INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

TECHNIQUES FOR TIME SERIES PREDICTION: - A REVIEW

D. K. Sahoo[1], A.Patra[2], S.N.Mishra[3], M. R Senapati[4]
[1*]D Department of Electronics and Communication Engineering, Centurion Institute of Technology
[2]Department of Computer Science Engineering, Gandhi Institute for Education and Technology
[3]Department of Computer Science Engineering, Indira Gandhi Institute of Technology, Saranga
[4]Department of Computer Science and Engineering Centurion Institute of Technology
Correspondence Author: dksahoo@cutm.ac.in

Keywords: Knowledge Management, Knowledge Creation, Organizational Performance, Human Capital, Bank Personnel

Abstract
This paper shows a recent literature survey on stock market prediction with the help of machine learning techniques. Neural network based models identified as the technique which is used by most of the researcher working on currency exchange rate prediction. Neural networks like artificial neural network(ANN) ,functional link artificial neural network (FLANN)radial basis function network(RBFN), wavelets neural network (WNN) ,Psi Sigma neural networks etc are ensemble with many other learning techniques such as least mean square(LMS) ,genetic algorithm(GA), particle swarm optimization (PSO)etc to improve the accuracy and the efficiency. Given currency exchange market model uncertainty, soft computing techniques are viable candidates to capture stock market nonlinear relations returning significant forecasting results.

Introduction
Financial markets data present a challenging and complex problem to understand and forecast. Forecasting is also a key element of financial and managerial decision making. The main purpose of forecasting is to reduce the risk in decision making that is important for financial organizations, firm and private investors. The common financial time series that need forecasting are stock prices, interest rates, price indices and currency exchange rates. These time series are known to be complex, difficult for econometric modeling, non-stationary, very noisy and badly fitted by linear models. The problem of economic and financial forecasting has recently drawn the attention of many researchers. [2] The stock market is a very complicated dynamic system. Big disturbance, serious non- linearity and blindness of investor all make the stock market prediction very complicated and very hard.

For example, it is a hard work to forecast the stock market, interest rate, exchange rate and bankruptcy. It has been observed to be a potential field of research due to its importance in financial and managerial decision making.

Many efforts of methods have been focused on traditional statistical economics. Limitation of traditional statistical economics consists of inflexibility in dynamic situation and complexity of modeling. Stock price, however, can be affected by various factors. In the long term, we have to admit that there are some essential fixed elements under change tendency as well as uncertainty in certain time. Change of stock price is perhaps a nonstandard economic fact.

So, forecasting of stock market very well is a very interesting problem for researchers and security analyzers [3] Traditional statistical economics could only generally predict stock price in certain period time. Comparatively, artificial neural network, a massively parallel processing non-linear system with self-learning ability and adjustability, modeling on inherent relationship between data, has obtained satisfactory achievement on short term prediction on stock price [4]

Different researchers have used different techniques for predicting the fluctuations in the stock market. The year 2000 Great Scholar proposed new hybrid model of ANN and GA for feature discretization. Feature discretization is to transform continuous values into discrete ones in accordance with certain thresholds. Properly discretized data can simplify the process of learning and may improve the generalizability of the learned results.[5]

In 2009 R Majhi et al[2] have proposed two nonlinear adaptive model ; FLANN and CFLANN for prediction of three different exchange rates for one, three, six and twelve months' ahead. Their performance has been assessed through simulation study and compared with that of LMS model.

In the year 2009 the researcher Feng Li, analyzes principles of stock prediction based on BP neural network, provides prediction model for stock market by utilizing three-layered feed forward neural networks, presents topology of network, principles of determining the number of hidden layers, selection and pretreatment of sample data and determination of preliminary parameters, In order to avoid local extreme and promote convergence speed [6].

Jui Chung Hung proposed the model for prediction on volatility of stock markets using adaptive fuzzy-GARCH and particle swarm optimization (PSO) parameter estimation algorithm to derive the optimal solution for the model. The author determined out of sample forecast volatility of financial stock by using the RLS adaptive algorithm to track changes in volatility in year 2011[7].

In this paper the author in year 2011, Wei Shen select a radial basis function neural network (RBFNN) to train data and forecast the stock indices of the shanghai stock exchange. Author introduces the artificial fish swarm algorithm (AFSA) to optimize RBF. To increase forecasting efficiency, a K-means clustering algorithm is optimized by AFSA in the learning process of RBF [8].

In the year 2012, A.A. Adebiyi predicted the model based on technical and fundamental indicators, and experts' opinions using

INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

neural network architecture is proposed. The aim is to yield more accurate results in stock price prediction. Based on the idea behind technical analysis of investment trading, it is assumed that the behavior of stock market in the future could be predicted with previous information [9].

In the year 2013 the author Mu-Yen-Chen proposes a hybrid ANFIS model to forecast 160 electronics companies listed by the Taiwan stock exchange corporation. Author calculated the volatility for financial ratios with one order and multi order momentum and again optimized the subtractive clustering parameters using the PSO, again optimized the fuzzy inference system parameters using the ANFIS model [10].

Techniques for predicting stock market:

Artificial neural network:

An artificial neural network (NN) is a computational structure modeled loosely on biological processes. NNs explore many competing hypotheses simultaneously using a massively parallel network composed of non-linear relatively computational elements interconnected by links with variable weights. It is this interconnected set of weights that contains the knowledge generated by the NN. NNs have been successfully used for low-level cognitive tasks such as speech recognition and character recognition. They are being explored for decision support and knowledge induction [11], [12], [13].

In general, NN models are specified by network topology, node characteristics, and training or learning rules. NNs are composed of a large number of simple processing units, each interacting with others via excitatory or inhibitory connections. Distributed representation over a large number of units, together with interconnectedness among processing units, provides a fault tolerance. Learning is achieved through a rule that adapts connection weights in response to input patterns. Alterations in the weights associated with the connections permits adaptability to new situations surveys the wide variety of topologies that are used to implement NNs.[14],[15],[16]

Radial Basis Function (RBF) and Artificial fish swarm algorithm (AFSA):

Artificial fish swarm algorithm (AFSA) which based on research in intelligent behavior of animal groups is a new bionic optimization algorithm [17] in 2002. Upon this algorithm, in this paper get an algorithm which is RBF Neural network artificial fish swarm algorithm. This algorithm is mainly used to optimize the RBFNN [18] in hidden layer node position and width value, this artificial fish need to determine the encoding and initialed, calculated to determine the fitness value of artificial fish behavior, calculate the hidden layer to the output layer weights to determine RBFNN output error, the paper is divided into steps to address these issues.

Radial Basis function neural network is a three layer feed forward network. It consists of input layer, hidden layer and output layer. The input layer contains units of signal force, and the second layer is hidden layer. A number of units on the hidden layer is determined by necessity. The third layer is an output layer which reacts to input model. The movement from input layer to hidden layer is non linear and that from hidden layer to output layer is linear. Activation function of the units in hidden layer is RBF[19].

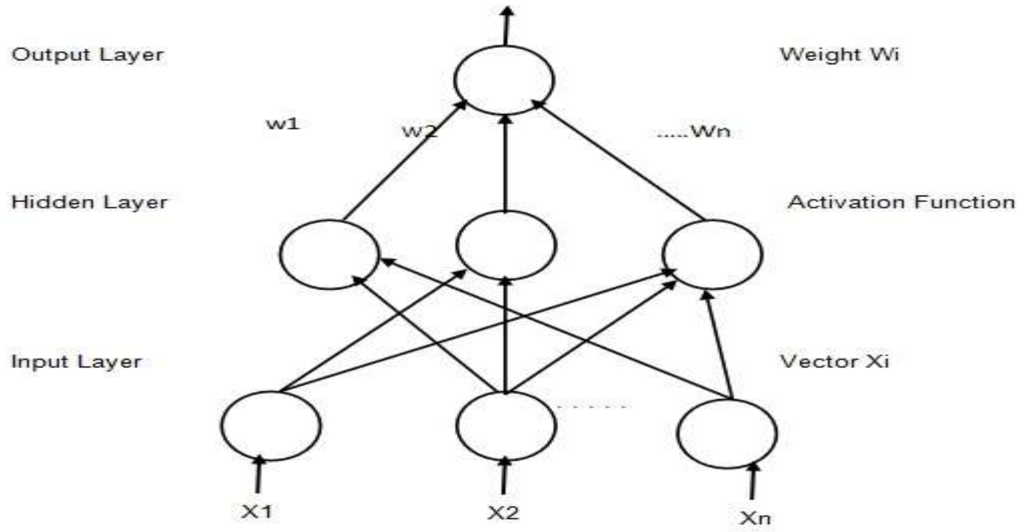

Fig-1 Structure of RBF

INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

Development of functional link artificial neural network (FLANN) model:

The FLANN [20] is a single layered neural network with nonlinear input and a single neuron at the output. This network is a useful substitute of multilayer artificial neural network (MLANN) [15]. However, it is structurally simple and involves less computations compared to those of MLANN. It is also reported that for some applications [21] the FLANN performs better than the MLANN. The nonlinear input is generated by functionally expanding the input vector in a nonlinear manner. Different nonlinear expansions may be employed. These are trigonometric(sine and cosine), Chebyshev and power series. In this paper the trigonometric expansion based financial model is developed for exchange rate prediction as it offers better performance compared to when other expansions are used.

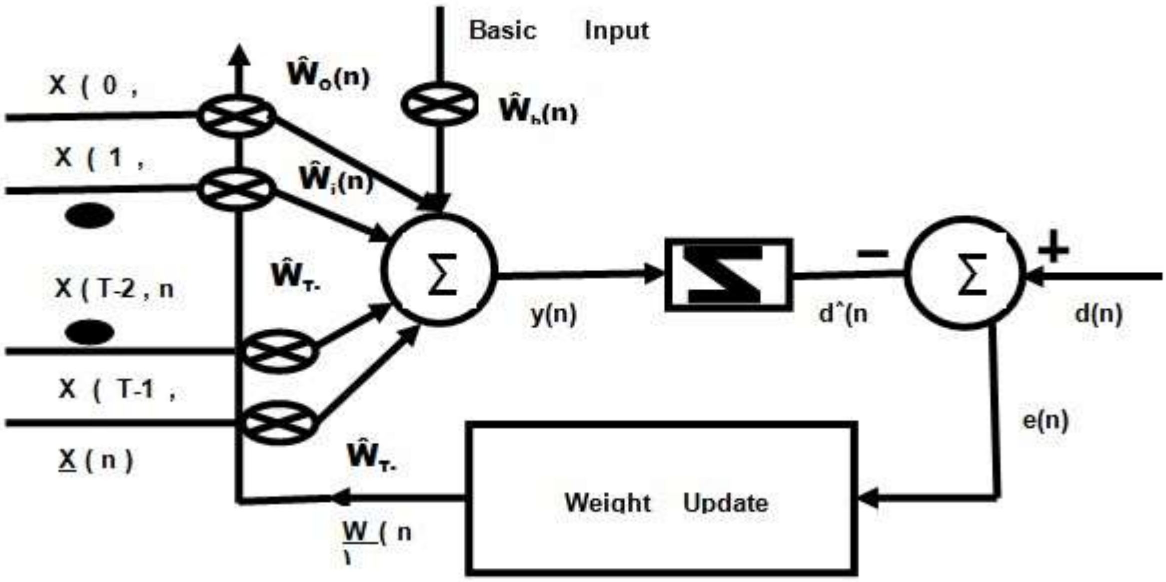

Fig-2 detailed structure of FLANN based forecasting model.

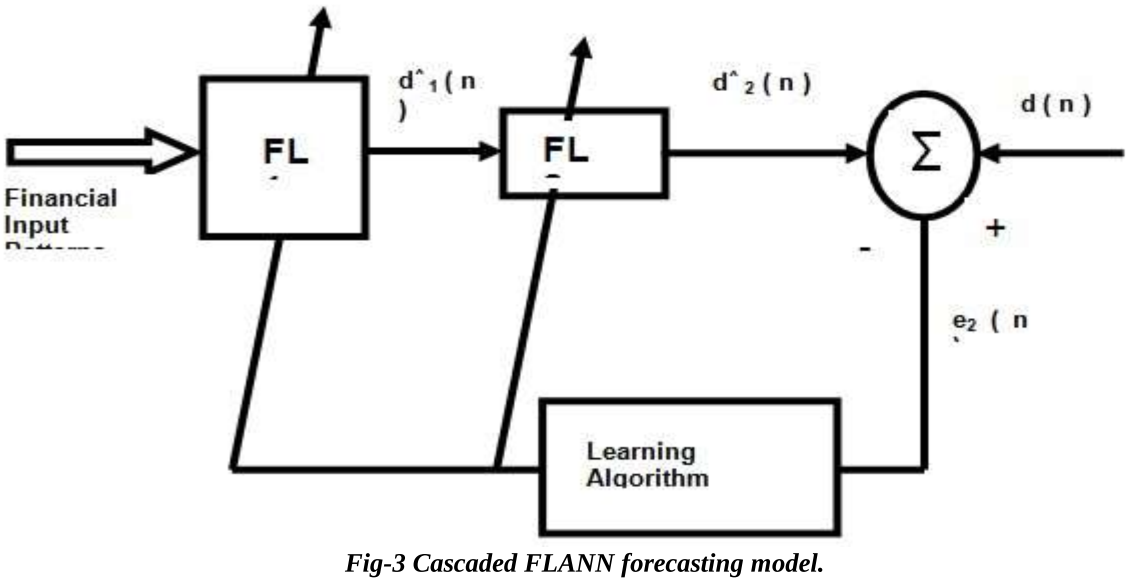

Fig-3 Cascaded FLANN forecasting model.

In this model two single layer FLANN structures are cascaded in series. Fig. 3 shows the block diagram representation of a two stage cascaded FLANN adaptive model. In this figure FL1 represents the first FLANN model as detailed in Fig. 2. This estimated

INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

value once again is expanded nonlinearly either by using trigonometric or exponential functions. As stated earlier this expansion introduces additional nonlinearity to the model. In this paper, we have employed only trigonometric expansion. Unlike multilayer ANN (MLANN), there is only one neuron in each of FL1 and FL2 blocks. Hence the computational complexity of CFLANN model is significantly reduced compared to the MLANN model. The number of inputs to FL1 is equal to the number of extracted features obtained from the given financial time series, which is only three in the present case. However, since FL1 provides only one output, the expanded values in FL2 is only due to one input. In our model five trigonometric expansions are used in FL2.

Back propagation neural network (BPNN):

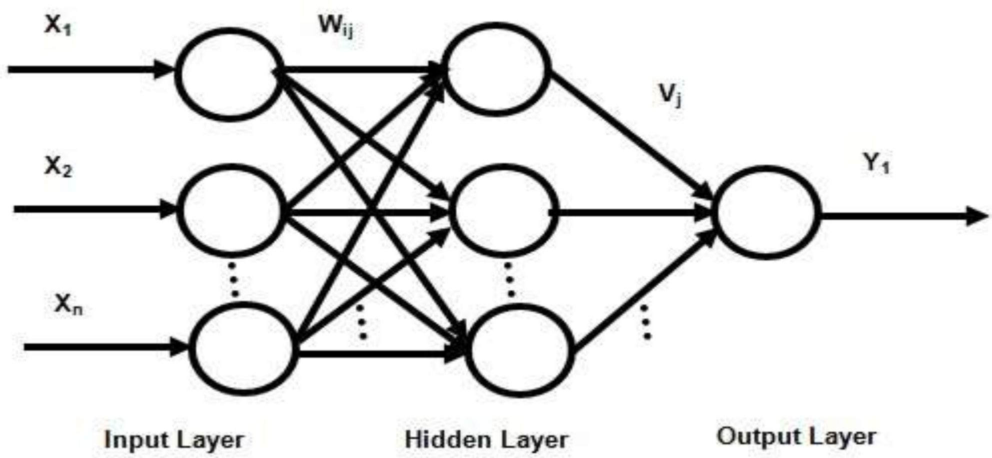

Fig-4 N-R-1 Neural Network Structure

Through training, we can get weight matrix and then form a three-layer BP network prediction model. The chosen macro-economic standard has the following affects on stock price: social retail goods, on the one hand represents purchasing power. On the other hand, with the increase of retail goods, saleroom of each company will do the same, together with the increase of interests, and benefits of investors. Furthermore, stock price will multiply. Consumer price index reflects level of price which significantly influences investors, whereas rising of price causes more expenditure on durable consumable products and lower benefits for investors which results in the decrease of stock price. On the other hand, price increasing could give enterprises more benefits, so as to the investors' income, and then stock price will multiply.

Genetic Algorithm

GA is a search algorithm based on survival of the fittest among string structure [22]. GA has been investigated recently and shown to be effective in exploring a complex space in an adaptive way, guided by the biological evolution mechanisms of reproduction, crossover, and mutation[23].

The following describe the basic steps of GA. The first step is problem representation. The problems must be represented as a suitable form to be handled by GA. GA often works with a form of binary coding. If the problems are coded as chromosomes, the populations are initialized. Each chromosome within the population is gradually evolved by biological operations. Larger populations ensure greater diversity but require more computational burden. Once the population size is chosen, the initial population is randomly generated. After the initialization step, each chromosome is evaluated by fitness function. According to the value of the fitness function, the chromosomes associated with the fittest individuals will be reproduced more often than those associated unfit individuals.

Crossover allows the search to fan out in diverse direction looking for attractive solutions and permits chromosomal material from different parents to be combined in a single child. In addition, mutation arbitrarily alters one or more components of a selected chromosome.GA tense to converge on optimal or near optimal solutions. GA is usually employed to improve the performance of AI techniques. For ANN, GA is popularly used to select NN topology such as optimizing relevant feature subset, determining the optimal number of hidden layers and processing elements.

GA approach to feature discretization (GAFD) for ANN. GA supports simultaneous optimization of connection weights and feature discretization. GAFD takes into consideration the dependent feature by fitness function in GA. GA iterates the evolution of the population to maximize the fitness function. GAFD simultaneously discretizes all features into the intervals at the exact thresholds. GAFD determines the maximal number of thresholds automatically.

GA approach to feature discretization for ANN:

Many fund managers and investors in the stock market generally accept and use certain criteria for technical indicators as the signal of future market trends. Even if a feature represents a continuous measure, the experts usually interpret the values in qualitative terms such as low, medium, and high[24]. For 'Stochastic %K', one of the most popular technical indicators, the value

INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

of 75 is basically accepted by stock market analysts as a strong signal if the value exceeds 75, the market is regarded as an overbought situation or a bullish market. On the other hand, if it drops below 25, it is considered as an oversold situation or the signal of a bearish market. When the value of 'Stochastic %K' is placed between 25 and 75, it is regarded as the signal of a neutral market [25]. The reasoning process of ANN may be like that of human experts.

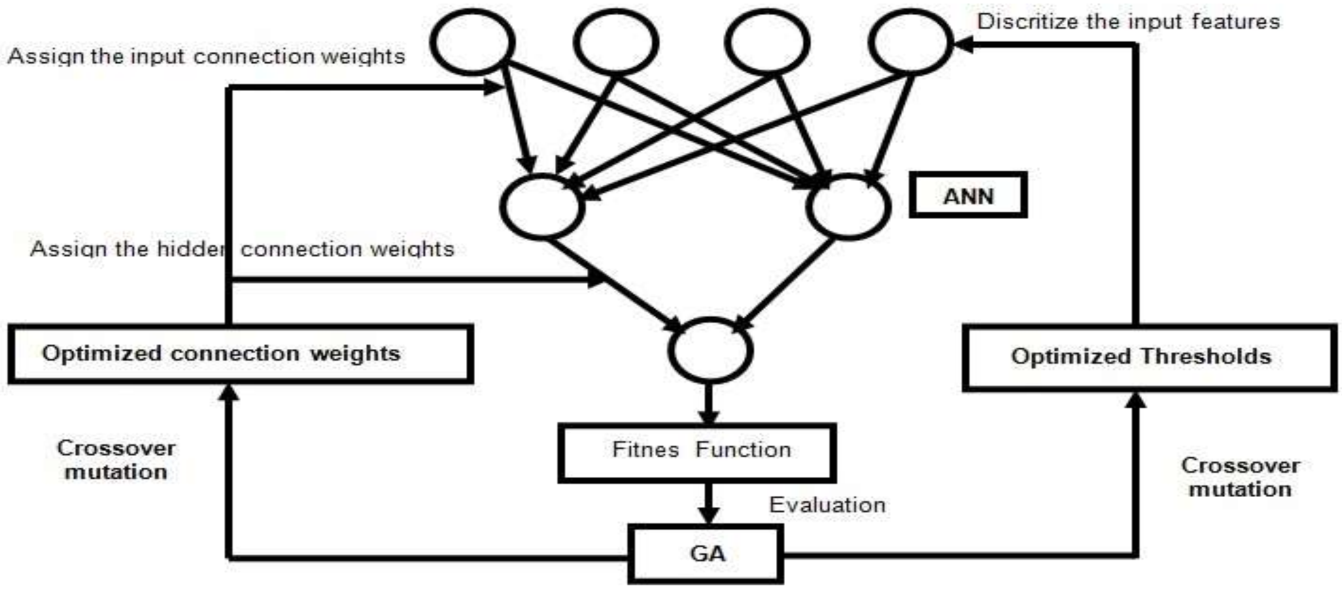

Fig-5 Framework of GAFD

Properly discretized data can simplify the process of learning and may improve the generalizability of the learned results because it may effectively reduce the noisy and redundant data. Feature discretization needs relevant and rational discretizing thresholds. However, the thresholds may vary depending on the securities being analyzed and the overall market condition [26]. There are no general guidelines to discretize for this reason. We may search the thresholds for discretizing a continuous measure into a qualitative norm to capture the domain-specific knowledge. As mentioned earlier, although some studies suggested various methods for discretizing features, this paper proposes the optimization of discretizing thresholds based on GA. GAFD may find optimal or near-optimal thresholds of discretization for maximum predictive performance because GA searches the optimal or near-optimal parameters to maximize the fitness function. The overall framework of GAFD is shown in Fig. 5.

The algorithms of GAFD consist of three phases. Phase1: - In the first phase, GA searches optimal or near optimal connection weights and thresholds for feature discretization. The populations, the connection weights and the thresholds for feature discretization, are initialized into random values before the search process. The parameters for searching must be encoded on chromosomes. This study needs three sets of parameters. The first set is the set of connection weights between the input layer and the hidden layer of the network. The second set is the set of connection weights between the hidden layer and the output layer. As mentioned earlier, the above two sets may mitigate the limitation of the gradient descent algorithm. These sets were incorporated in the studies of them [27], [28], [29], [30] and [31]. The third set represents the thresholds for feature discretization.

The fuzzy graph model with PSO:

The adaptive Fuzzy-GARCH model refers to both GARCH models and the parameters of membership functions, which are determined by the characteristics of market itself. Here, we present an iterative algorithm based on PSO to estimate the parameters of the membership functions. The PSO method aims to achieve a global optimal solution with a rapid convergence rate. Fuzzy Systems are universal approximation that can uniformly estimate non linear continuous function with arbitrary accuracy. The fuzzy GARCH model is described by IF- THEN rules and is employed to ensure that GARCH model can appropriately simulate the fluctuation of stock market.

PSO is a population based stochastic optimization algorithm that attempts to generate optimal solutions to improve fitness. Each swarm consists of many particles, and each particle represents a possible solutions. Each particle has its own position and velocity with initial random values. In order to find an optimal solution each particle evolves in search space following three ways. (1) By moving towards in the previous directions; (2) By moving towards the optimum, which is its present location; or (3) By moving towards the best solution for the entire population. The result is an improvement in the performance of the next iteration. This type of algorithm is more suitable for the fuzzy- GARCH model problem

INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

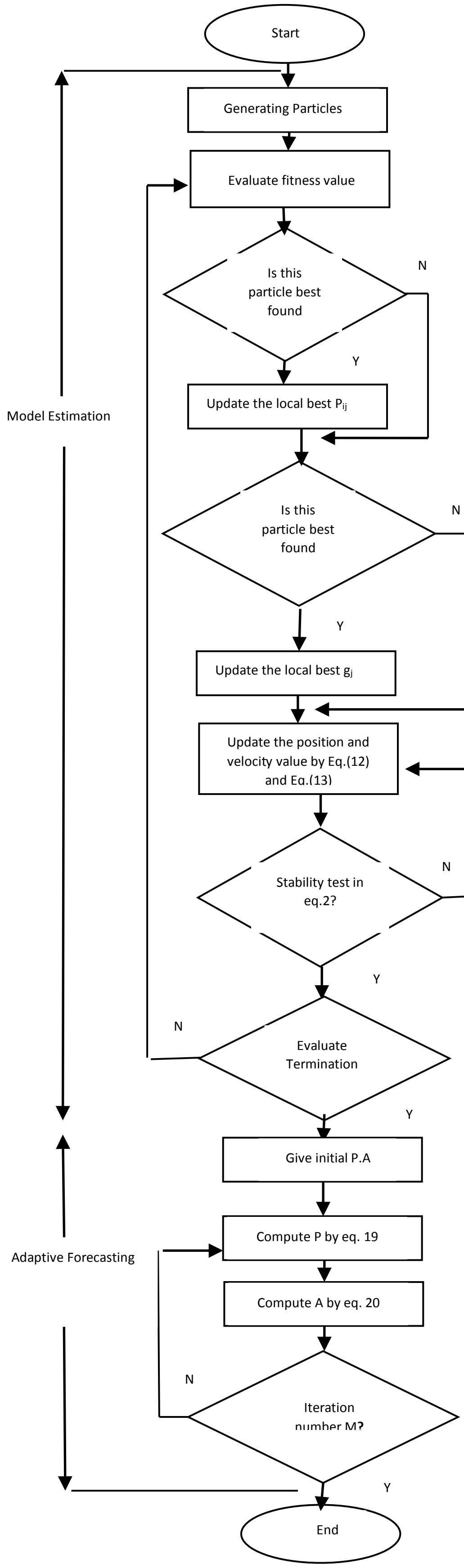

Fig-6

INTERNATIONAL JOURNAL OF RESEARCH SCIENCE & MANAGEMENT

Conclusion:

This study has surveyed articles that have applied neural networks hybridised with different learning methods and also with different statistical and machine learning methods to predict stock market values. The study has focused on input data, forecasting methodology, model comparisons and measures used for performance evaluation. The observation is that neural networks ensemble with statistical models with differential evolution as the learning algorithms are most suitable for stock market forecasting. Experiments demonstrate that machine learning techniques outperform conventional models in most cases. They return better results as trading systems of accuracy and efficiency. However,difficulties arise when defining the structure of the model (the number of hidden layer neurons etc, even if it does not affect that much).

References

1. Bodyanskiy, Y., & Popov, S. (2006). Neural network approach to forecasting of quasiperiodic financial time series. European Journal of Operational Research, 175(3), 1357–1366.
2. Ritanjali Majhi ,G. Panda ,G. Sahoo ,Efficient prediction of exchange rates with low complexity artificial neural network models, Expert Systems with Applications 36 (2009) 181–189, doi:10.1016/j.eswa.2007.09.005, ELSEVIER
3. W. Gao 2008, New Neural Network Based on Ant Colony Algorithm for Financial Data Forecasting , International Conference on Intelligent Information Hiding and Multimedia Signal Processing ,1437-1440,DOI 10.1109/IIH-MSP.2008.126, IEEE.
4. Feng Li, Cheng Li Application Study of BP Neural Network on Stock Market Prediction .2009 Ninth International Conference on Hybrid Intelligent Systems, 174-177,DOI 10.1109/HIS.2009.248 , IEEE
5. Kyoung Jae kim ,Ingo Han ,Genetic algorithms approach to feature discretization in ANN for prediction of stock price index. Expert Systems with Applications ,2000, Elsevier, 0957-4174/00027-0
6. Feng Li, Cheng Liu, 2009 Ninth International Conference on Hybrid Intelligent Systems, IEEE, DOI 10.1109/HIS.2009.248
7. Jui Chung Hung,2011,Adaptive fuzzy –GARCH model applied to forecasting the volatility of stock markets using particle swarm optimization ,page-4673-4683 information sciences,ELSEVIER,DOI:10.1016/j.ins.2011.02.027
8. Wei Shen,2011,Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm,page-378-385,doi:10.1016,knowledge based systems, ELSEVIER.
9. A.A. Adebiyi,2012, An Improved Stock Price Prediction using Hybrid Market Indicators ,page-124-135,ISSN-2006-1781,IEEE.
10. Mu-Yen-Chen ,2013,A hybrid ANFIS model for business failure prediction using PSO and subtractiveclustering,page-180-195,information sciences, doi:10.1016, ELSEVIER.
11. Schocken, S. and Ariav, G.,Neural networks for decision support: problems and opportunities', Decision Support Systems, 11 (1994), 393-414.
12. Dutta, S., Shekhar, S. and Wong, W. Y.,`Decision support in non-conservative domains: generalization with neural networks', Decision Support Systems, 11 (1994), 527-544.
13. Yoon, Y., Guimaraes, T. and Swales, G., Integrating artificial neural networks with rule-based expert systems', Decision Support Systems, 11 (1994), 497-507.
14. Ralston, A. and Reilly, E. D. Encyclopedia of Computer Science, New York.
15. Lippmann, R. P., `An introduction to computing with neural nets', IEEE ASSP Magazine, (1987), 4-21.
16. MONICA ADYA and FRED COLLOPY, How Effective are Neural Networks at Forecasting and Prediction? A Review and Evaluation, Journal of Forecasting J. Forecast. 17, 481-495 (1998).
17. LI Xiao-Lei, Xue Yun-can, Lu Fei, et,al. Parameter Estimation Method Based-on ArtificialschoolAlgorithm[J], Journal of Shandong University: Engineering Science,2004,34(3):84-87.
18. LIU Yao-nian,LI Ying-hong,, LIU Jun-feng et,al. RBF neural Network Based on Artifical Fish Swarm Algorithm[J], Journal of Northeast Dianli University: Natural Science Edition, 2006,26(4):23-27.
19. Qing-kunSong, Meng-meng Xu, Yi Liu, RBF Neural Network Controller Research Based on AFSA algorithm, http://dx.doi.org/10.14257/astl.2013.31.29.
20. Pao, Y. H. (1989). Adaptive pattern recognition and neural network. Addison-Wesley Publishing Company, Int.
21. Patra, J. C., Pal, R. N., Chatterji, B. N., & Panda, G. (1999). Identification of nonlinear dynamic systems using functional link artificial neural networks. IEEE Transactions on systems, man and cybernetics-Part B: Cybernetics, 29(2), 254–262.
22. Goldberg, D. E. (1989). Genetic algorithms in search, optimization, and machine learning, Reading, MA: Addison-Wesley.
23. Adeli, H., & Hung, S. (1995). Machine learning: neural networks, genetic algorithms , and fuzzy systems, New York: Wiley.
24. Slowinski, R., & Zopounidis, C. (1995). Application of the rough set approach to evaluation of bankruptcy risk. International Journal of Intelligent Systems in Accounting, Finance and Management, 4 (1), 27–41.
25. Edwards, R. D., & Magee, J. (1997). Technical analysis of stock trends, Chicago, IL: John Magee.
26. Achelis, S. B. (1995). Technical analysis from A to Z, Chicago: Probus Publishing.

27. Dorsey, R., & Sexton, R. (1998). The use of parsimonious neural networks for forecasting financial time series. Journal of Computational Intelligence in Finance, 6 (1), 24–31.
28. Gupta, J. N. D., & Sexton, R. S. (1999). Comparing backpropagation with a genetic algorithm for neural network training. Omega, 27 (6), 679–684.
29. Ignizio, J. P., & Soltys, R. (1996). Simultaneous design and training of ontogenic neural network classifiers. Computers and Operations Research, 23 (6), 535–546.
30. Sexton, R. S., Dorsey, R. E., & Johnson, J. D. (1998b). Toward global optimization of neural networks: a comparison of the genetic algorithm and backpropagation. Decision Support Systems, 22 (2), 171–185. Sexton, R. S., Dorsey, R. E., & Johnson, J. D. (1999). Optimization of neural networks: a comparative analysis of the genetic algorithm and simulated annealing. European Journal of Operational Research, 114(3), 589–601.
31. Shin, K., Shin, T., & Han, I. (1998). Neuro-genetic approach for bankruptcy prediction: a comparison to back-propagation algorithms. Proceedings of the International Conference of the Korea Society of Management Information Systems (pp. 585–597). Seoul, South Korea.