Stock Market Trading Assistant

Er.Puneet Kaur, Barenya Behara, Lakshay Verma, Goutom Kumar, Abhay Bansal, Disha Singh
Dept. of Computer Science Chandigarh University Mohali, India

Abstract*—Our paper illustrates the development of a Stock Market Trading Bot that uses machine learning techniques to predict stock market trends and assist traders in making informed decisions. The bot leverages historical market data, candlestick and news forecasts to predict future price movements. It integrates two machine learning models: XGBoost Regressor for gradient boosting with Long Short Term Memory(LSTM prediciton technique) for seasonality, combining their outputs for improved accuracy. The bot processes real-time data from Yahoo Finance and provides predictions through a user-friendly web application. The system's hybrid approach enhances forecast reliability and aids investors in navigating the complexities of stock market fluctuations. This paper highlights the importance of combining traditional technical analysis methods with modern learning algorithms to improve prediction accuracy and help traders make better financial decisions.

Index Terms*—Market Prediction, Machine Learned Models, XGBoost Regressor, Long short term memory methods(LSTM), Financial Forecasting, Gradient Boosting, Time Series Analysis.

I. INTRODUCTION

Trading has existed since the agricultural revolution, or for as long as human civilization. The type of trading is different across different communities. Trading basically means the exchange of goods and services between two parties. Stock trading takes place in the stock market. With investments and online trading, stock markets are now available to a larger section of people. The ability to forecast stock prices and market trends is extremely significant in the rapidly evolv- ing financial markets of today, where billions of dollars are transcated daily. Traders, Financial institutions, and investors always look for ways to get a competitive advantage. These techniques include using cutting-edge technology, conducting in-depth market research, or utilizing the most recent data. But the stock market is a complicated and frequently erratic setting. Numerous things impact it, such as trends in trading data, past price fluctuations, and worldwide events covered by the updates. The problem is in interpreting this massive volume of data to forecast what might occur after that. This study serves as an introduction to our project, which aims to create a stock market prediction bot. Market data is essentially the back- bone of stock trading. The Candlestick patterns are another critical aspect of our prediction model. This bot's prediction of future market patterns is intended to assist investors and traders in making more precise decisions. The bot does this by examining three primary sources of data: current market data, candlestick patterns, and forecasts for upcoming news. Together, these elements create a potent system for negotiating the complicated workings of the stock market. Each of these elements are essential to the bot's prediction-making process.

II. LITERATURE SURVEY

The literature review of the several approaches to stock market trading bots using machine leaning(ML) that has already been put into practice is mentioned below. It provides an overview of the software and the systems currently in use. The existing systems with Machine Learning methods includes Convolutional Network of neurons (CNNs), Recur- rent Network of neurons(RNNs), LSTMs, Genetic Algorithms, Artificial Networks of neurons, Support Vector Mechanisam, Random Forests, judgment Trees, Deep Reinforcement Learn- ing's(DRL's). Proposal provided in the research paper by Medha Mathur, Satyam Mhadalekar, Sahil Mhatre, and Vanita Mane is the use and creation of an algorithmic bot that can automatically make trading decisions on financial markets. It is based upon creating trading techniques that rely on previous data analysis, real-time market conditions, and sophisticated algorithms. Here we apply moving average convergence di- vergence indicator (MACD), index of relative strength (RSI), and maybe more complex methods using machine learning approaches.[2]. A study done by Atharva Shah, Maharshi Gor, Meet Sagar, and Manan Shah suggest a framework for stock market trading that makes use of deep learning algorithms to enhance trading judgements. Different transformer-based models, recurrent networks of neurons(RNNs), convolutional networks of neurons(CNN's), and networks with long short termed memory's(LSTMs) should be implemented and eval- uated in order to find a solution[3]. Some study suggests a deep

network of neurons based stock bots that makes use of parameters for evolutionary optimization in technical analysis to generate buy-sell points. The approach involves using genetic algorithms to optimize Simple Moving Average(SMA) and Index of Relative Strength(RSI), which are then used as inputs for a deep Multilayer Perceptron(MLP) neutral network to predict buy-sell-hold decisions[1]. In order to predict stock prices and making trading decisions research suggests an artificial network of neurons based stock trading environ- ment that makes use of technical analysis indicators and big data frameworks. The technique uses technical indicators like Index of Relative Strength(RSI), MACD, and Williams %R to first create buy-sell-hold-signals from financial time series data. Using the Apache Spark big data platform, an artificial network of neurons model called Multilayered Per- ceptron's(MLP's) is trained using daily stock values. After training, the model is evaluated with data spanning[2]. The research by Hiransha M, Gopalakrishnan E.A., Vijay Krishna Menon, and Soman K.P. aims to improve Market stock price forcast. The testing phase assesses the models and mea- sures accuracy using Mean Absolute Percentage Error which uses deep learning models like Networks of Recurrent Neu- ral's(RNN), Neural Networks using Convolutions(CNN), Net- works with Long Short Termed Memory's(LSTM) ,and Mul- tilayered Perceptron's(MLP). Their performance is compared with traditional linear models like Average Movement(MA), Autoregressive(AR), Self-regressive (AR), shifting average with Self-regressive Properties (ARMA), and selfegressive Integrated Moving Average (ARIMA) [9]. As researched by S.Z. Mahfooz, Iftikhar Ali, and Muhammand N. Khan aims to enhance stock trend prediction by utilizing a network of neurons using Long Short termed Memory's(LSTM) integrated with a four-layered analytical approach. This approach in- cludes technical pattern recognition, professional trader in- sights, Natural Language Processing(NLP), and event impact assessment[10]. Ibn Asyraf Zulkidfey, Mohd Edil Abh Sukor, Ali Fayyaz Munir, and Muhammad Hakimi, Mohd Shafiai's research investigates the application of artificial intelligence (AI) methodologies for spotting stock market manipulation, including spoofing , quote stuffing, and pump and dump schemes. The project aims to enhance conventional rule- based systems by utilising cutting-edge AI techniques like Judgement Trees(DT), Supported Vector Machine's(SVMs), Artificial Network of neurons(ANNs), and clustering. These AI models simulate manipulation situations, analyse trading volumes, examine regulatory gaps, and stock prices using supervised learning[12]. Using a combination of Probit Re- gression Models and Random Forest, the paper by Chihab Younes, Zineb Bousbaa, Marouane Chihab, Omar Bencharef, and Soumia Ziti provide an algorithmic trading approach for intraweek FX speculation. The system outperformed the Probit Regression Models and Random Forest individually, achieving a true positive rate of 78%[8]. In order to automate stock trad- ing, a complex trading bot combines sentiment analysis with deep evolutionary techniques was proposed by Ashish Bali and Archit Madan in their work. The bot processes this sequential data using a recurrent network of neurons (RNN) known as a Long Short termed Memory's(LSTM) model. It does this by using Natural Language Processing(NLP) techniques to anal- yse sentiment data from Twitter and other sites. To increase trading choices, the bot also uses deep evolutionary techniques. The bot is programmed to invest sensibly in equities that are expected to return at least 1% of the initial investment[9]. The reserch by Yasmeen, Sadaf Yasmin and Seungmin Rho provides a system for automatic stock decision making using Deep Reinforced Learning technique(DRL). This system dy- namically adjusts trading strategies by analysing both past and future stock trends, integrating a Gated Recurrent Unit(GRU)- based forecasting model for improved decision-making. The approach continuously updates based on market behaviour, enhancing profitability and adaptability compared to static strategies[3].

III. PROPOSED MODEL

The four steps of the method we suggested are data col- lection and cleaning, data scaling and splitting, training, and execution.

A. Data Gathering and Cleansing

The first and most crucial stage, referred to as preprocessing, is to use Yahoo Finance to get up-to-date stock information. Yahoo Finance provides a CSV file with the price data for the designated stocks. Seven features are retained for the collected data: Date, open, high, low, volume, and adjusted close. Furthermore to the transferring date, opening price, closing price, price at closing, and traded shares, these data also provide the stock's closing price at which investors get their dividends. Yahoo Finance is an open-source website that provides real-time stock

quotes, charts, and market data for many assets. It also offers the most up-to-date financial news analysis for the global market[15]. After gathering the data, we cleaned it up using a range of machine learning reduction approaches, removing irrelevant, repetitive, spammy, and non- sensical material. The preprocessing phase also includes the following[9]:

- Remove missing values: If there are just missing values, we decide to remove specific rows or columns.
- Imputation: Used more sophisticated methods like K- Nearest Neighbours (KNN) imputation, or substituted missing values with a statistical measure like the mean, median and mode.
- Eliminate Duplicates: Find and eliminate entries that duplicates that might skew our analysis or model training.
- Cleaning time series data involves addressing nonexistent timestamps by filling in gaps or interpolating missing periods, as well as smoothing to reduce noise using methods like moving average. This produces a consistent and dependable dataset for analysis and modelling.

B. Data Scaling and Splitting

Out of the entire data set, we are removing two features that are essential to our model: ”Close” and ”Volume.” It is crucial to scale the data after the preprocessing stage. Scaling data is crucial for many machine learning algorithms and models to perform as well as possible. MaxMin Scaling and Standardization are the two most widely used methods for scaling the data. As previously said, scaling is a crucial component of the algorithm that decreases numerical stability, enhances model convergence, and prevents feature dominance. The Holdout method, a popular data-splitting approach, is used in the second portion to divide the cleaned and scaled data into training and testing sets. Using this strategy, data sets are separated into testing and training sets (e.g., 20% and 80%, respectively). Testing sets make it possible to evaluate the performance level of the model’s on fictitious data set.

C. Training Phase

We came up with the strategy of training two models since the most crucial stage is selecting the model and the appropri- ate training technique. Initially, the LSTM’s procedure will be used to derive the market using 100-day and 200-day moving averages in order to obtain a prediction for the following day. Second, we are using the xgb-regressor approach to train for gradient boosting. This will forecast whether the graph will rise or decrease depending on the feature data that we are using to train it. Our unique strategy combines the two to achieve optimal accuracy and forecasts[13]. The total forecast will be provided by:

$$\text{Combined\_preds} = \{ xgb_{prediction[min_{len}]} + lstm_{prediction[min_{len}]} \} / 2$$

Fig. 1. Hybrid Prediction Formula

Now we will see the working of the model for stock market prediction bot:

- LSTM ALGORITHM: Recurrent network of neurons (RNN) design with Long Short Termed Memory’s (LSTM) are especially well-suited for assessing and fore- casting time series or sequential data where the sequence of the data points matters. The ability of long termed dependencies in the data set to be capture by LSTM’s is a critical aspect, as correct predictions frequently depend on this capacity[6]. To depict short-term trends and evaluate the stock’s medium-term trend—which reacts faster to fluctuations in the market rate for that particular stock—we will employ the 100 moving average trend. The long-term trend of a stock is ascertained using the 200-day moving average (MA). It is less vulnerable to abrupt price fluctuations and provides a clearer view of the overall trend. i. A green marker indicates 100 averages. ii. A red marker indicates 200 moving averages.

- XGB REGRESSOR ALGORITHM: XGBoost Regres-sor, often called XGB Regressor, is a machine learning method that can be used to forecast stock prices or other

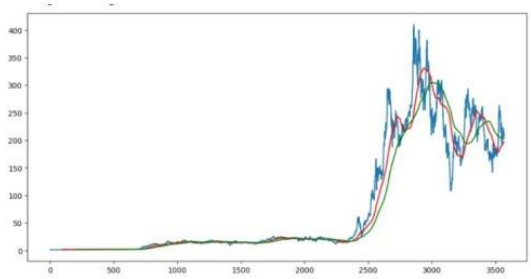

Fig. 2. Graph over 100ma and 200ma

continuous financial variables. Because of its efficiency, speed, and performance, Extreme Gradient Boosting, or XGBoost, is a popular gradient boosting method that performs well in large-scale prediction applications. Framework of Gradient Boosting: Gradient boosting, the basis of XGBoost, creates a powerful prediction model by combining several weak learners, usually decision trees. The process builds trees iteratively, focusing on correcting

the errors from the previous additions with each new tree. A powerful prediction model for financial market analysis may be created when sophisticated machine learning algorithms like XGBoost Regressor are combined with moving averages, such as the 100-day and 200-day mov- ing averages. This hybrid approach leverages the benefits of both traditional technical analysis and modern machine learning approaches to increase forecast resilience and accuracy[7].

D. Execution Phase

We have built a web application for the best user experience in which the customer will get the prediction of the model using the ticker or Alias of the particular stock the model will predict that the prices will fall or rise and customers can also see a visual representation of prediction over the original price rates. We also provide a live Stock market view with the latest updates.

IV.RESULT

The LSTM's (Long Short Termed Memory's) model and the XGBoost Regressor were the two machine learning models used in the stock market prediction bot's implementation. Both models were trained, and then their predictions were integrated to anticipate trends in stock prices. The model employed freshly cleaned and error-free real-time stock data from Yahoo Finance. On the basis of the 100-day and 200-day moving averages, the LSTM model assisted in the prediction of both short- and medium-term trends. In order to further improve the forecast, the XGBoost Regressor applied gradient boosting in addition to these forecasts. Combining the two models increased forecast accuracy to 93.80% and gave investors more trustworthy insights into future fluctuations in stock prices. In addition, a web application was created that lets users input stock tickers and get predictions in real time, as well as charts that show the past and projected patterns of the stock.

Table I: Evaluation Metrics

Metric Parameters	Value
Accuracy	93.80%
Precision	0.95
Recall	0.93

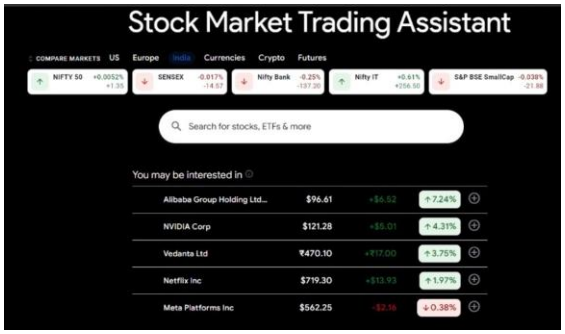

Fig. 3. user interface

V. FUTURE SCOPE

Various future developments in stock market trading bot include integration with advanced data sources, machine learn- ing and natural language processing techniques, expansion into new markets and various strategies for efficiency, risk management active management, and portfolio optimization strategies will all be important. In addition, changes to the user interface and experience, regulatory and compliance pro- cesses, and blockchain and cryptocurrency integration will be required. In addition, environmental, social and governance issues, including impact closure strategies and social respon- sibility activities will add value.

VI.CONCLUSION

A successful strategy for stock market prediction is the fusion of cutting-edge machine learning techniques (LSTM's and XGBoost) with conventional technical analysis methods (moving averages). Investor decision-making can be improved and forecast accuracy can be increased by this hybrid model. The system will be a useful tool for traders Because of its ability to interpret real-time data and provide visual forecasts via an intuitive interface. The bot exhibits potential for making well informed trading decisions; however, more sophisticated algorithms can be integrated and the dataset can be expanded for increased accuracy in a variety of market scenarios.

ACKNOWLEDGMENT

We would like to sincerely thank our supervisor, Mrs. Puneet Kaur, for her crucial advice and assistance during this research. Special thanks to Yahoo Finance for providing real- time stock data, which was essential to the development of our prediction model. Additionally, we appreciate the support from the academic community and the valuable feedback from our peers, which helped in refining this work.

REFERENCES

[1] Sezer, O.B., Ozbayoglu, M. and Dogdu, E., 2017. A deep neural-network based stock trading system based on evolutionary optimized technical analysis parameters. Procedia computer science, 114, pp.473-480.

[2] Sezer, O.B., Ozbayoglu, A.M. and Dogdu, E., 2017, April. An artificial neural network-based stock trading system using technical analysis and big data framework. In proceedings of the southeast conference (pp. 223-226).

[3] Ansari, Y., Yasmin, S., Naz, S., Zaffar, H., Ali, Z., Moon, J. and Rho, S., 2022. A deep reinforcement learning-based decision support system for automated stock market trading. IEEE Access, 10, pp.127469-127501.

[4] Chihab, Y., Bousbaa, Z., Chihab, M., Bencharef, O. and Ziti, S., 2019. Algo-Trading Strategy for intraweek foreign exchange speculation based on random forest and probit regression. Applied Computational Intelligence and Soft Computing, p 2019.

[5] Bali, A., Madan, A., Upadhyay, A., Sah, P. and Nehra, V., 2021. Development of Trading Bot for Stock Prediction Using Evolution Strategy. Preprint, p.6739.

[6] Azhikodan, A.R., Bhat, A.G. and Jadhav, M.V., 2019. Stock trading bot using deep reinforcement learning. In Innovations in Computer Science and Engineering: Proceedings of the Fifth ICICSE 2017 (pp. 41-49). Springer Singapore.

[7] Mathur, M., Mhadalekar, S., Mhatre, S. and Mane, V., 2021. Algorithmic trading bot. In ITM Web of Conferences (Vol. 40, p. 03041). EDP Sciences.

[8] Shah, A., Gor, M., Sagar, M. and Shah, M., 2022. A stock market trading framework based on deep learning architectures. Multimedia Tools and Applications, 81(10), pp.14153-14171.

[9] Zulkifley, M.A., Abd Sukor, M.E., Munir, A.F. and Shafiai, M.H.M., 2021, December. Stock market manipulation detection using artificial intelligence: A concise review. In 2021 international conference on decision aid sciences and application (DASA) (pp. 165-169). IEEE.

[10] Mahfooz, S.Z., Ali, I. and Khan, M.N., 2022. Improving stock trend prediction using LSTM neural network trained on a complex trading strategy. International Journal for Research in Applied Science and Engineering Technology, 10(7), pp.4361-4371.

[11] Gerlein, E.A., McGinnity, M., Belatreche, A. and Coleman, S., 2016. Evaluating machine learning classification for financial trading: An empirical approach. Expert Systems with Applications, 54, pp.193-207.

[12] Sachdeva, A., Jethwani, G., Manjunath, C., Balamurugan, M. and Krishna, A.V., 2019, March. An effective time series analysis for equity market prediction using deep learning model. In 2019 International Conference on Data Science and Communication (IconDSC) (pp. 1-5). IEEE.

[13] Hiransha, M.E.A.G., Gopalakrishnan, E.A., Menon, V.K. and Soman, K.P., 2018. NSE stock market prediction using deep-learning mod- els. Procedia computer science, 132, pp.1351-1362.

[14] Khang, P.Q., Hernes, M., Kuziak, K., Rot, A. and Gryncewicz, W., 2020. Liquidity prediction on Vietnamese stock market using deep learning. Procedia Computer Science, 176, pp.2050-2058.

[15] Rouf, N., Malik, M.B., Arif, T., Sharma, S., Singh, S., Aich, S. and Kim, H.C., 2021. Stock market prediction using machine learning techniques: a decade survey on methodologies, recent developments, and future directions. Electronics, 10(21), p.2717.