iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52756

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Stock Market Predictions through Deep Belief Neural Network

Vishal Kiran Bankar[1], Vaishnav Nandkishr Patil[2], Omkar Satishkumar Shirure[3], Samruddha Satish Dambalkar[4], Prof. Dheeraj Patil[5]
[1, 2, 3, 4, 5]*IT Dept, PCET's NMIET, Pune, India*

Abstract: ***Owing to the huge potential returns, the proportionate degree of exposure, and the adaptability of the transaction, a big number of people choose to invest their money in stocks. This is due to the fact that stocks come with both of these factors. If investors are able to successfully foresee changes in stock values, they have the opportunity to generate considerable returns from their financial investments. The value of a stock may be affected by a wide variety of variables, including the current state of the market and the macroeconomic environment as a whole, significant events in the economic and social spheres, investor attitude, and management actions made by the firm. Predicting the value of a company's shares has always been considered as the most challenging and essential component of academic research. This view continues to hold true today. Utilizing mathematical and statistical models is a common component of the traditional approaches used for predicting stock values. These methods have been around for quite some time. Despite this, these strategies are not enough for dealing with the intricate dynamics of the stock market, which are always evolving. The fast development of computer technology over the last several decades has led to an increase in the prevalence of the use of machine learning among academics. Therefore, this paper outlines an effective strategy for stock market prediction that utilizes K Nearest Neighbors, Linear Regression, along with Deep Belief Network and Decision making.***

Keywords: Stock market prediction, K Nearest Neighbors, Linear Regression, Deep Belief Network and Decision making.

I. INTRODUCTION

Experts in the area of economics have consistently shown a high level of skill in their ability to properly forecast the behavior of stock markets throughout history. On the other hand, as a result of the advancements that have been achieved in learning approaches, data analysts are starting to address the issues that are involved with predictive modeling. In addition, computer scientists have included techniques from the area of machine learning to improve the effectiveness of approaches to estimating and to refine the quality of forecasts. This has enabled the computer scientists to make more accurate predictions. The use of deep learning served as the next stage in the process of building a forecasting model that was more accurate than its predecessors. Data scientists have difficulty when constructing forecasting models owing to the inherent complexity involved with predicting the stock market, which is characterized by a high degree of uncertainty and unpredictability. These inherent complications may be attributed to the intrinsic volatility of the stock market. Complexity and non-linearity are two main problems that are significantly related to many important elements, including the unpredictable behavior of the share market as well as the link between investment psychology and market reaction. These key aspects are related considerably to two major worries.

Forecasting the course that stock prices will take is, and will continue to be, one of the most difficult challenges faced by experts working in the domains of analytics and economics. This forecast is based mostly on two investment strategies: one that involves purchasing stocks with the expectation that their prices would rise, and the other involves investing in them with the expectation that their prices will fall. Both of these strategies are expected to be profitable. When it comes to making projections about the state of the stock market, there are often two opposing points of view. An additional kind of evaluation is known as fundamental analysis, and it is predicated on the processes and basic data of a firm. These factors might include the company's portion of the market, its costs, and its yearly growth rates. The second approach, which is known as "technical indicators," focuses more of an emphasis on past stock prices and values as opposed to those that are occurring at the current time. The current study conducts an analysis of previous charts and patterns in order to provide forecasts on future price.

The stock market is a complicated time series phenomena that displays features of dynamic equilibrium. As soon as market operations begin, a wide variety of responsive trading activities are anticipated to take place with regard to equities, which will result in swings in the market value of those stocks. In addition, the value of shares is susceptible to a wide range of unpredictability factors, which leads to the characteristic non-stationary quality of time-series data for share prices.

Within the field of predictive inquiry, estimating the value of stocks is considered to be a very difficult and complex undertaking. In the most recent decades, academics have devoted a significant amount of time and energy to researching the topic of predicting stock values from a variety of angles. The classification of model components and the improvement of estimating procedures are two key directions that have emerged from this body of research as a result of its accumulation.

The progression of the stock markets is susceptible to the effect of a number of different factors, many of which cannot be predicted with any degree of accuracy at any one moment. These factors include the portrayal of companies in the public eye by the media as well as the political climate in which countries now find themselves. Because of this, one can deduce that the trajectory of the price of shares as well as the index may be forecast by conducting an effective preprocessing of data about asset values and making use of algorithms that are appropriate for the task. Deep learning and machine learning are two approaches to data analysis that have the potential to be of assistance to customers and dealers alike in the decision-making processes that are associated with stock market predictions and classification. The aforementioned methodologies were developed with the express purpose of automatically discovering and cultivating commonalities among vast amounts of data. A significant amount of autonomy is shown by algorithms, and they have the ability to take on the role of forecasting market swings in order to enhance trading tactics.

When one examines the research that serves as the basis for these algorithms, it becomes abundantly evident that each of them is equipped with the capability to provide accurate answers to issues pertaining to stock forecasting challenges. It is of the utmost importance to keep in mind that each one of them is constrained by a unique set of constraints. The outcomes of the prediction are not only influenced by the format in which the essential data is given; rather, they are also dependent on the approach that was used in order to produce the forecast. Additionally, it is feasible to considerably improve the dependability of estimating approaches by employing just the most important characteristics and identifying those features as the input information, as opposed to combining all of the components. This may be done by designating the most important features as the input information.

Kun Huang [1] has proposed a model that utilizes a multilayer graph-connected neural network to predict fluctuations in stock prices. The authors have developed a neural concentration network-based approach that incorporates data pertaining to the financial sector, media, and commercial connections through a unique feature extraction component. This is done to make up for the fact that they can't use their prior experience with conventional methods of stock market forecasting. The ML-GAT methodology utilizes a range of algorithms at different stages to dynamically filter diverse forms of input, with the aim of constructing a merged graph that can facilitate the comprehension of network categorization models, thereby aiding in prediction tasks. The objective of this endeavor is to enhance comprehension of the categorization of networks. The utilization of graph-based learning techniques for the purpose of improving prediction performance is currently a topic of ongoing scholarly investigation. The authors conducted a comparative analysis between ML-GAT and established production models that have been developed using publicly available datasets as standard benchmarks, in order to assess the effectiveness of ML-GAT.

Suman Saha's [2] graph-based algorithm, optimum loss function, recommended unique performance assessment metric, and network encapsulation approach have facilitated the prediction of top stocks. The encapsulation technique employed by a network is contingent upon the individual components therein. This analysis yields three significant conclusions. Commencing the discussion, it is possible to achieve noteworthy efficacy by reducing the ranking loss by a factor of three, while maintaining the identical embedding approach. The graph available on Wikidata indicates that the aggregate point-wise plus pair-wise loss outperforms both the aggregate total loss and the list-wise losses. The inclusion of the Industry chart is highly beneficial for both the NASDAQ and the NYSE. The findings of this research indicate that an error function which accounts for losses in a list-wise manner is more advantageous compared to one that accounts for losses in a point-by-point or pair-by-pair manner.

According to Mojtaba Nabipour [3], the authors of the study were driven by the objective of ascertaining the capacity of machine learning and deep learning-based algorithms to accurately forecast the conduct of the stock market. The dataset was created using 10 physical criteria and data spanning 10 decades. The implementation of the aforementioned was carried out in the basic metals, petrochemicals, non-metallic commodities, and differential profitability sharemarkets on the Tehran Stock Exchange. Furthermore, the study employed nine distinct machine learning algorithms and two distinct deep learning methodologies as classifiers. (nemaly RNN and LSTM). The authors have presented two alternatives for the input parameters of algorithms, namely, a data analysis technique that produces versus binary files. The evaluators utilized three distinct classifications to assess the items. The experimental findings indicate that the substitution of continuous variables with bit streams in the models leads to a significant enhancement in their efficiency.

The Literature Survey chapter of this research paper examines previous work. Section 3 delves into the approach in depth, while section 4 focuses on the outcomes evaluation. Finally, Section 5 brings this report to a close and gives some hints for future research.

II. LITERATURE SURVEY

As per the explanation provided by Mohammad Alsulmi [4], numerous techniques have been devised for the manual annotation of stock data. Stock datasets are frequently utilized in research environments to categorize them, with the aim of instructing algorithms to anticipate future market trends. The objective of this study was to provide a clear and explicit description of the metadata tagging problem as an NP-hard problem. Alternative automated labeling techniques encompassed the utilization of metaheuristic search algorithms such as hill-climbing and simulated annealing. In comparison to conventional, labor-intensive techniques, this approach exhibits considerable potential and achievement. The present study does not address the correlation between enhanced labeling productivity and improved prediction accuracy while constructing a machine learning forecasting model from the given data.

Salah Bouktif [5] has contributed to the ongoing discourse regarding the efficacy of sentiment analysis in predicting forthcoming trends in the stock market. The decision to utilize Twitter as a primary data source for forecasting the growth and decline of six prominent NASDAQ companies was motivated by the abundance of valuable information disseminated through social media platforms. The authors of the study contend that a holistic approach to textual integrated sentiment assessment could lead to enhanced stock market prediction. The present study proposes a methodology to enhance the expression of emotions by dissecting various textual elements to offer a comprehensive portrayal. Subsequently, it employs diverse feature-selection methodologies to ascertain the most appropriate attributes for the given problem. Ultimately, the approach involves the stacking of models to generate direction classifiers with the highest possible level of precision, based on the available stock data.

A novel three-step process for stock market forecasting was introduced by Saud S. Alotaibi [6]. In recent times, a novel methodology has been introduced that involves the process of feature extraction, optimal feature extraction, and forecasting. We were able to recover a number of features, such as the statistical and SOTI features. The classifiers utilized in the classification phase were trained using acquired characteristics. Selecting the most crucial attributes holds greater significance in achieving optimal anticipated results. RDAGW selected the most relevant characteristics from those that were retrieved, which were deemed significant in the present scenario. An ensemble-based clustering method was utilized to make predictions. This technique incorporated various models such as optimal neural networks, Random Forests 1 and 2, and support vector machines. The study employed the Optimal neural network to obtain the anticipated results, and the precision of the network was enhanced by adjusting its weights through the use of RDAGW.

The authors of a recent study on stock market prediction, as described by Zeynep Hilal Kilimci [7], have adopted a distinct approach by utilizing Turkish statistical models sourced from a network of social networks and websites that offer scientific analysis and news. The study aims to evaluate high-priced stocks by analyzing financial perspectives. The authors of the recent study have approached the challenge of predicting stock market fluctuations in a unique manner. The user's crawler provides them with access to a diverse range of textual resources, such as Twitter, the open records platform KAP, the personal opinions and authentic news site Mynet Finans, and the stock price movements tool Bigpara. Upon completion of the data collection, cleaning, and organization processes, it is possible to employ Embeddings techniques and Deep Learning algorithms to predict the future movements of BIST 100. The proposed model could potentially benefit from implementing various preprocessing techniques such as splitting, elimination of stop words, removal of hashtags, and elimination of URLs.

Yaohu Lin [8] developed a forecasting system that incorporates multiple machine learning techniques to effectively consider the diverse range of daily k-line patterns. The study's findings indicate that the framework provided possesses significant predictive capabilities, and that an investment portfolio constructed using the prediction models is more efficient. The results of this inquiry reveal significant aspects of the matter at hand. The present study enhances prior research on stock market prediction by incorporating traditional candlestick analysis alongside contemporary advancements in artificial intelligence. This scholarly article amalgamates two divergent schools of thought by utilizing diverse machine learning-based methodologies to evaluate the prognostic potential of individual one-day candlestick patterns for the subsequent price movement. The authors of the study have observed a potential predictive impact of specific candlestick patterns on the financial instruments market.

Xianghui Yuan [9] employed diverse feature extraction methods and classifiers for stock value movement analysis to evaluate the profitability of various stock selection strategies. Various techniques for feature selection are utilized to filter the unprocessed attributes. The utilization of a temporal feature extraction mechanism for cross-validation is better suited for practical investment operations, as it enables the identification of the attributes of share value tendency forecasting models. Cross-validation is employed to isolate the features of models that are capable of predicting movements in stock prices. Random forests have demonstrated exceptional performance in feature extraction and stock price forecasting, as evidenced by empirical evidence from real-world scenarios. The study demonstrated that the optimal approach results in the highest profits when selecting stocks within the uppermost percentile of the market, irrespective of the quantity of stocks employed in constructing the model.

Xuan Ji [10] has proposed a new methodology to enhance the accuracy of future stock value predictions. The methodology employed in this study involves the utilization of Doc2Vec algorithm to generate text feature maps by amalgamating data from the domains of social networking and finance. The utilization of SAE is employed to reduce the text representations in order to mitigate a significant disparity between the linguistic attributes and the economic characteristics. The Haar wavelet method is utilized by the authors to eliminate superfluous data and produce a more precise time series of share prices. The reduction of the influence of extraneous fluctuations in stock price data on the predictive model is observed. The study concludes with the researchers training a Long Short-Term Memory (LSTM) model to produce forecasts for stock prices by utilizing a blend of textual input and financial data. The results of the experiment demonstrate that the proposed method exhibits significantly lower MAE, RMSE, and R2 values in comparison to the methods that were previously employed. The results of this study offer backing for the notion that an amalgamation of textual elements from multiple sources could be employed to generate more accurate predictions of stock prices.

According to Nagaraj Naik [11], the heightened instability of the market renders it challenging to ascertain the occurrence of a potential stock market disaster. The fluctuations in the stock market can be attributed to a plethora of factors, such as alterations in corporate financials, political volatility, fluctuations in bond market rates, and global market trends. The price of a stock may experience volatility in the event of a change in management or the distribution of a substantial dividend or incentive. In the realm of securities trading, stock prices exhibit significant responsiveness to information emanating from diverse sources. Integrating heterogeneous pieces of information to form a coherent entity is a challenging issue to address. The study's authors observed that there was no discernible difference between the predictive capabilities of XGBoost and deep neural network models in forecasting the future performance of YES Bank shares. Based on the hypothesis testing, the aforementioned conclusion was drawn as the deep neural network models and XGBoost classification approach produced outcomes that were statistically significant.

Xingqi Wang [12] has argued that accurately predicting future stock values solely based on past data is a desirable yet challenging objective. The researchers initiated the study by employing KMSD grouping as a means of categorizing the stocks into clusters that share comparable attributes. Furthermore, the dataset of stocks possessing similar characteristics is utilized for the purpose of training Hierarchical Temporal Memory. The results of the experiment indicate that the utilization of Hierarchical Temporal Memory in conjunction with KMSD segmentation is adequate for the purpose of short-term prediction of stock prices. Furthermore, it has been observed that the precision of this approach can be considerably improved. Moreover, the findings substantiate the inference that such prognostication is legitimate. The ultimate objective is to develop algorithms that can examine a broader spectrum of variables in order to identify stocks with similar attributes. This endeavor will aid in enhancing the model's prognostic capabilities and expanding its versatility.

III. PROPOSED WORK

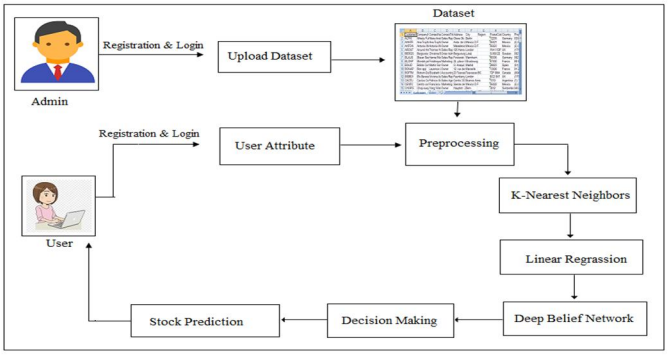


Figure 1: The Proposed System model overview

The proposed work for the stock market prediction through Deep learning model is explained in the detail with the below mentioned steps.

1) *Step 1: Data Gathering* – The proposed methodology for stock market prediction commences with an interactive web-based platform that is specifically developed to enable the administrator to submit a dataset containing stock-related data. The web application was developed by writing Java code employing a JSP page and hosting it on a Glassfish server located on the local host.

The admin utilizes this web page to sign up into system using the sign up option. Once this option is selected, the webpage with various attributes of the admin are displayed in a form. This form is filled by providing the relevant information such as username, mobile number, email address, password etc. After these details have been validated and stored into the database, the admin can utilize the sign in option to gain access into the system the system provides the ability to the admin to provide the input dataset to the system that will be discussed in the next step of the approach.

2) *Step 2: Dataset Reading and Preprocessing* – The file picker is activated and is currently in a state of the capacity for the incorporation of the dataset into the system. This process involves the utilization of the input dataset obtained from the Yahoo Finance service. The Yahoo Finance platform facilitates the efficient retrieval of stock information. The dataset comprises up-to-date information on chosen stocks utilized to make predictions. The dataset is obtained in the format of a comma-separated values file, which is subsequently furnished to the system for additional analysis.

The Java program utilizes the JXL libraries to access and interpret the dataset in workbook format. The information contained within the dataset is obtained and subsequently transformed into a list format, which is then subjected to preprocessing procedures. For the specific purpose at hand, a subset of attributes from the dataset has been selected for extraction, while the remaining attributes have been removed. The aforementioned characteristics comprise stock market, factor, objective, duration, investmonitor, and initial and concluding values.

Once the dataset has been furnished to the system and suitably preprocessed and transformed into a functional format for the code, the user can employ the interactive web interface to input user attributes. The initial step for the user entails registering through the designated sign-up page on the website. The user utilizes the web form to input the user details such as username, DOB, mobile number, email address and password which are validated and stored in the database. The credentials can be employed by the user to access the system via the webpage. Upon successful authentication of the login credentials, the system grants the user access to input their user attributes for the purpose of prediction.

The aforementioned attributes and their corresponding values are chosen and forwarded to the subsequent stage with the aim of categorizing them against the user-provided input values from the preceding step via the interactive web page.

3) *Step 3: K-nearest Neighbor Clustering* – The present stage of the methodology involves the reception of user input and preprocessed dataset from the preceding stages. The aforementioned data is being utilized for the purpose of approximating the spatial separation between the user's input and every individual row within the preprocessed dataset. The distance is computed using the equation 1 as presented.

$$ED = \sqrt{\sum(ATi - ATj)^2} \quad \text{_____ (1)}$$

Where,
ED=Euclidian Distance
A_{Ti}=Attribute at index i
A_{Tj}= Attribute at index j

After computing and adding the distance from each row to its corresponding row, the complete list is sorted in ascending order through the utilization of the bubble sort methodology. The value of k in this instance is set to 2, leading to the formation of two clusters. The initial cluster containing the data of high quality is subsequently subjected to regression analysis utilizing the Linear Regression technique. The entire process of KNN is summarized in the algorithm 1 given below.

ALGORITHM 1: KNN Classified Cluster Formation

//Input : Sorted Distance List $SD_{L,}$
//Output:Cluster List K_{CL}
1: Start
2: $I_L = \emptyset$ [Inner Layer] $O_L = \emptyset$ [Outer Layer], $K_{CL}=\emptyset$
3: MIN= 0 , MAX=SDL_{SIZE-1}
4: K= (MAX-MIN) /2

5: K=MIN+K
6: ***for*** i=0 to Size of SD_L
7: $R = SD_{L[i]}$
8: ***if***(i<=K), ***then***
9: $I_{L=} I_{L+R}$
10: ***else***
11: $O_{L =} O_{L+R}$
12: *end for*
13: $K_{CL[0]} = I_L$
14: $K_{CL[1]} = O_L$
15: return K_{CL}
16: *Stop*

4) *Step 4: Linear Regression* – Linear regression is utilized on the attributes that were previously chosen. The aforementioned attributes are utilized in the regression assessment of the dataset parameters in comparison to the input provided by the user. Linear regression is calculated for each user input with respect to the other variables in the dataset.

The array x[] contains the list of user input that was formerly obtained, while the array y[] stores the characteristics of the dataset. Equation 2 demonstrates the calculation of the variables m and b, which are determined through the application of linear regression to the aforementioned arrays. The determination of the value of the intercept can be achieved through the resolution of equations 3 and 4.

$Y=Mx + B$ _____ (2)

$M = \frac{N \sum(xy) - \sum x \sum y}{N \sum(x^2) - (\sum x)^2}$ _____ (3)

$B = \frac{\sum y - M \sum x}{N}$ _____ (4)

Where:
x = dataset attribute value
y = user input value
M = Slope or Gradient
B = the Y Intercept
N= Size of the array
Y=Intercept value

The list format is utilized to record all of the estimated intercepts for each row in linear regression. The outcomes are subsequently arranged in a descending sequence based on the regression coefficients, with the most significant coefficients being advanced to the subsequent stage of the methodology for integration into Deep Belief Networks.

5) *Step 5: Deep Belief Network* – The present computational step pertains to the introduction of the Deep Belief network. The aforementioned network employs the input list of the DBN subsystem that has been established recently. The DBN approach employs instance learning to assess the output layer along with hidden layer variables for the purpose of predicting the stock market. The process of assigning values to the output alongside hidden layers requires the assignment of arbitrary weights, along with a pair of bias parameters to offset the arbitrary weights.

The initial measures in the attributes are designated as the target values, and the output as well as hidden layer factors are obtained using the sigmoid activation function, as demonstrated in formulas 5 and 6.

$T=(\sum_{k=0}^{n} AT * W) + B$____ (5)

$H_{LV}=2 \left(\frac{1}{(1+\exp(-T))} X\ 2T \right) -1$ _____ (6)

Where,
n- Number of attributes

A_T- Attribute Values
W- Random Weight
B- Bias Weight
H_{LV} – Hidden Layer Value
The formation of the neural network probability list involves the aggregation of both the target measures and the output layer values that have been achieved. The probability list generated by the neural network is arranged in a descending order and subsequently utilized in the segregation process through the application of decision-making techniques. In the case of the Decision making input list, the leading half of the sorted list is selected and passed on to the subsequent step.

6) *Step 6: Decision Making* – This step takes the input of the sorted list containing the probability scores achieved from the DBN execution. These probability scores cannot be utilized as is as the user output as they do not represent the proper outcome. Therefore, these probability scores need to be classified effectively to extract the relevant stock market prediction. Specific if-then rules of the decision making approach are deployed for these probability scores to classify them and emerge the proper prediction. The classified predictions are provided to the user on the web page GUI.

IV. RESULTS AND DISCUSSIONS

The present study outlines a proposed methodology for achieving Stock Market Prediction by means of the utilization of machine learning techniques. The methodology has been developed utilizing the Java programming language. The development machine is comprised of a conventional setup, featuring an Intel Core i5 central processing unit, 8 gigabytes of random access memory, and 1 terabyte of hard disc drive storage. The technique has been devised to utilize a dataset of stock trends as input and apply algorithms based on machine learning to generate projections.

A. Performance Evaluation based on Precision and Recall

The evaluation of a component's execution accuracy within a framework can be effectively measured through the utilization of precision and recall metrics. The relative accuracy of a component is contingent upon its precision and encompasses a broad range of dependability.
The precision metric was determined in this approach by calculating the proportion of accurate Stock Market Predictions obtained relative to the overall number of completed trials. The recall requirements serve as a supplement to the precision metric and assist in ascertaining the general reliability of the Deep Belief Network module.
This method calculates the recall by dividing the number of accurate Stock Market Predictions by the total number of incorrect stock market predictions. The following formulas provide quantitative expansion of the aforementioned concept.
Precision and Recall can be depicted as below:
1) X = The number of accurate Stock Market Predictions
2) Y= The number of inaccurate Stock Market Predictions
3) Z = The number of accurate Stock Market Predictions not done

So, precision can be defined as
Precision = (X / (X+ Y)) *100
Recall = (X / (X+ Z)) *100
Table 1 unveils the experimental results obtained through the aforementioned formula in an accurate manner. The statistical parameters are employed to offer a visual representation, as exemplified in Figure 2.

Table 1: Precision and Recall Measurement Table

No. of Trials	Accurate Stock Market Predictions (A)	Inaccurate Stock Market Predictions (B)	Accurate Stock Market Predictions not done (C)	Precision	Recall
1	1	0	0	100	100
4	3	0	1	100	75
7	6	1	0	85.71429	100
10	8	2	0	80	100
13	10	1	2	90.90909	83.33333
16	13	2	1	86.66667	92.85714
19	17	1	1	94.44444	94.44444
22	17	3	2	85	89.47368
25	20	3	2	86.95652	90.90909
28	24	2	2	92.30769	92.30769

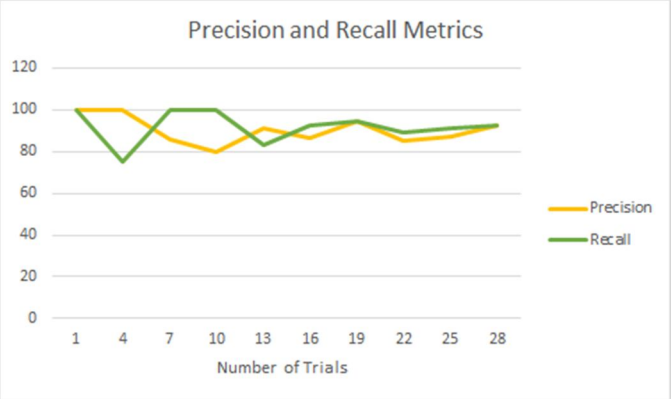


Figure 2: Comparison of Precision and Recall

The depicted graph illustrates the performance of the Deep Belief Network and its corresponding recommendations based on input data across multiple trial counts. The extraordinary accuracy of the approach is demonstrated by precision and recall levels of 90.19 percent along with 91.83 percent, correspondingly. The numerical values hold considerable importance in the preliminary execution of the aforementioned procedure and have yielded an acceptable outcome.

V. CONCLUSION AND FUTURE SCOPE

This research article outlines the suggested technique for achieving precise and exact stock market predictions. The predictions were executed utilizing machine learning methodologies. The methodology commences with the administrator accessing the system by utilizing valid credentials and submitting the dataset pertaining to the stock market trends of Yahoo Finance as the input. The dataset has been preprocessed to eliminate unnecessary components and insufficient data points. Subsequently, the preprocessed dataset is made available for subsequent assessment. Subsequently, the user proceeds to access the system by utilizing the appropriate login credentials that were provided during the registration process. Upon successful authentication, the user submits their user attributes to facilitate prediction. These attributes are then combined with the dataset supplied by the administrator and subjected to the K-Nearest Neighbors algorithm. Prior to any analysis, it is necessary to cluster the data. Subsequently, the K nearest Neighbors algorithm will cluster both the preprocessed dataset along with user attribute data. The clusters obtained in the preceding stage are employed as an input for the Deep Belief Network, which attains the probability scores that require effective classification to achieve accurate predictions. The Decision Making methodology involves categorizing the potential results, which are subsequently presented to the user as a forecast for the stock market. The methodology has undergone an assessment for forecasting inaccuracy through the utilization of precision and recall metrics, and the results fall within acceptable parameters.

To advance the research's future direction, it is recommended to augment the approach to operate in a real-time setting on a cloud-based platform, thereby further enhancing the user experience.

REFERENCES

[1] K. Huang, X. Li, F. Liu, X. Yang and W. Yu, "ML-GAT:A Multilevel Graph Attention Model for Stock Prediction," in IEEE Access, vol. 10, pp. 86408-86422, 2022, doi: 10.1109/ACCESS.2022.3199008.

[2] S. Saha, J. Gao and R. Gerlach, "Stock Ranking Prediction Using List-Wise Approach and Node Embedding Technique," in IEEE Access, vol. 9, pp. 88981-88996, 2021, doi: 10.1109/ACCESS.2021.3090834.

[3] M. Nabipour, P. Nayyeri, H. Jabani, S. S. and A. Mosavi, "Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data; a Comparative Analysis," in IEEE Access, vol. 8, pp. 150199-150212, 2020, doi: 10.1109/ACCESS.2020.3015966.

[4] M. Alsulmi, "Reducing Manual Effort to Label Stock Market Data by Applying a Metaheuristic Search: A Case Study From the Saudi Stock Market," in IEEE Access, vol. 9, pp. 110493-110504, 2021, doi: 10.1109/ACCESS.2021.3101952.

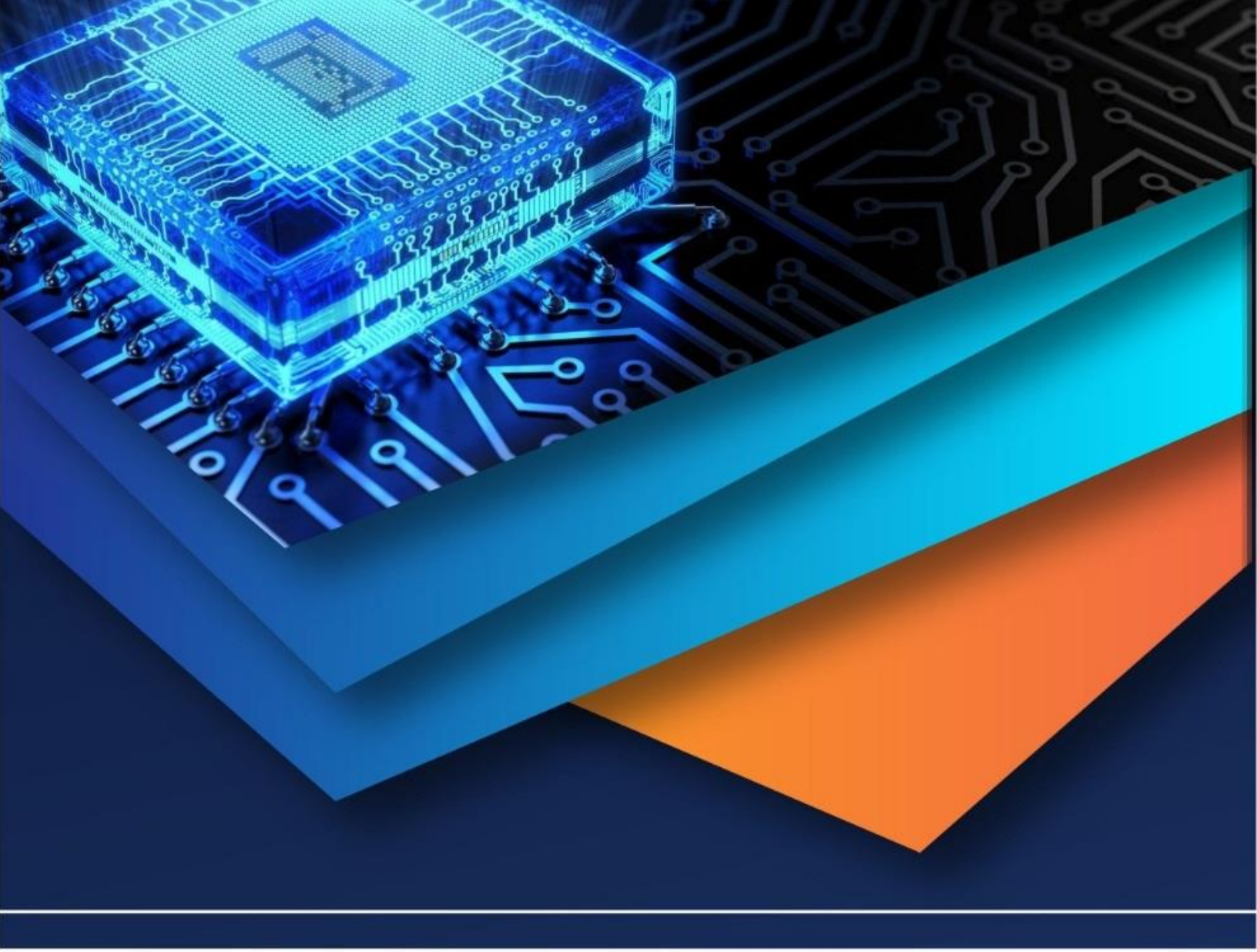
[5] S. Bouktif, A. Fiaz and M. Awad, "Augmented Textual Features-Based Stock Market Prediction," in IEEE Access, vol. 8, pp. 40269-40282, 2020, doi: 10.1109/ACCESS.2020.2976725.

[6] S. S. Alotaibi, "Ensemble Technique With Optimal Feature Selection for Saudi Stock Market Prediction: A Novel Hybrid Red Deer-Grey Algorithm," in IEEE Access, vol. 9, pp. 64929-64944, 2021, doi: 10.1109/ACCESS.2021.3073507.

[7] Z. H. Kilimci and R. Duvar, "An Efficient Word Embedding and Deep Learning Based Model to Forecast the Direction of Stock Exchange Market Using Twitter and Financial News Sites: A Case of Istanbul Stock Exchange (BIST 100)," in IEEE Access, vol. 8, pp. 188186-188198, 2020, doi: 10.1109/ACCESS.2020.3029860.

[8] Y. Lin, S. Liu, H. Yang and H. Wu, "Stock Trend Prediction Using Candlestick Charting and Ensemble Machine Learning Techniques With a Novelty Feature Engineering Scheme," in IEEE Access, vol. 9, pp. 101433-101446, 2021, doi: 10.1109/ACCESS.2021.3096825.

[9] X. Yuan, J. Yuan, T. Jiang and Q. U. Ain, "Integrated Long-Term Stock Selection Models Based on Feature Selection and Machine Learning Algorithms for China Stock Market," in IEEE Access, vol. 8, pp. 22672-22685, 2020, doi: 10.1109/ACCESS.2020.2969293.
[10] X. Ji, J. Wang and Z. Yan, "A stock price prediction method based on deep learning technology," in International Journal of Crowd Science, vol. 5, no. 1, pp. 55-72, April 2021, doi: 10.1108/IJCS-05-2020-0012.
[11] N. Naik and B. R. Mohan, "Novel Stock Crisis Prediction Technique—A Study on Indian Stock Market," in IEEE Access, vol. 9, pp. 86230-86242, 2021, doi: 10.1109/ACCESS.2021.3088999
[12] X. Wang, K. Yang and T. Liu, "Stock Price Prediction Based on Morphological Similarity Clustering and Hierarchical Temporal Memory," in IEEE Access, vol. 9, pp. 67241-67248, 2021, doi: 10.1109/ACCESS.2021.3077004.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)