STOCK FEE PREDICTION USING MULTI -LAYER PERCEPTRON

MRS.L. SHERIN BEEVI [1], MS. C.A. DAPHINE DESONA CLEMENCY[2]

[1] Assistant Professor, Department of Computer Science and Engineering, DMI College of Engineering. Tamil Nadu, India.
[2] Department of Computer Science and Engineering, DMI College of Engineering. Tamil Nadu, India.

ABSTRACT- Stock costs are thought-about to be Chaotic and Unpredictable. The Economical Market Hypothesis deems stock prices to follow the stochastic process Model. This Terrible drawback is challenged by securities market Prediction Models. Predicting the long run stock prices of economic commodities or prediction the coming stock market trends will change the investors to gainer take advantage of their mercantilism by taking calculated risks supported reliable trading strategies. There are Numerous Number of approaches based on that a prediction model is developed and implemented. This Work concentrate to use suitable forecasting techniques to predict future stock returns supported past returns and numerical news indicators to construct a google stocks so as to diversify the risk. By applying supervised learning strategies for stock worth prediction by deciphering the apparently chaotic market data.

Keywords - Artificial Neural Network (ANN); Multi-Layer Perceptron (MLP); Stock Market; Technical Analysis; Time Series Analysis.

1. INTRODUCTION

Investors can track the changes in the value index over time. They can also utilize it as a benchmark against which to make comparison of the returns from the portfolio. Stock market indices evaluate the value and merits of groups of different stocks. As the stock market being chaotic, nonlinear, and dynamic in nature it is very difficult to understand because of its volatility, hence it is of great importance for the investors to understand behavior which might help for his or her effective investment in it. Artificial Neural Network (ANN) has the capability to identify the chaotic and Nonlinear inter relation within the input data set without a priori presumption of knowledge of correlation between the input data and the output. Hence Artificial Neural Networks suits higher than different models in statement the stock market returns [10]. In existing the piecewise linear representation which mixes turning points and maximum absolute deviation points available statistic to extracts sequence features. First and foremost, the proposed technique discovers defining moments that fulfill the condition given during this paper, and characterizes the reason distance equation to ascertain the fitting mistakes between the aftereffect segment and the fitted straight line, whose average value is set as threshold P and the subsequence length is as the threshold d [11].

The Task of Prediction of Securities Market is incredibly difficult and of nice interest for researchers because the actual fact that stock market could be an extremely volatile in its behavior. All kinds of major events in the society may affect the trend of economic development, which in turn causes stock prices to fluctuate in the trading market [1]-[5]. Therefore, the stock is the most active part in the securities market. Stock trading activity is affected by the data generated for daily trading, and stock investors also buy and sell stocks according to the economic rules reflected by these data. How to effectively utilize these data to obtain valuable information and provide scientific guidance for investors has become a hot research topic [6]-[9]. Prediction of stock market is substantial in finance and it's drawing a lot of attention, because of the fact that the investors could also be higher radio-controlled if the direction and trend of the stock market is foreseen successfully. A stock market index is that the total value obtained by the mix of various stocks or many different investment instruments altogether. Stock's indices reflect an overall stock market and trace the market's changes over time.

2. RELATED WORK

The Time series refers to the sequence of corresponding changes with time. For the characteristics of the statistic itself, like high dimension and enormous scale, Fallouts et al. [12] They believe that when the sequence dimension decreases, the mining effect is more obvious in recent years, researches [13],[14]. In existing the piecewise linear representation which mixes turning points and maximum absolute deviation points available statistic to extract sequence features [11]. A voluminous literature has examined the flexibility of AI-based techniques to predict the longer-term stock movements. Kimoto used standard Neural Networks to identify the relationships among completely different market factors and developed a prediction system for the Tokio stock price level (TOPIX). Moreover, they compared their projected model to Multiple Regressions. It absolutely was terminated that NNs end in higher coefficients compared to Regressions. Their analysis on exchange prediction is one in all the works within which the matter of overfitting has been tested by victimization cross-

validations. Chang and Liu [4] tried to integrate the result of qualitative factors with technical indices trough a Genetic Algorithm-based Fuzzy Neural Network (GFNN). They evaluated their prediction system on information obtained from Taiwan securities market (TSE) Associate terminated that the projected system outperforms the ANNs that solely contemplate the quantitative factors. Boyacioglu and Avci [3] examined the foregone conclusion of the urban center securities market index by victimization six political economy variables and 3 indices as input of their reconciling Network-based Fuzzy reasoning System (ANFIS). The authors showed that their proposed model is capable of predicting the monthly returns of ISE national one hundred indices with an accuracy of 98.3%.

2.1 LITERATURE VIEW

In [15], the stock market prediction has become a lot of vital issue within the gift time. One amongst the ways utilized is technical analysis, however such methods don't continuously yield correct results. Thus, it's important to develop methods for a more accurate prediction. Generally, investments are created using predictions that are obtained from the stock worth once considering all the factors which may have an effect on it. Since monetary stock marks generate enormous amounts of data at any given time a great volume of data needs to undergo analysis before a prediction can be made. Each of the techniques listed under regression has its own advantages and limitations over its other counterparts. One of the noteworthy techniques that were mentioned was linear regression. The way rectilinear regression models work is that they're often fitted using the smallest amount squares approach, but they'll alternatively be even be fitted in other ways, such as by diminishing the "lack of fit" in some other norm, or by diminishing a handicapped version of the least sq. loss function.

In [16], the use of machine learning and AI techniques to predict the costs of the stock is an increasing trend. Additional and more researchers invest their time on a daily basis in springing up with ways to attain techniques which will any improve the accuracy of the stock prediction model. Thanks to the immense range of choices available, there'll be n number of how on the thanks to predict the worth of the stock, however all ways don't work an equivalent means. The output varies for every technique though the same information set is being applied. Within the cited paper the stock price prediction has been dole out by mistreatment the random forest algorithmic program is getting used to predict the worth of the stock using monetary magnitude relations type the previous quarter. This can be only 1 way of viewing the matter by approaching it employing a prophetic model, using the random forest to predict the longer term price of the stock from historical data. By using the financial ratio along with a model which will effectively analyze sentiments the accuracy of the stock worth prediction model will be increased stock In [17], accurately predicting the securities market may be a difficult task, however the trendy internet has verified to be a awfully useful gizmo in creating this task easier. Thanks to the interconnected format of data, it's straightforward to extract bound sentiments therefore making it easier to determine relationships between numerous variable and roughly scope out a pattern of investment. Investment pattern from various companies show sign of similarity, and also the key to with success predicting the securities market is to use these same consistencies between the {information} sets. The means stock market information will be expected successfully is by mistreatment over simply technical historical data, and using other ways just like the use of sentiment instrument to derive a very important affiliation between people's emotions and the way they're influenced by investment in specific stocks. An additional important phase of the prediction method was the extraction of important events from internet news to ascertain how it affected stock prices.

3. PROPOSED METHOD

A multilayer perceptron (MLP) might be a class of feedforward counterfeit neural organization (ANN). The term MLP is utilized vaguely, once in a while freely to any feedforward ANN, at times carefully to ask networks made out of different layers of insights (with limit actuation) Terminology. Multilayer perceptron's are here and there conversationally alluded to as "vanilla" neural organizations, particularly when they have a solitary concealed layer. Regardless of whether a neuron isn't reacting or a snippet of data is feeling the loss of, the organization can recognize the deficiency and still produce the yield. This examination is driven in two phases, preparing stage, and testing stage.

Steps from these two phases are given underneath:

1) Training Stage

Stage 1: Read the preparation information. Stage2. Adjust Date.
Stage 3: Apply windowing administrator to change the time arrangement information into a nonexclusive dataset. This progression will change over the last column of a windowing inside the time arrangement into a mark or target variable. Last factor is treated as mark.
Stage 4: Perform a sliding windowing approval cycle of the delivered mark from windowing administrator to take care of them as contributions to ANN model.
Stage 5: Select preparing cycles and extraordinary boundaries of ANN (learning rate, energy, mistake epsilon).
Stage 6: Run the model and notice the exhibition (precision).
Stage 7: If the precision is acceptable than go to stage 8

or go to step 4.(As the primary thought process of the examination was to improve precision of Stock Price expectation, the best boundary blend ought to be set. Thus, if the outcome found in this progression isn't acceptable enough; the entire cycle ought to be done from stage 4 once more) Stage 8: Exit from the preparation arrange and apply prepared model to the testing dataset. Shrouded Layer 1

2) Testing Stage

Stage1: Read the testing dataset. Stage2: Apply the preparation model to test information Stage3: Produce the anticipated cost and market patterns.

4. METHODOLOGY

Let us enumerate some available forecasting methods in predicting the stock prices.

4.1 DATA DESCRIPTION AND PRE-PROCESSING: The Data Pre-processing is a technique that's accustomed to convert the information into clean data. Information is gathered from completely different sources is collected in raw format that isn't possible for the analysis. The Information Preprocessing is an Integral step in machine learning because the quality of knowledge and also the helpful information which will be derived from it directly affects the flexibility of our Model to learn; therefore, it's extremely necessary to preprocess the data. The Entries are gift within the dataset. The null values are removed victimization df = df.dropna() wherever df is that the data frame. The specific attributes (Date, High, Low, Close, Adjacent value) are born-again into numeric victimization Label Encoder. The Data are taken from real time Google Dataset.

Date	Open	High	Low	Close	Volume	Ex-Dividend	Split Ratio	Adj. Open	Adj. High	Adj. Low	Adj. Close	Adj. Volume
2004-08-19	100.01	104.06	95.96	100.335	44659000.0	0.0	1.0	50.159839	52.191109	48.128568	50.322842	44659000.0
2004-08-20	101.01	109.08	100.50	108.310	22834300.0	0.0	1.0	50.661387	54.708881	50.405597	54.322689	22834300.0
2004-08-23	110.76	113.48	109.05	109.400	18256100.0	0.0	1.0	55.551482	56.915693	54.693835	54.869377	18256100.0
2004-08-24	111.24	111.60	103.57	104.870	15247300.0	0.0	1.0	55.792225	55.972783	51.945350	52.597363	15247300.0
2004-08-25	104.76	108.00	103.88	106.000	9188600.0	0.0	1.0	52.542193	54.167209	52.100830	53.164113	9188600.0
2004-08-26	104.95	107.95	104.66	107.910	7094800.0	0.0	1.0	52.637487	54.142132	52.492038	54.122070	7094800.0
2004-08-27	108.10	108.62	105.69	106.150	6211700.0	0.0	1.0	54.217364	54.478169	53.008633	53.239345	6211700.0
2004-08-30	105.28	105.49	102.01	102.010	5196700.0	0.0	1.0	52.802986	52.908323	51.162935	51.162935	5196700.0
2004-08-31	102.32	103.71	102.16	102.370	4917800.0	0.0	1.0	51.318416	52.015567	51.238167	51.343492	4917800.0
2004-09-01	102.70	102.97	99.67	100.250	9138200.0	0.0	1.0	51.509003	51.644421	49.988312	50.280210	9138200.0
2004-09-02	99.09	102.37	98.94	101.510	15118600.0	0.0	1.0	49.698414	51.343492	49.623182	50.912161	15118600.0
2004-09-03	100.95	101.74	99.32	100.010	5152400.0	0.0	1.0	50.631294	51.027517	49.813770	50.159839	5152400.0
2004-09-07	101.01	102.00	99.61	101.580	5847500.0	0.0	1.0	50.661387	51.157920	49.959219	50.947269	5847500.0
2004-09-08	100.74	103.03	100.50	102.300	4985600.0	0.0	1.0	50.525969	51.674514	50.405597	51.308384	4985600.0

Fig 1 Input Dataset

4.2 FEATURE GERENATION

Features choice is that the method of reducing the number of input variable once developing models. Feature choice is one altogether the core ideas in machine learning that massively impacts the performance of your model. The knowledge options that you simply just use to teach your machine learning models have an outsized influence on the performance you'll achieve. Unsuitable or partly relevant features can negatively impact model performance. Feature selection and Data cleaning should be the primary and most vital step of your model designing, which may be wont to build the model. The attributes used for feature selection are Date, Price, Adj close, Forecast X coordinate, Y coordinate, Latitude, Longitude, Hour and month.

4.3 TRAINING MODEL:

After feature choice location and month attribute are used for training.

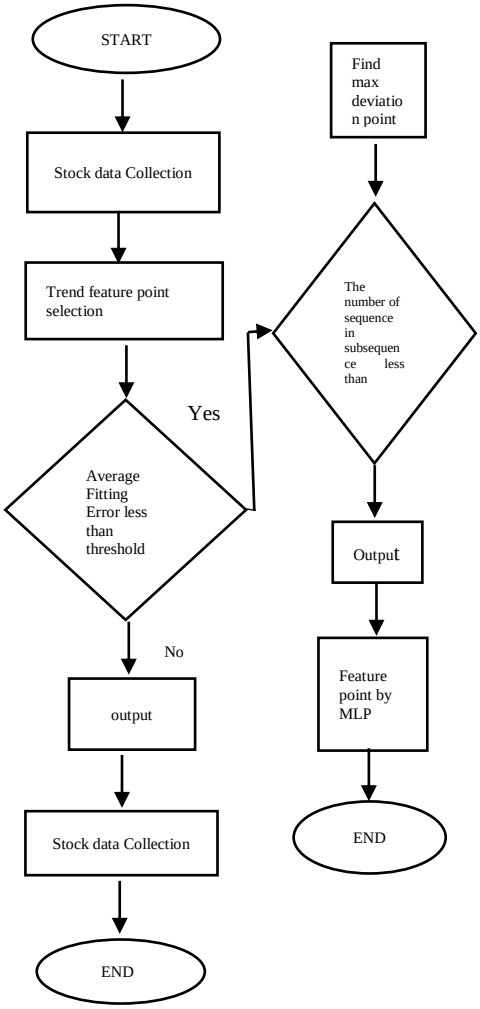

Fig 2. Data Flow Diagram

The dataset is split into combine of xtrain, ytrain and xtest, y test. The Algorithmic programs model is foreign kind skleran. Building model is finished exploitation model. work (xtrain, ytrain). This part would involve supervised classification strategies like rectilinear regression, Ensemble classifiers (like Adaboost) Training a model merely suggests that learning (determining) smart values for all the weights and also the bias from tagged examples. In supervised learning, a machine learning algorithm builds a model by examining several examples and making an effort to hunt out a model that minimizes loss; this process is named empirical risk minimization. Loss is that the penalty for a nasty prediction. That is, loss may be a number indicating how bad the model's prediction was on one example. If the model's prediction is ideal, the loss is zero; otherwise, the loss is bigger.

4.3.1 MULTI LEVEL PRECEPTRON

Multilayer feedforward network with back propagation algorithm. The prognostic capability of neural networks comes from the ranked or multi-layered structure of the networks. Weighted input signals associate degree turn-out an signaling exploitation an activation function. MLP may be a feedforward neural network withone or more layers between Input and output layer. Feedforward implies that informationflows in one direction from input to output layer. MLP has three layers; an input layer, an additional hidden layers and output layer. The input data are fed to then neurons within the input layer and when processing within the individual neurons of the input layer the output values are forwarded to neurons in the hidden layer and at last to the neurons within the output layer.

Similarly, the ReLU function performs a operation on each input element where all values but zero are set to zero. Thus, the ReLU is represented as:

$$F(y) = \max(o,y) = \begin{cases} y_i \text{ if } y_i \geq 0 \\ 0 \text{ if } y_i < \end{cases} \quad \begin{matrix} \text{-1.14} \\ 0 \end{matrix}$$

The weights are Incremented in the outer layer.

$$W_{mn}(y+1) = W_{mn}(y) + \Delta w_{mn}(y) \quad \text{----(3)}$$

The weights in hidden layer will depend on the weights of output layer to calculate the error gradient δ and is calculated by,

$$\Delta w_{mn}(p) = \propto_i . X_i(y).\partial_j(y) \quad \text{-----(4)}$$

MLPs are hugely used in pattern classification, recognition, prediction and approximation. Connections among the neurons are associated by weights and dynamical the loads in a particular manner results to learning of the associated network. The procedure by that the weight changes take place in the network is termed learning or coaching algorithm. The backpropagation algorithm is that the most ordinarily used learning technique.

ANN contains a few points of interest nonetheless a standout amongst the foremost perceived of those is that the method that it will very gain from look info sets. Along these lines, ANN is used as associate impulsive capability estimate instrument [18]. These forms of devices gauge the foremost effective and ideal systems for touching base at arrangements whereas characterizing calculation capacities or disseminations. ANN takes info tests rather than whole information sets to the touch base at arrangements, that spares each time and money. ANNs are thought of genuinely simple numerical models to boost existing information investigation innovations. ANNs have 3 layers that are interconnected. the first layer contains of information neurons. Those neurons send information on to the second layer, which so sends the yield neurons to the third layer [19]. The study utilized three-layer (a hidden layer) perception model (a feed forward neural network) ready with back propagation calculation. Authentic stock prices of distinctive organizations were

taken from distributed stock info on the Web. the educational capability or the initiation work that was utilized is sigmoid equation. Neural Network gets numerous inputs (either from the distinctive information or from the yield of various somatic cells within the neural network). each information comes through associate association that contains a quality (or weight); these weights relate to colligation adequacy in an organic neuron. each neuron in addition has solitary limit esteem. The weighted mixture of the inputs shaped, and therefore the limit subtracted, to create the initiation out of the somatic cell (otherwise known as the post-synaptic potential or PSP, of the neuron). The activation signal is then {passed through skilled older more matured more experienced more responsible more established seasoned knowledgeable versed capable competent skillful well-versed tried associated true gone through had undergone saw felt responded to suffered} an activation operates.

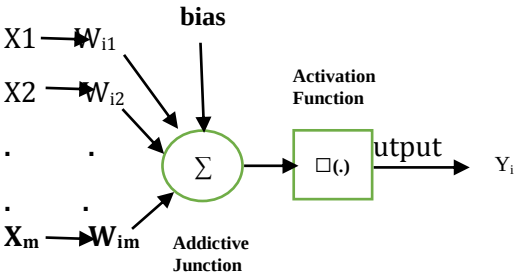

Fig 3 Artificial Neural Network

4.4 VISUALIZATION

Using the mathematical Poltlib library from sklearn. The Stock data set is analyzed by drawing various graphs. Data visualization is that the discipline to do to grasp knowledge by inserting it during a visual context so patterns, trends and correlations that may not well be detected can be exposed can be exposed.

5. EXPERIMENT AND RESULT WORK.

Assume that xi is that the data series of stock price, wherever i = 1,2,3,..,N and N is the number of knowledge of elect companies under city Stock Exchange.

In the current study Multilayer feedforward network (MLP) with dynamic backpropagation learning has been used. Each the network contains an input layer, one hidden layer and one output layer. The Number of neurons in the input layer is 3 that are the three consecutive past costs of the particular stock beneath study. The Number of hidden layers and also the number of neurons selected in the study has been done heuristically. Within the study it had been started with a fewer number of neurons, however thanks to the inaccuracy in the initial predictions the number of neurons were magnified in the hidden layer [20]. To live the performance of the neural network model used, Mean Absolute Percentage Error (MAPE), Mean Absolute Deviation (MAD) and Root Mean Squared Error (RMSE) were calculated.

Suppose 1 2 3 (a,a,a ,...,) an are actual values and 1 2 3 (y ,y ,y ,...,y)n are the predicted values then the MAPE, MAD and RMSE can be calculated by using the formula(4),(5)and (6)

$$\text{MAPE} = 100*1/n\sum\left|\frac{x_i-y_i}{x_i}\right| \quad \text{---(4)}$$

$$\text{MAD} = \frac{\sum|x_i-y_i|}{n} \quad \text{--- (5)}$$

$$\text{RMSE} = \sqrt{mean(xi-yi)^2} \quad \text{--(6)}$$

In this examination NEUROPH has been utilized to conjecture stock costs of chosen organizations and areas recorded.

Step1: Divide the entire data into two teams 60% for training set and 40% for testingset.

Step 2: 75% data has been used to train the network. The normalization for the input is done using the formula where x denotes the value that should be normalized; N x denotes the normalized value of x; min x represents the minimum value of x; max x represents the maximum value of x. After the standardization the info (stock prices) are going to be within the vary of [0,1].

Step 3: A most network error of 0.01, the momentum rate and learning rate range between 0.1 and 0.9 are utilized in the training

Step4: The trained network obtained in step 3 has been tested with randomly selected data from the 25% testing sets

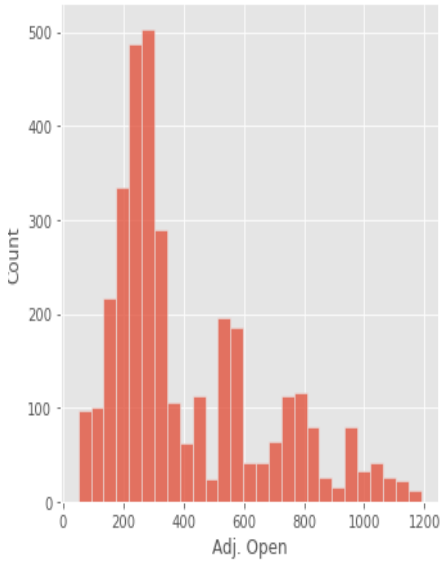

CHART.1 ADJACENT OPEN

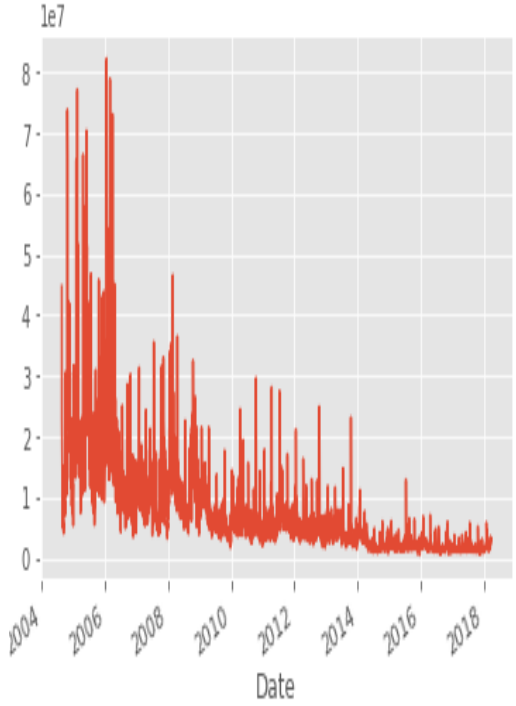

CHART 2 ADJACENT VOLUME

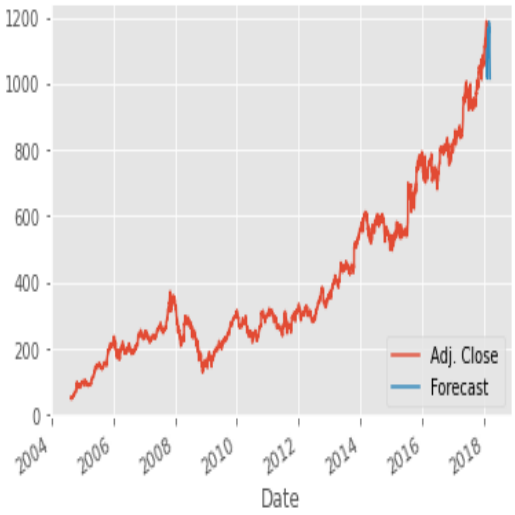

CHART 3.STOCK PRICEPREDICTION

6. CONCLUSION

Financial exchange information are exceptionally time-variation and are regularly in a nonlinear example, anticipating the future cost of a stock is profoundly testing. Stock market knowledge are highly time-variant and are usually in a very nonlinear pattern, predicting the long - term worth of a stock is extremely challenging. Prediction provides knowledgeable information concerning this standing of the stock price movement. It detected that Artificial Neural Network technique is very helpful in predicting stock indices further as stock price of specific company. Many various algorithms are used with neural network. Feedforward MLP neural network technique is taken into account to predict the stock price of corporations listed beneath. From the result are often all over that MLP neural network technique provides the satisfactory output.

In Future Analysis Work, in order to higher predict the movements of the stock market, will seek for some ways in which change intelligent choice of thresholds supported data, and additionally begin from the activity finance in social media to predict the stock trend by deciphering the emotions within the text data and mixing stock historical data. The Textual Information regarding social media and on-line news are proved to be effective in predicting the long run trend of the stock market. By increasing textual feature information in the media, we are able to build the results a lot of reliable and greater edges market regulators and traders.

REFERENCES

[1]. R. J. Cebula and D. Capener, ‘‘The Impact of federal income tax rate cuts on the municipal bond market in the U.S.: A brief exploratory empirical note,’’ Quant. Finance Econ., vol. 2, no. 2, pp. 407–412, 2019.

[2]. Y. Liu, Y. Zheng, and B. M. Drakeford, ‘‘Reconstruction and dynamic dependence analysis of global economic policy uncertainty,’’ Quant. Finance Econ., vol. 3, no. 3, pp. 550–561, Sep. 2019.

[3]. J. Zhang, Y.-H. Shao, L.-W. Huang, J.-Y. Teng, Y.-T. Zhao, Z.-K. Yang, and X.-Y. Li, ‘‘Can the exchange rate be used to predict the shanghai composite index?’’ IEEE Access, vol. 8, pp. 2188–2199, 2020.

[4]. S. Zhou and J. Zhang, ‘‘Empirical test of size effect in China stock market,’’ in Proc. IEEE Int. Conf. Bus. Intell. Financial Eng., Beijing, China, Jul. 2009, pp. 691–694.

[5]. Y. Xia, ‘‘The effects of economic and political events on the behaviour of stock market index in China,’’ in Proc. 2nd Int. Conf. Artif. Intell., Manage. Sci. Electron. Commerce (AIMSEC), Aug. 2011, pp. 2524–2527.

[6]. J. Engel, M. Wahl, and R. Zagst, ‘‘Forecasting turbulence in the asian and European stock market using regime-switching models,’’ Quant. Finance Econ., vol. 2, no. 2, pp. 388–406, 2018.

[7]. X. Pang, Y. Zhou, P. Wang, W. Lin, and V. Chang, ‘‘An innovative neural network approach for stock market prediction,’’J.Supercompu., vol. 76, no. 3, pp. 2098–2118, Mar. 2020.

[8]. Y. Zhang, P. Shang, and H. Xiong, ‘‘Multivariate generalized information entropy of financial time

series,'' Phys. A, Stat. Mech. Appl., vol. 525, pp. 1212–1223, Jul. 2019.

[9]. J. R. Thompson and J. R. Wilson, ''Multifractal detrended fluctuation analysis: Practical applications to financial time series,'' Math. Comput. Simul., vol. 126, pp. 63–88, Aug. 2016

[10] Ravi, D. Pradeepkumar, and K. Deb, „Financial time series prediction using hybrids of chaos theory, multi-layer perceptron and multi-objective evolutionary algorithms", Swarm and Evolutionary Computation, no. January, pp. 1–14, 2017.

[11] M. Zhou, J. Yi, J. Yang and Y. Sima, "Characteristic Representation of Stock Time Series Based on Trend Feature Points," in IEEE Access, vol. 8, pp. 97016-97031, 2020.

[12] C. Faloutsos, M. Ranganathan, and Y.Manolopoulos, ''Fast subsequence matching in time-series databases,'' ACM SIGMOD Rec., vol. 23, no. 2, pp. 419–429, Jun. 2020.

[13] L. Fang, H. Zhao, P. Wang, M. Yu, J. Yan, W. Cheng, and P. Chen, ''Feature selection method based on mutual information and class separability for dimension reduction in multidimensional time series for clinical data,'' Biomed. Signal Process. Control, vol. 21, pp. 82–89, Aug. 2015.

[14] X. Zhong and D. Enke, ''Forecasting daily stock market return using dimensionality reduction,'' Expert Syst. Appl., vol. 67, pp. 126–139, Jan. 2017.

[15] Survey of Stock Market Prediction Using Machine Learning Approach Authors: Ashish Sharma ; Dinesh Bhuriya ; Upendra Singh, 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA)

[16] Impact of Financial Ratios and Technical Analysis on Stock Price Prediction Using Random Forests:Authors: K. S. Loke, 2017 International Conference on Computer and Drone Applications (IConDA)

[17] Stock Market Prediction via Multi-SourceMultiple Instance Learning Authors:XiZhang ; ;JieyunHuang ; Binxig Fang ; Philip Yu, 2018 IEEE Access.Chan, C. C., E. W. C. Lo, and Shen Weixiang. "The available capacity

[18] Computation model based on artificial neural network for leadacid batter-ies in electric vehicles." Journal of Power Sources 87.1 (2000): 201-204.

[19] Kim, Kyoung-jae, and Ingoo Han. "Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index." Expert systems with Applications 19.2 (2000): 125-132.

[20] S. Chaigusin, C. Chirathamjaree and J.M. Clayden, "The Use of Neural Networks in the Prediction of Stock Exchange of Thailand (SET) Index", Proceedings of International Conference on Computational Intelligence for Modelling, Control and Automation, Vienna, Austria, pp. 670-673., 2019.