Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 09, September-2023

"Machining the Market: A Comprehensive Review of Machine Learning in Stock Market Analysis"

Om Sohani
Dept. of Artificial Intelligence and Data Science
K.J. Somaiya Institute of Technology
Mumbai, India

Rushi Sangani
Dept. of Artificial Intelligence and Data Science
K.J. Somaiya Institute of Technology
Mumbai, India

Aryan Chadha
Dept. of Artificial Intelligence and Data Science
K.J. Somaiya Institute of Technology
Mumbai, India

***Abstract*— The stock market is a financial platform where securities are purchased and sold, and predicting it with ML requires analyzing historical data and patterns to forecast future price movements. In this thorough analysis, we explore the wide range of novel ideas and strategies used in stock market price prediction methodology. Selected research publications are examined, including topics such as XGBoost-enhanced predictions, stochastic gradient descent (SGD), long short-term memory (LSTM) models, and mixed swarm intelligence and artificial neural networks (ANN). Together, the papers provide insight into the dynamic field of stock market forecasting. They emphasize the use of historical data, sophisticated machine learning and artificial intelligence algorithms, and Python programming as effective methods for predicting stock market values. Review includes an examination of LSTM and recurrent neural network (RNN) architectures, demonstrating how well they may capture complex temporal connections. It also looks into how effective AI-driven models are, highlighting how they could increase prediction accuracy for the stock market. The "CPSMP_ML" framework is examined closely since it emphasizes closing price prediction, proving that machine learning models may be used in this situation. In addition, the combination of XGBoost and LSTM models is explored as a potential method to increase prediction accuracy, highlighting the complementary nature of deep learning tactics and conventional machine learning methods. For both scholars and practitioners, this study offers a comprehensive summary of recent research and insightful information about the complex field of stock market price prediction**.

Keywords—XGBoost, Machine Learning, LSTM, RNN, Stock Market, Artificial Neural Network.

I. INTRODUCTION

Dynamic financial hubs, stock markets facilitate the purchasing and selling of company shares. Predicting their intricate price movements is a complex challenge, often tackled using advanced data analytics and machine learning, seeking to provide investors with insights to make informed decisions in this ever-fluctuating landscape.

It is impossible to overestimate the importance of the field of stock market price prediction in the financial markets. However, it has always been a difficult one. A plethora of novel approaches and procedures have surfaced in recent times with the goal of improving the precision of stock price projections. This review paper sets out on an extensive exploration of this ever-changing terrain, where classical financial analysis meets artificial intelligence and machine learning.

The set of research articles under review is a diverse tapestry of methodologies, each offering special insights and tactics to the always changing field of stock market prediction. This review covers a wide range of approaches, from modeling long short-term memory (LSTM) and stochastic gradient descent (SGD) to utilizing artificial neural networks (ANN) and swarm intelligence collaborative intelligence. It also explores the usefulness of putting these ideas into practice, including using Python as a computer language to forecast stock market prices.

Temporal dependencies in historical stock market data have been a focus, and several articles show how well recurrent neural networks (RNN) and long short-term memory (LSTM) architectures can capture these complex patterns. The potential for AI-driven models to increase forecasting accuracy is demonstrated by the research' exploration of the theme of using AI to anticipate the stock market.

In addition, the paper includes a thorough analysis of the "CPSMP_ML" framework, which emphasizes closing price prediction and highlights how machine learning models can be tailored to various aspects of stock market research.

Using the potent gradient boosting algorithm XGBoost in conjunction with LSTM models is another noteworthy area of research that illustrates the mutually beneficial link between conventional machine learning methods and deep learning approaches in the context of stock market prediction.

This study seeks to provide a thorough overview of current developments in stock market price prediction by synthesizing the abundance of knowledge contained

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 09, September-2023

within these research publications. In doing so, it hopes to offer scholars, professionals, and interested parties a useful tool for comprehending the various strategies used to address one of the most difficult and financially significant jobs in the financial industry.

II. LITERATURE SURVEY

The field of stock market price prediction has undergone a notable transformation with the incorporation of artificial intelligence and machine learning methodologies. We examine a wide range of research publications in this overview of the literature, each of which offers a distinctive perspective and methodology to the effort of enhancing daily stock market price predictions.

1. Mr. Pankaj K. Bharne et al. In order to improve daily stock market price predictions, this study investigates the synergy between swarm intelligence and artificial neural networks (ANN). It aims to maximize forecast accuracy by utilizing collective intelligence, offering a fresh solution to this difficult issue.[1]

2. Harshvardhan Singh et al. In order to forecast stock values based on historical data, this research study uses long short-term memory (LSTM) and stochastic gradient descent (SGD) models. By combining these methods, one can gain important insights on how to record temporal dependencies in the data.[2]

3.Pujitha.J et al.Using Python as a programming language for stock market price prediction, this research highlights the practical application of machine learning models. It underlines the significance of easily available resources for efficient forecasting.[3]

4.Sahith Addagalla et al. This paper offers a thorough summary of the many machine learning methods used to forecast stock market prices. It is an invaluable tool for comprehending the terrain of accessible approaches.[4]

5. Priyanka Srivastava et al.In this research The mainstays of this research are long short-term memory (LSTM) structures and recurrent neural networks (RNN). It demonstrates how well they can extract intricate sequential patterns from stock market data.[5]

6.K. Ritwik Reddy et al. This research investigates the application of recurrent neural networks in stock market trend prediction, expanding on the RNN subject. It adds to the expanding corpus of research on recurrent models' efficacy in this field.[6]

7. Kanthimathi N et al. This research highlights artificial intelligence (AI) and shows how it might enhance stock market forecasts. The study emphasizes how AI-driven models have the potential to be revolutionary.[7]

8. Sanjiv Kumar et al. This research focuses on the efficacy and efficiency of stock market prediction by introducing an LSTM-based technique. It offers an alternate viewpoint on making use of LSTM structures.[8]

9.Diaa Salama Abd Elminaam.Using various machine learning models, the "CPSMP_ML" framework focuses on closing price prediction. It demonstrates how adaptable machine learning methods are for tackling particular facets of stock market research.[9]

10.Agustinus Bimo Gumelar. This study combines XGBoost with LSTM models in a hybrid deep learning approach to improve prediction accuracy. It emphasizes the possibility of combining various prediction techniques in a beneficial way.[10]

III. METHODOLOGY

A. Linear regression:Modeling the relationship between a dependent variable and one or more independent variables is a fundamental statistical technique used in statistics and machine learning. It computes coefficients that minimize the difference between anticipated and actual values under the assumption that variables have a linear relationship. For activities like forecasting stock prices, examining economic trends, and figuring out how factors affect results in diverse disciplines, linear regression is frequently used.

B. K-Nearest Neighbors (KNN):It is a straightforward but efficient machine learning technique used for regression and classification tasks. Based on the majority class or average of its k nearest neighbors in the feature space, it assigns labels or forecasts values. KNN is adaptable for a variety of applications, including recommendation systems and pattern recognition, as it uses distance metrics like Euclidean distance and is non-parametric in nature.

C. Random Forest (RF)It is an effective ensemble learning algorithm that is frequently employed in machine learning. To increase accuracy and lessen overfitting, it builds numerous decision trees during training and combines their predictions. The selection of features and data points is randomized by RF, which makes it robust and appropriate for classification and regression problems. It excels in a variety of fields, including image identification, finance, and healthcare.

D. Recurrent Neural Networks (RNNs): They are neural network designs created for the processing of sequential data. They are useful for tasks like time series prediction and natural language interpretation because they maintain a hidden state that allows them to recognize relationships and patterns in sequences. RNNs can, however, experience vanishing gradient problems, which are frequently resolved by cutting-edge variations like LSTMs and GRUs.

E. Long Short-Term Memory (LSTM):It is a kind of recurrent neural network (RNN) architecture that excels at detecting and keeping track of distant dependencies in sequential data. For applications like natural language processing, speech identification, and time series forecasting, LSTMs are extremely effective because they use specific memory cells and gating methods to govern input flow.

F. Gradient Boosting:This method of ensemble machine learning combines the skills of weak learners to

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 09, September-2023

create a strong prediction model by building models successively. By basing upcoming models' adjustments on the mistakes of earlier ones, it reduces inaccuracies. Widely utilized in tasks like classification, regression, and ranking, gradient boosting is well known for its remarkable predictive performance and adaptability. It is frequently implemented using algorithms like XGBoost, AdaBoost, and Gradient Boosted Trees.

G. XGBoost:Extreme Gradient Boosting, for which it is short, is a high-performance, scalable machine learning technique renowned for its astounding speed and accuracy. It minimizes overfitting by combining gradient boosting with tree-based models and applying regularization techniques. Many data-driven applications, including classification, regression, and ranking tasks, employ XGBoost extensively. Because of its effectiveness and efficiency in processing structured data, XGBoost has won multiple machine learning contests.

IV. RESULTS AND DISCUSSIONS

An extensive overview of the numerous subfields within stock market price prediction is given by the combination of research that is presented throughout these publications. The goal of improved accuracy and efficiency through various approaches runs across these works as a common theme. Swarm intelligence and artificial neural networks (ANN) are combined, providing a novel approach to model optimisation while demonstrating the potential of collective intelligence to optimise ANN parameters. For the purpose of predicting stocks based on historical data, stochastic gradient descent (SGD) and long short-term memory (LSTM) models are used, which highlights the need of sequential modelling for capturing temporal dependencies. Python's prominence as a useful tool in "Stock Market Price Predictions Using Python" emphasises the value of user-friendly programming languages for putting machine learning models into practise. The inclusion of artificial intelligence (AI) in "Stock Market"

V. CONCLUSION

As a result, the various approaches and conclusions generated from these research articles provide a thorough understanding of the dynamic environment of stock market price prediction. The use of cutting-edge algorithms like SGD and LSTM as well as the combination of swarm intelligence with artificial neural networks are all examples of how these tactics are advancing. The importance of sequential modelling using RNN and LSTM architectures is emphasised, while the practicality of Python implementations is highlighted. Deep learning and other AI-driven methods are being integrated, which supports the pursuit of higher accuracy. Inspiring more accurate and trustworthy daily stock market forecasts, these group initiatives show the field's tenacity and creativity.

VI. REFERENCES

[1]P. K. Bharne and S. S. Prabhune, "Survey on combined swarm intelligence and ANN for optimized daily stock market price," 2017 International Conference on Soft Computing and its Engineering Applications (icSoftComp), Changa, India, 2017.

[2]H. Singh, P. Bhatt, S. Jacob, A. Kaur, A. Vashist and D. Vij, "Stock Prediction on Historical Data based on SGD and LSTM," 2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM), Gautam Buddha Nagar, India, 2022.

[3]J. Pujitha and G. Mahesh, "Stock Market Price Predictions Using Python," 2021 International Conference on Forensics, Analytics, Big Data, Security (FABS), Bengaluru, India, 2021.

[4]S. Addagalla, S. Koppuravuri, R. Krosuri, M. S. Kunapareddy, S. Reddy Mallu and M. Rashmi, "Stock Market Price Prediction Using Machine Learning Techniques," 2023 4th International Conference for Emerging Technology (INCET), Belgaum, India, 2023.

[5]P. Srivastava and P. K. Mishra, "Stock Market Prediction Using RNN LSTM," 2021 2nd Global Conference for Advancement in Technology (GCAT), Bangalore, India, 2021.

[6]K. R. Reddy, B. T. Kumar, V. R. Ganesh, P. Swetha and P. K. Sarangi, "Stock Market Prediction Using Recurrent Neural Network," 2022 IEEE International Conference on Current Development in Engineering and Technology (CCET), Bhopal, India, 2022.

[7]F. G. D. C. Ferreira, A. H. Gandomi and R. T. N. Cardoso, "Artificial Intelligence Applied to Stock Market Trading: A Review," in IEEE Access.

[8]S. Kumar, U. Aggarwal, P. Gautam, A. Tuteja, P. Matta and S. Maurya, "LSTM Approach for Efficient Stock Market Prediction," 2023 3rd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2023.

[9]D. S. A. Elminaam, A. E. Tanany, M. A. Salam and M. A. E. Fattah, "CPSMP_ML: Closing price Prediction of Stock Market using Machine Learning Models," 2022 2nd International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC), Cairo, Egypt, 2022.

[10]A. B. Gumelar et al., "Boosting the Accuracy of Stock Market Prediction using XGBoost and Long Short-Term Memory," 2020 International Seminar on Application for Technology of Information and Communication (iSemantic), Semarang, Indonesia, 2020.