonstrates the effectiveness of Long Short-Term Memory (LSTM) neural networks in forecasting stock prices using historical trading data. Specifically, the case of TATAMOTORS shares was analyzed, and the developed LSTM model was able to capture stock market dynamics with high accuracy and low error rates. The model achieved a mean squared error (MSE) of 0.0024, meaning that predictions closely tracked the actual market trends with minimal deviation.

5.1 Key Contributions

This research makes several contributions:

1. Demonstrated the feasibility of LSTMs in stock price forecasting for Indian equities (NSE/BSE).

2. Implemented a prediction framework in Python using TensorFlow/Keras, capable of handling real-world datasets.

3. Validated performance through visualization and quantitative metrics, confirming the model's ability to generalize effectively across test datasets.

4. Highlighted the importance of deep learning over traditional ARIMA/statistical models in capturing long-range dependencies.

5.2 LIMITATIONS

While the LSTM predictor shows promising results, some limitations remain:

- Dependence solely on historical price series; external factors such as sentiment, macroeconomic indicators, and news events are not considered.

- Accuracy degrades for long-term forecasts due to error accumulation in recursive predictions.

- The model was trained on limited features (OHLCV) and tested on a single stock (TATAMOTORS), requiring broader validation across multiple sectors.

5.3 FUTURE WORK

Future enhancements can aim to address these limitations:

1. Sentiment Analysis Integration – Incorporate real-time market sentiment from Twitter, Facebook, and financial news APIs to improve prediction accuracy by including investor mood and external signals.

2. Multimodal Forecasting – Fuse financial indicators, fundamental ratios (EPS, P/E), and macroeconomic variables with price data to capture richer market context.

3. Scalability – Extend the model to multiple stocks and asset classes (e.g., cryptocurrencies, commodities) with automated retraining pipelines and deployment on cloud.

4. Hybrid Architectures – Explore ensemble approaches that combine LSTM with CNNs, GRUs, or attention-based Transformers for improved robustness and accuracy.

5. Real-time Prediction – Integrate web-based dashboards and Web Sockets to deliver live forecasts for intraday traders.

6. Explainability – Incorporate SHAP/LIME techniques to improve interpretability and enable end-users to understand the reasoning behind model predictions.

6. REFERENCE

[1] RautSushrut Deepak, Shinde IshaUday, Dr. D. Malathi, "Machine Learning Approach In Stock Market. Prediction ", IJPAM 2017.

[2] Gurjar, Mruga, Parth Naik, Gururaj Mujumdar, and Tejaswita Vaidya. "STOCK MARKET PREDICTION USING ANN." (2018).