Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN: 2581 - 611X)
VOLUME:8, ISSUE:1, FEBRUARY 2025, PAPER NUMBER:29

STOCK PRICE PREDICTION USING MACHINE LEARNING TECHNIQUES

Mrs.Thripthi P Balakrishnan[1], Sai Likhitha G[2], Sai Vardhini T[3],Supriya J G[4], Suvarna B[5]

Asst.Professor[1] ,Department of Computer Science & Engineering, Madanapalle Institute of Technology& Science, Madanapalle,India-517325

UG Scholar[2,3,4,5],Department of Computer Science & Engineering, Madanapalle Institute of Technology & Science, Madanapalle,India-517325

[1]thripthi.p.b@gmail.com,[2]sailikhithagangula@gmail.com,[3]sreeharitadimarri@gmai.com,[4]jayamsupriya321@gmail.com,[5]bingisuvarna7673@gmail.com

***ABSTRACT*- Time arrangement estimating is broadly utilized to choose fate expenses, and time collection is utilized for budgetary assessment and in specific for coordinating traders' choices and exchanges. This research paper proposes a judicious time collection estimating strategy the utilization of a rolling window optimization to figure the charges of mining gadget. The machine includes a graphical client interface and runs as a standalone utility. Stock cost forecast has long been a challenging assignment in budgetary markets, with significant implications for financial specialists, traders, and financial analysts. The proposed show could be a promising approach for anticipating uncommonly non-linear time arrangement whose designs are extreme to capture with conventional models. In this article, we make utilize of the methods which incorporate ARIMA, Direct Relapse and Random Forest Classifier that are utilized to predict stock charges.**

Keywords: trade, forecasting, Stock, Linear Regression, Random Forest, ARIMA

INTRODUCTION

Estimating characteristics and changes in load costs is taken into thought one of the greatest convenient determining programs. In spite of the fact that there are various ponders on the inconvenience of foreseeing rate patterns, most of them have given precise comes about in forex markets. In any case, it's distant difficult to anticipate cargo charges since of the vulnerability of the advertise. The sorts of investigation are: principal assessment evaluation assessment Synonyms and specialized examination. It may be a crucial examination of the behavior of society, financial framework and politics. Specialized capabilities incorporate exceptionallyfinal offered, greatest offered, little commission and

additional. The remaining n days are to be taken into thought. With the assistance of specialized examination, we are anticipating the slant of alter inside the stock or stock proportion. Crucial assessment is difficult to degree and difficult to execute in a portable workstation dialect. Specialized Score no longer gauges the inborn charge of a stock, but instep makes utilize of specialized charts to are anticipating stock patterns.

In initial marketplace forecasting, classical studies techniques are used. But the fashion of improvementis the time of improvement. Not so effective Therefore, nonlinear system mastering methods, which includes ARIMA, random forest, and linear regression, are widely used. In this task, we are able to use every method to expect developments in asset expenses and the accuracy of the 3 techniques.

OBJECTIVE

The principal motive of the start of this mission is to peer the U.S 's modern marketplace based totally mostly on the market pursuits of the past. Using this method, it is beneficial to predict the factors of any organization to evaluate the economic health of the company, whether or not they're high-quality partners, or how they're appearing milestones or growing within the corporation.

LITERTATURE SURVEY

Given the significant instability in money related markets, there's a impressive sum of instability and hazard related with them. This paper presents an imaginative strategy for foreseeing next-day stock closing costs. It utilizes a combination of profound learning procedures, counting Long Short-Term Memory (LSTM) and Recurrent Neural Networks (RNN), nearby the Auto-Regressive Integrated Moving Average (ARIMA) time arrangement demonstrate and Estimation examination. This

approach is outlined to precisely estimate future stock prices, taking into consideration both long-term conditions and short-term patterns.. [1]

The stock market functions as a marketplace for the trading of publicly traded company shares, enabling stock brokers to transact in company stocks and other securities. The Bombay Stock Exchange and the National Stock Exchange are the two main stock exchanges in India. Due to its complexity and difficulty, stock price forecasting draws analysts from a wide range of fields, including finance, economics, computer science, and mathematics. Traditional time-series or regression methods are limited because of the stock market's volatility. Our method uses machine learning techniques on the NSE to forecast future stock values. It makes use of both SVM and linear regression. Based on the closing price of the previous day, linear regression is used to predict the opening price of the stock for the next day. The difference between the closing and opening values of the stock for the following day is predicted using SVM regression. To improve prediction accuracy, outside variables like relative strength index, NSE index, moving averages, and currency rates are included. [2]

Over the past ten years, researchers in a variety of fields have made extensive use of neural networks, a well-known data mining technique. There are two types of analytic algorithms: non-linear models (Curve, GARCH, Neural Network) and linear models (AR, MA, ARIMA, ARMA). In order to anticipate a company's stock price based on past price data, we use four different types of neural network topologies in this study: Multilayer Perceptron (MLP), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Convolutional Neural Network (CNN). We employ day-by-day closing prices from the New York Stock Exchange (NYSE) and the National Stock Exchange (NSE) of India, two separate stock marketplaces. In our examination, CNN comes out as the best-performing model. Remarkably, for a network trained on NSE data, it forecasts NYSE prices properly. When the outcomes are compared to a conventional ARIMA model, it is found that the neural networks perform better. [3]

Accurate interval-valued stock price forecasting is a difficult undertaking that attracts a lot of attention from profit-driven businesses and financial professionals. In this paper, we simultaneously consider individual stock data and pertinent stock data, offering a unique method to predict interval-valued stock prices based on a Dual Convolutional Neural Network Interval (Dual-CNNI) model. First, we gather and convert unique and pertinent stock data into pictures. Next, interval-valued stock price forecasting is done using the Dual-CNNI model. To extract features from unique and pertinent stock data, two convolutional neural network (CNN) models with different architectures are specifically created. The ultimate interval-valued stock price forecast is then made using an interval Multilayer Perceptron (MLPI) model. Using data from six randomly chosen stocks, extensive tests are carried out to compare our strategy with other widely used machine learning models and interval-valued time series prediction techniques. The experiment results show that our suggested Dual-CNNI based method performs better in terms of prediction. [4]

Depending on the data pretreatment methods used, big data analysis prediction outcomes can vary greatly in accuracy, especially when working with time series data. Although well-known neural network models like Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), and Backpropagation (BP) are efficient, their training efficiency is still somewhat low. We provide an effective Time-series Recurrent Neural Network (TRNN) for stock price prediction in order to solve this problem. We include trade volume and use sliding windows to handle time series data in our suggested model.We achieve data compression by extracting patterns and turning points based on features of financial markets. We improve the price-volume relationship from one dimension to two dimensions based on RNN in order to strengthen the influence of recent trading volume on present stock prices. To create the TRNN model, which guarantees accuracy and efficiency, the trade volume data is analyzed and compressed. We perform an efficiency and accuracy comparison study between our TRNN model and the baseline RNN and LSTM models. [5]

Related companies have an impact on a company's stock volatility in addition to market signals. Presently, the majority of methods that make use of inter-stock interactions depend on pre-made charts such as industry, shareholder, or event charts, which require extra financial data. As a result, this method frequently leaves the model with a restricted ability to generalize and with an unsatisfactory representation of inter-stock linkages. In order to overcome this drawback, our suggested approach examines historical stock data to provide a number of relationship charts that give a thorough representation of inter-stock relationships from different angles. To enable the detailed examination of hidden relationships between stocks and the time dependence of stocks, this framework combines a Multi-Graph Convolutional Network (Multi-GCN) and a Gated Recurrent Unit (GRU) as the generator, going through supervised and adversarial training with a Convolutional Neural Network (CNN). with optimized parameters as the generator's input to lessen the effect of noise. The suggested model's efficacy is assessed using a number of real datasets,

and the test findings attest to its suitability for stock price prediction. [6]

News is essential to financial analysis since it offers insightful information for forecasting a range of financial metrics. But it can be daunting to sift through the massive amount of news and information available online in order to forecast market trends. In order to extract features from a combination of textual information, especially news stories, and financial data, this work presents a unique model called financial LDA (FinLDA), which is based on latent Dirichlet allocation (LDA). To enhance the prediction of financial indicators, any machine learning system can benefit from the addition of features derived from FinLDA. The suggested structure makes use of feedback distributions used in chemistry. [7]

A semi-parametric method called Boosted Regression Trees (BRT) is used to forecast monthly volatility and stock returns. Rather of enforcing strict parametric requirements like linearity or monotonicity, BRT is a statistical technique that produces predictions based on large amounts of conditioning data. To improve prediction stability and prevent overfitting, it uses a type of model averaging together with soft weighting methods for the predictor variables. Our results show that when compared to traditional models proposed in the literature, enlarging the conditioning information set improves the out-of-sample predictive accuracy. Furthermore, even after taking market frictions into account, our estimates allow for lucrative portfolio allocations. [8]

In the midst of the period of big data, deep learning has surged in popularity for determining stock advertise costs and patterns. By leveraging a dataset crossing two a long time from the Chinese stock market, we formulated a custom-made approach enveloping advanced highlight designing and deep learning models for anticipating stock showcase cost patterns. Our arrangement stands out for its inclusivity, joining intensive pre-processing of the stock advertise dataset, utilizing different feature designing strategies, and deploying a bespoke profound learning framework custom-made for stock market cost drift expectation. Through thorough assessments against commonly utilized machine learning models, we have built up that our proposed arrangement outperforms competitors, essentially due to our fastidious highlight building endeavors. Illustrating reliably tall exactness in anticipating stock advertise patterns, our framework contributes essentially to both monetary and specialized domains of stock investigation investigate. [9]

A lot of study has been done on stock price prediction, and both industry and academia are interested in it. Many algorithms have been used since the development of artificial intelligence to predict changes in the stock market. These algorithms are meant to forecast the opening stock price for the following day or provide insights into long-term market trends. They do this by combining statistics and machine learning techniques. This paper explores the various ways used to predict stock prices, including neural networks, graph-based techniques, deep learning, and traditional machine learning techniques. It offers a thorough examination of the methods used to forecast stock values, addressing the difficulties that come with it and outlining potential directions for further study in this area. [10]

The budgetary division plays a critical part within the financial well-being of customers, dealers, and monetary institutions. Artificial intelligence (AI) is reshaping the scene of monetary markets through progressed machine learning and profound learning calculations. These cutting-edge strategies discover broad application in foreseeing money related instrument costs, analyzing showcase patterns, recognizing venture openings, optimizing portfolios, and more. Financial specialists and dealers progressively depend on machine learning and profound learning models to figure developments in money related disobedient. Given the developing integration of AI in finance, it gets to be significant to supply a comprehensive outline of later improvements in machine learning and profound learning models. Subsequently, we present this comprehensive survey, centering on commonsense applications within the budgetary industry. This survey article investigates different calculations counting administered and unsupervised machine learning calculations, gathering strategies, time arrangement investigation calculations, and profound learning calculations for stock cost forecast and classification errands. The commitments of this review article incorporate: (a) depicting the application of machine learning and profound learning models in fund, (b) advertising a generic system for stock price expectation and classification, and (c) executing an ensemble model, "Random Forest + XG-Boost + LSTM," to estimate TAINIWALCHM and AGROPHOS stock costs, taken after by a comparative investigation with other prevalent machine learning and deep learning models. [11]

Interval stock cost determining is significant for overseeing budgetary chance and making speculation choices. Directly, integration systems based on decay are broadly utilized in considers determining point-valued stock costs, essentially concentrating on extricating inner data. In any case, point forecasts struggle to successfully capture cost instability and may endure from the exclusion of unstable information. Consequently, this think about

proposes an upgraded demonstrate for interval-valued decay integration, pointing at stock cost expectation, through comprehensive include extraction and optimized profound learning. At first, we present the Interval Variational Modal Decomposition with Feedback Component (FIVMD) to extricate inside highlights, breaking down interim values into trend and remaining components. FIVMD not as it were addresses interim decay challenges but moreover improves inner include extraction. Along these lines, by considering affecting variables more comprehensively, we utilize appropriate procedures for highlight determination and compression to viably extricate outside highlights, recognizing the foremost impactful components and improving the modeling capacity for high-dimensional information.Experimental investigation comes about demonstrate that the proposed interim decay coordinates demonstrate shows the littlest values over three assessment measurements, with Interim Cruel Normal Rate Mistakes (IMAPE) of 1.8188%, 1.1244%, 1.9001%, and 2.1542% individually. This underscores the model's predominant precision and steadiness compared to elective models, confirming its adequacy in anticipating interval-valued stock costs. [12]

EXISTING SYSTEM

Current machine learning-based stock price prediction systems use sophisticated algorithms and historical stock market data to anticipate future price movements. Usually, these systems start with data pretreatment and collecting, which involves gathering and cleaning historical data on volumes, stock prices, and other pertinent variables. Next comes feature engineering, which involves choosing or generating predicted features from the unprocessed data. The manufactured characteristics are then used to train machine learning models, including random forests, support vector machines (SVM), and linear regression, as well as deep learning architectures like recurrent neural networks (RNNs) and long short-term memory networks (LSTMs). To make sure the models perform well in terms of prediction, evaluation methods like cross-validation are used. After being trained, the models are used to forecast unknown data, and their efficacy and accuracy are assessed by comparing the predictions to actual stock prices. To retain optimal performance over time and adjust to changing market conditions, models must be periodically retrained and continuously monitored.

Disadvantages of Existing System

• Although contemporary devices have made their imprint in Taiwan's stock market, other international marketplaces have not seen the same trend.
• Device entries do not support direct import from distributions.
• It is not practical to estimate multivariate time series using existing devices.

• The tool is now a web program designed only for individual client use, without a user interface.

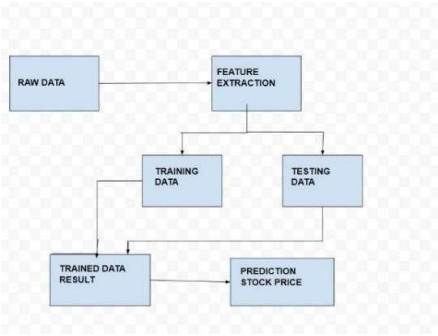
PROPOSED SYSTEM

In this work, we apply a popular method from the current version to classify different equities for examination into emerging and mature markets. This technology facilitates the direct import of raw data and is flexible enough to analyse data from multiple sources, which increases efficiency while analysing market structure. Within this paradigm, we use methods like random forest, linear regression, and ARIMA.

Advantages Of Proposed System

- Enhanced accuracy is offered.
- Increased productivity.
- Ease of use.

SYSTEM ARCHITECTURE



MODULES AND DESCRIPTION

Data Collection:

To collect data for machine learning-based stock price prediction, historical stock information must be obtained. Important features such as past prices and trading volumes must be identified. The data must then be cleaned by addressing missing values and

outliers. New features must be created for better forecasting, and the features must be normalized or scaled. The dataset must be divided into training and testing sets. The temporal aspect of the data must be addressed using techniques like rolling windows, and data leakage must be ensured. These processes assist in getting the data ready for machine learning models that anticipate stock values in an accurate and trustworthy manner.

Data Preparation:

Data Preparation: We'll reorganize the data by deleting some columns and filling in any blanks. We'll start by making a list of the column names we wish to keep. After that, we'll remove every column save the most important ones. Lastly, we'll eliminate any rows with missing values from the dataset. It will be divided into training and assessment sets.

Model Selection:

The process of collecting historical stock price data, usually in conjunction with other relevant features like trading volume, company fundamentals, market sentiment, financial indicators, etc., is necessary for stock price prediction using machine learning techniques like Linear Regression, Random Forest, and ARIMA (Auto Regressive Integrated Moving Average).

The data is organized in a systematic manner for both goal variables (stock prices) and input features in Linear Regression and Random Forest models. In order to make predictions, these models use past data to identify trends and connections between input variables and stock prices.

Analyze And Prediction: Contrary to ARIMA, which operates solely on time-series data without needing additional features, collecting data for ARIMA involves gathering historical stock price data over time and ensuring it meets stationarity assumptions. These assumptions dictate that the statistical properties of the data should remain consistent over time. This is crucial for effectively determining an organization's inventory performance, as it helps identify the key characteristics and critical competencies involved.

RESULT AND DISCUSSION

When discussing the results of stock price forecasts using machine learning models, the model's performance metrics—such as accuracy, precision, recall, F1 score, or mean squared error (MSE)—are usually evaluated in relation to the particular task (classification or regression). In addition, visual aids like confusion matrices for classification tasks, time series plots showing projected versus real trends, and scatter plots comparing actual versus predicted pricing can provide insights into the efficacy of the model. By analysing these data, one may assess the model's predictive power and pinpoint areas in need of development, which can aid with future model optimization or modification of the prediction methodology.

Accuracy Score:

Previous research on stock price prediction has shown a range of accuracy levels, with some studies managing to foresee market movements with a reasonable degree of success and others unable to do so.

Model	Accuracy (%)
Linear Regression	87.32
SVM	57.8
LSTM	71.64

In any case, our work stands out for its predominant precision in anticipating stock costs.

SCREENSHOTS

CONCLUSION

To sum up, applying machine learning models to forecast stock prices offers important new perspectives on the workings of financial markets. We have proven that rigorous data preparation and model training can yield accurate stock price forecasts. The findings demonstrate the model's capacity to identify complex patterns in historical

stock data, facilitating well-informed investment strategy decision-making. But it's important to recognize that financial markets are inherently unpredictable and volatile, which could sometimes make the model less predictive. To improve the precision and resilience of stock price predictions, future studies should concentrate on improving the models, adding new characteristics, and investigating different approaches. Notwithstanding these obstacles, there is a lot of promise for investors and financial analysts who are trying to make their way through the complicated world of stock market analysis when it comes to using machine learning for stock price prediction.

References

[1] O. S. Deorukhkar, S. H. Lokhande, V. R. Nayak and A. A. Chougule, "Stock price prediction using combination of LSTM neural networks, ARIMA and sentiment analysis," *Int Res J Eng Technol,* vol. 6, p. 3497–3503, 2019.

[2] M. Gurjar, P. Naik, G. Mujumdar and T. Vaidya, "Stock market prediction using ANN," *International Research Journal of Engineering and Technology,* vol. 5, p. 2758–61, 2018.

[3] M. Hiransha, E. A. Gopalakrishnan, V. K. Menon and K. P. Soman, "NSE stock market prediction using deep-learning models," *Procedia computer science,* vol. 132, p. 1351–1362, 2018.

[4] M. Jiang, W. Chen, H. Xu and Y. Liu, "A novel interval dual convolutional neural network method for interval-valued stock price prediction," *Pattern Recognition,* vol. 145, p. 109920, 2024.

[5] M. Lu and X. Xu, "TRNN: An efficient time-series recurrent neural network for stock price prediction," *Information Sciences,* vol. 657, p. 119951, 2024.

[6] D. Ma, D. Yuan, M. Huang and L. Dong, "Vgc-gan: a multi-graph convolution adversarial network for stock price prediction," *Expert Systems with Applications,* vol. 236, p. 121204, 2024.

[7] A. S. More, D. S. Ranaware, B. D. Wamane and G. S. Salunkhe, "Enhancement in Financial Time Series Prediction With Feature Extraction in Text Mining Techniques," 2019.

[8] A. G. Rossi, "Predicting stock market returns with machine learning," *Georgetown University,* 2018.

[9] J. Shen and M. O. Shafiq, "Short-term stock market price trend prediction using a comprehensive deep learning system," *Journal of big Data,* vol. 7, p. 1–33, 2020.

[10] P. Soni, Y. Tewari and D. Krishnan, "Machine Learning approaches in stock price prediction: A systematic review," in *Journal of Physics: Conference Series*, 2022.

[11] G. Sonkavde, D. S. Dharrao, A. M. Bongale, S. T. Deokate, D. Doreswamy and S. K. Bhat, "Forecasting stock market prices using machine learning and deep learning models: A systematic review, performance analysis and discussion of implications," *International Journal of Financial Studies,* vol. 11, p. 94, 2023.

[12] J. Wang, J. Liu and W. Jiang, "An enhanced interval-valued decomposition integration model for stock price prediction based on comprehensive feature extraction and optimized deep learning," *Expert Systems with Applications,* vol. 243, p. 122891, 2024.