International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Survey on Stock Market Prediction using Deep Learning

1) Prof V.M. Sardeshmukh, E&TC Dept, SAE Kondhawa, pune
2) Saurabh Gund, E&TC Dept, SAE Kondhawa, pune
3) Prathamesh Mhalsekar, E&TC Dept, SAE Kondhawa, pune
4) Santhipriyan Nair, E&TC Dept, SAE Kondhawa, pune

Abstract - The share market is a very unpredictable entity that is extremely sensitive to different sorts of criteria. Substantial research has been done in the field of financial markets projection to discover the many characteristics and elements connected with price volatility. The routine movements of the financial markets is exceedingly turbulent and dynamic. It is impacted by a large number of underlying elements and intricate interactions, resulting in a substantial percentage of non linearity. According to certain beliefs on the behavior of the financial markets, stock exchange progression is entirely unpredictable. Therefore, to achieve effective prediction of the stock markets through the use of machine learning approaches a number of related works have been analyzed. This analysis has allowed for the development of a useful and highly effective methodology that will be discussed in the upcoming editions of this research article.

Key Words: *K Nearest Neighbors, Long Short Term Memory, Stock Market Prediction.*

1.INTRODUCTION

The predictive analysis aims to estimate market return and offers the general population with information about the commodity on the exchange as well as its share prices. It moreover represents the country's entire economy. The share market generates a vast volume of information in a relatively brief period of time and fluctuates at all times, rendering it unpredictable. Stock prices are determined by firms purchasing and exchanging shares of publicly listed firms, and information on opening, closing, peak, as well as other stock statistics is disclosed on a daily basis. This information is accessible to the public and is safeguarded by a transparency rule. Share value forecasting is critical for everybody, but particularly for investors and company owners who make substantial purchases in these equities.

Because of the unpredictable characteristics of the share market, people are going to lose and win money at random just about all the time. Whenever an investor invests in stocks, he anticipates the cash he put to provide significant returns, and then when he wants to sell, he intends to make the same reward. Due to the obvious high risk connected with the high rewards, significant numbers of individuals are participating, and as a result, there's still an increased demand for shares which will only result in a lucrative conclusion. The only answer to this dilemma is to examine historical stock market statistics for the firm and comprehend the trend based on many factors linked with that specific organization. Because the data is huge and cannot be evaluated instantaneously, machine learning algorithms are employed to interpret and forecast such results.

For a long time, stock market statistics was thought to be unpredictable. Any abnormalities or patterns that emerge in a data sample averaged out throughout the entire dataset. Advancement in the development of neural network models has been remarkable, and these innovations have gained widespread implementation since they can be implemented with a high degree of competence and confidence. There are several possibilities to do this, however, because numerous elements may perform an influence in determining the valuation of a particular stock at a specific instance, and varying levels of vigilance to the multiple variables (e.g. estimation progression centered on market data but instead of journalism and shareholder details) also play a significant role in determining the performance of the models.

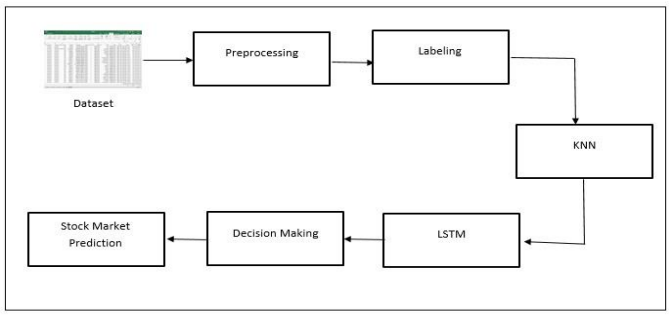
At just one side, forecasting the economics or share market prices is a vast study topic, yet forecasting with high accuracy remains tough. Machine learning, on the other extreme, as a useful technology, can cope with large amounts of data and challenging categorization issues to create high predictions. Because online news and social media do have large impacts on stock markets nowadays, classic economic systems are no longer acceptable for forecasting. Predicting the recommendations for future stock markets is a problem that has received a lot of attention in a variety of industries, notably trading, finance, statistics, and computer science. The goal for this is, of course, to forecast the direction of upcoming values so that stocks may be purchased and sold at lucrative levels. Fundamental and/or technical examination is commonly used by professional traders to examine equities and generate investing choices.

This literature survey paper segregates the section 2 for the evaluation of the past work in the configuration of a literature survey, and finally, section 3 provides the conclusion and the future work.

2.Methodology

- **Module A: Pre-processing**
  - Input : Dataset
  - Process : Attribute identification and Selection
  - Output : Pre-processed list
- **Module B: Labelling**
  - Input: Pre-processed list
  - Process: unique list and index allocation
  - Output: Labelled list
- **Module C: K-NN**
  - Input: Labelled List
  - Process: Euclidean Distance and Centroid Estimation
  - Output: Classified nearest neighbour list
- **Module D: LSTM**
  - Input: Classified nearest neighbour list
  - Process: Neuron formation and Hidden layer
  - Output: Stock market prediction

Block Diagram :



3.Algorithm Explanation:

Preprocessing:

1.Preprocessing is used to identify and handle missing values in the dataset .

2.The technique for handling missing values includes imputation where missing values are replaced with estimated values(e.g., mean, median, or mode of the column), or deletion, where rows or columns containing missing values are removed from the dataset.

3.It helps to clean and analysis and ensuring that the predictive model effectively lean from the data to make accurate predictions.

Labelling:

1. Labelling in context of stock market prediction involves creating the target variable that the predict to model will learn to predict .

2. The target represent what the model aim to forecast weather it's future stock price or market tends to binary multiplex classification of price movement

3. What's the target variable is chosen a time horizon is established to determine how far into the future prediction is made.

4. labels are assigned to historical data point based on the time horizon.

Finally the labelled data is verified and validate to ensure occurrency and representativeness

KNN :

1.K-Nearest Neighbour (KNN) is a simple and effective algorithm used for classification and regression task

2.In the context of stock Market Prediction KNN is applied to forecast future stock price or market trends based on historical data.

3.The algorithm works by finding the K nearest data points to new data points in future space.

Long Short Term Memory(LSTM):

1. Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) architecture that is specifically designed to model sequential data and overcome the vanishing gradient problem associated with traditional RNNs

2. The Long Short-Term Memory (LSTM) deep learning model is employed in this block for stock market prediction.

3. The LSTM model can capture sequential patterns in time series data, making it suitable for stock price forecasting

4. They can capture long term dependency in data

4.RELATED WORKS

Yi Wei [1] emphasizes that one of the primary essential aspect of the stock time - series data is the price movement's directional cues. The extensive experiments reveal that the existing projection error used to validate performance of the model can only represent the model's anticipated price and the current market price, but not the directional cues of market movements. The non-negativity of actual number is the primary issue of this phenomenon. As a result, judging the success of the stock market prediction employing regression neural network relying just on the minimum and maximum values of forecasting deviation is inaccurate.

Aditya Menon [2] states that after analyzing the approaches accessible for our issue, the authors may conclude that while deep learning techniques alone exhibit potential, they can indeed be rendered increasingly precise by combining them with a text-based processing technique. The researchers might also advocate using financial news as a resource instead of a community since economic news seems to be more true and accurate. The key restrictions involve additional study on the issue of the LSTM model for predictions and economic news as a resource in association with stock market data and the consultants' usage of extremely specialized prediction algorithms. The authors also suggested that they could try to accomplish the same thing in Python on a relatively small scale.

Mingze Shi [3] narrates that as per the outcomes of simulation investing, this concept of investing tactic is highly effective for locating the true golden cross and assisting traders in making huge profits. However for a handful equities, the influence is negative; nonetheless, for the majority of shares in the selections, this approach will provide a return after investing in a few years. Turning points with timeframes of several days are appropriate for various stocks. Stocks with substantial relatively brief changes should not employ a turning point of less than a week and should not be held for an extended period of time. Stocks with moderate curves, on the other hand, should employ a greater turning point or retain for a prolonged period of time.

Zixuan Liu [4] expresses that in the recent years there has been a considerable increase in the volatility of the stock markets across the world and especially in China. This is one of the most problematic occurrences that leads to the instability in the predictions of the stock market and its prices. Therefore, the traditional approaches have been ineffective in the realizations of the precise and accurate predictions in such scenarios as it can be highly difficult to predict the future prices with such volatility. To provide a solution for this, the authors in this approach have proposed the use of Support Vector Machines for the purpose of prediction of the stock prices that leads to a much better and precise implementation of the prediction.

Sayavong Lounnapha [5] discusses that the Convolutional neural network stock market estimation mechanism in this publication has elevated reliability and computational application value. The authors recommended using CNN as the core technique in comparable estimation and forecasting to achieve more precise outcomes. Although the authors' central emphasis in this paper is on equities represented on the Thai Stock Exchange, the created methodology can be applied to any other trading platform where a significant number of daily price history is accessible. The researchers will intend to utilize this approach to other equities in the coming years, as well as integrate other technologies to create a more appropriate forecasting methodology.

Md. Mujibur Rahman Majumder [6] describes that according to the evaluation of the responses, the share price of the Dhaka Stock Exchange in Bangladesh tends to fall under prediction with time series data. All techniques have a high average prediction performance. In compared to arima, which has comparable accuracy ratings, the Feed- Forward network has a substantially higher maximal average precision. Furthermore, when evaluating a single investment, the forecast precision of the shares is demonstrated to be in the region of 3%. After all, depending on the research in this report, time series prediction is quite excellent for anticipating Bangladeshi stocks, but significant gains cannot be steady improvement to the unpredictability and complex behavior.

Du Peng [7] explains that, in light of the present volatility scenario in China’s share market, shareholder attitude and behavior markers are built, and a share market volatility index relying on a vast data approach is assembled to indicate the fluctuations in the stock market. The scheme's novelty combines big data with measurable index collection and model creation. It recognizes the potential of vast online big data implementation, breaks through the obstruction of the fundamental source of data of investor confidence characterization, and gives a new technique to quantify fluctuation risk.

Sahil Vazirani [8] states that within this work, researchers employ two ways to estimate the value, one using a singular algorithm while the other employing a hybrid model. Researchers computed the difference between both the projected and actual prices utilizing MAE, MSE, and RMSE. The closer the number was to 0, the more precise the algorithm, and so linear regression

produced the much more consistent readings amongst random forest decision tree, support vector machine, and KNN, with support vector machine coming in second. The chart generated by linear regression and SVM successfully predicted the actual cost. It may be inferred that a hybrid approach that combines linear regression between another linear regressions produces more economical, precise, and superior outcomes than random forest, decision tree, SVM, and KNN.

Hemil N. Shah [9] the stock market's development has been a strong point, but projecting this has been a huge issue since stock prices fluctuate, rendering it an uncertain undertaking. Many individuals participate in the financial market in order to increase their assets. There are several approaches for predicting the performance of a stock market. The ANFIS model is used to forecast the Istanbul stock market. ANFIS has three steps wherein wavelet decomposition and recurrent technique are coupled like in the artificial bee methodology.

Sneh Kalra [10] states that the applied model investigates the impact on the stock market of assessing various sorts of stock related story with numerical historical information. An endeavor is needed to develop a viable framework for forecasting future market movements. The reliability of future price movements is increased by taking into account quantitative sentiment score and numerical history data. KNN achieves the best predictive performance. The suggested technique's results are consistent with earlier studies that show a substantial link among stock-related announcements and changes in stock price. In contrast to other applicable technologies, KNN fared the best on the given dataset.

Samuel Olusegun Ojo [11] suggests utilising NASDAQ Composite data to anticipate stock market behaviour leveraging a layered LSTM network model. The model has been trained, and the results demonstrate that it can forecast share price behaviour with absolute precision. An ongoing problem is that share market volatility can sometimes be reduced just by analysing historical data; current affairs in the worlds of economics and politics, for example, may alter the investor behavior and, as a result, the behavior of stock prices.

Rajni Jindal [12] argues that due to the increasing incidence of COVID-19 infections worldwide, the market has never been more unpredictable. As a consequence, numerous standard pattern forecasting techniques have performed poorly since they do not compensate for the influence of the epidemic on share market patterns. The researchers have added several characteristics in the suggested strategy to accommodate for the epidemic, so that the algorithms can learn the pattern and provide more appropriate findings. The researchers validated the approach by testing it with stock market data while using various standard stock market forecasting techniques to make forecasts. The machine learning algorithms utilized consists of Support Vector Regressor, Random Forest Regressor, and Decision Tree Regressor.

Warren Landis [13] explains that, having picked LSTM as the preferable backup prediction technique, several adjustments must be performed to adequately prepare it for particular use case while also addressing its troublesome overfitting characteristic. Overfitting with LSTM is caused mostly by its manner of functioning inside a closed loop, monitoring and learning through patterns. Numerous external influences intervene on the stock market, causing its values to fluctuate. As a result, this seems to be an open system which is not particularly suitable for LSTM application. To begin with, even outside the use case, introducing more LSTM layers might significantly increase its performance.

5.Conclusions and Future Scope:

This survey article is aimed at analysis of the previous researches on the topic of stock market prediction to achieve our methodology for the estimation. The stock markets are essential for the economy of a particular nation as it allows for free markets to flourish and achieve their goals for funding. This is crucial for the development and overall growth of a country that allows for the various companies to carry out their operations for stability and improvement of the financial conditions. There are stock brokers and investors that are looking for opportunities to understand the behavior of the market volatility and the herd mentality of the populous investing in the market. This is quite complex to evaluate as it seems to be very random and dependent on a variety of external factors. Therefore, the analysis of previous and existing systems for the prediction of the stock market has been effective in realization of our approach which will be elaborated further in the upcoming editions of this research.

6.REFERENCES :

[1] Y. Wei and V. Chaudhary, "The Directionality Function Defect of Performance Evaluation Method in Regression Neural Network for Stock Price Prediction," 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA), 2020, pp. 769-770, doi: 10.1109/DSAA49011.2020.00108.

[2] A. Menon, S. Singh and H. Parekh, "A Review of Stock Market Prediction Using Neural Networks," 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), 2019, pp. 1-6, doi: 10.1109/ICSCAN.2019.8878682.

[3] M. Shi and Q. Zhao, "Stock Market Trend Prediction and Investment Strategy by Deep Neural Networks," 2020 11th International Conference on Awareness Science and Technology (iCAST), 2020, pp. 1-6, doi: 10.1109/iCAST51195.2020.9319488.

[4] Z. Liu, Z. Dang and J. Yu, "Stock Price Prediction Model Based on RBF-SVM Algorithm," 2020 International Conference on Computer Engineering and Intelligent Control (ICCEIC),2020,pp.124-127,doi: 10.1109/ICCEIC51584.2020.00032.

[5] L. Sayavong, Z. Wu and S. Chalita, "Research on Stock Price Prediction Method Based on Convolutional Neural Network," 2019 International Conference on Virtual Reality and Intelligent Systems (ICVRIS), 2019, pp. 173-176, doi: 10.1109/ICVRIS.2019.00050.

[6] M. M. R. Majumder, M. I. Hossain and M. K. Hasan, "Indices prediction of Bangladeshi stock by using time series forecasting and performance analysis,"2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2019, pp. 1-5, doi: 10.1109/ECACE.2019.8679480.

[7] D. Peng, "Analysis of Investor Sentiment and Stock Market Volatility Trend Based on Big Data Strategy," 2019 International Conference on Robots & Intelligent System (ICRIS), 2019, pp. 269-272, doi: 10.1109/ICRIS.2019.00077.

[8] S. Vazirani, A. Sharma and P. Sharma, "Analysis of various machine learning algorithm and hybrid model for stock market prediction using python," 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), 2020, pp. 203-207, doi: 10.1109/ICSTCEE49637.2020.9276859.

[9] H. N. Shah, "Prediction of Stock Market Using Artificial Intelligence," 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), 2019, pp. 1-6, doi: 10.1109/I2CT45611.2019.9033776.

[10] S. Kalra and J. S. Prasad, "Efficacy of News Sentiment for Stock Market Prediction," 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing(COMITCon),2019,pp.491-496, doi: 10.1109/COMITCon.2019.8862265.

[11] S. O. Ojo, P. A. Owolawi, M. Mphahlele and J. A. Adisa, "Stock Market Behaviour Prediction using Stacked LSTM Networks," 2019 International Multidisciplinary Information Technology and Engineering Conference (IMITEC), 2019, pp. 1-5, doi: 10.1109/IMITEC45504.2019.9015840.

[12] R. Jindal, N. Bansal, N. Chawla and S. Singhal, "Improving Traditional Stock Market Prediction Algorithms using Covid-19 Analysis," 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), 2021, pp. 374-379, doi: 10.1109/ESCI50559.2021.9396887.

[13] W. Landis, S. Cha and M. Shaalan, "On Optimization of Stock Market Prediction Methods," 2019 IEEE International Conference on Big Data (Big Data), 2019, pp. 6116-6118, doi: 10.1109/BigData47090.2019.9005612.