iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.61019

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue IV Apr 2024- Available at www.ijraset.com

Stock Market Prediction and Visualisation

Arpita Singh[1], Basar Imam Mazhari[2], Karan Gupta[3], Yash Veer Singh[4]
[1,2,3]*Student*, [4]*Professor*, *Computer Science And Engineering Galgotias College Of Engineering And Technology* Greater Noida, India

Abstract: ***Stock price forecasting is a widely discussed and significant topic in both financial and academic circles. The stock market is inherently unpredictable, lacking clear rules for estimating or predicting share prices. Various methods, including technical analysis, fundamental analysis, time series analysis, and statistical analysis, have been employed to forecast stock prices. However, none of these methods consistently serve as reliable prediction tools. In this paper, we explore the implementation, prediction, and analysis of stock market prices. Artificial Neural Networks and Machine Learning prove effective for forecasting stock prices, returns, and modeling stock behavior. By conducting statistical analyses, we establish relationships between selected factors and share prices, contributing to more accurate predictions. While the stock market remains inherently uncertain, this paper aims to apply data analysis and prediction concepts to forecast stock prices.***
In the era of global digitization, stock market prediction has undergone significant technological advancements, transforming traditional trading models. As market capitalization continues to rise, stock trading becomes a focal point for financial investors. Analysts and researchers have developed tools and techniques to predict stock price movements, aiding decision-making. Advanced models leverage non-traditional textual data from social platforms for market prediction. Machine learning approaches, including text data analytics and ensemble methods, have significantly improved prediction accuracy. However, analyzing and predicting stock markets remains challenging due to dynamic, erratic, and chaotic data. This study explores machine learning-based approaches for stock market prediction, emphasizing a generic framework. By critically analyzing findings from the last decade (2011–2021) from digital libraries like ACM and Scopus, we provide insights for emerging researchers to delve into this promising area.

I. INTRODUCTION

The stock market is dynamic, unpredictable, and complex. Forecasting stock prices is an intricate task influenced by a multitude of factors such as the global economy, company financials, and overall performance. Historically, two primary methods have been utilized for predicting a company's stock price. The first method, technical analysis, relies on historical stock prices, including opening and closing prices, trade volume, and previous closing values, to forecast future stock prices.

The second method, qualitative analysis, considers external factors such as the company's profile, market conditions, economic and political influences, as well as narrative information from financial news articles, social media, and economic analysts' blogs.

The stock market is a critical area of focus for investors, making the prediction of stock price trends a perennially relevant subject for researchers in both the financial and technical sectors. This research aims to develop a cutting-edge model for predicting stock price trends, with a particular emphasis on short-term predictions.

Investment firms, hedge funds, and even individual investors rely on financial models to gain insights into market behavior and make profitable trading decisions. A wealth of stock data is available for analysis and processing. But is using machine learning for stock price prediction truly effective? Investors make informed guesses by analyzing data, studying company histories, industry trends, and other relevant variables. While prevailing theories suggest that stock prices are entirely random and unpredictable, top firms like Morgan Stanley and Citigroup still hire quantitative analysts to build predictive models.

This paper explores the application of Deep Learning models, specifically Long Short-Term Memory (LSTM) Neural Networks, for stock price prediction. While recurrent neural networks (RNNs) are suitable for time-series data, recent research indicates that LSTMs are the most popular and useful RNN variants. It's essential to recognize that businesses can be vulnerable to market fluctuations beyond their control, including market sentiment, economic conditions, and sector-specific developments.

II. RELATED WORK

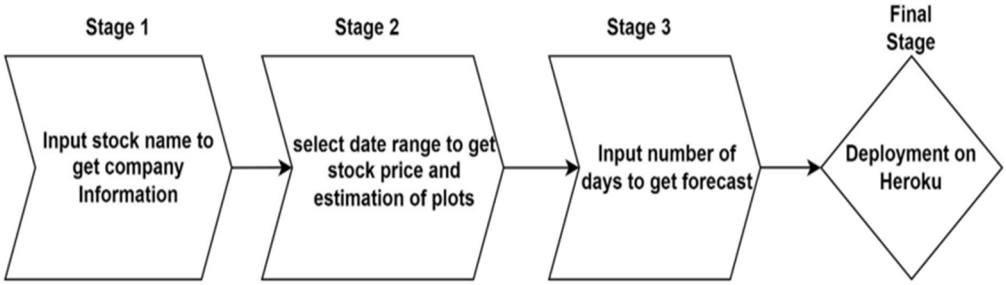
Conventional methods for analyzing the stock market and predicting stock prices encompass fundamental analysis, which evaluates a company's historical performance and overall reputation, and statistical analysis, which focuses on data analysis and pattern

©IJRASET: All Rights are Reserved

recognition in stock price fluctuations. Predictive outcomes were previously derived using Genetic Algorithms (GA) or Artificial Neural Networks (ANNs), yet these techniques often overlook the long-term correlations among stock prices.

A significant challenge in employing basic ANNs for stock prediction is the exploding/vanishing gradient problem, where network weights become excessively large or small, impeding the optimization process. This issue arises from random weight initialization and disproportionate changes in weights towards the network's output layer.

An alternative strategy for stock market analysis involves dimensionality reduction of input data and the utilization of feature selection algorithms to identify key features (like GDP, oil prices, inflation rates, etc.) that significantly influence stock prices or currency exchange rates. However, this approach does not account for long-term trading strategies, as it neglects the complete historical trend data, and lacks mechanisms for outlier detection.



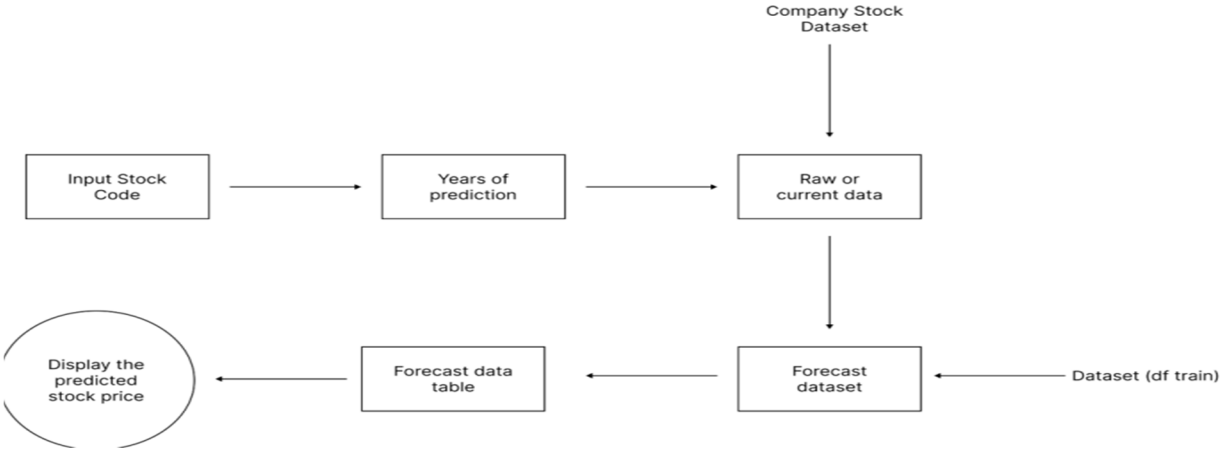
Predicting stock prices is an intricate and challenging task due to their behavior resembling a random walk and being time-varying. Extensive literature surveys reveal that machine learning techniques are increasingly applied to stock market prediction worldwide. These techniques offer greater accuracy and speed compared to existing prediction methods. Previous research in stock prediction provides valuable insights, particularly concerning historical, social media, and financial news data. Notably, the work by Edgar P. Torres, Myriam HernándezÁlvarez, Edgar A. Torres Hernández, and Sang Guun Yoo focuses on forecasting and predicting future stock market data using artificial intelligence techniques, specifically machine learning algorithms.

III. PROPOSED SYSTEM

We have introduced an online web-based application that utilizes a learning model to predict stock prices. The primary challenge of this project lies in accurately forecasting the future closing value of a specific stock over a defined time horizon. To achieve this, we employ a Long Short-Term Memory (LSTM) network, commonly referred to as 'LSTMs,' which leverages a dataset of historical stock prices to predict the closing price of the S&P 500.

IV. PROPOSED ANALYTIC MODEL

Our project involves the development of a web-based application powered by a learning model, designed to forecast stock prices. The core challenge of this initiative is to provide precise predictions for the future closing values of a selected stock over a forthcoming period. For this purpose, we have utilized a Long Short-Term Memory network, often abbreviated as 'LSTMs', which uses historical price data to estimate the closing price of the S&P 500.



Our initial approach involved applying linear regression to forecast stock prices. Subsequently, we shifted to employing Long Short-Term Memory networks, commonly known as 'LSTMs', to determine the closing price of the S&P 500, based on historical pricing data. To evaluate performance, we utilized Mean Squared Error (MSE) and calculated Root Mean Squared Error (RMSE), which measures the discrepancy between the predicted and actual adjusted closing values of the target stock. Additionally, we compared the effectiveness of our primary deep learning model against the benchmark linear regression model.

A. Exploring The Stock Prices

In this study, we utilize Google's historical data spanning from January 1, 2005, to March 20, 2019. The dataset consists of a time-ordered series of data points, forming a time series. Our objective is to predict the closing price for any given date after training. To ensure reproducibility and ease of use, we sourced all data from the Yahoo Finance Python API. Specifically, we focus on predicting the closing (adjusted closing) price. Since Yahoo Finance already adjusts closing prices, our prediction targets the 'CLOSE' price. The complete dataset can be found in the 'Google.csv' file located in the project's root folder.

Table -1: Head of data set

Date	Open	High	Low	Close	Volume
30-Jun-17	943.9	945.0	929.6	929.6	2287662
29-Jun-17	951.3	951.6	929.6	937.8	3206674
28-Jun-17	950.6	963.2	936.1	961.0	2745568

The mean, standard deviation, maximum and minimum of the data was found to be following:

Table -2: The following table contains the statistical information about the data

Feature	Open	High	Low	Close	Volume
Mean	382.5	385.87	378.73	382.35	42057
Std	213.4	214.60	212.08	213.43	38483
Max	1005.	1008.6	1008.6	1004.2	41182
Min	87.74	89.29	86.37	87.58	52114

Upon analyzing the dataset, we observe that features such as date, high, and low values do not significantly impact the data. However, the relationship between open and close prices is evident. The opening price and closing price of a stock are crucial factors. If the closing price exceeds the opening price, it indicates profit; otherwise, losses are incurred. Additionally, share volume plays a vital role. In a rising market, volume should also increase. Conversely, decreasing volume alongside price fluctuations signals a lack of interest and potential reversal. Notably, a substantial price drop (or rise) accompanied by high volume suggests a fundamental change in the stock. During data processing, we excluded date, high, and low features. The processed data's mean, standard deviation, maximum, and minimum values are as follows:

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

Table -3: The mean, std, max and min of normalized data set

Mean	Std	Max	Min
Open	0.3212	0.23261	1.0
Close	0.3215	0.2328	1.0
Volume	0.09061	0.0953	1.0

B. Exploratory Visualization to visualize the data

For the preliminary visualization of the dataset, we employed the Matplotlib library in Python. The data presented is historical and plotted to scale, focusing on the number of days and the corresponding opening price for each day.

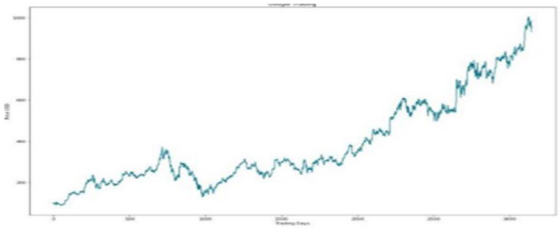


Fig -1: Visualization of processed hysterical data fetched from the API.

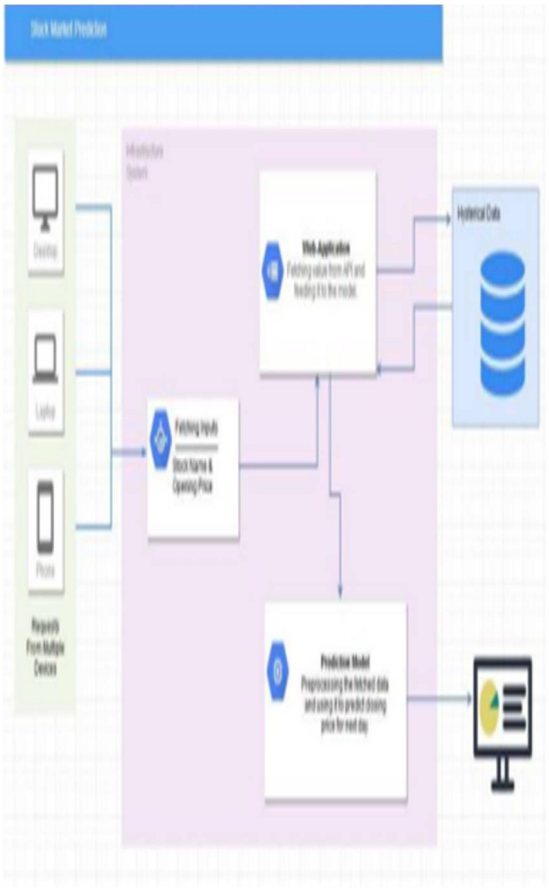


Fig 2: Flow of the system

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

V. METHODOLOGY

The current system employs a sliding window algorithm that does not incorporate a dropout process, leading to the inclusion of irrelevant data. This results in inefficient predictions and unnecessary consumption of time and resources. The sliding window approach is particularly inadequate for processing nonlinear data. In contrast, our proposed method utilizes Long Short-Term Memory (LSTM) networks, which surpass the sliding window algorithm in efficiency. Our machine learning-based stock price prediction model leverages LSTM to estimate future stock values. LSTMs excel at adapting to changes in stock price behavior over specified periods. While traditional Artificial Neural Networks (ANNs) struggle to achieve satisfactory accuracy, LSTMs, along with other techniques like Restricted Boltzmann Machines (RBMs) and ANNs, have been explored. Given that stock prices are influenced by historical data, Recurrent Neural Networks (RNNs) are deemed most appropriate for such predictions. However, RNNs have limitations in retaining states across numerous hidden layers, affecting accuracy. Thus, this paper advocates for the use of LSTM networks for more precise stock price forecasting.

VI. ALGORITAMS AND TECHNIQUES USED

This research was focused on examining time-series data to identify the most effective methods for precise stock price prediction. Our findings indicate that Recurrent Neural Networks (RNNs), which are tailored for learning sequences and patterns, are networks that contain loops, enabling them to retain information and thus memorize data with accuracy. However, RNNs encounter the issue of vanishing gradients, which hinders their ability to learn from historical data. This problem is addressed by Long-Short Term Memory Networks (LSTMs), a unique type of RNN designed to learn long-term dependencies. To enhance the neural network's architecture, we can adjust a comprehensive set of parameters to refine the prediction model, including:

1) Input Parameters:

- Data preprocessing and normalization.
- Neural network architecture.
- Number of layers (we utilized 3 layers of nodes in the model).
- Number of nodes per layer.

2) Training Parameters:

- Training/test split ratio (maintained at 71% training and 29% testing for benchmarks and LSTM model).
- Batch size (the number of time steps included in a single training iteration).
- Optimizer function (a mean operation performed at the end of the prediction to adjust the output back to the expected range).
- Epochs (the total number of training cycles).

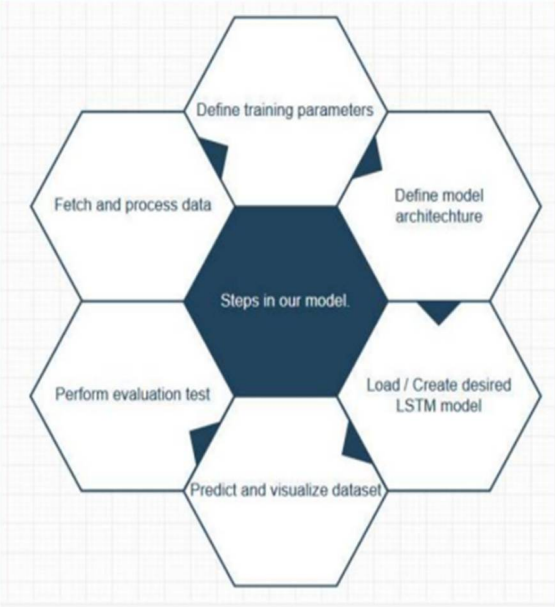


Fig -3: Steps in our model.

3) Software functionalities

- Company Selection: Users can choose a company from the dropdown menu, and the system will retrieve the stock code as input.
- Years of Prediction: Users have the flexibility to select the desired prediction timeframe, ranging from 1 to 4 years.
- Current Stock Price Analysis: Users can interactively explore the current stock price using a slider.
- Predicted Stock Price Analysis: After reviewing the current stock price, users can examine the predicted stock price and observe how it may fluctuate over time.
- Predicted Forecast Components: Finally, users can explore the various factors and components that contribute to the stock price forecast.

VII. RESULTS

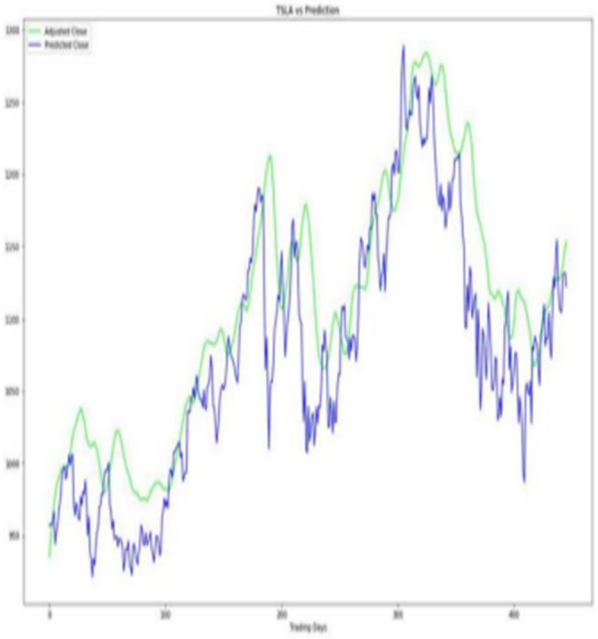


Fig -4: Here's a graph displaying the predicted pattern from our model alongside the actual pattern observed in the closing prices dataset.

To evaluate the robustness of our model, we tested it with previously unused data, specifically Facebook's data from January 1, 2011, to April 16, 2019. The performance of the model on this new data was satisfactory. The following are the results: The model forecasted a closing price of $178.27 for Facebook on April 17, 2019, while the actual closing price on that day was $178.88.

Model Evaluation outputs

During training, the model achieved a Root Mean Square Error (RMSE) of 0.0417 and a Mean Square Error (MSE) of 0.001739. When tested on unseen data, the RMSE was 0.1061, and the MSE was 0.011277.

VIII. CONCLUSIONS & FUTURE WORK

We have applied Long Short-Term Memory (LSTM) networks to predict stock market behavior and identify key features. Our initial analysis reveals significant correlations among various input parameters. The results obtained in both cases were reasonably accurate. As demonstrated by the figures above, predictions hold well unless there are substantial and abrupt variations in the actual data. Interestingly, this reinforces the hypothesis that stock markets are inherently unpredictable. Following the prediction and analysis phase, users will receive the results displayed on web pages.

The goal of this study is to explore potential directions for future research in machine learning-based stock market prediction, informed by an analysis of existing literature. By examining ML-related systems, problem contexts, and findings from selected articles, we can draw several conclusions about our current understanding in this field.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

1) *Method-Problem Fit:* There exists a strong correlation between ML methods and the specific prediction problems they address. This relationship resembles the concept of task-technology fit (Goodhue and Thompson, 1995), where system performance depends on aligning tasks with appropriate technologies.
2) *Method Suitability:*

- Artificial Neural Networks (ANNs): These are well-suited for predicting numerical stock market index values.
- Support Vector Machines (SVMs): Best applied to classification problems, such as determining whether the overall stock market index will rise or fall.
- Genetic Algorithms (GAs): Employ an evolutionary problem-solving approach to identify high-quality system inputs or predict optimal stock portfolios for maximum returns.

3) *Limitations of Single-Method Applications:* While each study demonstrates effective application of specific methods, relying solely on a single approach has limitations.
4) *Hybrid Approaches:* Combining multiple machine learning techniques can mitigate some of these limitations and enhance prediction accuracy.

The complexity of systems can reach a point where they become impractical. This theoretical and practical challenge warrants further exploration in future studies. Another key finding from the review of existing research is the need to enhance the generalizability of findings. Many studies evaluate their ML systems using data from a single market or time period, without considering their effectiveness in other contexts. To address this, we propose three enhancements for experimental system assessment:

- Market Diversity: While several studies focus on Asian stock markets, it's essential to test these systems on US or European markets during the same time periods. This cross-market evaluation will provide valuable insights.
- Market Environment Variation: Evaluating systems during both rising and declining market conditions is crucial. For instance, can an approach accurately predict US market values during the 2008-2009 financial crisis and also during the recent growth period from 2018-2019? Assessing predictions for market growth and contraction is equally important.
- Risk and Volatility Considerations: Systems should be tested on stock market indices comprising only small firms versus only large firms. Understanding their effectiveness under different risk and volatility scenarios is essential.

Lastly, it's evident that financial investment theory should strongly inform the inputs, algorithms, and performance measures of ML systems. Without this alignment, results may lack practical utility and be merely random.

Numerous studies have applied techniques without adequately considering the extensive body of financial theory established over many years. It would be beneficial to report instances where these techniques fail to enhance predictive accuracy. Such reports are scarce, making it difficult to discern when a specific stock market prediction issue does not align with a machine learning approach. A paradox within this field of research is that for investors, it represents a zero-sum scenario. Should an optimal machine learning technique for stock market prediction be discovered and universally adopted, it would negate any competitive advantage. Consequently, major investment firms engaged in identifying superior machine learning strategies are unlikely to share their findings publicly.

When evaluating the standard Linear Regression model against the enhanced LSTM model, there was a notable improvement in Mean Squared Error. Adjusting the mean on the processed LSTM data contributed to superior outcomes and more precise pattern recognition within historical datasets. The endeavor of forecasting stock market prices is fraught with uncertainty and can frequently result in imprecise predictions due to the myriad of influencing factors. Future enhancements to this project could involve augmenting the model with additional features and incorporating key non-numerical attributes with guidance from industry experts.

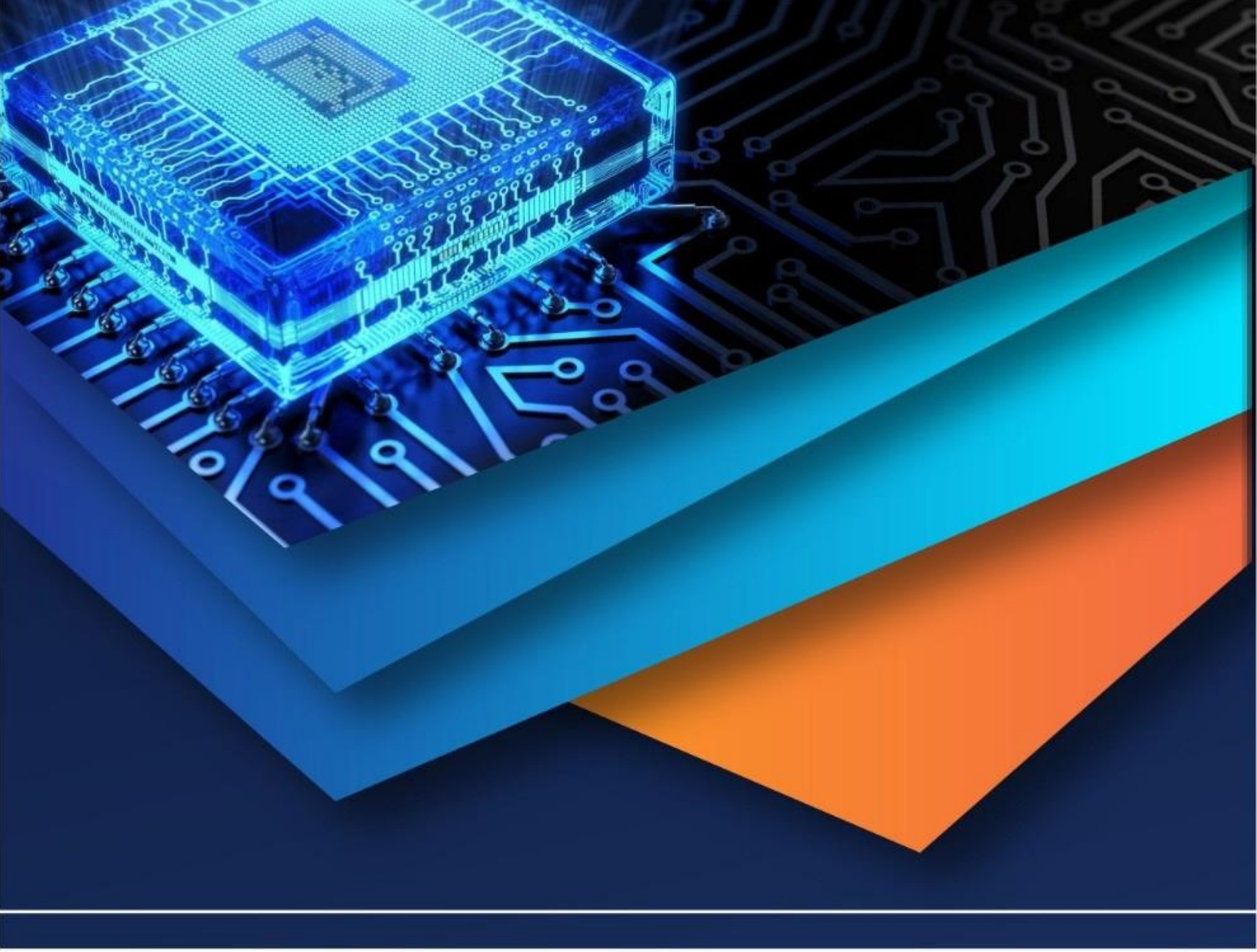
The process of stock market prediction is arduous and time-intensive. Nonetheless, the advent of Machine Learning and its diverse algorithms has ushered in new methodologies for analyzing stock market data. Through evaluating the precision of various algorithms, it was determined that the Long-Short Term Memory (LSTM) algorithm is the most effective for predicting stock prices based on historical data points. This algorithm is invaluable for brokers and investors, having been refined through extensive training on historical data and validated against a sample dataset. The project showcases a machine learning model that surpasses other models in predicting stock values. Prospective improvements to the stock market prediction system include leveraging larger datasets with greater computational power and optimizing the number of training epochs to enhance prediction accuracy. Additionally, exploring other machine learning models for their accuracy and integrating sentiment analysis from social media with LSTM could refine weight training and boost precision. This leads to the conclusion that machine learning techniques can predict stock market trends with heightened accuracy and efficiency.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

REFERENCES

[1] Poonam Somani, Shreyas Talele, Suraj Sawant, (2014)“ Stock market prediction using Hidden Markov Model,”. IEEE 7th Joint International Information Technology and Artificial Intelligence Conference.
[2] Dinesh Bhuriya, Girish Kaushal, Ashish Sharma, Upandra Singh, (2017)”Stock market predication using a linear regression”. International Conference on Electronics, Communication and Aerospace Technology ICECA
[3] Dinesh Bhuriya, Girish Kaushal, Ashish Sharma, Upandra Singh, (2017)”Survey of Stock Market Prediction Using Machine Learning Approach”. International Conference on Electronics, Communication and Aerospace Technology ICECA
[4] Rajat-Dhyani, Stock Price Predictor, Github.
[5] V Kranthi Sai Reddy, (2018) “Stock Market Prediction Using Machine Learning”, International Research Journal of Engineering and Technology (IRJET), Volume 05 Issue: 10
[6] Alice Zheng, Jack Jin. Stanford University, (2017) “Using AI to Make Predictions on Stock Market”, Proceedings on the International Conference on Artificial Intelligence (ICAI).
[7] Prof. S .P. Pimpalkar, Jenish Karia, Muskaan Khan, Satyam Anand, Tushar Mukherjee.(2017) “Stock Market Prediction using Machine Learning”, International Journal of Advance Engineering and Research Development, Volume 4,
[8] Taewook Kim, Ha Young Kim.(2019) “Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data”, PLOS ONE Journals,
[9] Mruga Gurjar, Parth Naik, Gururaj Mujumdar.(2018) “STOCK MARKET PREDICTION USING ANN”, International Research Journal of Engineering and Technology (IRJET), [10]Nirbhey Singh Pahwa, Neeha Khalfay, Vidhi Soni, Deepali Vora.(2017) “Stock Prediction using Machine Learning a Review Paper”, International Journal of Computer Applications (0975 – 8887) Volume163 – No 5.
[10] R. Singh, and S. Srivastava, “Stock prediction using deep learning”, Multimedia Tools and Applications, 76(18), pp.18569- 18584, 2017.
[11] T. Fischer, and C. Krauss, “Deep learning with long short-term memory networks for financial market predictions”, European Journal of Operational Research, 270(2), pp.654-669, 2018.
[12] C. Zhu, J. Yin, and Q. Li, “A stock decision support system based on dbns”. Journal of Computational Information Systems, 10, pp. 883- 893, 2014.
[13] X. Zhong, and D. Enke, “Forecasting daily stock market return using dimensionality reduction”, Expert Systems with Applications, 67, pp. 126-139, 2017.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)