Efficient Forecasting of Exchange rates with Recurrent FLANN

[1]Ajit Kumar Rout, [2]Sai Prasad Setty

[1,2]*Computer Science & Engineering GMR Institute of Technology India*

Abstract: *In this paper, Functional link artificial recurrent neural network (FLARNN) ensemble forecasting model is proposed for foreign exchange rates prediction. In the survey of existing literature, it is revealed that there is need to develop efficient forecasting models involving less computational load and fast forecasting capability. We compare the FLARNN model performance with some existing neural network models such as LMS and FLANN in terms of the exchange rates between US Dollar and other three currencies (Indian Rupees, British pound, Japanese Yen). Experimental results reveal that the predictions using the proposed approach are consistently better than the above mentioned methods on the basis of various parameters including error convergence and the Mean Average Percentage Error (MAPE).*

Keywords: *Artificial neural networks; Recurrent; Currency Exchange rate Forecasting; FLANN; FLARNN.*

I. Introduction

Financial Forecasting is one of the hottest fields of research lately due to its commercial applications owing to the high stakes and the kind of attractive benefits that it has to offer. Prediction of foreign exchange (FX) rates and stock prices has been an active area of research in computational intelligence. Forecasting the price movements in exchange rate markets has been a major challenge for common investors, businesses, brokers and speculators. As more and more money is being invested the investors get anxious of the future trends of the exchange rate in the market. The FX market is the largest and most lucrative financial markets [1].

Researchers have devoted a great deal of effort in order to find a good explanation (model) of the movement of FX rates between major currencies during the last decade [2, 3]. It is widely known that FX rates are affected by many co-integrated, macro-economic, political and even psychological factors. So far, there is lack of modeling technique that can completely accommodate all the factors and their complex interactions. Therefore, modeling and forecasting FX rate movements poses a great challenge in today's global economy.

Since financial time series are complex, the researchers consider the processes, which change the time series data as a black box and they just study the fluctuations of the series (Anastasakis & rt, 2009).In order to predict nonlinear time series such as exchange rate time series, it is better to employ neural networks that are nonlinear in nature. Neural networks can be divided into two major types, feed forward and recurrent networks. Feed forward neural networks, such as the Multilayer Perceptron (MLP) and the Radial Basis Function (RBF) net, have been successfully used for time series prediction (Lee and Haykin, 1999-4). However, MLPs utilize computationally intensive training algorithms (such as the error back-propagation (Rumelhart et al., 1986-5) and can get stuck in local minima. In addition, these networks have problems in dealing with large amounts of training data, while demonstrating poor interpolation properties, when using reduced training sets. In the case of RBFs, the networks are trained only once on a large example set taken from the signal such that the dynamics of the underlying system can be captured. Therefore, the networks produce sequential outputs in response for newly arriving data. This means that such a system can be used when the dynamics of the time series does not change considerably over time, a condition which is usually contravened in practice (Lee and Haykin, 1999-4).

Recurrent neural networks have advantages over feed forward networks in that they have the ability to store previous state information and prevent the need to predict the model order of the time series (Pao et al., 1992-6).

This study proposes a Functional Link or FLANN architecture based model to predict the movements of prices in the foreign exchange rates. The functional link ANN is a novel single neuron based architecture first proposed by Pao [4]. It has been shown that this network may be conveniently used for functional approximation and pattern classification with faster convergence rate and lesser computational load than a Multi-layer Perceptron (MLP) structure.

This paper shows the comparison three models such as RBF, FLANN and FLARNN for prediction of three different exchange rates for one , five, ten ,thirty and sixty days ahead. There performance has been assessed through simulation study with MAPE parameters. The rest of this paper is organized as follows. The development of FLARNN model dealt in section 2. The design of input data of these models is covered in section 4. In section 5 the simulation of experiment is carried out. Finally, conclusions are given in Section 6

II.   Introduction To Flarnn Based Model For Exchange Rate Conversion.

This study proposes a FLARNN architecture based model to predict the movements of prices in the exchange rates between US Dollar and other three currencies. The FLARNN is a novel single neuron based architecture first proposed by Pao [5]. It has been shown that this network may be conveniently used for functional approximation and pattern classification with faster convergence rate and lesser computational load than a Multi-layer Perceptron (MLP) structure. It is a flat net without any need for a hidden layer. Therefore, the computations as well as learning algorithm used in this network are simple. The functional expansion of the input to the network effectively increases the dimensionality of the input vector and hence the hyper planes generated by the FLARNN provide greater discrimination capability in the input pattern space. Various system identifications, control of nonlinear systems, noise cancellation and image classification systems [6] have been reported in recent times. These experiments have proven the ability of FLARNN to give out satisfactory results to problems with highly non-linear and dynamic data [7]. Further the ability of the FLARNN architecture based model to predict exchange rate movements, both for short term (next day) and medium term (one month and two months) prediction using statistical parameters consisting of well known technical indicators based on historical index data is shown and analyzed.

2.1 Structure of Functional Linked ARNN

FLARNN is a single layer, single neuron recurrent architecture, first proposed by Haykin and Li [8], which has the exceptional capability to form complex decision regions by creating non-linear decision boundaries. The architecture of the FLARNN is different from the linear weighting of the input pattern produced by the linear links of the better known Multi Layer Perceptron (MLP). In a FLARNN, each input to the network undergoes functional expansion through a set of basic functions as well as Technical indicator (TE). The functional link acts on an element or the entire pattern itself by generating a set of linearly independent functions.

The inputs expanded by a set of linearly independent functions in the function expansion block, causes an increase in the input vector dimensionality. In our experiment, the functional expansion block comprises of a set of trigonometric function.
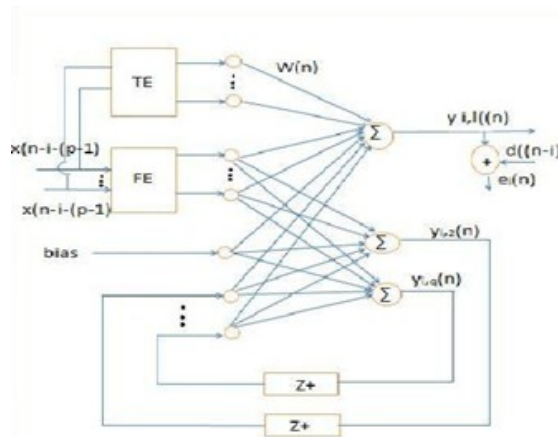

Fig-1. Structure of FLARNN with single output

The basic function for the FLARNN, $B = \{\phi_i \in L(A)\}_{i \in x}$ is to be selected keeping the following properties into consideration 1)$\emptyset_i = 1$, 2) the subset $b_j = \{\emptyset_i \in B\}_{i=1}^{j}$ is a linearly independent set, i.e., if $\sum_{i=0}^{n} w_i \emptyset_i = 0$, $w_i = 0$ for all i=1,2,3…….j, and 3)$\sup_j \lfloor \sum_{i=1}^{n} \parallel \theta_i \parallel_A^2 \rfloor^{\frac{1}{2}} < \infty$.

Let $B_N = \{\emptyset\}_{i=1}^{N}$ be a set of basic functions to be considered to the FLARNN as shown in fig. 1. Thus, the FLARNN consists of N basis functions $\{\emptyset_1, \emptyset_2, \emptyset_3, ……..,\emptyset_N\} \in B_N$ with following input-output relationship for the *j*th output:

$$\hat{y}_f = p(S_f);$$

$$Sf = \sum_{i=1}^{N} w_{ij} \phi_i(x) \quad (1)$$

Where $X \in A \subset R^n$ i.e. x $= [x_1 x_2 …… x_n]^t$ is the input pattern vector, $\hat{y} \in R^m$ i.e., $\hat{y} = [\hat{y}_1 \ \hat{y}_2 ….\hat{y}_m]^T$ is the output vector and $W_j = [w_{j1} \ w_{j2} …. w_{jn}]$ is the weight vector associated with the

jth output of the network. The non–linear function considered in this case $\rho(\cdot) = \tanh(.)$ Considering the m-dimensional input vector, equ(1) can be written as $S = W\Phi$ Where W is $(m \, X \, N)$ Weight matrix of FLARNN given by, $W = [w_1 w_2 \ldots \ldots w_m]^T$, $\emptyset = [\emptyset_1(X)\emptyset_2(X) \ldots \ldots \emptyset_N(X)]^T$ is the basic function vector, and $S = [S_1 S_2 \ldots \ldots S_n]^T$ is a matrix of linear outputs of FLARNN. The m–dimensional output vector $\hat{y}$ may be given by

$$\hat{y} = \rho(S) = f_w(X) \qquad (2)$$

2.2 Learning with Functional Linked ARNN

The learning of ARNN can be described as approximating a continuous, multivariate function f(X) by an approximating function fw(X) .Given a function the objective of the learning algorithm is to find the optimum weights such that fw (X) obtained approximates f(X) within an error e. This is achieved by recursively updating the weights. Let the training sequence be denoted by {xk ,yk } and the weight of the network be W(k), where k is the discrete time index given by k= κ +λ K where λ=0,1,2,….., and κ = 0,1,2,…,K. From (1) the j th output of FLARNN at a given time k can be given as:

$$\hat{y}_j = \rho\left(\sum_{i=1}^{N} w_{ji}(k)\phi_i(x_k)\right)$$
$$= \rho(w_j(k)\phi_i(x_k)) \qquad (3)$$

For all $X \in A$ and j =1,2,3,……,m where $\phi = [\phi_1(X_k)\phi_2(X_k)\cdots\phi_N(X_k)]^T$ .Let the

corresponding error be denoted by $e_j(k) = y_j(k) - \hat{y}_i(k)$ .The Least Mean Square (LMS) update rule for

all the weights of the FLARNN is given by

$$W(k+1) = W(k) + \mu\delta(k)\phi(x_k) \qquad (4)$$

Where , $W = [w_1(k)w_2(k)\cdots w_m(k)]^T$ is the $M \times N$ dimensional weight matrix of the FLARNN at

the k-th time instant is $\delta(k) = [\delta_1(k)\delta_2(k)\cdots\delta_m(k)]^T$ and $\delta_j(k) = \left(1 - \hat{y}_j(k)^2\right)e_j(k)$ (5).

Similarly the Recursive Least Square (RLS) update rule for all weights of the FLARNN is given by
W(k +1) =W(k) + ej(k)zzk '(k)   (6)
Where, zzk(k) = z(k) /(1+ q) ,
q = X (k).zk(k) and zk(k) = R(k).X (k)   (7)
The autocorrelation matrix R(k) is updated with the equation,
R(k +1) = R(k) − zzk(k).zk(k) '   (8)
Which is initialized using the expression, R(0) =η.I, where I is the identity matrix and η is a constant. The motivations for using trigonometric polynomials in the functional expansion stage are explained below. Of all the polynomials of N-th order with respect to an ortho-normal system {φ i(x)}I N =1 gives the best approximation in the metric space L2 is given by the N-th partial sum of its Fourier series with respect to this system.

Thus, the trigonometric polynomial basis functions given by {1,cos(π x),sin(π x),cos(2π x),sin(2π x),....,cos(Nπ x),sin(Nπ x)} provide a compact representation of the function in the mean square sense. However, when the outer product terms are used along with the trigonometric polynomials for function expansion, better results were obtained in the case of learning of a two-variable function.

III.   Network Input Selection Data Preprocessing

The data for the exchange rate prediction experiment has been collected for three different currency converter indices namely 1 US $ to Rupees, Pound and Yen from Canada bank currency converter.. The proposed forecasting model is developed to forecast the closing price of the index in each day of the forecasting period. Different technical and fundamental indicators are used as inputs to the network. Technical indicators are any class of metrics whose value is derived from generic price activity in a currency. Technical indicators look to predict the future price levels, or simply the general price direction, of a security by looking at past patterns. Out of the many technical indicators used by traders, 10 indicators (Table-1) have been chosen as input to the network which has been used before by many researchers for exchange rate forecasting problems.

IV.   Experiment Model Setup

Data samples are collected from the historical values of 10-year currency converter- Bank of Canada. We employ models using Radial Basis Function and Functional Link Neural Network architecture (FLANN) and FLARNN structure where the parameters of each of the structure is updated using LMS algorithm learning.

4.1 Modeling FLARNN structure

For our experiment, the total input to the single neuron FLARNN is 30 plus a bias. This gives us 31 weights that are to be trained using a suitable adaptive algorithm for a particular exchange rate index. The neuron adds up the input weight products and bias. The sum is then taken up by a suitable activation function to give the output of the network. For this particular case we used the tan hyperbolic activation function. In the FLARNN model prediction, four trigonometric functions namely Cos πx, Cos 2πx, Sin πx and Sin 2πx were used along with the technical indictor. An optimum value of the convergence coefficient was taken as 0.1 for all the prediction experiments. The inputs are normalized to values between +1 and -1. This can be done by a number of normalization techniques. One of the popular techniques we used was expressing the data in terms of the maximum and minimum of the data set. All the values are normalized by using the following equation

$Y = (2*X – (Max + Min))/ (Max + Min)$

Where Y: - normalized values, X: - present value.

Table-1

Technical Indicators	Formula
Simple Moving Average (SMA)	$\dfrac{1}{N}\sum_{i=1}^{N} x_i$ N = No. of Days. x_i = today's price
Exponential Moving Average (EMA)	$(P \times A) + (\text{Previous EMA} \times (1 - A))$; A=2/(N+1) P – Current Price, A- Smoothing factor, N-Time Period
Accumulation/ Distribution Oscillator (ADO)	$\dfrac{(C.P - L.P) - (H.P - C.P))}{(H.P - L.P) \times (\text{Period's Volume})}$ C.P – Closing Price, H.P – Highest price, L.P – Lowest price
Stochastic Indicator (STOC)	$\%K = \dfrac{(\text{Today's Close - Lowest Low in K period})}{(\text{Highest High in K period - Lowest Low in K period})} \times 100$ $\%D$ = SMA of $\%K$ for the Period.
On Balance Volume (OBV)	If Today's Close > Yesterday's Close OBV = Yesterday's OBV + Today's Volume If Today's Close < Yesterday's Close OBV = Yesterday's OBV – Today's Volume
WILLIAM's %R	$\%R = \dfrac{(\text{Highest High in n period - Today's Close})}{(\text{Highest High in n period - Lowest Low in n period})} \times 100$
Relative Strength Index (RSI)	$RSI = 100 - \dfrac{100}{1 + (U/D)}$
Price Rate Of Change (PROC)	$\dfrac{(\text{Today's Close - Close X-period ago})}{(\text{Close X-period ago})} \times 100$
Closing Price Acceleration (CPAcc.)	$\dfrac{(\text{Close Price - Close Price N-period ago})}{(\text{Close Price N-period ago})} \times 100$
High Price Acceleration (HPAcc.)	$\dfrac{(\text{High Price - High Price N-period ago})}{(\text{High Price N-period ago})} \times 100$

4.2 Training Process

The training of the network takes place in the following fashion. The weight update is epoch based. Since we have three different networks namely RBF, FLANN and FLARNN. As mentioned in section 4.1, altogether 31 weights of FLARNN model need to be updated by learning algorithm. The weights remain unchanged till all of the training data set is fed into the network, compared with the desired output and their respective error stored. The mean error for the entire epoch is calculated, and then the adaptive weight update takes place. The Least Mean Square (LMS) update algorithm is used in our experiment updates the weights by adding the product of the convergence constant, the respective input with the mean error for the epoch to the weights of the previous epoch. The cost function for the training process is the Mean Square Error (MSE). It is suitable to end the training of the network when the minimum level of the cost function is observed. The number of iteration is decided upon by gradient of the MSE curve. If it is observed that there is no significant decrease in the MSE then the training experiment can be stopped. There exists a trade-off between the time taken and quality of training. High number of iterations tends to give better training of the network at the cost of time taken to train

4.3 Testing Process

At the end of the training process of the network, the weights are frozen for testing the network on inputs that were set apart from the training set. The testing set patterns are the input to the network and the output is compared with desired output. The percentage of error is recorded for each data set. The Mean Absolute Percentage Error (MAPE) is used to gauge the performance of the trained prediction model for the test data. In our simulation, we have calculated both MSE and MAPE, but the analysis and comparison is done on the basis of MAPE only. The effort is to minimize the MAPE for testing patterns in the quest for finding a better model for forecasting exchange rate movements. The MAPE is given as

$$MAPE = \frac{1}{N}\sum_{j=1}^{N}\left|\frac{y_j - \hat{y}_j}{y_j}\right| \times 100$$

V.   Simulation Results

5.1 Simulation Results for 1 US$ to Rupees

Out of around 3000 days data of 1 US$ to Rupees it was found that only 1000 days data is Sufficient enough to train the various models for 1 day, 5 day, and 10 days ahead prediction. For 30 and 60 day ahead prediction, up to 2000/2500 days data is used to train the network. Model is tested with fresh 600 days data, out of which only 100 are shown for clarity.
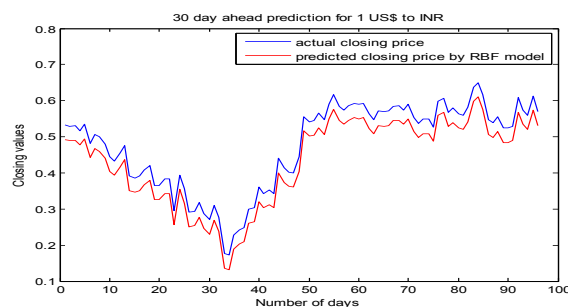
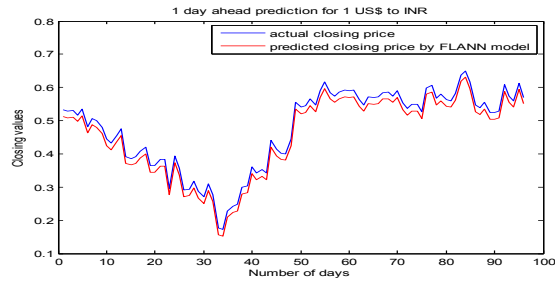Testing for RBF model**:**

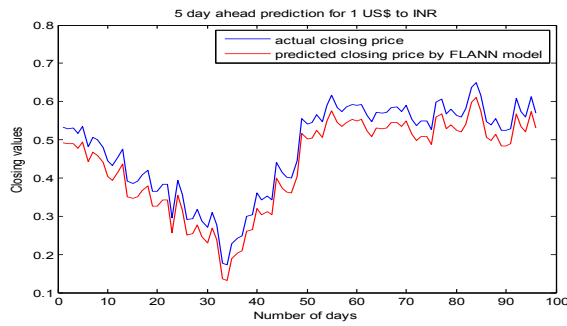For 1 day ahead prediction
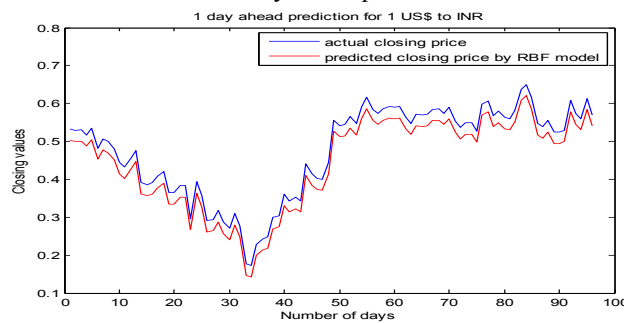
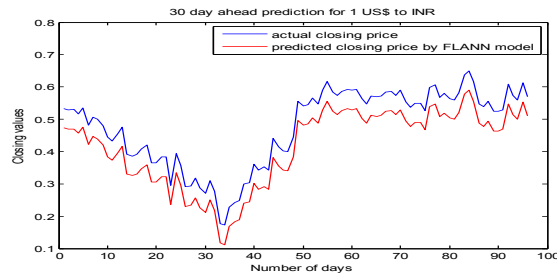For 5days ahead prediction
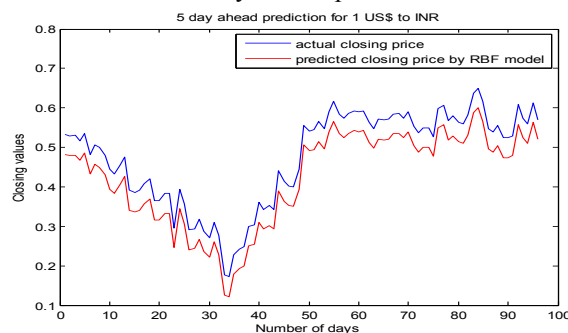
For 30 days ahead prediction

Testing for FLANN model:
For 1 days ahead prediction
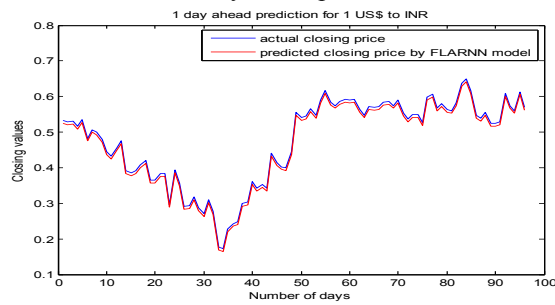
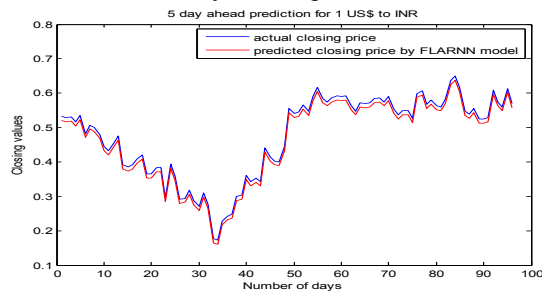For 5days ahead prediction

For 30 days ahead prediction
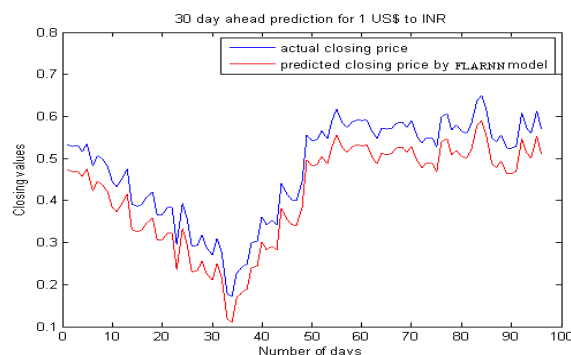

Testing of FLARNN model

For 1 day ahead prediction

For 5 days ahead prediction

For 30 days ahead prediction

Model Comparison via MAPE for 1US $ to Rupees

Ahead prediction	RBF MAPE	FLANN MAPE	FLARNN MAPE	Training done for
1 day	3.7736	2.349	**1.991**	1000 day
5 day	4.992	3.878	**2.897**	1000 day
10 day	6.145	**5.678**	4.543	1000day
	5.987	3.998	**3.212**	2000 day
30 day	6.783	**5.198**	7.327	1000day
	6.342	**5.02**	6.789	2000 day

Model Comparison via MAPE for **1 US$ to Yen**

Ahead prediction	RBF MAPE	FLANN MAPE(%)	FLARNN MAPE	Training done for
1 day	1.24	1.349	**0.991**	1000 day
5 day	3.7390	3.878	**2.197**	1000 day
10 day	5.87	6.437	**3.543**	1000day
	4.861	5.498	**2.212**	2000 day
30 day	6.25	**5.498**	6.7327	1000day
	5.89	**4.5**	5.289	2000 day
60 day	9.4478	**6.22**	7.7321	1000day
	8.8178	**5.227**	6.1569	2000 day

VI. Discussion on simulation results

Starting from the calculation of various parameters associated with the exchange rate data, in this paper we have formulated a comparison between various models such as RBF, FLANN and FLARNN and results are compiled. In our lucid simulations, we tested our algorithms on three types of exchanges rates namely 1 US$ to Rupees, pound, Yen. The results are compiled along with the response plots. LMS has the advantage of faster convergence but the probability of getting stuck in local optima is high.

VII. Conclusion

The Recurrent Functional Link Artificial Neural Network and Radial basis function exchange rate prediction model is introduced. With the use of FLARNN, FLANN and RBF, the model for prediction of exchange rate indices becomes simpler and involves lesser computations compared to other such model reported earlier. Experiments show that in case of lower number of ahead prediction, FLARNN parameters updated with LMS algorithm along with technical indicator gives the best result. While for higher number of days ahead prediction, RBF parameters updated with LMS algorithm works the best.

References

[1]. R.Baillie,P.McMahon, TheForeignExchangeMarket:Theory and Econo-metric Evidence,Cambridge,UK,1989.
[2]. A.Meese, K.Rogoff, Empiricalexchangeratemodelsoftheseventies: dothey fitoutofsample? ,Journal of International Economics14(1983)3–24.
[3]. L.Kilian, M.P.Taylor, Whyisitsodifficulttobeatrandomwalkforecastof exchangerates, Journal of International Economics60 (2003)85–107.
[4]. Y-H. Pao, "Adaptive Pattern Recognition & Neural Networks", Reading, MA;Addison-Wesley, 1989.
[5]. Y-H. Pao, "Adaptive Pattern Recognition & Neural Networks", Reading, MA; Addison-Wesley, 1989.
[6]. Y-H. Pao, S.M. Phillips, D.J. Sobajic," Neural Net Computing and intelligent controlsystems", Int. J. Contr. Vol. 56, No. 2, P 263-289, 1992
[7]. J.C. Patra, R.N. Pal, B.N. Chatterji, G. Panda, "Identification of non- linear & dynamicsystem using functional link artificial neural network ", IEEE Transactions on System, Man &Cybernetics – Part B; Cybernetics, Vol. 29, No. 2, April 1999.
[8]. Lee, P. and Haykin, S. (1999) 'A dynamic regularised radial basis function network for nonlinear,nonstationary time series prediction', IEEE Transactions on Signal Processing, Vol. 47, No. 9,**pp**.2503–2521.

[9]   Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986) 'Learning presentation byback-propagating errors', Nature, Vol. 323, pp.533–536.
[10]   Rao, S., Sethuraman, S. and Ramamurthi, V. (1992) 'A recurrent neural network for nonlinear time series prediction–a comparative study', IEEE 1992 Workshop on Neural Networks for Signal Processing (NNSP '92), Copenhagen, Denmark, pp.531–539.
[11].   J. Hanm, N. Kamber (2001), "Data Mining: Concepts & Techniques, San Francisco;Morgan Kanfmann Publishers.
[12].   E.W. Saad , D.V. Prokhorov, D.C. Wunsch (1998), " Comparative study of stock trendprediction using time delay, recurrent and probabilistic neural networks", IEEE Transactions of Neural Network, 9(6), 1456-1470.
[13]   H. Ni, H. Yin, "Exchange rate prediction using hybrid neural networks and trading indicators", Neurocomputing 72(2009) 2815-2823.
[14]   R. 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.
[15]   H Zhao and J Zhang, "Pipelined Chebyshev Functional link artificial   recurrent neural network for Non linear adaptive filter", IEEE Transaction on systems, man and cybernetics, Vol.40, NO, 1, February 2010.