International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

An Experiment Using Innovative Ml Model to Predict Stock & Index Prices

[1]Abhishek Kumar, [2]Dr. Naveen Kumar V.

22MBAR0854 || CMS Business School(JU)

ABSTRACT

Stock price prediction is crucial for investors and financial institutions, offering potential benefits through machine learning (ML) techniques. ML can analyse large datasets to uncover intricate patterns, aiding in forecasting future stock movements by examining historical data such as stock prices and trading volumes. Various ML algorithms have been effective in this domain, including linear regression, random forests, and deep learning methods like recurrent neural networks. However, market volatility, data quality, and model complexity affect prediction accuracy. Researchers must pre-process data, select relevant features, and adjust model parameters regularly to adapt to changing market conditions. Despite these challenges, ML-based predictions offer significant advantages, aiding decision-making for buying, selling, and holding stocks while mitigating risks. They also help identify trading opportunities and optimise investment strategies. In summary, ML-driven stock price prediction has the potential to revolutionize stock market analysis, empowering investors to make informed decisions and potentially improve investment outcomes.

Keywords: Stock Price Prediction, Machine Learning, Historical Data

INTRODUCTION

In the dynamic world of finance, the ability to predict stock prices with precision is the holy grail for investors and financial institutions. The advent of machine learning has revolutionized this domain, offering a powerful lens through which vast datasets can be analyzed to reveal intricate patterns that elude traditional analytical methods. This paper delves into the transformative potential of machine learning in stock price prediction, exploring its capabilities to parse historical data and project future market trends with an unprecedented level of sophistication.

The utilization of machine learning techniques—ranging from linear regression and random forests to advanced deep learning networks—has shown promising results in deciphering the complex interplay of factors influencing stock prices. By training on historical stock data, these algorithms unearth correlations and trends that serve as harbingers of future price movements. However, the volatile nature of the stock market, influenced by an array of economic, geopolitical, and psychological factors, presents a formidable challenge to even the most advanced predictive models.

This research paper aims to navigate the intricacies of machine learning applications in stock price prediction, acknowledging the hurdles such as data quality, feature selection, and model complexity while highlighting the profound advantages they hold for market participants. Through meticulous data preprocessing, judicious feature selection, and continuous model refinement, machine learning stands as a beacon of innovation, poised to redefine investment strategies and decision-making processes in the financial sector.

STATEMENT OF PROBLEM

The goal of the research project is to create and improve a strong machine learning-based framework for stock price prediction while taking into account all of the difficulties and complexity that come with it. This comprises collecting pertinent features and previous stock closing price data, pre-processing and cleaning the data to deal with missing values, choosing relevant features, and dividing the data into training and testing sets. Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) are two well-known metrics that are used to assess the performance of machine learning models after they have been trained and fine-tuned. Furthermore, by highlighting the necessity of continual model adaption and refining, the research seeks to address the persistent problem of properly projecting stock values in the context of the dynamic and turbulent character of financial markets. The final objective is to implement a dependable and efficient model that can offer insightful analysis and forecasts for in-the-moment decision-making in investing and stock trading scenarios.

REVIEW OF LITERATURE

For stock market prediction, the authors of [1] examined the behaviour of the market and identified the best fir model among several well-known machine learning techniques, such as Random Forest (RF), Support Vector Machine (SVM), Naïve Bayes, K-Nearest Neighbour (KNN), and SoftMax.

The methodologies were compared by the authors, who also used several technical indicators to the data collected from various sources, such as Yahoo and NSE-India. When the accuracy of each model was examined, it was found that for large datasets, RF produced the most gratifying results. In contrast, Naïve Bayesian produced the highest accuracy for small datasets. Another finding was that the model's accuracy dropped as the number of technical indications decreased.

The paper [2] forecasted stock prices for the next day using TF-IDF characteristics and data from multiple news channels. The word score was calculated by the authors using TF-IDF weights. Ultimately, the probabilities of switching values were included in an HMM model that was created to determine the probability of a sequence. The authors of this model saw a trend of partially matching positive and negative predictions with an error of 0.2 to 4%; however, higher accuracy can be attained by expanding the dataset, using different machine learning algorithms or adding more technical indicators and input features.

Historically, share price forecasting has exclusively used historical data. However, analysts now understand that many other factors play a significant role in influencing the stock price, so depending just on previous data isn't reliable. The authors of the research [3] examine and use a variety of techniques to forecast stock prices, however even after examining the key variables influencing the price of the company, they are unable to reach a high degree of accuracy. Major techniques like SVM, Regression, Random Forest, etc. have been studied by the authors, and they have also examined hybrid models that combine two or more techniques. The authors claim that certain models perform better on historical data than on sentiment data. Results from fusion algorithms had better predictions.

The supervised learning method of linear regression is used in the study [4] to forecast stock prices. The suggested research project essentially describes how to use a dataset to predict the closing value by analysing the GOOGL stock and extracting data spanning over 14 years.

After doing Principal Component Analysis (PCA) on the data to identify the most pertinent components, Meghna Misra et al.'s article concludes that predictions generated using the Linear Regression Models have a higher accuracy rate [5]. A high accuracy rate was seen on a binary classification model using the Random Forest Approach, and the Multilayer Perceptron (MLP) produced the least amount of error when making predictions. SVM shows high accuracy on non-linear classification data. Linear regression is preferred for linear data due to its high confidence value.

As the authors in the paper [6] conclude by studying the application of machine learning models to analyse financial trading and to design optimal strategies, many of the techniques are not just limited to stock price prediction but can be used broadly in the financial markets. Following a quantitative examination of several methods, the authors advise going deeper into behavioural finance to assess investor or market psychology and comprehend changes in the market. The authors suggest using machine learning and text mining techniques to track user behaviour on digital financial trading platforms.

RESEARCH GAP

Although machine learning-based stock price prediction research has advanced significantly, there are still several important areas that need more study. Integrating unstructured data sources, including news articles and social media sentiment, into predictive models is one important area of weakness. While most of the research that has been done so far has focused on structured data—such as past price and volume—including unstructured data might improve forecast accuracy. Interpretable models that provide stakeholders with clear explanations of projections are also necessary. Moreover, the dynamic character of financial markets is frequently ignored in the present research, which makes the creation of models that are flexible enough to adjust to shifting market circumstances necessary. Furthermore, there is a dearth of studies on risk management, quantifying uncertainty, and ethical issues in machine learning- based stock prediction. Finally, there is a lack of real-time prediction systems for dynamic trading settings and long-term forecasting models that can sustain predictive accuracy over lengthy time horizons. Filling up these research voids will advance the field and make stock price prediction models stronger and more accurate.

OBJECTIVE OF THE STUDY

- To leverage historical stock price data to train the machine learning model.
- To develop a machine learning model that can accurately predict future stock prices.

Test of Proposed Model

- To test and compare the accuracy of models under various gradient descent.

RESEARCH METHODOLOGY

This study describes the creation and use of a machine learning model that makes use of the Long Short-Term Memory (LSTM) deep learning algorithm to improve the predictive analytics of stock market values. This study's sequential and methodical methodology starts with the careful gathering of historical stock price data, which is a necessary component of the model's input. This dataset includes important characteristics including opening, closing, high, and low prices in addition to trading volume and other relevant information that may affect stock prices. The data goes through a thorough preparation step to remove any errors or missing values. To guarantee consistency, numerical values are normalised, and to make computational manipulation easier, category attributes are encoded. To identify market trends, the Moving Average indicator is used to analyse the closing price, which has been established as a major predictor for this research. The feature selection procedure is crucial.

After that, a division of the gathered data is carried out to produce separate training and testing sets. The testing set functions as a standard to assess the predictive ability of the LSTM model, whereas the training set provides the basis for model development and refinement. The LSTM algorithm is selected due to its exceptional ability in time series forecasting, a crucial factor in stock price prediction.

The LSTM model is exposed to training data iteratively throughout model training, which enables it to pick up on trends. Hyperparameters are fine-tuned using methods like batch size optimisation to improve the model's efficiency. The testing data is used to thoroughly evaluate the model once it has been trained. To measure the correctness of the model, metrics like Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) are calculated.

This process culminates in the prediction step when future stock prices are predicted using the LSTM model, which is now proficient at identifying patterns and trends. The effectiveness of the model is graphically represented by contrasting the predictions with the actual values. Understanding that the stock market is dynamic, the model is constantly adjusted. To improve forecast accuracy, hyperparameters are recalculated, new characteristics can be added, and other methods are taken into consideration.

When the model's prediction accuracy is high enough, it is used to forecast stock prices in real time. It is crucial to recognise that because there are so many variables at play in the stock market, stock price prediction is by its very nature complicated and volatile. Although machine learning models, such as the one used in this study, can provide insightful analysis and predictions, perfect accuracy cannot be assured. As a result, to respond to changes in the market and enhance the model's performance over time, it is necessary to continue a continuous process of monitoring and updating it. The objective of this study is to contribute to the area by offering a methodological framework that may be improved and repeated for comparable prediction projects.

ANALYSIS AND INTERPRETATION OF THE MODEL

Based on past price data, the LSTM (Long Short-Term Memory) model used for this study seeks to forecast Reliance Industries Limited's (RELIANCE.NS) stock prices. Because LSTM

neural networks are recognised for their capacity to grasp temporal relationships in sequential data, such as time series, the model makes use of deep learning techniques. **Preparing Data and Creating Models:**

Using the Yahoo Finance API, historical stock price data is retrieved, moving averages are computed, training and testing sets are divided, and Min-Max scaling is applied to the data. Eighty percent of the data is the training set and the remaining twenty percent is put aside for testing.

The architecture of the LSTM model is intended to capture intricate temporal patterns in the input sequences. It is made up of many LSTM layers layered on top of one another, with dropout layers added for regularisation to avoid overfitting. Each LSTM layer has a progressively larger number of neurons, which enables the model to learn hierarchical characteristics from the input sequences.

Model Development and Assessment:

The prediction performance of the model is tested using the testing data once it has been trained. To calculate assessment measures like Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and accuracy score, the model's predictions are compared to the actual target values.

The training data is used to train the constructed LSTM model. The model is fed the input sequences (x) and associated target values (y) for a predetermined number of epochs. In the process of training, the model minimises the mean squared error loss function by adjusting its weights using the Adam optimizer.

Findings and Analysis:

The assessment metrics shed light on how well the LSTM model performs:

Mean Absolute Error (MAE): The average absolute difference between the expected and actual stock prices is measured by the Mean Absolute Error (MAE). The mean absolute error (MAE) in this study is roughly 59.53, meaning that the model's predictions often differ from the actual values by about Rs. 59.53.

Mean Squared Error (MSE): The average squared difference between the expected and actual prices is quantified by the Mean Squared Error (MSE). The model's predictions show deviation from the actual prices, with an MSE of around 5478.90.

Root Mean Squared Error (RMSE): The standard deviation of the prediction errors is represented by the RMSE, which is the square root of the MSE. An RMSE of around 74.02 indicates that the standard deviation between the model's predictions and the actual values is approximately Rs. 74.02.

Accuracy Score: This metric assesses how well the model performs in comparison to the actual price mean. The model's performance in forecasting stock prices is strong, with an accuracy score of about 0.8851, or 88.51%.

CONCLUSION

In summary, our study has shown how well machine learning techniques, in particular LSTM neural networks, may be used to predict stock prices. The LSTM model performed admirably when used to analyse past stock price data, correctly predicting future price movements by identifying underlying patterns and temporal relationships in the data.

The results demonstrated how well the LSTM model handled the complexity of financial time series data and how well it generalised to new data, underscoring its resilience and flexibility. Notwithstanding its inherent drawbacks, which include the possibility of overfitting and market volatility, the LSTM model was effective in giving portfolio managers, financial analysts, and investors useful information.

The integration of machine learning into finance theory has theoretical ramifications that underscore a paradigm shift towards data-driven techniques in comprehending market dynamics. Furthermore, the managerial implications delineated the revolutionary capacity of machine learning across multiple domains within the financial industry, such as regulatory compliance, risk reduction, portfolio management, and investment decision-making.

It is imperative to recognise the limits of this research, which include issues with interpretability of the model, overfitting, market uncertainty, and reliance on feature engineering. These drawbacks highlight the necessity of ongoing study and improvement to improve the accuracy and dependability of stock price prediction algorithms.

The paper essentially represents how the field of financial modelling is changing and how machine learning techniques provide new ways to enhance decision-making and obtain an understanding of market dynamics. Stakeholders can more adeptly navigate the intricacies of financial markets by embracing technological developments in machine learning and addressing existing restrictions, which will ultimately result in improved investment results and risk management techniques.

References

[1] Kumar, I., Dogra, K., Utreja, C. and Yadav, P., 2018, April. A comparative study of supervised machine learning algorithms for stock market trend prediction. In 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT) (pp. 1003- 1007). IEEE.

[2] Ingle, V. and Deshmukh, S., 2016, August. Hidden Markov model implementation for prediction of stock prices with TF-IDF features. In Proceedings of the International Conference on Advances in Information Communication Technology & Computing (pp. 1-6).

[3] Sukhman, Tarun Kumar Madan, J. Kumar and A. Singh. “Stock Market Forecasting using Machine Learning: Today and Tomorrow.” 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT) 1 (2019): 738-745.

[4] Pahwa, K. and Agarwal, N., 2019, February. Stock market analysis using supervised machine learning. In 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon) (pp. 197-200). IEEE.

[5] Misra, Meghna, Ajay Prakash Yadav, and Harkiran Kaur. “Stock Market Prediction using Machine Learning Algorithms: A Classification Study.” 2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE) (2018): 2475-2478.

[6] Vats, P. and Samdani, K., 2019, March. Study on Machine Learning Techniques In Financial Markets. In the 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN) (pp. 1-5). IEEE.