Stock Market Analysis & Prediction

1st B. Shivani
Computer Science and Engineering
GRIET
Hyderabad Telangana, India.

2nd Dr. S. P Govinda Rao
Computer Science and Engineering
GRIET
Hyderabad TelanganaIndia.

Abstract*— Stock exchanges are an essential part of all global economies. Organizations can acquire capital this way to be able to perform their daily activities. By exchanging protection, securities, and values, stock intermediaries benefit from a market. Dealers, financial backers, and retailers can purchase and sell stocks once a company has been listed on the stock exchange. As of late, a great deal of work has been done to foresee the development of the stock market. Gauging the development of the financial exchange is acquiring force among different professionals, contributing networks, and followers as it gives better direction regard to contributing. Consistency is one of the main considerations on which, the benefit of exchanging stock and contributing is reliant. The benefits procured by speculation and exchanging the stock exchange rely upon the consistency of the stock, generally. On the off chance that any framework is created which can reliably foresee the patterns of the unique financial exchange, would make the proprietor of the framework wealthy. In addition, the anticipated patterns of the market will assist the controllers with taking remedial measures to balance out the market. Numerous master specialists and analysts have advanced a few models utilizing different specialized, principal, and scientific methods to give an expectation on the financial exchange design. During the most recent quite a while, a ton of studies have been done to predict stock exchange patterns utilizing Classical, AI, ML, and Deep learning methods. This study will help the pursuers and analysts in choosing algorithms that might be valuable for foreseeing the stock's performance. This survey of different algorithms and their boundaries for stock exchange prediction is included in this research.

Keywords*— AI, ML, Deep Learning, Stock Exchange

I. INTRODUCTION

The stock exchange plays a crucial part in the country's financial development just as the personal economy also. Figuring out the opportune chance to purchase and trade the stocks is subject to anticipating the patterns in the stock exchange. The procedure for the most precise expectation is to gain from past occurrences and plan a model to do this by utilizing conventional and AI algorithms [1]. The Stock exchange pattern differs because of a few factors like political, financial matters, climate, society, and so forth [2]. There are two sorts of stock examination. One is an essential examination, which requires investigation of the organization's nuts and bolts, for example, accounting report, costs and incomes, yearly returns, organization's profile, and position, and so on the other one is a specialized examination, which manages contemplating the insights produced by market exercises like verifiable information, past cost, and volume [3].

On the stock exchange, investors and stockholders can exchange the stocks of all organizations. With a limited focus time, this is an advantageous method of making a good income. Nevertheless, investors believe that it's exceptionally difficult to predict market prices. The costs are not only affected by monetary events and an organization's exhibition, however also by the brains of the people who put resources into the market. Currently, there are various factual methods for determining prices on securities exchanges. However, stock forecasts are difficult to achieve with common measurable strategies due to high instability and non-direct information relating to it.

Nowadays, Artificial Intelligence (AI) methods are offering promising results in expectations regarding financial exchanges as a result of the current status of the foundation [4]. There are two theories that can be used to predict securities exchange development: (I) the Random Walk approach and (ii) the Efficient Market Hypothesis (EMH). In 1970, FAMA recommended that EMH be constructed. It has been predicted by the EMH that the current market contains all the data it needs. As more information is accumulated, the market consumes it as well as getting its cost back. A financial backer cannot anticipate the stock market with some other method, it implies. A further division of EMH is frail EMH, semisolid EMH, and solid EMH. The weak EMH predicts the market based on previous verifiable information alone, the semi-solid EMH relies on chronicled information and publicly available data, and the solid EMH makes predictions based on information already recorded, public data, and other privately available data. Moreover, Random walk also states that current and chronicled stock costs, which are exceptionally unstable and independent of each other, cannot predict future stock costs. An alternative hypothesis, called Inefficient Market Hypothesis (IMH), holds that consistent showcases are not generally effective, as the stock price data alone does not provide sufficient information.

The future development of stock costs can be analyzed by utilizing other components that IMH suggests. It is possible to predict the stock market in two different ways: (i) by studying its fundamentals, and (ii) by studying its technical indicators. Foreseeing the stock price of an organization requires a fundamental analysis, which utilizes more market information like annual reports, financial records, and reviewer's reports. Utilizing time arrangement graphs, the specialized analysis only uses recorded stock value information of the organization. It has been mentioned before that the stock market is a highly unpredictable and dynamic one, so it is possible to use the above-mentioned methods as a basis for trading. However, AI is evidence showing that it is capable of predicting stock market costs in recent weeks. Access to information and the development of algorithms have made this possible. A review of recent work in the field

978-1-6654-2005-1/21/$31.00 ©2021 IEEE

of stock market forecasting is presented in this paper. A review of the strategies employed and outcomes achieved by each specialist has been conducted. This paper also attempts to address the gaps in their work.

II. Literature Survey

In 2007, SUI Xue-shen et al., proposed a data mining technique using SVM algorithm in which feature selection can eliminate irrelevant or redundant attributes and increase the density of samples in feature spaces that can improve classification performance [4]. In this proposed framework they used two new co-efficient which is NDEM-neighborhood decision error and ND- neighbor dependence is used to calculate the complexity of classification in order to do feature selection. To estimate the complexity of the input data classification they followed four measures which is ASNN, ASH, ND, NDER and specified 15 technical indicators which contains features using the four selected features sets of four measures is used as input data for Support Vector Algorithm to predict the moving trend for the next five days. In this experiment they concluded that NDEM pick the minimal set of features and achieved 63.23% of accuracy and ASNN got 62.93%, ASH and ND got 60.75% accuracies [4].

In 2009, Qinghua Wen et al., presented an experiment of Artificial intelligence-based stock market trading system based on oscillation box prediction by combining support vector machines and box theory of stock investment [5]. The SVM algorithm which predicts the lower bound and upper bound of the candle sticks respectively. This trading system is based on candle system rule. To analyze the possibility, they tested the framework on regular stock developments and S&P 500 segments. The tests show a promising exhibition of the framework and it beats the purchase and holds procedure. After doing the procedure they came to a conclusion which is; there is a complexity of the movement of stock because of using prediction algorithms with the combination of single algorithm which makes the system fragile. They have achieved the prediction accuracy for the lower and upper circuits is SVM max 96.91% and SVM min 95.93% for 3 stocks when it comes to 50 stocks the success rate is gradually decreased which is SVM max 56.29% and SVM min 46.28%.

In 2009, Ling-Bing Tang et al., Proposed a Manifold Wavelet Support Vector Machine (MWSVM) to predict the future returns of the stock market. This manifold wavelet kernel is composed of manifold theory which incorporated with Wavelet technique which is in SVM. This Wavelet technique can give the features output which describes the stock market time series of various locations at different time series. This method can accurately find out the nonlinear function and predict the stock market returns accurately. They have analysed the prediction performance using the 1. Normalised Mean Square Error (NMSE), 2. Mean absolute error (MAE) and 3. Root Mean Square Error (RMSE). Ling et al., experimented the model on different datasets (DAXINDX, FTSE100, JAPDOWA and SPCOMP) [6] and achieved these results.

Manifold Wavelet Kernel		
NMSE	RMSE	MAE
1.3309	0.2565	0.1974
0.8642	0.2412	0.1889
0.9034	0.1868	0.1468
0.8816	0.1818	0.1427
0.9252	0.3005	0.2412

Table 1: NMSE, RMSE, MAE Accuracies

In 2010, A. S. M. Shihavuddin et al., [7] proposed a Naïve Bayes classifier to predict the stock price by analysing the local economic trends and online financial news. In their research they represented the datamining algorithms which is tested on the stock market data which holds the advantage to interpret the present and future stocks. Naïve Bayes classifier is used to train the previous data present on the website and also news articles on the internet. They trained the classifier using the dataset which is collected from April 2008 to October 2008 which is a total of 140 and 35 datasets from October 2008 to November 2008 and tested. They find the performance varied at different datasets: they achieved an accuracy of 58.62% at Test-1, 66.62% in Test-2, and 79.4% in Test-3 but after adding the heuristics they have achieved 59.91% in Test-1, 79.12% in Test-2 and 87.22% in Test-3 [7].

In 2011, Hyun Joon Jung and J. K. Aggarwal from University of Texas at Austin come up with a Binary Stock Event Model (BSEM) for stock market prediction. A feature set is generated using this BSEM to forecast the future stock trend. The two learning models used in this method are Naive Bayes and Support Vector Machines. Predictions in this research are made based on binary classifications. As a result, a diagonal sum from the confusion matrix is used to compute the accuracy of prediction, which achieved 70% to 80% accuracy for the next day's stock trend on average, and it is observed that as the period increases, the accuracy falls as low as 55% [9]. This clearly demonstrates the benefits of BSEM features. Although the BSEM is showing 73*74 outcomes versus 55%*57% for the features-based solution. There is a critical contrast between highlights and shadows as a result. According to BSEM [9], stock pattern events are strongly related to stock occasions. However, an asset value is merely an estimate based on a technical indicator that doesn't emphatically reflect an asset pattern. Testing the prediction model was done by backtesting the strategy they used to evaluate it earlier [9]. In addition to the creation of a BSEM, an important contribution is the development of a feature generator. Despite its expectations and back-testing, it performed well on a genuine dataset. Additionally, they suggest a few intriguing points: Bayesian Naive Classifiers have the advantage of high computational complexity, while SVM-based methodologies have advantages and disadvantages, such as high forecast accuracy and expensive computation. According to their research, significant exchanging techniques like moving averages and relative strength files do not guarantee exceptional market returns [8].

In 2016, W. Lertyingyod and N. Benjamas proposed a method for predicting stock value expectations by reducing the number of features. The project proposes a novel information mining approach to determine stock costs. They rely on Thailand Stock Exchange for their work. With the Gain Ratio attribute and Wrapper Selection with Greedy stepwise pursuit technique, the features were determined using the Ranker Search and Ranker Attribute strategies. With the Wrapper Subset Evaluation with Greedy Algorithm, the ascribes were reduced from 14 to 6 (57.14%). Results of this study showed that the predictive model for week by week (5 to 10 days) stock value forecasting improved using Artificial Neural Networks and reached the highest accurate prediction of 93.89% using just six selected bonds [9].

In this research, M. R. Vargas et al., in 2018 used deep learning models to forecast a stock value utilizing financial news titles and specialized pointers as information. Two distinct, specialized markers are compared, set-1: stochastic (K%), rate of change, momentum, accumulation/distribution (A/D) oscillator, and dispersion 5, stochastic %D, Williams (R%). The second set consists of Moving Average Convergence-Divergence, Exponential Moving Average, On Balance Volume, Relative Strength Index, and Bollinger Bands. Using deep learning techniques, the information can be sorted out and examined in a complex way, allowing for more accurate conversion. Researchers have demonstrated that Convolutional Neural Networks (CNN) can be better at getting semantic information from writings than Recurrent Neural Networks (RNN), an RNN is better at retrieving setting information and displaying complex attributes for securities exchange determining. In this paper, we examine two models: a crossover model using CNN for the monetary news and Long Short-Term Memory (LSTM) for markers, called SI-RCNN, and an LSTM network designed for pointers only, known as I-RNN. When a model predicts that it will rise in cost a trade specialist purchases stock, sells it the following day, and then buys it on the following day when the model forecasts that it will decrease. The specialist buys stock on the current day and sells it the following day. The proposed strategy shows a significant part of monetary news in settling the outcomes and practically no improvement when looking at changed arrangements of specialized pointers [10].

Model	Training ACC (%)	Validation ACC (%)	Test ACC (%)
I-RNN	55.22	55.97	52.52
SI-RCNN	84.08	60.45	56.84
I-RNN-2	59.08	50.74	48.92
SI-RCNN-2	88.31	61.19	51.08

Table 2: Multiple Accuracies

In 2018, D Soni et al., proposed a Optimised prediction model for stock market trend analysis. The primary objective of this research is to apply the basic understanding of analysing the stock market trend using machine learning techniques which are Naïve bayes, Decision tree, PSO, Black Hole techniques [11]. Initially the bifurcated share market data into test and train datasets and calculates the absolute values of change in stock price overt the time interval and mean value of it. Likewise, they set up a grid wherein if the estimation of offer cost increments on a specific day regarding the past one they denoted that field in the lattice as 1 in any case assuming the offer value falls, the field is set apart as 0. Likewise, each line of the network compares to the day number. Later they tally the quantity of 1s and 0s in the readied framework. On the off chance that the quantity of 1s is more than the number of 0s, we group the share as a decent share else it would be assigned as a terrible share. After that they arranged a forecast framework and appointed it the qualities 1 or 0 as per the expectation and named as y. In the event that it is a decent offer, the estimation of the offer cost is expanded by the change determined before and the main line of the recently made forecast network is set apart as 1. Something else, assuming the offer is a terrible offer, the worth is diminished and correspondingly the primary column of the expectation framework is set apart as 0. Later again arranged a comparative network for the testing set too as far as 1s and 0s based on genuine ascent and fall in the cost of an offer and named as x. after that the following arrangement of days for forecast and would rehash the accompanying strides for each expectation from this time forward. This proposed model with the same dataset and tested it on a number of conditions. The results were obtained and the accuracy of this model is as follows [11].

S. No	1	2	3	4	5
Algorithm	Black Hole	PSO	Decision Tree	Naive Bayes	Proposed Model
Mean Accuracy	95.1	83.57	81.16	85.56	96.60

Table 3: Algorithm Accuracy

In 2018, Lyhyaoui et al., developed a stock forecasting model for the Casablanca stock exchange (CSE). Africa's CSE stock market is among the largest and most established in the world. It is additionally one of the most unpredictable stock trades, which can leave the financial backers with either great gains or enormous losses. In association with a model, the creators had predicted market developments for five Moroccan banks. There will be significant boundaries that will affect a stock value a great deal and plenty of variables may influence this. (I) Dimensionality decreases. Taking away dimensions will create a distinct boundary and it will simplify preparing a model because less measurement will be required. The paper uses a Kinetic Principal Component Analysis (KPCA). As a result, 23 informational boundaries have been identified. During the next stage, stock expectations are calculated using the Support Vector Machine (SVR). When using KPCA+SVR, the market forecast is more exact than when using SVR alone [12].

III. Conclusion

This Survey paper presumes that benefits in the stock exchange may be amplifying through various methodologies and strategies, yet every strategy has its own benefits and restrictions. We have considered an assortment of strategies

or procedures for financial exchange expectation and infer that it is feasible to deliver another and cross strategy to estimate the forthcoming development of the offer market or to conjecture the monetary state of an organization through any of the talked technique or through various access methods. But on the other hand, it's crucial to plan according to the structure by which the framework can be augmented with precision and the exhibition with less computational confusion. All things considered, we accept that stock expectation is that best in class work and various components ought to be viewed as a lot of exactness and effectiveness to anticipate the market.

Aurthor	Algorithm	Accuracy
SUI Xue-shen et al.,	ASNN, ASH, ND, NDER (SVM)	62.93%, 60.75%, 63.23%
Qinghua Wen et al.,	SVM Min and max for 3 stocks	96.91% to 95.93%
Qinghua Wen et al.,	SVM Min and max for 50 stocks	56.29% to 46.28%
Ling-Bing Tang et al.,	Manifold Wavelet Support Vector Machine (MWSVM)	Test: 1. 58.62% 2. 66.62% 3. 79.4%
Ling-Bing Tang et al.,	Manifold Wavelet Support Vector Machine (MWSVM) With Heuristics	Test: 4. 59.91% 5. 79.12% 6. 87.22%
Hyun Joon Jung and J. K. Aggarwal	Binary Stock Event Model (BSEM) Using SVM	70% to 80%
W. Lertyingyod and N. Benjamas	Wrapper Subset Evaluation with Greedy Algorithm (ANN)	93.89%
M. R. Vargas et al.,	I-RNN SI-RCNN I- RNN SI- RCNN	52.52% 56.84% 48.92% 51.08%
D Soni et al.,	Naïve Bayes PSO Black Hole Decision RNN	85.56% 83.57% 95.10% 81.16% 96.60%

Table 1 Comparison Table

IV REFERENCES

1. "Cao, Lijuan & Tay, Francis. (2001). Financial Forecasting Using Support Vector Machines. Neural Computing and Applications. 10. 184-192. 10.1007/s005210170010.".
2. "Tang, Jinqi and Xiong Chen. “Stock Market Prediction Based on Historic Prices and News Titles.” ICMLT '18 (2018).".
3. "Puri, Nischal. (2018). A Survey on Machine Learning Approach for Stock Market Prediction. HELIX. 8. 3705-3709. 10.29042/2018-3705-3709.".
4. "S. Xue-shen, Q. Zhong-ying, Y. Da-ren, H. Qing-hua and Z. Hui, "A Novel Feature Selection Approach Using Classification Complexity for SVM of Stock Market Trend Prediction," 2007 International Conference on Management Science and Engineering, Harbin, Chin".
5. "Q. Wen, Z. Yang, Y. Song and Peifa Jia, "Intelligent stock trading system based on SVM algorithm and oscillation box prediction," 2009 International Joint Conference on Neural Networks, Atlanta, GA, USA, 2009, pp. 3341-3347, doi: 10.1109/IJCNN.2009.517884".
6. "L. Tang, H. Sheng and L. Tang, "Stock Returns Prediction Using Manifold Wavelet Kernel," 2009 International Conference on Electronic Commerce and Business Intelligence, Beijing, China, 2009, pp. 306-309, doi: 10.1109/ECBI.2009.67.".
7. "A. S. M. Shihavuddin, Mir Nahidul Ambia, Mir Mohammad Nazmul Arefin, Mokarrom Hossain and Adnan Anwar, "Prediction of stock price analyzing the online financial news using Naive Bayes classifier and local economic trends," 2010 3rd International Conferenc".
8. "H. J. Jung and J. K. Aggarwal, "A Binary Stock Event Model for stock trends forecasting: Forecasting stock trends via a simple and accurate approach with machine learning," 2011 11th International Conference on Intelligent Systems Design and Applications,".
9. "W. Lertyingyod and N. Benjamas, "Stock price trend prediction using Artificial Neural Network techniques: Case study: Thailand stock exchange," 2016 International Computer Science and Engineering Conference (ICSEC), Chiang Mai, Thailand, 2016, pp. 1-6, do".
10. "M. R. Vargas, C. E. M. dos Anjos, G. L. G. Bichara and A. G. Evsukoff, "Deep Leaming for Stock Market Prediction Using Technical Indicators and Financial News Articles," 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazi".
11. "D. Soni, S. Agarwal, T. Agarwel, P. Arora and K. Gupta, "Optimised Prediction Model for Stock Market Trend Analysis," 2018 Eleventh International Conference on Contemporary Computing (IC3), Noida, India, 2018, pp. 1-3, doi: 10.1109/IC3.2018.8530457.".
12. "Nahil, Anass & Lyhyaoui, Abdelouahid. (2018). Short-term stock price forecasting using kernel principal component analysis and support vector machines: the case of Casablanca stock exchange. Procedia Computer Science. 127. 161-169. 10.1016/j.procs.2018.01".
13. "N. Budhani, C. K. Jha and S. K. Budhani, "Prediction of stock market using artificial neural network," 2014 International Conference of Soft Computing Techniques for Engineering and

Technology (ICSCTET), Bhimtal, India, 2014, pp.
1-8, doi: 10.1109/ICSCTET".

978-1-6654-2005-1/21/$31.00 ©2021 IEEE