AIRO JOURNALS | Peer Reviewed Multidisciplinary International

ISSN:2320-3714
Volume:4 Issue:3
December 2024
Impact Factor: 10.2
Subject: Computer Science

Using AI To Make Accurate Stock Market Predictions: An Analysis from The Perspective of Machine Learning

Swagnik Dey
Research Scholar
(BTECH Computer Science and Business Systems)
Heritage Institute of Technology Kolkata
email ID : swagnik.dey.official@gmail.com
Mayukh Paul
Student (BCA)
IEM KOLKATA
Using AI to Make Accurate
Stock Market Predictions An
Analysis from the Perspective
of Machine Learning

DECLARATION: I AS AN AUTHOR OF THIS PAPER /ARTICLE, HERE BY DECLARE THAT THE PAPER SUBMITTED BY ME FOR PUBLICATION IN THE JOURNAL IS COMPLETELY MY OWN GENUINE PAPER. IF ANY ISSUE REGARDING COPYRIGHT/PATENT/OTHER REAL AUTHOR ARISES, THE PUBLISHER WILL NOT BE LEGALLY RESPONSIBLE. IF ANY OF SUCH MATTERS OCCUR PUBLISHER MAY REMOVE MY CONTENT FROM THE JOURNAL WEBSITE. FOR THE REASON OF CONTENT AMENDMENT /OR ANY TECHNICAL ISSUE WITH NO VISIBILITY ON WEBSITE /UPDATES, I HAVE RESUBMITTED THIS PAPER FOR THE PUBLICATION.FOR ANY PUBLICATION MATTERS OR ANY INFORMATION INTENTIONALLY HIDDEN BY ME OR OTHERWISE, I SHALL BE LEGALLY RESPONSIBLE. (COMPLETE DECLARATION OF THE AUTHOR AT THE LAST PAGE OF THIS PAPER/ARTICLE

Abstract:

Aim:This research work aims at finding how ML models are used for the short-term prediction of stock market based on how past data is used for forecasting patterns of stock price movements..

Purpose:The purpose of this research is to compare different machine learning techniques for the prediction of stock price movements. Techniques used include Logistic Regression, Bayesian Networks, Support Vector Machines with radial basis kernels, and Stochastic Neural Networks. Technical indicators will also be discussed on how they impact prediction accuracy.

Method:The study collected historical stock market data on 80 NYSE stocks using the Alpha Vantage API, thus capturing open, high, low, close prices, volume, and other time-series data. A total of 20 predictors, incorporating 15 commonly used technical indicators, were used for training the models. Models were evaluated based on criteria such as mean squared error MSE and prediction accuracy. It utilized different methods of machine learning and emphasized predicting the trend of stock prices.

Results:The preliminary results of the "MSFT" stock demonstrated that technical indicators are indeed valuable for improving the accuracy of predictions. Among the models under test, Support Vector Regression outperformed linear models in terms of classification success and reached 69.5%. Long Short-Term Memory networks were unable to converge and had poor accuracy, although promising at first.

Conclusion:This study concludes that machine learning models, especially SVR, are effective for predicting the stock price trend. However, this study also pointed out some challenges in scaling complex algorithms to large datasets and issues with the performance of some models, like LSTM.

Keywords: ***AI-based Stock Market Predictions, Machine Learning in Finance, Predictive Analytics, Stock Market Forecasting, Algorithmic Trading***

ISSN:2320-3714
Volume:4 Issue:3
December 2024
Impact Factor:10.2
Subject:Computer Science

I. INTRODUCTION

Investors have always faced a difficult environment in the stock market due to the inherent complexity and dynamic nature of the market. When it comes to capturing the intricate patterns and sharp shifts in market sentiment, traditional methods of analysis frequently fall short [1]. When it comes to predicting the movements of the stock market, there has been a significant increase in both interest and investment in the utilization of artificial intelligence (AI) in recent years. It is possible that artificial intelligence will be able to make predictions that are more accurate and timely because of its capacity to process vast amounts of data, recognize patterns, and adapt to changing conditions [2].

In this article, we investigate the ways in which artificial intelligence is revolutionizing the prediction of stock market movements and the potential advantages that it may bring to investors.

A. Background of Stock Market Prediction

Since the beginning of time, economists, financial analysts, and scholars have been interested in the issue of stock market prediction due to the relevance it holds in terms of investment decision-making and the stability of the economy. In the beginning, there were ways for anticipating stock movements based on historical data and market indicators. These early methods included traditional forecasting methodologies such as fundamental analysis and technical analysis. The complexity and fluidity of the financial markets, on the other hand, were frequently not well captured by these methodologies. The field of financial forecasting entered a new age through the introduction of machine learning, which made it possible to gain insights that were more accurate and based on data [3].

Throughout this part, we will investigate the historical development of stock market forecasting, the limitations of traditional methods, and the increasing significance of machine learning in addressing these difficulties.

FIGURE 1: STOCK PRICE PREDICTION

B. The Evolution of Artificial Intelligence in Financial Markets

The financial markets have been completely transformed by artificial intelligence (AI), which has progressed from simple computing tools to complex frameworks that are able to analyze large-scale datasets that contain several dimensions. In the beginning, artificial intelligence was utilized in algorithmic trading and risk management; but, with the advent of machine learning, its capabilities have significantly grown even more. The capacity to forecast stock prices has been significantly improved by the application of machine learning methods, which include support vector machines, neural networks, and ensemble techniques. These algorithms are able to uncover previously hidden patterns and trends in market data [4].

FIGURE 2: THE EVOLUTION OF ARTIFICIAL INTELLIGENCE IN FINANCIAL MARKETS

This section covers the milestones in the development of artificial intelligence for financial markets, beginning with the early statistical models and progressing all the way up to the incorporation of deep learning and natural language processing

ISSN:2320-3714
Volume:4 Issue:3
December 2024
Impact Factor: 10.2
Subject: Computer Science

techniques, which have further strengthened prediction abilities and market analysis [6].

C. Importance of Accurate Stock Market Predictions

The purpose of avoiding financial risks and ensuring that decisions are made in an informed manner, accurate stock market predictions are fundamental [7]. The purpose of this section is to investigate the far-reaching economic ramifications of accurate market forecasting, paying particular attention to the role that it plays in maintaining economic stability, directing investment strategies, and building market confidence. It is possible for traders and investors to greatly improve portfolio management and risk avoidance by making correct predictions. More efficient use of resources and improved resource allocation are also beneficial to financial organizations. It is imperative that advanced machine learning techniques be utilized in order to obtain a higher degree of accuracy in forecasts, which will eventually promote sustainable financial growth. This is because the complexity of global financial markets is increasing at an alarming rate [8].

D. Overview of Machine Learning Models for Stock Market Prediction

A wide assortment of algorithms that are explicitly intended to address the intricacies of financial data are made available by machine learning, which has arisen as a distinct advantage in the field of stock market prediction. In this segment, a top to bottom outline of the machine learning models that are frequently utilized is introduced. These models include linear relapse, choice trees, irregular backwoods, and slope boosting machines. Utilizations of these models in forecasting are likewise talked about. Likewise, it separates among supervised and unsupervised learning, investigating the manners by which these two techniques for processing financial data are distinct from each other. Besides, the incorporation of deep learning and neural networks, which can perceive nonlinear connections and intricate patterns, is being promoted as a huge forward-moving step in the development of market forecasting [9].

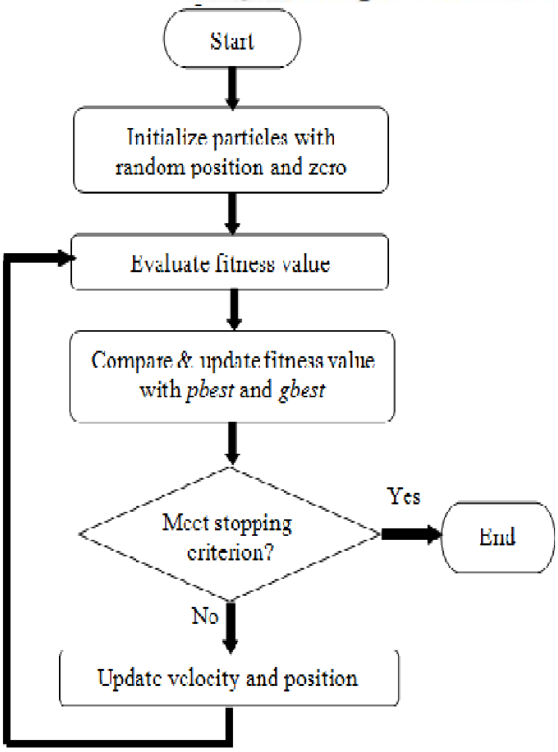


FIGURE 3: A MODEL OF MACHINE LEARNING FOR PREDICTION OF STOCK MARKETS

E. Challenges in Stock Market Prediction

Predicting the stock market continues to be troublesome because of the volatility and capriciousness of financial markets, notwithstanding the headway that has been made in artificial intelligence and machine learning. The restrictions of conventional models, which don't consider the elements of the market progressively, are talked about in this part, along with the inherent challenges that are related with predicting. With regards to data quality, it handles issues like as inconsistencies, missing qualities, and commotion, which can all be negative to the presentation of machine learning algorithms. Likewise, the trouble of pre-processing financial data in request to ensure that it is reasonable for analysis is underlined as a fundamental stage during the time spent overcoming these issues and enhancing the reliability of predictions [10].

F. Scope of the study

The motivation behind this study is to investigate a utilization of machine learning techniques to foresee short-term changes in stock prices. The review centers around predicting stock price trends based on authentic data, including open, high, and low prices, as well as trading volume, and it does as such by using daily time series data of eighty

Peer Reviewed Multidisciplinary International
AIRO JOURNALS
ISSN:2320-3714
Volume:4 Issue:3
December 2024
Impact Factor: 10.2
Subject: Computer Science

stocks that were picked aimlessly from the New York Stock Trade (NYSE). To determine whether the stock price will go up or down over the following 'n' days, the examination dissects various different machine learning models, including Strategic Relapse, Bayesian Networks, Neural Networks, and Backing Vector Machines (SVM). Likewise, the exploration investigates the impacts of combining fifteen specialized indicators, which are gotten from notable models utilized in stock market analysis, into predictive models in request to work on the accuracy of predictions. Among the models that are being examined are Linear Relapse, Tether, Edge Relapse, and Backing Vector Relapse (SVR) [13]. The motivation behind this evaluation is to perceive the way that well these models can estimate changes in stock prices and whether they can decrease prediction blunder. Also, the degree includes the investigation of the computational hardships that emerge while employing specific models, especially Backing Vector Relapse and Long Short-Term Memory (LSTM) networks, two models that require the successful management of enormous datasets. To determine how well these models sum up and adjust to various data volumes, the presentation of these models is assessed using key measurements like mean squared mistake and prediction accuracy [14].

G. Aim and research Objectives

This study's main objective is to assess and differentiate how well unique machine learning models —, for example, Bayesian networks, linear relapse, support vector machines, basic neural networks, Edge and Tether relapse, and others — anticipate stock market trends, with a specific accentuation on short-term variances in stock prices[15]. Using time-series data and specialized indicators to work on conjecture accuracy, this analysis is done from a machine learning standpoint.

The following objectives are recorded beneath:

1. To evaluate the predictive power of different machine learning models, such as Bayesian networks, support vector machines, neural networks, and logistic regression, in predicting short-term stock price movements based on historical trading volumes and stock prices.
2. To evaluate by adding technical indicators, such as momentum and moving averages, can increase the precision of machine learning models used to predict stock prices.
3. To examine the well-suited advanced models—like Support Vector Regression and Long Short-Term Memory (LSTM) networks—are for predicting stock values in contrast to more straightforward linear regression methods.

II. LITERATURE REVIEW

Rouf et al. (2021) [17]brought to light the fact that conventional trading methods were being revolutionized by the emergence of cutting-edge technology, which was ushering in a new age of stock market prediction. Many wealthy people turned to stock trading as a primary means of investing in the ever-increasing market capitalization. In order to help investors, make educated judgments, researchers and analysts had created a number of tools and approaches for predicting the movement of stock prices. Modern trading algorithms made it possible to forecast market movements using social media posts and other non-traditional textual data. Text data analytics and ensemble methods are two examples of machine learning technologies that greatly improved prediction accuracies. Using a generic framework as its focal point, the research examined the systematic methods of machine learning for stock market prediction. Results from 2011–2021, culled from resources including Scopus and the ACM Digital Library, were subjected to a rigorous evaluation. Furthermore, in order to determine the major trends and future directions of this discipline, a thorough comparative analysis was carried out.

Kumar et al. (2022) [12]brought to light the fact that patterns in stock market prediction are now seen as an important activity that leads to better decisions and possible gains. Stock market predictions have been very difficult for investors to make because of static and noisy data. Therefore, one of the biggest obstacles for profit-maximizing investors has been predicting stock market developments. When it comes to predicting the stock market, the authors stressed the importance of mathematical methods and learning tools. They summarized 30 studies and offered recommendations for approaches including computation methodologies, ML algorithms, performance metrics, and prestigious journals. The chosen research contributed to a better understanding of the ML methods and datasets utilized for stock market forecasting. In order to get reliable stock market predictions, the majority of people use ANN and NN, according to the study. Many restrictions persisted in the most recent stock market forecasting approaches, despite significant advancements in the field. In order to improve accuracy, the study recommended seeing stock market forecasting as an integrated process and putting more emphasis on unique parameters.

Strader et al. (2020) [19]analyzed the many facets of stock market investment methods that depend on analyzing large datasets. Examining the potential benefits of machine learning techniques over more conventional approaches to market forecasting has garnered increasing attention in recent years. The purpose of this study was to conduct a comprehensive literature analysis on the topic of machine learning-based stock market prediction in order to determine potential areas for future research. Reviewing publications published in peer-reviewed journals during the last 20 years and classifying them according to research methodologies and settings was the methodology. Research using artificial neural networks, support vector machines, hybrid or other AI approaches, and studies integrating genetic algorithms with other methodologies formed four distinct groups. Every category has its own set of restrictions, common and unique results, and places that needed additional research brought to light by the review. In its last section, the study summed up its findings and offered suggestions for further investigation.

Jiang (2021) [11] emphasized the traditionally difficult nature of stock market prediction, which has garnered a large amount of interest from professionals in the fields of economics and computer science. Both linear and machine learning tools have been investigated throughout the course of the last few decades in an effort to produce forecasting models that are accurate. Deep learning models have recently emerged as a new frontier on account of the significant breakthroughs that have been made in the discipline itself. For the purpose of providing a complete evaluation of the most recent efforts on deep learning for stock market prediction, this survey was conducted. In addition to addressing implementation and repeatability, the review categorized a variety of data sources, neural network architectures, and evaluation criteria using a classification system. Facilitating the replication of prior studies as baselines and providing assistance to researchers in keeping up with recent advancements were the two primary goals of this project. The article also identified a number of prospective avenues that could be pursued by researchers in the field in the future.

Basak et al. (2019)[5]conducted and published in The North American Journal of Economics and Finance, they investigated the possibility of using tree-based classifiers to forecast the direction in which stock market prices would move. In order to forecast whether or not the prices of stocks will go up or down, the authors concentrated on applying machine learning techniques, namely decision trees. Compared to more conventional statistical models, they intended to achieve a higher level of accuracy in their forecasts by employing a variety of tree-based techniques. Tree-based classifiers were found to be effective in capturing non-linear correlations and giving improved prediction performance in financial markets, as demonstrated by their extensive investigation. By highlighting the practical applications of such algorithms in stock market analysis, this work made a contribution to the expanding corpus of research on the application of machine learning to financial forecasting.

Sheth and Shah (2023) [18] explored the application of machine learning techniques to predict stock market trends, especially on the accuracy and effectiveness of these methods in forecasting future stock prices. In their study, they discussed several machine learning algorithms and compared their performance in stock price prediction. They pointed out that some machine learning models, like regression, decision trees, and neural networks, are promising for increasing the accuracy of stock market predictions. The authors highlighted the fact that even though such models are promising, volatility in the markets and the complexity of the data make it challenging to overcome such hurdles. In addition, they surveyed existing literature, and earlier research had indicated mixed results wherein some works were very successful in achieving accuracy while others failed mainly because of the dynamic financial markets. In conclusion, Sheth and Shah concluded that machine learning is promising for stock market prediction but further refinement of models and consideration of market-specific factors are essential to achieve consistent and reliable results.

Alharbi (2024) [2]investigated the use of machine learning techniques in predicting the stock market prices in advanced global markets. The literature review showed an increasing interest in applying machine learning models to the financial market forecasting problem. It has used methodologies including support vector machines, neural networks, and decision trees. Previous studies suggested that such prediction of stock prices is difficult due to market fluctuation and noise. Meanwhile, discovering complex patterns hidden in such large datasets can be identified using machine learning, according to an emerging solution. Alharbi extends these findings in his analytical work, focusing on evaluating different machine learning algorithms toward predicting accuracy and efficiency to compare them while using real data from global stock markets for comparison.

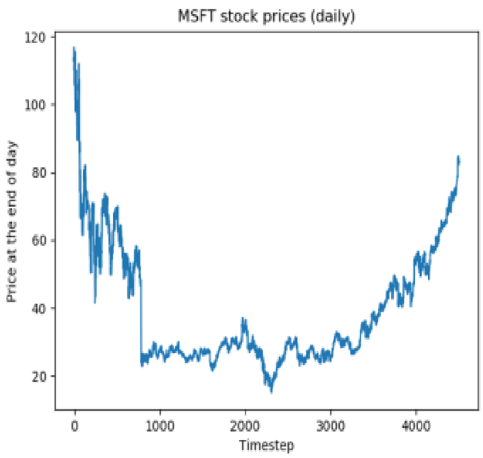
AIRO JOURNALS
Peer Reviewed Multidisciplinary International
ISSN:2320-3714
Volume:4 Issue:3
December 2024
Impact Factor: 10.2
Subject: Computer Science

Research Gap

This paper is developed by referencing existing literature with ML applications in the predictions of stock markets, such as that relating to various algorithms using Bayesian networks, SVMs, NNs, and decision trees. However, it was noted that much research lacks the comparative study between these models for using them specifically for short-term predictions of stock prices that incorporate historical trading volumes and stock prices. The above studies also refer to some technical indicators such as momentum and moving averages, but their influences on enhancing the accuracy of the ML models are yet unexplored. Moreover, there is a lack of comprehensive analysis regarding the comparative effectiveness of advanced models like Support Vector Regression (SVR) and Long Short-Term Memory (LSTM) networks compared to more traditional methods such as linear regression in the context of forecasting stock markets, especially concerning market volatility and data noise. This research gap opens up the opportunity to delve deeper into how combining these models and features can improve predictive performance.

III. DATASET AND FEATURES

Using the Alpha Vantage Programming interface, we are able to access the time series data of 80 stocks that are traded on the New York Stock Transfer. If you know how to use the application programming interface (API), you can access time series that are interday, daily, weekly, and monthly. We have chosen to start with daily time series data, which contains the opening price, high price, low price, close price, and volume for each day. Possible explanation: we're just interested in making predictions for the near future [16].



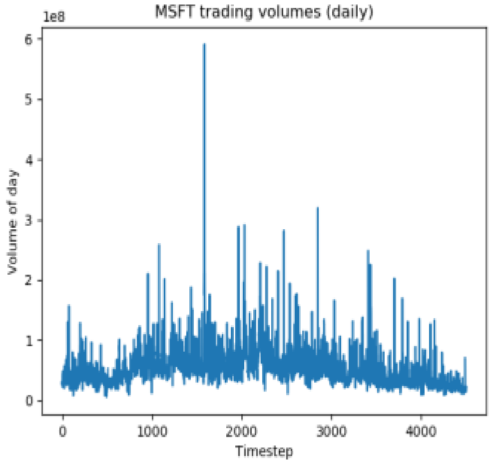


FIGURE 4: DAILY PRICES AND STOCK "MSFT" TRADING VOLUME"

Following that, we looked at a number of websites that provide information about investments. In order to obtain 15 technical indicators from the models, we took great attention in selecting ten models for the processing of time series data [18]. We chose the models based on how well-known they are and how many reviews they have received on investment education websites [20].

APO-SMA Absolute price oscillator values with SMA	$PO = SlowMA(price) - FastMA(price)$	Moving Average Convergence Divergence (**MACD**)	$shortema = 0.15 \times price + 0.85 * shortema_{[-1]}$ $longema = 0.075 \times price + 0.925 * longema_{[-1]}$ $MACD = shortema - longema$
APO-EMA		Stochastic oscillator (**STOCH**)	$\%K = 100 \times \frac{close - LowestLow_{[last\ n\ period]}}{HighestHigh_{[last\ n\ period]} - LowestLow_{[last\ n\ period]}}$ $\%D = MovingAverage(\%K)$

CCI Commodity channel index values	$CCI = \frac{TP - ATP}{0.015 \times MD}$
	$TP = \frac{high_n + low_n + close}{3}$
	TP = Typical Price
	$high_n$ = Highest high in the last n time periods
	low_n = Lowest low in the last n time periods
	$ATV = SimpleMovingAverage(TV)$
	$MDTV = MeanDeviation(TV)$
		Relative strength index (**RSI**) values	If $close > close_{[-1]}$ then
			$up = close - close_{[-1]}$
			$dn = 0$
			else
			$up = 0$
			$dn = close_{[-1]} - close$
			$upavg = \frac{upavg \times (n-1) + up}{n}$
			$dnavg = \frac{dnavg \times (n-1) + dn}{n}$
			$RMI = 100 \times \frac{upavg}{upavg + dnavg}$

AROON	$AroonUp = 100 \times \left(\frac{n - PeriodsSinceHighestHigh}{n}\right)$ $AroonDown = 100 \times \left(\frac{n - PeriodsSinceLowestLow}{n}\right)$		
Bollinger bands (**BBANDS**) values	$TP = \frac{high + low + close}{3}$ $MidBand = SimpleMovingAverage(TP)$ $UpperBand = MidBand + F \times \sigma(TP)$ $LowerBand = MidBand - F \times \sigma(TP)$		

Chaikin A/D line (**AD**) values.	$CLV = \left(\frac{(close - low) - (high - close)}{(high - low)}\right)$
	$AD = AD_{[-1]} + CLV \times volume$
		On balance volume (**OBV**) values.	If $close > close_{[-1]}$ then
			$OBV = OBV_{[-1]} + volume$
			elseIf $close < close_{[-1]}$ then
			$OBV = OBV_{[-1]} - volume$
			else
			$OBV = OBV_{[-1]}$

Average directional movement index (**ADX**) values.	$ADX = \frac{ADX_{[-1]} \times (n-1) + DX}{n}$		

We now have 260k data samples from eighty different companies after combining the stock price at the conclusion of the day and the beginning of each period, the volume of the same day, and the

ISSN:2320-3714
Volume:4 Issue:3
December 2024
Impact Factor: 10.2
Subject: Computer Science

technical indicators that were discussed before. We have a total of twenty predictors, which comprises the daily trade volume as well as the price.

IV. METHODS

Our goal is to identify a challenging one, therefore we began by fixing an improved problem: predicting the next n days' worth of price changes using the stock prices and volumes from the previous m days. We used the sklearn library to create Strategies for Relapse, Bayesian Organization, Straightforward Neural Organization, and Support Vector Machine with RBF component for this characterisation task. Next, we used these models to analyze the price of a specific stock, which we'll call "MSFT." We integrated the specialist indicators into the indicator after collecting our early findings and sought to evaluate the exact change in pricing that will occur over the next n days. To showcase our results, we've preserved the "AEB" stock, while data from several stocks is utilized for training and testing purposes. We randomly split the data into a training set and a test set at a ratio of nine to one for every sample. We use the mean squared error and the forecast mistake rate as metrics to compare the various relapse models. You can see if the price is going up or down by looking at the predicted mistake rate. We typically utilize the primary measurement to find out if we dislike over-fitting because it isn't as simple and can't be applied to other data with different alterations. To compare our results to our initial findings, we employ the following metric, which is our major model. We began with Linear Relapse and then tested two different regularizations of it, Tether and Edge, to see if we could improve our results. But we don't think over-fitting will be a big deal, therefore we don't think regularization like this will make our model much better than the essential one. On the other hand, they can help us find out if our indicators are meaningful.

Then, we used Help Vector Relapse (SVR) using the spiral premise component. Despite its excellent performance, we are unable to run this model on the entire dataset due to our limited computational resources. A well-known problem is that the space complexity required to do the piece stunt is O (nobservations2). In addition, the model's time complexity is extremely high. This is why we only used a subset of our total data when applying the model.

V. EXPERIMENTS AND DISCUSSION

A number of different machine learning models, such as logistic regression, Bayesian networks, basic neural networks, and support vector machines, are evaluated for their ability to accurately forecast the direction of stock price trends. With regard to time-series forecasting, we also study the influence of using technical indicators and sophisticated models such as ridge regression, lasso regression, and support vector regression, as well as the difficulties that are encountered when applying LSTM-based neural networks.

*A. **Preliminary Experiments***

Prior to tackling the subject at hand, we begin with a simplified question: is it possible to accurately forecast price trends? The results of our implementation of various models from the Sklearn package on the stock "MFTS" are as follows:

TABLE 1: RATES OF TEST ERROR FOR DIFFERENT MODELS

(m.n.)	(20.0)	(20.5)	(10.5)	(10.0)	(5.0)
Logistic regression	50.30%	50.00%	50.30%	45.23%	49.23%
Bayesian Network	51.80%	50.90%	41.12%	49.03%	45.12%
Simple Neutral Network	45.09%	44.90%	44.16%	40.23%	40.19%
SVM with rbf kernel	48.20%	44.70%	43.01%	45.12%	40.36%

Interpretation: In the table the test error rates for different models used to forecast stock price movements for the ticker symbol "MFTS" are shown in the table. The figures in parenthesis (20.0, 20.5, 10.5, 10.0, and 5.0) in the table represent the various parameters that were used to evaluate the various models. Support Vector Machine (SVM) with an RBF kernel, Bayesian Network, Simple Neural Network, and Logistic Regression are among the models that were put to the test. The Simple Neural Network performed best at 40.19% when the parameter was set to 5.0, and it consistently displayed the lowest error rates among

ISSN:2320-3714
Volume:4 Issue:3
December 2024
Impact Factor:10.2
Subject:Computer Science

the models across all parameter choices. At the same parameter setting, the SVM with RBF kernel achieved its lowest error rate of 40.36%, trailing closely behind. The error rates of the Bayesian Network and Logistic Regression were greater, with Logistic Regression showing the worst performance at 50.30% when the parameters were set to 20.0. Overall, the findings indicate that the SVM and Simple Neural Network models performed better than the other models for predicting stock price trends; nevertheless, more experimentation and improvement are required for increased accuracy.

B. Linear Regression with Technical Indicators

We decided to incorporate more predictors into the problem because our preliminary findings were adequate to demonstrate that historical prices by themselves are not a reliable indicator of the issue that we would like to address. By conducting research, we found that individuals in the field of trading every now and again use specialized indicators to make decisions. Accordingly, we have settled on the choice to incorporate these indications into our model. Along with the stock prices, we have a sum of 19 indicators for time-step, and we pick the 13 indicators that are used the most in the industry. Assuming that these 19 indicators are exact, we will make an endeavor to estimate the price change throughout the following n days.

The findings of our initial investigation drove us to the end that linear relapse would be the best technique in any case.

In the segment on techniques, we tended to the main measurement, and we found that the test misfortune changes somewhere in the range of 0.2 and 0.8 for n = 1 with different arbitrary parcels of the data. Then again, the training misfortune remains in the scope of 0.5 to 0.6. Based on this, apparently our model approves of overfitting, and it additionally sums up extremely very well.

For the second metric, we were able to reach a correct prediction rate of approximately 55% when n was greater than or equal to 10, and as high as 61.5% when n was equal to 1. The accurate categorization rate is depicted in the graphic that can be found in Figure 5.

The findings are in agreement with the findings of our preliminary research, and when we compare them to the findings, we discover that the prediction is improved more when the value of n is large, but the improvement is not as significant for values of n that are less.

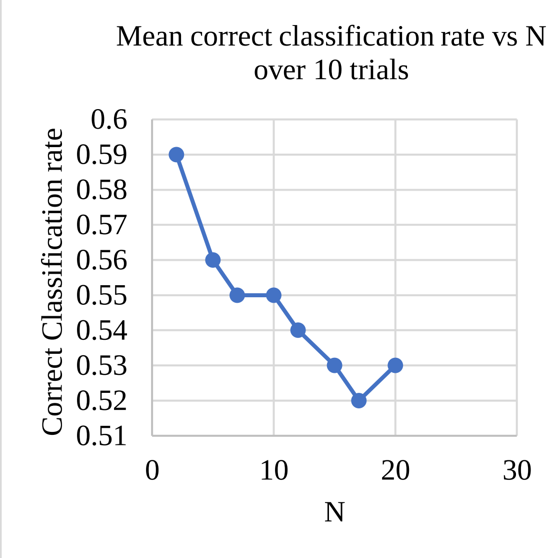


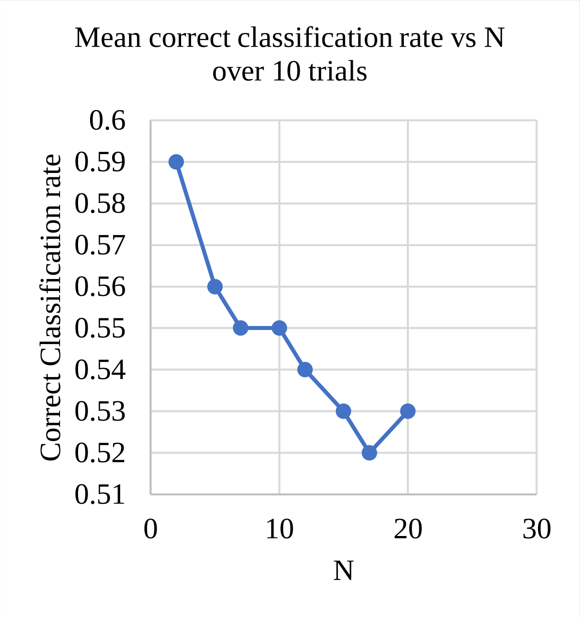
FIGURE 5: CORRECT CLASSIFICATION RATE FOR LINEAR REGRESSION

Interpretation: The connection that exists between the number of observations (N) and the rate at which information is correctly classified (CCR). Specifically, it demonstrates that the CCR generally decreases as the number of observations (N) grows, which indicates a downward trend in classification accuracy with larger datasets. It is at 0.59 that the CCR is at its highest for lesser values of N (for example, N=2). However, as N increases to 20, the CCR experiences some modest fluctuations but eventually settles around 0.53. This shows that the model's capacity to accurately classify decreases as the dataset grows, which may bring to light problems such as the model being underfit, the complexity of the data increasing, or noise having an effect on the performance as the number of observations increases.

For the purpose of providing an explanation for our choice of variables, we furthermore utilized ridge regression and lasso regression in addition to the more traditional linear regression. There is not a single coefficient that has significantly lowered, as seen by the output of the ridge regression coefficients. Taking into consideration the test error rates that are presented in Figure 6, this indicates that all of our predictors are significant, which is once again justified.

AIRO JOURNALS
Peer Reviewed Multidisciplinary International

ISSN:2320-3714
Volume:4 Issue:3
December 2024
Impact Factor: 10.2
Subject: Computer Science



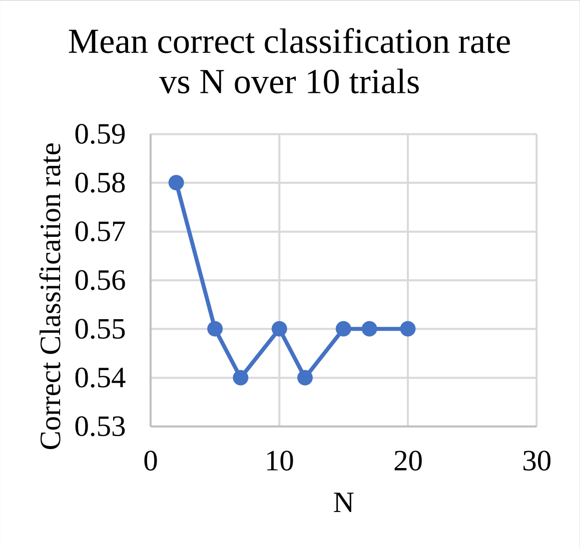


FIGURE 6: RIDGE REGRESSION AND LASSO REGRESSION

C. Support Vector Regression

A help vector machine (SVM) with a Gaussian piece is known to perform well, and subsequently, we have pursued the choice to apply support vector relapse to the data in request to determine whether we can accomplish results that are better than those obtained using linear relapse. Then again, because of the design of the model, we were constrained to deal with a more restricted training set during the execution stage: We are constrained to utilize the bit approach in light of the fact that the Gaussian portion changes our indicators into an infinite-layered feature vector. This outcomes in a reality intricacy of O (n2 tests), which is a lot of work. Consequently, working with multiple hundred thousand examples isn't plausible. At the point when we lead our trial, we pick five percent of our training data at irregular to act as the data for the relapse. The consequences of our examination are introduced in Figure 7, which shows the blunder rates that we showed up at.

The blunder rates vary, rather than linear relapse, since we select a more modest training set, which brings about a more noteworthy level of variety in the model that we have fitted. One trademark that is shared, in any case, is that we notice a similar pattern as n increases. This is on the grounds that the exact rate diminishes around the time when n rises to twenty. The way that we can arrive at a right pace of 69.5% for the highests and 68.5% for the lowests is a critical improvement over linear relapse, and likewise better than numerous different ventures are tantamount.

D. Other Attempt

A neural organization with a secret layer that was based on LSTM (long-term-short-term memory) and a thick layer for yield was likewise executed by us notwithstanding the models that we portrayed before. Considering that the model incorporates storing memory of past data in a time series, it is conceivable that it very well may be powerful for stock price prediction. The model is much of the time utilized for an assortment of normal language processing applications, for example, voice processing. We saw that the improvement MSE merges around the change of the advancement set, and the combined model created predictions that were close to nothing and had a legitimate prediction pace of 50%. Be that as it may, tragically, our execution of the model gave no indications of finding lasting success.

ISSN:2320-3714
Volume:4 Issue:3
December 2024
Impact Factor: 10.2
Subject: Computer Science

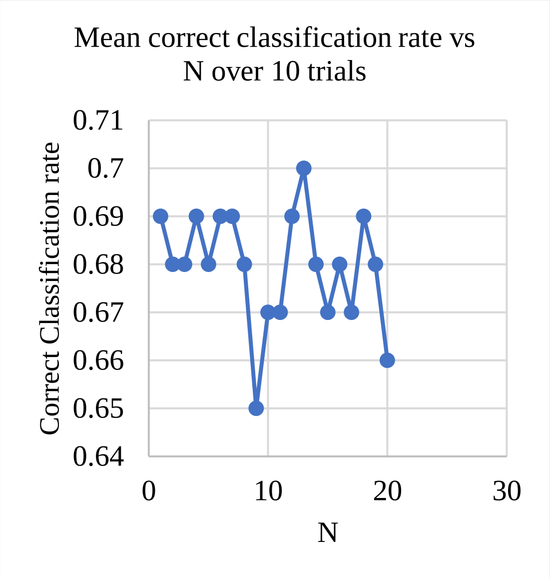


Figure 7: Support Vector Regression

The fluctuations in the performance of the classification function as the sample size rises. A minor decrease in classification accuracy is observed at the beginning of the study; however, this is followed by a recovery at specific sample sizes, which indicates that there is variability rather than a constant trend. Certain times in time result in an improvement in performance, which is followed by a subsequent decline as the sample size increases. Given this pattern, it appears that the accuracy of the model is affected by the sample size in a non-linear manner. This could be the result of variations in the distribution of the data, the flexibility of the model, or external factors that affect the efficiency of categorization.

D. Impact analysis

Practical Applications

The applications of machine learning models in predicting the stock market have many practical usages in finance. For example, employing techniques such as SVR and LSTM networks could make more informed decisions based on predictions about stock price trends, helping investors make better choices to trade on potential movements of the market, manage their portfolios better, and adjust trading strategies accordingly. Additionally, incorporating technical indicators like momentum and moving averages can enhance model accuracy, offering practical tools for day-to-day stock trading. Automated trading systems based on these models can also execute trades faster and more efficiently, capitalizing on market trends in real-time.

Financial Risks

Although machine learning models can give insights into the trends of stock prices, they do come with financial risks. The uncertainty and volatility in the stock markets mean that even highly accurate models may still be wrong in some predictions. This can result in potential financial losses for investors relying heavily on such models for trading decisions. Additionally, there is the problem of overfitting where the models may work well with historical data but fail to generalize to unseen data and hence make poor predictions for the future. Models such as SVM and LSTM are highly demanding in terms of computational resources, and working with smaller datasets may not allow accuracy in prediction, which again amplifies the risk. Investors must keep this in mind and apply them with caution.

Ethical Considerations and Biases

Use of machine learning in predicting stock markets raises ethical issues about the bias of the models. Since training data is used in creating these models, and it may be biased either because certain market conditions or economic events are underrepresented, predictions from such models could end up being skewed and may bring about unfair or unethical trading results. For example, models that are trained based on historical data may reflect past dynamics of the market that may not apply in the present economic scenario. Furthermore, a reliance on technical indicators would disadvantage investors who do not have access to sophisticated machine learning tools. Ethical considerations must include transparency in how the models are developed, fairness in the predictions, and ensure that the models do not perpetuate harmful biases in the behavior of the stock market.

Model performancefor Stock Market Prediction

Model	Error Rate (%) (Best Observed)	Key Parameters	Performance Insights	Strengths	Limitations
Logistic	45.23	Parameter =	Moderate	Simplicity,	Limited by

Regression		10.0	performance, struggles with capturing non-linear trends.	interpretable.	linearity; higher error rates.
Bayesian Network	41.12	Parameter = 10.5	Performs better with moderate parameter tuning.	Handles uncertainty well.	Sensitive to noise in large datasets.
Simple Neural Network	40.19	Parameter = 5.0	Best overall accuracy with lower parameter values.	Captures non-linear relationships effectively.	May underfit with minimal tuning.
Support Vector Machines	40.36	Kernel = RBF; Parameter = 5.0	Competitive accuracy; requires significant computational resources.	Robust against overfitting in smaller datasets.	Computationally expensive with large datasets.
Linear Regression	61.50 (CCR)	n = 1	Works well with technical indicators for small n.	Simple and interpretable.	Poor performance on larger, complex datasets.
Ridge/Lasso Regression	Comparable to Linear Regression	Regularization applied	Highlights the importance of technical indicators.	Avoids overfitting effectively.	Does not offer substantial accuracy gains.
Support Vector Regression	69.50 (CCR)	Kernel = Gaussian; n = 20	Outperforms others but computational limits prevent large-scale application.	Excellent for complex relationships.	High space and time complexity.
LSTM Neural Network	50.00	LSTM layers; dense output	Struggles with convergence and performance compared to simpler models.	Retains memory of sequential data effectively.	Convergence issues, less robust in this study.

ISSN:2320-3714
Volume:4 Issue:3
December 2024
Impact Factor: 10.2
Subject: Computer Science

Key takeaways: Combining technical indicators enhances model performance, though simpler models often struggle with the complexity of stock market data.

Simple Neural Networks outperformed traditional models: While the models listed above, the simplest neural network model with parameters of 40.19 percent error rates had the best overall accuracy with very low parameter values, capturing the non-linear associations in stock market data correctly. It shows that the high complexity of models is required to obtain the best outcomes.

The performance of SVR is good but requires more computational demands. SVR with a Gaussian kernel provided the best prediction performance in this study at an error rate of 69.50%. This indicates that this approach is less applicable in larger applications since it consumes a lot of computational resources, which in turn degrades its efficiency.

Linear models such as Logistic and Linear Regression are rather underperformance models: Both logistic regression, with a 45.23% error rate, and linear regression, with a 61.50% error rate, have a moderate to low performance in more complex data sets. They provide the simplicity and interpretability required but cannot capture the nature of the stock market data which is more non-linearly patterned and complex.

VI. CONCLUSION

In this study, the researchers highlight the effectiveness of machine learning models in predicting short-term trends in the stock market. They place a special emphasis on the usefulness of technical indicators. When compared to other methods, such as Logistic Regression and Bayesian Networks, Support Vector Regression (SVR) with a radial basis kernel fared the best among the models that were evaluated. It achieved a significant improvement in classification accuracy. In addition to stock prices and volumes, the incorporation of technical indicators resulted in the provision of valuable additional predictors, which increased the effectiveness of the models. The implementation of more advanced approaches, such as Long Short-Term Memory (LSTM) networks, however, presented difficulties, as these networks exhibited convergence problems and a lower level of accuracy than was anticipated. In addition, the computational limitations of SVR, notably the space and time complexity associated with the kernel trick, resulted in a limitation in its ability to scale to huge datasets. The study implies that machine learning, and more specifically Support Vector Machines, holds promise for stock market prediction, despite the limitations that have been presented. The improvements in model performance in real-world financial scenarios could be the result of future work that focuses on strengthening LSTM models and investigating hybrid techniques to solve the limits that were identified in this research.

VII. FUTURE SCOPE

The future scope of this research lies in refining the models and enhancing their predictive capabilities for forecasting in the stock market. Although the results here show the great potential of machine learning models such as Support Vector Regression (SVR) and Simple Neural Networks, the scope for improvement is highly significant. Future work may focus on the application of more advanced techniques like deep reinforcement learning to capture complex market dynamics and integrate real-time data to make better predictions. It may also be extended to a larger dataset that will include a broader range of stocks and the integration of macroeconomic indicators, news source sentiment analysis, and social media data to boost the accuracy of the models. Hybrid models which merge traditional statistical methods with the approaches of machine learning could be more effective. Furthermore, handling the problem of computational complexity by making algorithms more efficient for high-dimensional data and testing feature engineering techniques could make models more efficient and robust enough to open up the path for more reliable and scalable stock market prediction systems.

REFERENCES

1. *Abunasser, B. S., Daud, S. M., & Abu-Naser, S. S. (2023). Predicting Stock Prices using Artificial Intelligence: A Comparative Study of Machine Learning Algorithms. International Journal of Advances in Soft Computing & Its Applications, 15(3).*
2. *Alharbi, M. H. (2024). Prediction of the Price of Advanced Global Stock Markets Using Machine Learning: Comparative Analysis. Journal of Financial Risk Management, 13(4), 689-702.*
3. *Alpha Vantage. (2017). Alpha Vantage - Free APIs for real-time and historical financial data, technical analysis, charting, and more!*
4. *Awad, A. L., Elkaffas, S. M., & Fakhr, M. W. (2023). Stock market prediction using deep reinforcement learning. Applied System Innovation, 6(6), 106.*

ISSN:2320-3714
Volume:4 Issue:3
December 2024
Impact Factor: 10.2
Subject: Computer Science

5. *Basak, S., Kar, S., Saha, S., Khaidem, L., & Dey, S. R. (2019). Predicting the direction of stock market prices using tree-based classifiers. The North American Journal of Economics and Finance, 47, 552-567.*
6. *Bonde, G., & Khaled, R. (2012). Extracting the best features for predicting stock prices using machine learning. Proceedings of the International Conference on Artificial Intelligence (ICAI). The Steering Committee of The World Congress in Computer Science, Computer Engineering, and Applied Computing.*
7. *Chatzis, S. P., Siakoulis, V., Petropoulos, A., Stavroulakis, E., &Vlachogiannakis, N. (2018). Forecasting stock market crisis events using deep and statistical machine learning techniques. Expert systems with applications, 112, 353-371.*
8. *Chen, K., Zhou, Y., & Dai, F. (2015). A LSTM-based method for stock returns prediction: A case study of China stock market. 2015 IEEE International Conference on Big Data (Big Data), 2823–2824. IEEE.*
9. *Chhajer, P., Shah, M., & Kshirsagar, A. (2022). The applications of artificial neural networks, support vector machines, and long–short term memory for stock market prediction. Decision Analytics Journal, 2, 100015.*
10. *Hu, H., Tang, L., Zhang, S., & Wang, H. (2018). Predicting the direction of stock markets using optimized neural networks with Google Trends. Neurocomputing, 285, 188-195.*
11. *Jiang, W. (2021). Applications of deep learning in stock market prediction: recent progress. Expert Systems with Applications, 184, 115537.*
12. *Kumar, D., Sarangi, P. K., & Verma, R. (2022). A systematic review of stock market prediction using machine learning and statistical techniques. Materials Today: Proceedings, 49, 3187-3191.*
13. *Leung, C. K.-S., MacKinnon, R. K., & Wang, Y. (2014). A machine learning approach for stock price prediction. Proceedings of the 18th International Database Engineering & Applications Symposium. ACM.*
14. *Lin, C. Y., & Marques, J. A. L. (2024). Stock market prediction using Artificial Intelligence: A systematic review of Systematic Reviews. Social Sciences & Humanities Open, 9, 100864.*
15. *Mokhtari, S., Yen, K. K., & Liu, J. (2021). Effectiveness of artificial intelligence in stock market prediction based on machine learning. arXiv preprint arXiv:2107.01031.*
16. *Nti, I. K., Adekoya, A. F., &Weyori, B. A. (2020). A systematic review of fundamental and technical analysis of stock market predictions. Artificial Intelligence Review, 53(4), 3007-3057.*
17. *Rouf, N., Malik, M. B., Arif, T., Sharma, S., Singh, S., Aich, S., & Kim, H. C. (2021). Stock market prediction using machine learning techniques: a decade survey on methodologies, recent developments, and future directions. Electronics, 10(21), 2717.*
18. *Sheth, D., & Shah, M. (2023). Predicting stock market using machine learning: best and accurate way to know future stock prices. International Journal of System Assurance Engineering and Management, 14(1), 1-18.*
19. *Strader, T. J., Rozycki, J. J., Root, T. H., & Huang, Y. H. J. (2020). Machine learning stock market prediction studies: review and research directions. Journal of International Technology and Information Management, 28(4), 63-83.*
20. *Wen, M., Li, P., Zhang, L., & Chen, Y. (2019). Stock market trend prediction using high-order information of time series. Ieee Access, 7, 28299-28308.*

Author's Declaration

I as an author of the above research paper/article, here by, declare that the content of this paper is prepared by me and if any person having copyright issue or patent or anything otherwise related to the content, I shall always be legally responsible for any issue. For the reason of invisibility of my research paper on the website /amendments /updates, I have resubmitted my paper for publication on the same date. If any data or information given by me is not correct, I shall always be legally responsible. With my whole responsibility legally and formally have intimated the publisher (Publisher) that my paper has been checked by my guide (if any) or expert

ISSN:2320-3714
Volume:4 Issue:3
December 2024
Impact Factor:10.2
Subject: Computer Science

to make it sure that paper is technically right and there is no unaccepted plagiarism and hentriacontane is genuinely mine. If any issue arises related to Plagiarism/ Guide Name/ Educational Qualification /Designation /Address of my university/ college/institution/ Structure or Formatting/ Resubmission /Submission /Copyright /Patent /Submission for any higher degree or Job/Primary Data/Secondary Data Issues. I will be solely/entirely responsible for any legal issues. I have been informed that the most of the data from the website is invisible or shuffled or vanished from the database due to some technical fault or hacking and therefore the process of resubmission is there for the scholars/students who finds trouble in getting their paper on the website. At the time of resubmission of my paper I take all the legal and formal responsibilities, If I hide or do not submit the copy of my original documents (Andhra/Driving License/Any Identity Proof and Photo) in spite of demand from the publisher then my paper maybe rejected or removed from the website anytime and may not be consider for verification. I accept the fact that as the content of this paper and the resubmission legal responsibilities and reasons are only mine then the Publisher (Airo International Journal/Airo National Research Journal) is never responsible. I also declare that if publisher finds Any complication or error or anything hidden or implemented otherwise, my paper maybe removed from the website or the watermark of remark/actuality maybe mentioned on my paper. Even if anything is found illegal publisher may also take legal action against me.

Swagnik Dey
Mayukh Paul

******