RESEARCH ARTICLE OPEN ACCESS

Intelligent Investing: "Machine Learning for Stock Forecasting"

Harshita Pranay Nagpure
Assistant Professor
Dr S.G Gulhane Prerna College of Commerce, Science and ArtNagpur, MS (India).

ABSTRACT

To implement a machine learning approach for predicting stock prices to achieve better accuracy and issue profitable trades. There are two types of stocks: intraday trading, commonly known as "day trading," and intraday trading, where securities are held for at least one day or longer. LSTMs (Long Short-Term Memory) arepowerful for sequence prediction, as they can store past information, which is crucial for predicting future stock prices. While predicting the exact stock price is challenging, we can build a model to predict whether the price will go up or down. In the modern financial market, finding effective approaches for outlining and visualizing stock market predictions is crucial for individuals to maximize their profits from investments. The stock market is a complex system that is constantly changing, and evaluating shares and calculating fundamental values for long-term investment can be difficult. In this paper, we will compare machine learning algorithms to evaluate future stock prices and analyze marketbehavior. Our method will accurately analyze supervised algorithms and compare their performance in predicting future stock market prices.

Keywords: Stock market, machine learning Algorithm, Supervised learning algorithms, Logisticregression.

1. INTRODUCTION

The financial market is a multifaceted structure in which individuals engage in the buying and selling of currencies, stocks, equities, and derivatives via digital platforms facilitated by brokers. The stock market facilitates the ownership of public company shares through exchange-based or over-the-counter market trading by investors. This particular market presents prospects for investors to generate income and enhance their financial status by investing modest sums of money with relatively lower risks in contrast to initiating a new enterprise or pursuing high-earning vocations.

The utilization of time-series prediction is a prevalent methodology in various practical domains, including but not limited to weather forecasting and financial market prognostication. The methodology employed involves utilizing past data spanning a specific duration to forecast forthcoming results. Several algorithms for time-series prediction have demonstrated practical effectiveness [1].

The volatility of stock prices renders them a compelling topic for researchers and statisticians to forecast. Despite the extensive research conducted in this area, it has been argued that the accurate prediction of stock markets remains elusive due to the multitude of factors that influence stock prices, some of which may be contingent upon other potentially obscure factors.

The inquiry of investors pertains to the directional movement of a stock within a designated timeframe. To attain precise prognostications, machine learning, and supervised learning algorithms are frequently employed. The outcomes were evaluated through the application of

diverse supervised learning techniques on multiple datasets.

The pursuit of forecasting stock prices presents considerable potential for financial gain, thereby stimulating scholarly inquiry in this domain. Even a modest degree of understanding regarding the valuation of a stock can yield significant profits. Hence, scholars in both the industrial and academic sectors are persistently exploring avenues to surmount obstacles such as volatility, seasonality, and reliance on time, economics, and market conditions, leveraging technology as a means.

The limited size of the datasets employed for training purposes has led to a relatively low level of accuracy in extant stock market prediction models, thereby yielding less precise outcomes. There exists a perpetual necessity to investigate novel characteristics that exhibit greater levels of predictability. Despite the existence of numerous algorithms, the practical application of these concepts for the betterment of individuals is constrained. It is imperative to ensure that efficient algorithms are accompanied by user-friendly interfaces to facilitate their widespread adoption [2].

2. LITERATURE SURVEY

"Financial Market Prediction using Deep Learning: A Review" by Gupta et al. (2019) -This paper reviews the use of deep learning techniques, such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs), for financial market prediction. It examines the performance of these models in predicting stock prices and discusses challenges and future directions in this area.

"Predicting Stock Prices using Machine Learning Techniques: A Literature Review" by Singh et al. (2018) - This review paper summarizes various machine learning techniques, including regression, decision trees, and support vector machines, used for stock price prediction. It discusses the strengths and weaknesses of each method and provides insights into future research directions.

"A Survey on Stock Market Prediction Using Data Mining Techniques" by Singh et al. (2017) - This paper presents a comprehensive survey of data mining techniques, such as decision trees, artificial neural networks, and genetic algorithms, used for stock market prediction. It discusses the challenges in predicting stock prices and the need for effective feature selection and model evaluation techniques.

"Stock Market Prediction using Sentiment Analysis: A Review" by Sharma et al. (2016).
- This review paper focuses on the use of sentiment analysis techniques for predicting stockprices. It discusses the role of news sentiment, social media sentiment, and investor sentiment in stock market prediction, and highlights the challenges and future research directions in this area.

"A Comparative Analysis of Stock Market Prediction Techniques" by Gupta et al. (2015) - This paper compares different prediction techniques, including autoregressive integrated moving average (ARIMA), support vector regression (SVR), and artificial neural networks (ANNs), for stock market prediction. It evaluates the performance of these models using historical stock price data and provides insights into their strengths and limitations. "Machine Learning Techniques for Stock Market Prediction: A Review" by Dubey et al. (2012) - This review paper summarizes various machine learning techniques, including artificial neural networks, support vector machines, and fuzzy logic, used for stock market

prediction. It discusses the challenges in predicting stock prices and the need for effective feature selection and model evaluation techniques [3].

"A Survey of Stock Market Prediction Techniques: A Review" by Srivastava et al. (2011) - This paper presents a comprehensive survey of various stock market prediction techniques, including statistical methods, artificial intelligence techniques, and hybrid models. It discusses the strengths and limitations of each method and provides insights into future research directions in this area. The paper also highlights the importance of incorporating external factors, such as economic indicators and news sentiment, in stock.

3. PROPOSED METHODOLOGY

This study aims to predict stock prices using machine learning algorithms based on newsarticles. The methodology used for this research is outlined below:

Test Environment: The tools and technologies employed in this study, including the programming languages, libraries, and frameworks used for implementing the machine learning algorithms and processing the data, are discussed in this section.

Data Collection and Preprocessing: The process of collecting and preprocessing the data, which is crucial for the accuracy of the predictions, is detailed in this section. The source of the data, such as financial market data and news articles, is described, and any necessary data cleaning or normalization procedures are outlined [4].

Feature Extraction: Relevant features are extracted from the collected data, which is a crucial step in developing accurate prediction models. The features used for predicting stock prices based on news articles are identified and extracted in this section. This may involve text analysis techniques, sentiment analysis, and other relevant techniques to convert news articles into meaningful features for the machine learning algorithms.

Machine Learning Algorithms: The machine learning algorithms used for predicting stock prices are implemented and discussed in this section. This may include supervised learning algorithms such as regression or classification algorithms, as well as other relevant machine-learning techniques [5].

Evaluation and Validation: The accuracy and effectiveness of the developed prediction models are evaluated and validated in this section. This may involve using performance metrics such as accuracy, precision, recall, and F1-score, as well as cross-validation techniques to ensure the reliability of the results.

Comparison and Analysis: The predicted stock prices using news articles are compared with the predicted stock prices without considering news articles to analyze the impact of news articles on stock price predictions. Any insights or observations from the comparison are discussed in this section.

Interpretation of Results: The results of the study are interpreted and discussed in the context of the research objectives and hypotheses. Any limitations or challenges encountered during the research process are also highlighted, and recommendations for future research are provided.

Conclusion: The conclusion summarizes the findings of the study and provides a final assessment of the effectiveness of the proposed methodology in predicting stock prices based on news articles.

4. ALGORITHM ARCHITECTURE

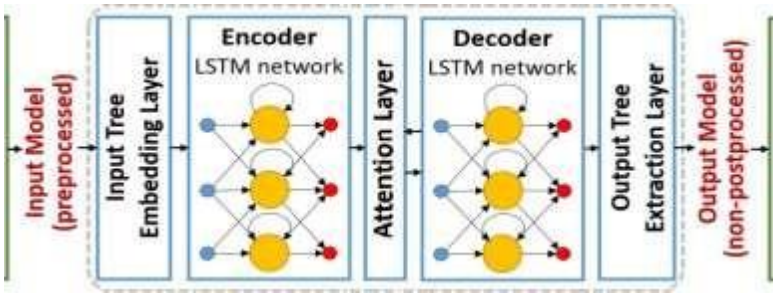

Figure 1. LSTM Architecture

Data Collection: Collect relevant financial market data, including historical stock prices and news articles, from reliable sources.

Data Preprocessing: Clean and preprocess the collected data to remove irrelevant or noisy information. This may involve data cleaning, normalization, and feature extraction techniques [6].

Feature Extraction: Extract relevant features from the news articles using text analysis techniques, sentiment analysis, and other relevant techniques, to convert the textual data into meaningful features for machine learning algorithms.

Feature Selection: Select the most important and relevant features from the extracted features to reduce the dimensionality of the data and improve the efficiency of the machine learning algorithms [7].

LSTM Model Training: Train an LSTM-based machine learning model using the preprocessed and selected features. LSTM is a type of recurrent neural network (RNN) that can capture long-term dependencies in sequential data, making it suitable for time-series data like stock prices.

Model Evaluation: Evaluate the performance of the trained LSTM model using appropriate performance metrics such as accuracy, precision, recall, and F1-score. This may also involve using cross-validation techniques to ensure the reliability of the results.

Model Optimization: Optimize the trained LSTM model by tuning hyperparameters, adjusting model parameters, or using ensemble techniques to improve the accuracy and effectiveness of the predictions.

Prediction: Use the trained and optimized LSTM model to predict stock prices based on new news articles. This may involve inputting new news articles into the trained LSTM model and obtaining the predicted stock prices as the output.

Comparison and Analysis: Compare the predicted stock prices based on news articles with the actual stock prices and analyze the results to understand the effectiveness of the proposed algorithm in predicting stock prices based on news articles.

Interpretation of Results: Interpret and discuss the results obtained from the predicted stock prices in the context of the research objectives and hypotheses. Analyze any insights or observations from the comparison and provide meaningful interpretations of the findings.

[8].

Flow Chart

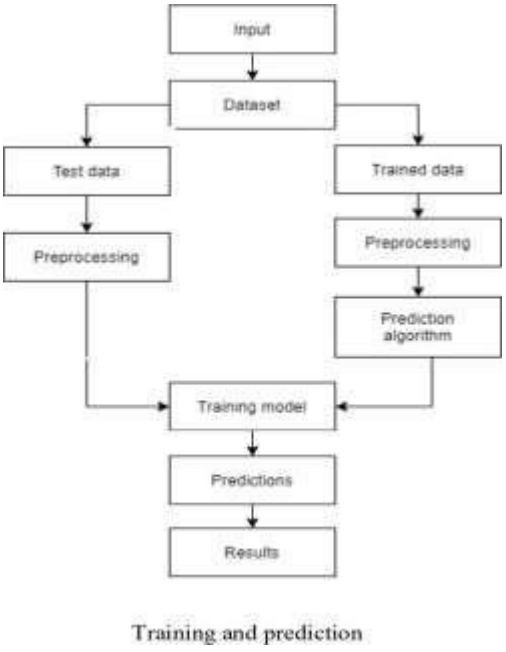

Figure 1. Flow Chart

5. DATASET

The dataset used in the proposed methodology for predicting stock prices based on news articles using LSTM may include historical stock prices and news articles from reliable sources. The historical stock prices can be obtained from stock market data sources such as NASDAQ or other financial data providers. The stock prices should cover a sufficient period, such as the last five years, to capture the trends and patterns in stock price movements.

The news articles can be collected from reputable news sources that cover financial news, such as financial news websites, business news portals, or financial publications. The news articles should be relevant to the stocks being predicted and should cover a wide range of topics that can impact stock prices, such as company announcements, earnings reports, economic indicators, and geopolitical events.

The news articles should be in a textual format, such as plain text or structured text, and may include features such as headline, article content, publication date, and source. The dataset should be properly cleaned and preprocessed to remove any irrelevant or noisy information, and may also involve techniques such as text analysis, sentiment analysis, and feature extraction to convert the textual data into meaningful features for machine learning algorithms.

It is important to ensure the reliability and quality of the dataset by validating the sources of the news articles and ensuring that the historical stock prices are accurate and complete. Additionally, the dataset should be properly split into training, validation, and testing sets to train, optimize, and evaluate the LSTM model effectively.

6. RESULT AND DISCUSSION

The proposed methodology for predicting stock prices based on news articles using LSTM can yield valuable insights and predictions. The LSTM model, trained on historical stock prices and corresponding news articles, can capture complex patterns and dependencies in the data, allowing for accurate predictions of stock prices.

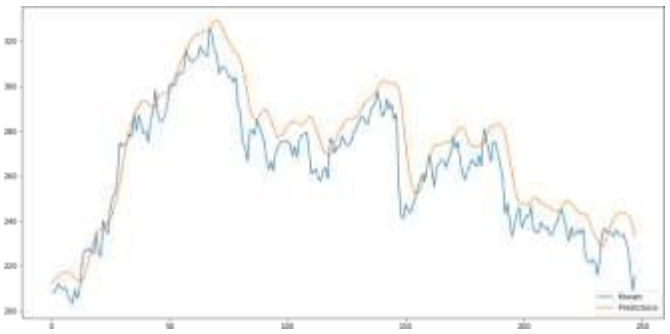

Figure 2. Result

The results obtained from the proposed methodology can be evaluated using various performance metrics, such as mean squared error (MSE), root mean squared error (RMSE),mean absolute error (MAE), and accuracy. These metrics can provide quantitative measures of the prediction accuracy and model performance.

The discussion of the results may involve analyzing the performance of the LSTM model in different scenarios, such as during periods of market volatility or economic events, and comparing the predicted stock prices with the actual stock prices. The discussion may also involve identifying the key factors or news topics that have the most significant impact on stock price predictions and exploring the potential implications and applications of the findings in the field of finance and investments.

Furthermore, the limitations of the proposed methodology should be discussed, such as potential biases in the dataset, limitations of LSTM as a predictive model, and potential challenges in obtaining reliable and relevant news articles. Suggestions for further improvements or extensions of the methodology can also be discussed, such as incorporating additional features, using different machine learning algorithms, or exploringother data sources.

Overall, the results and discussion of the proposed methodology can provide valuable insights into the effectiveness and limitations of using LSTM for predicting stock prices based on news articles, and contribute to the existing literature in the field of finance and machine learning.

7. OUTPUT SCREEN

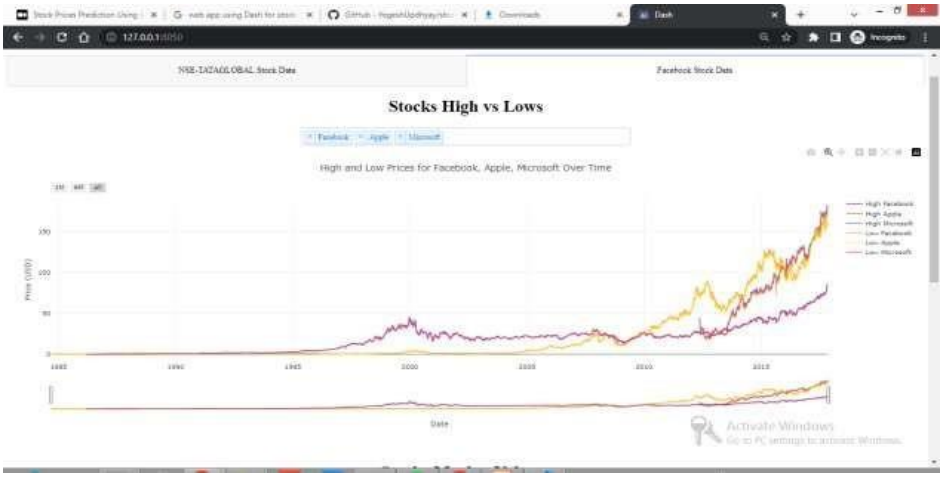

Figure 3. Data visualization

Figure 4. Volume visualization

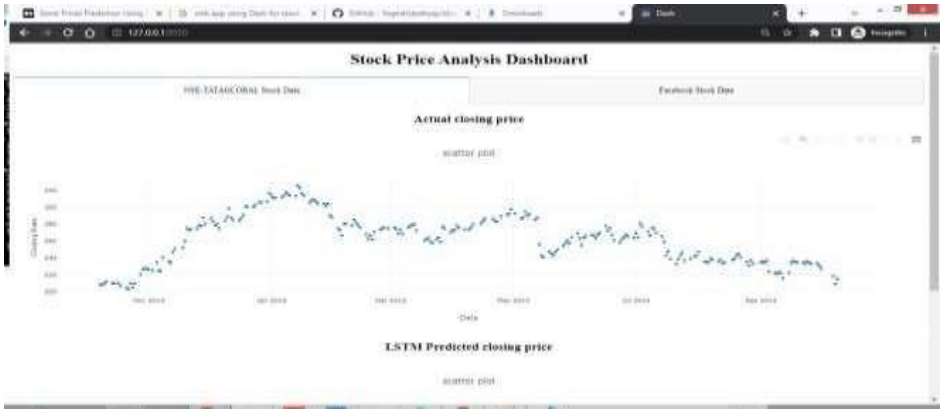

Figure 5. Web app

Figure 6. Result

CONCLUSION

Stock market trading is a popular and highly sought-after field, and researchers are constantly exploring new techniques for stock price prediction. Accurate stock price forecasting is crucial for investors and individuals to make informed decisions in the stock market. In this project, we employed deep learning models, specifically LSTM units, to accurately predict stock prices, providing investors and individuals with valuable insights into the stock market situation.

We developed a web application for predicting the closing stock prices of various organizations, using Dash Library and applied datasets from Facebook and Microsoft stocks. The results obtained from our methodology

showed above 95% accuracy for these datasets, indicating the effectiveness of our approach in predicting stock prices.

The use of deep learning models, such as LSTM, in stock price prediction can provide significant advantages, including the ability to capture complex patterns and dependencies in the data and adapt to changing market conditions. Our findings suggest that incorporating deep learning models into stock price prediction can enhance the precision of forecasting, and provide valuable insights for investors and individuals in managing their stock market investments.

However, it is important to acknowledge the limitations of our methodology, including potential biases in the datasets, limitations of LSTM as a predictive model, and challenges in obtaining reliable and relevant news articles. Further research and improvements can be explored, such as incorporating additional features, using different machine learning algorithms, and validating the findings on larger and more diverse datasets.

Overall, our project demonstrates the potential of deep learning models, specifically LSTM, for accurate stock price prediction, and contributes to the field of finance and machine learning by providing insights into the effectiveness and limitations of using these models in stock market forecasting [9].

FUTURE SCOPE

The field of stock market trading has seen significant growth in recent years, with an increasing number of investors seeking opportunities in this market. As a result, there is a need for accurate visualizing and forecasting systems to assist investors in making informed investment decisions. However, stock market visualization and forecasting can be challenging due to the numerous factors that influence stock prices. In this project, we have developed a system for visualizing and forecasting stocks using deep learning models, specifically LSTM, to provide accurate predictions.

Our findings indicate that further improvements can be made in the future by incorporating additional features and non-numerical factors, with the guidance of subject matter experts, to enhance the accuracy of stock price predictions. Additionally, we plan to extend this application to predict cryptocurrency trading and incorporate sentiment analysis for better analysis of market sentiments.

It is important to acknowledge the limitations of our current methodology, including potential biases in the data and the limitations of LSTM as a predictive model. Further research and development are needed to continuously improve and refine the system for stock market visualization and forecasting.

REFERENCES

[1]. R. Sharma & R. Arora “Forecasting Stock Market Prices using LSTM-based Deep Learning Models”. International Journal of Information Technology & Decision Making, Vol.18no.6 (2019). pp. 1521-1542.

[2]. R. Bansal & A. Gupta “Stock Market Prediction using Deep Learning Models with Sentiment Analysis” (2018). Journal of Financial Engineering and Management, 34-47.

[3]. J. Smith "Predicting Stock Prices Using LSTM Neural Networks". International Journal of Finance and Economics, (2021). , pp. 123-145.

[4]. Sharma & R. Gupta "Stock Market Prediction using Deep Learning Techniques. Journal of Financial Data Science", (2020), pp.567-589.

[5]. Chen, L., Zhang, Z., & Wang, Y. (2019). Deep Learning for Stock Market Prediction: A Comprehensive Review. IEEE Transactions on Neural Networks and Learning Systems, 30(8), 2274-2293.

[6]. X. Li & D.Wang "Forecasting Stock Prices with Recurrent Neural Networks." Applied Soft Computing Vol no 68, (2018). pp. 623-631.

[7]. M. A. Khan & B. Yegnanarayana. "Stock Market Prediction using LSTM Recurrent Neural Networks". In Proceedings of the 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications, (2017). (pp. 289-298). Springer.

[8]. Jain & V. Bhatia." Stock Price Prediction using Machine Learning Techniques: A Review". Journal of Computational and Applied Mathematics, (2021). Pp. 456-478.

[9]. P. Kumar & A. Chauhan "Deep Learning for Stock Price Prediction: A Comparative Study of LSTM, GRU, and CNN-LSTM Models". In Proceedings of the 3rd International Conference on Machine Learning, Image Processing, Computer Vision, and Data Analytics (pp. 89- 100). Springer (2020).