International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 04 | Apr 2019 www.irjet.net p-ISSN: 2395-0072

Novel based Stock Value Prediction method

Sarabjeet singh, Rekhit Gupta, Arihant Kumar Jain, Satyam Agarwal, S. Pandiaraj

---***---

Abstract - *In security exchanges, a considerable amount of unpredictability occurs due to the price movement or the stock market. Hence if we can predict this financial price changes it can bring a huge amount of financial advantages on us. This can only be done through a company review, and this process is also called a fundamental analysis. We can produce a predictive algorithm using machine learning, which is also a method for prediction and at this moment under research. To do that, before we start training our machines, then we can use our algorithm for the short-term trading decisions made before. The newest innovation in machine learning is Deep Neural Networks. A short-term prediction model was developed. We have tried in this paper to predict the short-term inventory prices. For this review, we use 10 unique stocks recorded. We focus on short-term investment prices in this review to forecast short-term stock prices. We use technical analyzes to inform our frameworks and to predict future stock projections on historical price patterns. In this report we discussed two different types of neural artificial networks, the feed networks and the recurring neural systems. The review found that multilayer Feed Forwards Perceptron is superior to long - term stock prices.*

I. INTRODUCTION

The prediction of "living data ' victimizing networks ' prediction capacity for several extremely volatile and client stocks[3],[4] and[5] is employed by time - delayed, repeated and probabilistic neural network. Trade systems utilize the chaotic modeling methodology to create various prediction models that are supported by technical, adaptive and applied mathematical models [6]. Therefore, the AN Intelligent Commerce Call Network [7] has been developed, which can be used to predict the shopping of short-term and long-term trends of victimization rules based NNs and merchandising signals. Neural grids and genetic algorithms were used to analyze exchange costs[8] as a smart telephone (STN) network. In BSS or cracking exchanges, DSS often predict values four weeks into the longer term, and suggests purchasing and selling selections, supported by the common value of the projected high and lower value. The same technique was used to recognize the "bull flag" pattern and to identify commercial rules on the Composite Index (NYSE-CI) big apple security market value and volume[9]. In the meantime, the [10] rules were used to integrate basic and technical analysis for the prediction of monetary performance.

.

II. LITERATURE SURVEY

The area of analysis in high space is increased by financial engineering. Each researchers and money dealers have been significantly interested in marketing systems supporting machine intelligence or cash-plus management, notably in equity mercantilism and risk management for derivatives such as choice and swaps. In the stock market prediction [1],[2] neural network was used.

To forecast the monetary statistic Artificial neural networks are extensively used. In [11], the potential of your time delay, continual and probabilistic neural networks for stock market prediction trends are supported and previous knowledge sets has been explicit. The information set consists of damage of the daily shares. The [12] technical indicators are taken as inputs to the neural network for stock market prediction. In [13], we create a choice neural network, were technical indicators and neural networks were utilized [14] to predict the U.S. greenback Vs British pound exchange rates. The framework are given by automatic commercialism In [15], a rear propagation neuro network will not estimate the buy - sell value for an stock by using a Dynamic window enhancement prediction accuracy. In [16] a survey over hundreds of articles that used neural networks and neuro-fuzzy models for stock market prediction are given. It absolutely determined that soft computing techniques embarks typical models in most cases. In [17], review of information mining applications available markets was given. [18] We used a two-layer bias call tree with technical indicators feature to make recommendations which makes a choice rule once to shop for a stock and once to not pass.

III. METHODOLOGY

A. **Reconstruction of Phase Space**

For statistical forecasting, the statistics are usually swollen to use a lot of information in higher dimensions. In the case of space reconstruction, an adequate m (lag) integration and time delay must be selected [19]. For the operational sensitivity analysis, embedded parameters are determined. The Input Time Delay vectors are reconstructed in all X cases with the Input matrix and Y as a matrix of output. Research results are based on the data we tend to verify long - term value from input and operation of previous values.

(1)

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \xrightarrow{m} X = \begin{bmatrix} x_1 & x_2 & \cdots & x_{m-1} & x_m \\ x_2 & x_3 & \cdots & x_m & x_{m+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{N-m-\tau+1} & x_{N-m-\tau+2} & \cdots & x_{N-\tau-1} & x_{N-\tau} \end{bmatrix}, \quad Y = \begin{bmatrix} x_{m+\tau} \\ x_{m+1+\tau} \\ \vdots \\ x_N \end{bmatrix}$$

Statistics are usually enlarged to include additional data in higher dimensions in the statistical prediction. For space reconstruction, it is very important to select an acceptable attempt to integrate dimension m (lag) and time delay [19]. Embedded parameters are determined for victimization sensitivity analysis. Take a statistical into account. In any case where X is the matrix and Y the corresponding output matrix the time - delay vectors are reconstructed. The analysis output is retrieved from previous values of input and victimization.

This is Xp+1 in the series i.e. The previous value of the X1 is changed from the window buffer for our second check and therefore the latest value of Xp+1 is extra, so that we tend to keep window length constantly slippery in the future. Xp+2. Xp+2. Until the data set tip is reached, the window can still slide. When the number of observations is "N," then it would be (n-p) the whole range of validations.

B. Machine regression learner in metaheuristic optimization

L.S.S.V.R. given by Siukens et al. (2002)[20] has many great characteristics and supports highly generalized fast computing to get a series of linear equations to reduce the computational costs, the minimum square cost function can be used. A L.S.S.V.R. training process is used to find a solution by efficiently resolving several linear equations, e.g. the combination gradient method. A quadratic loss function [21] is used in the regression model for this project to reduce the L.S.S.V.R computational to prevent the function. The optimization issue is formulated (2) in accordance with the L.S.S.V.R.

(2)

$$\min_{\omega,b,e} J(\omega,e)=\frac{1}{2}\|\omega\|^2+\frac{1}{2}C\sum_{k=1}^{N}e_k^2$$

Since this is an optimization problem, the differentiation function of Lagrange multipliers with constraints can be solved. Equation (3) is the result of the predictive L.S.S.V.R model.

(3)

$$f(x)=\sum_{k=1}^{N}\alpha_k K(x,x_k)+b$$

Where the Lagrange multiplier is constant with αk and b, where the bias term is b and the kernel function is K(x, xk). This project uses a Radial Function Kernel (RBF). For detailed technical explanation of this approach, refer to the work of Chou et al. (2015)[22]. L.S.S.V.R's predictive accuracy depends heavily on finding its optimized hyper parameters. And we use the enhanced F.A as part of this project. An algorithm designed to optimize the L.S.S.V.R. hyper parameters, i.e. the R.B.F. kernel (β) and the R.B.F. kernel sigma.

C. Hyper parameters tuning : metaheuristic optimization algorithm and swarm

The F.A., created by the Yang [23], is the smart swarm method most successful. This algorithm is inspired by the blinking pattern. The luminosity for any maximization problem is proportional to the value of the objective function. The attraction of a firefly is equally commensurate with the intensity of its light which is apparent to adjacent fires.

(4)

$$\beta=\beta_0 e^{-\gamma r^2 1}$$

The β is the firefly's proportionality, 0β is the firefly's proportionality on r=0, where r is the distance of interest between the firefly. E is a constant coefficient; and π is a coefficient of absorption. The F.A. in detail. Procedure set out in[37]. Despite the F.A. In most problems, it is highly efficient, often becoming a local optimum trap. Furthermore, it is another challenge to set tuning parameters that improve the FA convergence. It's the F.A. To balance exploration, the control parameter should be optimized. That's why Meta F.A. Includes three metaheuristic components in the conventional F.A., namely chaotic map, adaptive inertia weight, and Levy flight. Increase its ability to search and optimize. Fig.2 shows the Meta F.A .- L.S.S.V.R. pseudocodes. Pattern. Model.

Chaotic logistic map to increase the initial population

This is F.A. This is the F.A. A typical approach produces a random first solution. There are two main drawbacks to this approach, firstly because of the slow convergence and secondly because it tends to be trapped in local optima because of a decreased population diversity. Initial response diversity and initial population quality are further improved using a logistical chaotic map. It produces a very diverse range of population fireflies at the initial stage.

Chaotic map of gauss / mouse to maintain attractiveness

The map of Gauss / mouse is the best way to adjust the attractive parameter for FA (β). For a detailed technical explanation, readers can refer to the work of He et al. (2001) in this study.

Randomization adjustment with adaptive weight inertia

Random reduction as progress in iterations increases the convergent efficiency of a swarm-based algorithm. In the early stages of the search process New area, the high inertia weight can boost global exploration performance. Reducing inertia weight enhances local exploration in every final phase (fine tuning of the search area currently in use). Weight of inertia is essential to converging the optimum known response to the world optimum, and the weight of inertia increases simulation time.

Movement control with the levy

Random walking theory plays an important role in modern swarm intelligence and evolutionary optimization algorithms [20]. Levy flights are a random walk in the direction of the Levy distribution. There is no characteristic dimension to the step lengths, because the second or even the first moment can differ, and the distribution shows auto-affinity properties. Random flights are used in two steps to generate random numbers: random direction selection and steps to be taken to comply with the selected distribution of Levy. In this work, directions have been generated with uniform distributions. A stable Levy symmetric distribution is made by using the most popular Mantegna algorithm.

D. **Intelligent time series prediction system by optimizing sliding windows metaheuristics**

Matlab Guide and the MATLAB Runtime Developing Tools used in this study are all based on the software MATLAB and developed on Intel Core i3 and 2 GB RAM machines in a Window environment A User-Friendly Interface is designed with MATLAB GUIDE. The first step, sliding values-time series parameters, lag and size of the window are found. Learning data in the sliding window is included in the input and output matrix. The second phase is used to predict a step ahead of us: Meta F.A-L.S.S.V.R model. The number of test data is the equal number of validations performed. Every validation involves a one-stage prediction. The window continues and validation is completed. This process is repeated up to the completion of all validations. Evaluation and Forecast are two modules available.

The other module controls the performance of the sliding Meta F.A-L.S.S.V.R and model L.S.S.V.R. Users can choose from a number of evaluation variances which are used to open data files, use the test file, hold and slide the control window. The Predict module makes a single or multi-day prediction in advance. A performance evaluation of the predicted values will then be reproduced. The system also provides the feature allowing the user to save the model after it runs, allowing the user to reuse it later.

(a) Sliding-window representation

(b) (b). Reconstruction of space for analysis of time series.

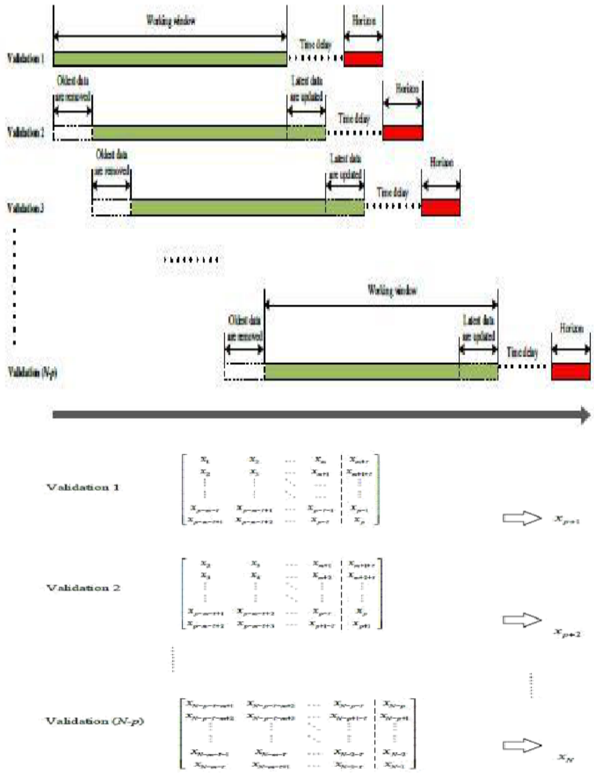


Fig. 1. Space phase reconstruction and sliding window for time series analysis.

$$f(m) = objective\_function_{Validation-data}$$

Equation. Equation. (5) acts like the F.A.-L.S.S.V.R. Meta. Fitness function. Fitness function. The M.A is the main objective where root mean square (RMS) error can be specified. Multi-relationship error, M.A.P error, non-linear regression and mean square error. In addition, if the dataset contains actual zeros (M.A.P), it means that the main function cannot make absolute percentage errors.

IV. ALGORITHM (PSEUDOCODE)

```
Perform objective function f(x), x = (x_1,...,x_d)^T
Set search space and number of generations
Generate initial population of fireflies x_i (i = 1,2,...,n) using logistic chaotic
map
Determine light intensity I_i at x_i by f(x_i)
Define light absorption coefficient γ
Generate initial population, k = 0
1. While (t ≤ MaxGeneration) do
  (1) Update the generation number, k = k + 1
  (2) Tune randomization parameter α by adaptive inertia weight
  (α = α_0 0.9^t)
  (3) Tune attractiveness parameter β by using Gauss/mouse chaotic map
     for i = 1: No. fireflies
        for j = 1: No. fireflies
           if (I_j > I_i)
              Move firefly i toward j in d-dimension by Lévy flight;
              end if
           Vary attractiveness with distance r via exp[-γ * r]
          Evaluate new solutions and update light intensity
        end for j
     end for i
    Rank the fireflies and find the current best
  end while
2. MetaFA-LSSVR function validation
    Set kernel (rbf) and loss-function (least-square) parameters
    Train model with hyperparameters (C,σ)
    Evaluate trained LSSVR model
    Evaluate fitness function f(m), and go to step 1
3. Has the stopping criterion been met?
  If the criterion has been met,
      Go to step 4
  else
      go to step 1
  End
4. Optimized LSSVR model
5. Postprocess results and visualization
```

Fig 2. Meta F.A.-L.S.S.V.R model pseudocode.

E. Methods Evaluation of performance

The performance measurements used for assessing the proposed system's predictive accuracy included root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), mean square error (MSE), correlation coefficient (R), non-linear regression multiple correlation coefficient (R2), and synthesis index (SI). These indexes are used to determine whether the values predicted are close with the actual values.

F. Methods for Data Collection

The two ways the stock market can be analyzed are fundamental analysis and technical analysis. Fundamental analysis are based on overall financial terms and conditions for a company's potential for future growth that are revenue, future, earning market growth, and profit margins and other aspects. Technical analysis, on the other hand, does not help in measuring a stock security value. This method assesses threats and securities that help to predict stock future by analyzing previous market activity statistics. The technical analysis has a great advantage over the fundamental analysis. These contain: Previous data from past years such as stock volumes and inventory prices, which could easily be obtained, and many indicators utilized to carry out technical analysis can be reconstructed from this information.

Algorithm applicability in this model

Technical analysis may be performed for any stock, because the company does what is based on historical data without an understanding. However, one has to assume more in fundamental analysis and the process of analysis can vary in different cases. There are many technical analytical setbacks because it is based essentially on market results and cannot consider some external facts.

1. Data from Time Series

The derived form of stock prices like moving average, oscillators are typical time series. The chronological order of the observations on any particular variable is called a Time series. To discover a pattern in historical data, Time series are mostly used for preparing a forecast and it mostly consist of four main component.

Irregular Variations: Various variations in a time series that follow a pattern that can not be recognized or defined.
Trend: The period of time that is characterized by its ups and down.
Seasonal changes: periodical patterns in a time series.
Cycles: The up and down recurring movements in trend level
In the prediction of stock market, the main aim is to obtain information and identify patterns at an initial stage for creating an investment strategy until any evidence suggests that pattern is not followed. Most of the times large irregular pattern introduce uncertainties and reduce the efficiency of prediction. The Moving Average technique is employed to show the trend by reducing cyclical price fluctuations and to identify irregularity of patterns through crossover. For finding the course of safety movement other techniques such as the percentage price oscillating (PPO), the absolute price oscillating (APO) and the moving average convergence divergence (MACD) are good.

2. Selection of data by Fuzzy Neural Network (FNN).

The selection of Fuzzy neural network data (input or output) shall be as follows:
FNN Entry Clustering: Due to the constraint of clustering technique of FNN, the input range should be stable over time. It suggests that the input raw prices are not right and can increase or decrease to any level in training data.
Dependences of input / output: Since the noisy data does not give significant data and learning patterns, but confuses the FNN, fewer dependencies between input and outputs should be found.
Data Disposability: The availability of data can help us to move to the approach of technical analysis. Since the P.P.O. has a stable value, it can build model with historical information.
P.P.O. is also less noisy than M.A.C.D. or price series and is thought to have more regular and predictive patterns. The problem with the FNN is the detection of early trading signals, described below:
1. Since the model is based on P.P.O. i.e. if it goes long only when P.P.O.>0 and if it goes short only when P.P.O.<0.

2. The input to the FNN are the previous values of n P.P.O. value: P.P.O.(t0),...(t0-1).........PPO(t0n+1).

3. The Fuzzy Neural Network are used to predict P.P.O. in assumption that the trading signal can be detected. So it improves the performance of the pure P.P.O.

4. The output of the Fuzzy neural networks are next m values: P.P.O.(t0+1)...(t0+2)........(t0+m). The new predicted values is used in the place of existing P.P.O.
value.

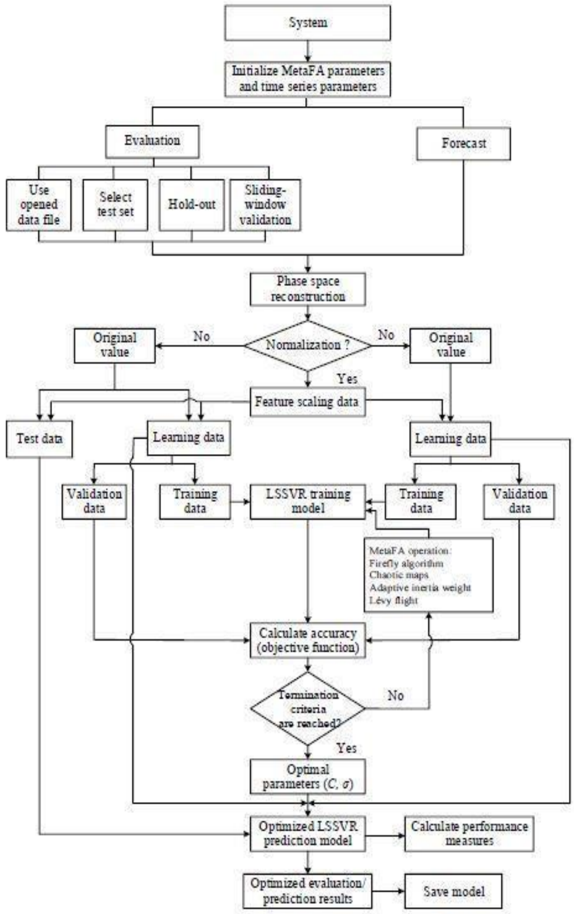
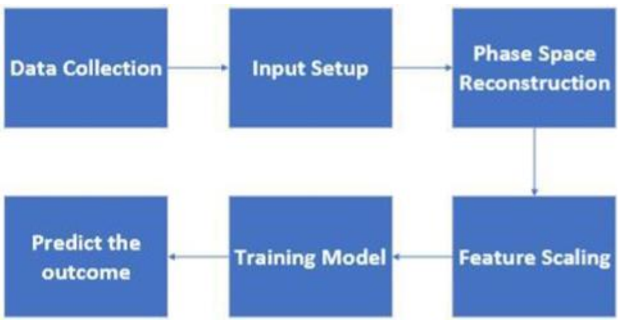


Fig 3. Flow chart of the System

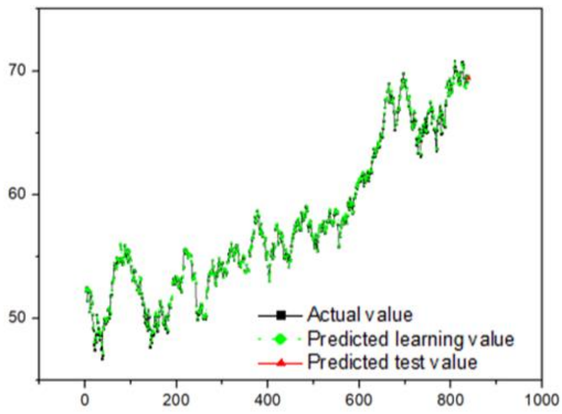
The System architecture model basically shows the behavior of the system in the model. The large models are divided into small subsystems to make them avail same set of services. The architecture include different parts such as collection of data with the input of setup and then reconstruction of phase followed by scaling of features and followed by training of models and at last predicting of result.



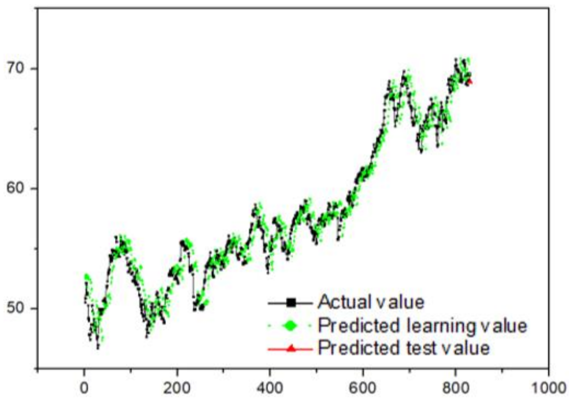
(a) System Architecture (Block Diagram)

V. RESULT

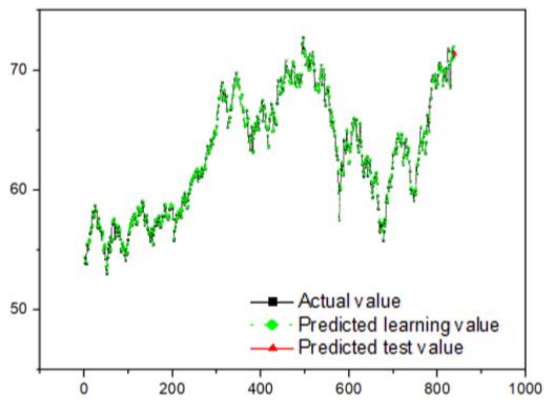
In this project we can see that we can predict the short term stock prices only. This proposed model is a short term prediction model which can predict the values only till next 1 to 30 days only correctly. If we try to predict the data after that then the predicted data will not be so precised as shown in below given results.



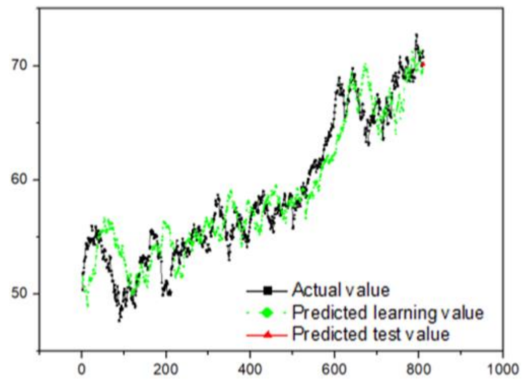
(a) Predicted values of first day



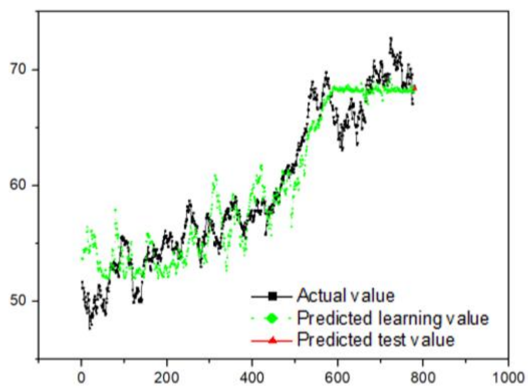
(b) Predicted values of next 10 days



(c) Predicted values of next 15 days



(d) Predicted value of next 1 month



(e) Predicted values of next 2 months

In above we can see that the stock prices can be predicted for the short time give the best results and as the prediction duration increases it becomes hard to analyze the data and produce the output.

Hence from the above results we can say that that the accuracy of the predicted output depends upon the duration of the period for which data has to been predicted.

Accuracy inversely proportional to the duration of data to be predicted

VI. CONCLUSION

In this project we have developed a model that can predict the forecasting of stock market prices based on the historical data and technical analysis of stock market data available using data mining technique. The determined data obtain show us the stock prices available and their peak and low value. This data can guide the investors by telling them whether to invest in a particular company or not on short term basis and allow them to make a profitable investment decision on their shares.

VII. REFERENCES

[1] H. White, "Economic Prediction Using Neural Networks: The Case of IBM Daily Stock Returns", Proc. Second Annual IEEE Conf. on Neural Networks, Vol II, pp.451-458, 1988.

[2] W.-C. Chiang, L. Urban and G. W. Baldridge, "A Neural Network Approach to Mutual Fund Asset [3] Value Forecasting", Omega, Int. J. Mgmt Sci. 24(2), 205-215, 1996.

[3] E. W. Saad, D. V. Prokhorov, and D. C. Wunsch, II, "Comparative Study of Stock Trend Prediction Using Time Delay, Recurrent and Probabilistic Neural Networks", IEEE Trans. Neural Networks, 9(6), 1456-1470, 1998.

[4] J. Moody and M. Saffell, "Learning to Trade via Direct Reinforcement", IEEE Trans. Neural Networks, 12(4), 875-889, 2001.

[5] A.-S. Chen, M. T. Leung and H. Daouk, "Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index", Computers & Operations Research, 30(6), 901-923, 2003.

[6] C. L. Wilson, "Self-Organizing Neural Network System for Trading Common Stocks." IEEE International Conference on Neural Networks, 3651- 3654. 27 June- 2 July, 1994.

[7] T. S-C Chou, C-C Yang, C-H Chen and F. Lai, "A Rule-based Neural Stock trading Decision Support System," IEEE/IAFE Conference on Computational Intelligence for Financial Engineering, New York, USA, pp.148-154, March 1996.

[8] N. Baba, N. Inoue and H. Asakawa, "Utilization of Neural Networks & Gas for Constructing Reliable Decision Support Systems to Deal Stocks" Working Paper of Information Science, Osaka-Kyoiku University, Asahiga-Oka, 4-698-1, Kashiwara City, Osaka Prefecture, 592-8582, Japan, 2000.

[9] W. Leigh, R. Purvis and J. M. Ragusa, "Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: a case study in romantic decision support", *Decision Support Systems*, 32(4), 361-377, 2002.

[10] M. Lam, "Neural Network Techniques for Financial Performance Prediction: Integrating Fundamental and Technical Analysis", Decision Support Systems, 37(4), pp. 567-581, 2004.

[11] Lee, C-T., and Chen,Y-P. 2007. The efficacy of neural networks and simple technical indicators in predicting stock markets. In Proceedings of the International Conference on Convergence Information Technology, pp.2292-2297.

[12] Kuo, M-H., and Chen, C-L. 2006. An ETF trading decision support system by using neural network and technical indicators. In Proceedings of the International Joint Conference on Neural Networks, pp. 2394-2401.

[13] Nagarajan,V., Wu,Y., Liu,M. and Wang Q-G. 2005 Forecast Studies for Financial Markets using Technical Analysis. In Proceedings of the International Conference on Control and Automation
(ICCA2005), pp. 259-264. [14] Chang, P.-C. , Liu, C.-H. , Lin, J.-L. , Fan, C.-Y.
, & Celeste, S.P. Ng. 2009 A neural network with a case based dynamic window for stock trading prediction. In Expert Systems with Applications, vol.36, pp.6889–6898.

[15] Atsalakis G. S., and Valavanis K. P., 2009 Surveying stock market forecasting techniques – part II. In soft computing methods Expert Systems with Applications.

[16] Setty, Venugopal D., Rangaswamy, T.M. and Subramanya, K.N. 2010. A Review on Data Mining Applications to the Performance of Stock Marketing. In International Journal of Computer Applications, vol. 1, no. 3, pp.33-43.

[17] Wang, J-L. , and Chan, S-H. 2006 Stock market trading rule discovery using two-layer bias decision tree. In Expert Systems with Applications, 30, pp.605– 611.

[18] Wu, M-C., Lin, S-Y., & Lin, C-H. 2006 An effective application of decision tree to stock trading. In Expert Systems with Applications, 31, pp.270–274. Regression and Swarm Intelligence," Applied Computational Intelligence and Soft Computing, vol. 2014, no. 2014, pp. 8, 2014.

[20] J. A. K. Suykens, T. V. Gestel, J. D. Brabanter, B. D. Moor, and J. Vandewalle, Least squares support vector machines: World Scientific, Singapore, 2002.

[21] J.-S. Chou, and N.-T. Ngo, "Time series analytics using sliding window metaheuristic optimization-based machine learning system for identifying building energy consumption patterns,"Applied Energy, vol. 177, pp. 751-770, 2016.

[22] J.-S. Chou, W. K. Chong, and D.-K. Bui, "Nature-Inspired Metaheuristic Regression System: Programming and Implementation for Civil Engineering Applications," Journal of Computing in Civil Engineering, vol. 30, no. 5, 2016.

[23] X.-S. Yang, Nature-Inspired Metaheuristic Algorithms: Luniver Press, 2008.

[24] J.-S. Chou, and N.-T. Ngo, "Engineering strength of fiber-reinforced soil estimated by swarm intelligence optimized regression system," Neural Computing and Applications, December 28, 2016.

[25] I. Fister, I. Fister Jr, X.-S. Yang, and J. Brest, "A comprehensive review of firefly algorithms," Swarm and Evolutionary Computation, vol. 13, pp. 34-46, 2013.

[26] J.-S. Chou, and A.-D. Pham, "Nature-inspired metaheuristic optimization in least squares support vector regression for obtaining bridge scour information," *Information Sciences,* vol. 399, pp. 64-80, 2017.

[27] D. He, C. He, L.-g. Jiang, H.-w. Zhu, and G.-r. [19] N.-D. Hoang, A.-D. Pham, and a. M.-T. Cao, "A Novel Time Series Prediction Approach Based on a Hybridization of Least Squares Support Vector.