

An Ensemble Model of Multiple Classifiers for Time Series Prediction

Dr. A. Chitra and S. Uma

Abstract—Time series forecasting is a challenging task in many fields. Due to the complex non-linear relationship between the multidimensional features of the time series data, improved time series forecasting requires a forecasting model that combines multiple prediction models. Ensemble learning performs better than single learning model and discovers regularities in dynamic and non-stationary data. In literature, single level neural network ensembles are used for the prediction problems [1- 5]. This paper introduces a novel two level ensemble learning approach based on Radial Basis Function networks (RBF), K - Nearest Neighbor (KNN) and Self Organizing Map (SOM) for time series prediction with the aim of increasing the prediction accuracy. The evaluation of the proposed Pattern Prediction Ensemble Model (PAPEM) using three input datasets such as, Mackey dataset, Sunspots dataset and Stock Price dataset shows that the proposed PAPEM model performs better than the individual classifiers.

Index Terms— Ensemble, KNN, PAPEM, RBF, SOM, Time Series Prediction.

I. INTRODUCTION

Time series analysis and prediction is an important task in all fields of science for applications like financial forecasting, weather forecasting, electricity power demand forecasting, process monitoring and control, research, medical sciences, etc., Artificial neural networks are widely used for solving pattern prediction problems[1-9]. Different types of neural networks exist and every neural network has its own benefits and limitations. Ensemble methods aim at leveraging the performance of a set of models to achieve better prediction accuracy than that of the individual models. In literature, single level ensemble models are used for prediction problems. Yet the prediction accuracy is limited. Combining multiple prediction models using two level ensembling will improve the prediction accuracy as the base learner itself is an ensemble. Hence, a hybrid Pattern Prediction Ensemble Model (PAPEM) using Radial basis function (RBF) networks, K- Nearest Network (KNN) and Self Organizing Maps (SOMs) is proposed.

II. RELATED WORKS AND MOTIVATIONS

From literature, it is seen that a number of pattern

recognition methods like neural networks, fuzzy logic, genetic algorithms, the moving average autoregressive exogenous (ARX) Auto-Regressive Moving Average (ARMA), Auto Regressive Integrated Moving Average (ARIMA), state space models, bilinear auto regressive (AR), Threshold Auto Regressive (TAR), KNN, SOMs, decision tree are available. Each method has its own advantages and disadvantages. Relying on the predictive power or the classifying power of a single pattern recognition method will not be successful always. Ensembling of neural networks was found to improve the generalization ability of neural network system [4]. Hence, combining the predicted results of a number of pattern recognition methods suitably will provide better results comparatively. Bagging, boosting are the popular ensemble methods used in literature [3,5]. Neural network ensembles are used in a variety of applications like face recognition, optical character recognition, financial decisions, etc.,

Considerable improvement in the classification of remote sensed images by using neural network ensembles was reported by K. Venugopal et al in [10] by studying the error diversity among the neural networks in an ensemble. It is said that fuzzy techniques can be used in combining the outputs of the neural networks in an ensemble.

The single layer nonlinear neural network ensemble model proposed in [1] used principle component analysis to determine the ensemble members.

In 2007, Bo Qian and Khaled Rasheed [5] examined Dow Jones Industrial Index using voting and stacking ensemble methods of back propagation neural networks, k-nearest neighbor and decision trees. Yet a prediction accuracy of only 65% was reported.

From literature, it was found that there are two main problems in the existing hybrid models and combined models namely, a linear combination approach is not necessarily appropriate for all the circumstances and it is not easy to determine the number of individual forecasting models. It was reported that, "The more, the best" rule cannot be applied effectively for all circumstances [11]. The number of individual models to be combined for forecasting has to be determined.

Hence, an alternative two level ensemble is proposed in this paper which is both diverse and accurate as the base learner itself is an ensemble.

III. EXISTING SYSTEMS

Several types of classifiers are available for pattern prediction. RBF, KNN and SOM are considered as the best methods for non-linear time series analysis and prediction.

Manuscript received October 24, 2009.

F. Dr. A. Chitra is with PSG College of Technology, Coimbatore 641 004, Tamil Nadu, INDIA as the professor of CSE Department. (phone : 98432 22273, e-mail: ac_psg@yahoo.com).

S. S. Uma is pursuing Doctoral Degree in PSG College of Technology, under Anna University, Coimbatore, Tamil Nadu, INDIA. (phone: 94439 13517; e-mail: umakaruna19@yahoo.com).

Hence, the PAPEM model is designed as a two level ensemble model of RBF, KNN and SOM. In literature, these classifiers are used mostly as standalone models for pattern prediction. A brief review of these classifiers is discussed in this section.

A. Radial Basis Function (RBF) Networks

RBF is a kind of artificial neural network that uses radial basis functions as activation functions (figure 1). Localization, function approximation, quasi-orthogonality, cluster modeling and interpolation are some of the potential properties of RBF networks that make it suitable for a wide variety of applications. It is best suited for time series prediction and analysis. It can learn using both positive and negative samples. It finds the input to output map using local approximators. Usually, the supervised segment is simply a linear combination of the approximators. These networks train extremely fast and require fewer training samples as linear combiners have few weights.

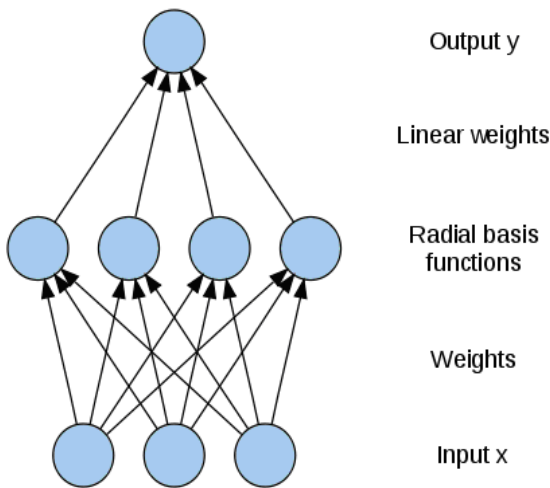


Figure 1. Radial Basis Function Neural Network

Though RBF has numerous advantages, it has some limitations like difficulty in using Gaussian functions to approximate constant values if a function has nearly constant values in some intervals and the classification error strongly depends on the selection of the number, centers and widths of basis functions[12-15].

B. K-Nearest Neighbor (KNN)

K-nearest neighbor is the fundamental and simplest classification technique when there is little or no prior knowledge about the distribution of the data. An example for the k-nearest neighbor classification is given in figure 2. The classification process is transparent, easy to implement and debug. KNN has the potential to generate precise local estimates and maps of the attributes. In contexts where an explanation of the output of the classifiers is useful, KNN can be very effective as the analysis of the neighbors can be given as an explanation. Some noise reduction techniques work only for KNN that can be effective in improving the accuracy of the classifier.

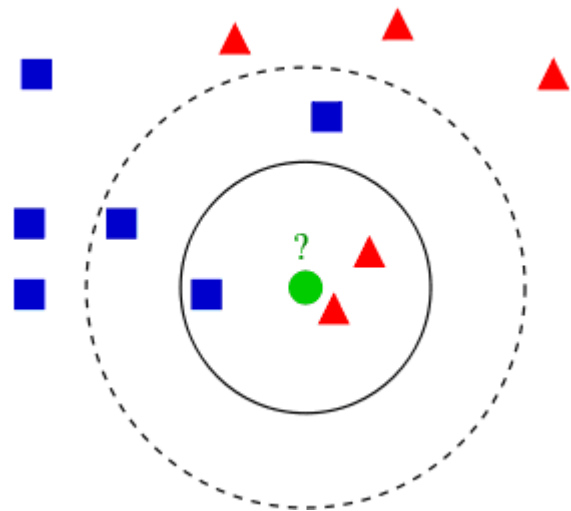


Figure 2. K-Nearest Neighbor Classification – an example

Despite the advantages given above, KNN has a few limitations. KNN can have poor run-time performance when the training set is large. It is very sensitive to irrelevant or redundant features because all features contribute to the similarity and thus to the classification. By careful feature selection or feature weighting, this can be avoided.[16-17]

C. Self Organizing Map (SOM)

Self Organizing Map is an unsupervised learning method in which training is entirely data driven and no information about the input data is required. A simple SOM network is given in figure 3. SOMs are good at non-linear time series prediction due to its topological preservation and easy visualization features. It is very easy to understand and very simple. It has good classifying ability. It can be easily evaluated to find how good a map is and how strong the similarities between the objects are.

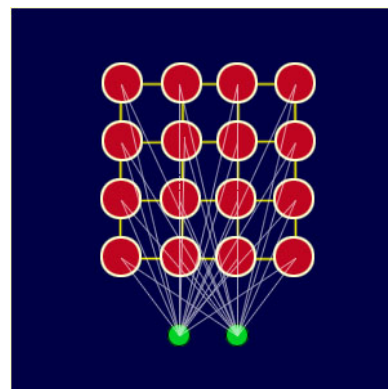


Figure 3. A Simple SOM Network

SOM has a few limitations like difficulty in generating a map if some data is missing. Every SOM is different and finds different similarities among the sample vectors. SOMs organize sample data so that in the final product, the samples are usually surrounded by similar samples. However, similar samples are not always near each other. So a lot of maps need to be constructed in order to get one final good map[18-20].

To overcome the limitations of the existing system and to take advantage of the benefits of the classifiers RBF, KNN and SOM, PAPEM, a two level ensembling model of radial

basis function networks, k nearest neighbor and self organizing map is done to discover regularities in dynamic and non-stationary data and increase the prediction accuracy compared to the standalone and single level ensemble learning models.

IV. PATTERN PREDICTION ENSEMBLE MODEL (PAPEM)

The block diagram of the PAPEM model is given in figure 4. In the PAPEM model, three types of classifiers namely radial basis function neural networks, k-nearest neighbor and self organizing maps are used. Ten diverse models of each of these classifiers are generated by varying the architecture. The number of hidden layers is varied in radial basis function neural network to obtain diverse models. For KNN classifier the number of nearest neighbors is varied to get ten diverse models and in self organizing map classifier the map size is varied to get ten diverse models.

In the first level, heterogeneous ensemble learning of different configurations of RBF, KNN Classifier and SOM classifiers are done and the best rules of these individual classifiers are used to train in the second level. The best rule out of the second level training is used for final pattern prediction. Out of the dataset taken for pattern prediction, 75% of the data is used for training and 25% is used for testing.

In this study, the pattern prediction of the benchmark time series data sets namely the Mackey glass time series data, sunspot data and stock data is examined. The Mackey glass time series data is generated using the following nonlinear differential equation (1) :

$$\frac{dx(t)}{dt} = -0.1 * x(t) + \frac{0.2 * x(t - \tau)}{1 + x^{10}(t - \tau)} \quad (1)$$

The values a=0.2, b=0.1 and x(0)=0.9 are assumed for the test bench and 3000 data are generated for the experiment. The rate of change of Mackey Glass function at a given time 't' depends on both the value function at that time t and at an older point in time (t-τ).

The sunspot time series data contains the average number of sunspots per month measured from January 1749 to July 2009. These data are available at <http://sidc.oma.be> ('RWC Belgium World Data Center for the Sunspot'). The sunspot time series data is chaotic and has local behaviors, noise and even unpredictable zones using the archived knowledge.

A sunspot is a region on the sun's surface called the photosphere, which is temporarily cool and dark compared to surrounding regions. The average surface temperature of the sun is 6000° Celsius and that the sunspots are about 1500° Celsius cooler than the area surrounding them, and can

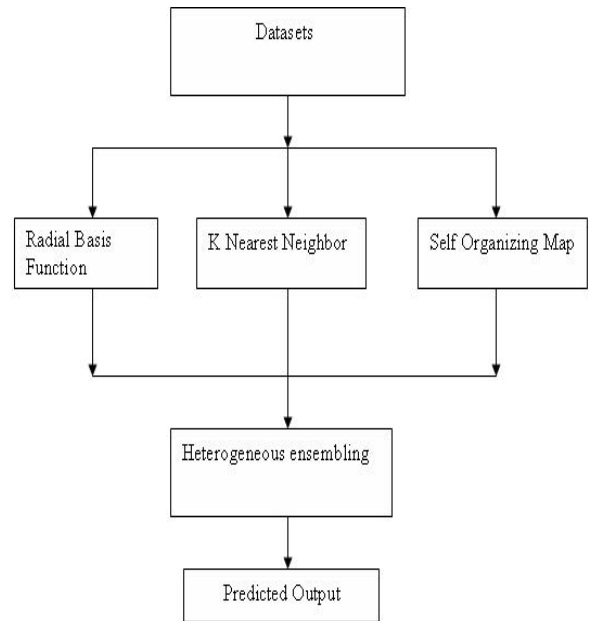


Figure 4. PAPEM Model

last from a few hours to a few months. Sunspots expand and contract as they move across the surface of the sun and can be as large as 80, 000 km in diameter. Sunspots are linked with other kinds of solar activity [21]. Hence, it can be used to predict the space weather and the state of the ionosphere. Thus sunspots are useful in the prediction of conditions of short-wave propagation or satellite communications.

The Mackey glass time series and sunspot time series are widely used by researchers for comparing the generalization ability of different models.

The generalization ability and the prediction accuracy of the PAPEM model was examined using the data set of several stocks. The experimentation results of the stock IBEX is discussed in this paper as an example. The stock data of IBEX is taken from 14th Jan 1992 to till date for the experimentation. Out of these patterns the first 75% of the patterns is used for developing models and the last 25% of patterns are used as out-of-data samples to test the performance of the final models. Four fold cross validation is run for ten times for each classifier. Out of the ten diverse models of each of these classifiers, the best models is selected for training the ensemble. In the second level the base learner itself is an ensemble. The performance of the classifiers are measured using Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Normalized Mean Square Error (NRMSE) which are calculated as given in the equations 2-4.

RMSE is a good measure of prediction accuracy. It is frequently used to measure the differences between values predicted by a model or an estimator and the values actually observed from the thing being modeled or estimated. These individual differences are also called residuals.

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2} \quad (2)$$

Mean Absolute Percentage Error (MAPE) is the measure of accuracy in a fitted time series value in statistics,

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (3)$$

where A_t is the actual value and F_t is the forecast value and n is the number of observations.

The normalized root mean squared error is the RMSE divided by the range of observed values, or:

$$NRMSE = \frac{RMSE}{x_{\max} - x_{\min}} \quad (4)$$

the value is often expressed as a percentage, where lower values indicate less residual variance.

Experiments were conducted using RBF, KNN, SOM, single level RBF ensemble, single level KNN ensemble, single level SOM ensemble and the proposed PAPEM classifiers on Mackey Glass, Sunspot and stock datasets. The results of the experimentation are given in section V.

V. EXPERIMENTAL RESULTS

A comparison of the root mean square error, mean absolute percentage error, normalized root mean square error for the classifiers RBF, KNN, SOM, single level RBF ensemble, single level KNN ensemble, single level SOM ensemble and the proposed PAPEM model for the benchmark non linear time series datasets Mackey glass, Sunspot time series and stock data are shown in the figure 5-7 respectively. A comparison of the prediction accuracy and the percentage change in prediction accuracy of these classifiers with the PAPEM are give in the charts of figure 8 and figure 9 respectively. The prediction accuracy of these classifiers and the PAPEM model is given in table I.

The experimental results have proved that the PAPEM model is better than the other standalone classifiers and single level ensembled classifiers obtained using RBF, KNN and SOM.

TABLE I. COMPARISON OF THE PREDICTION ACCURACY OF DIFFERENT CLASSIFIERS

Classifiers / Models Used	Prediction Accuracy (%)		
	Mackey Glass	Sunspot	Stock
RBF	82	78	78
KNN	61	51	60
SOM	81	80	77
RBF Ensemble	83	85	82
KNN Ensemble	69	54	64
SOM Ensemble	85	82	81
PAPEM	89	91	88

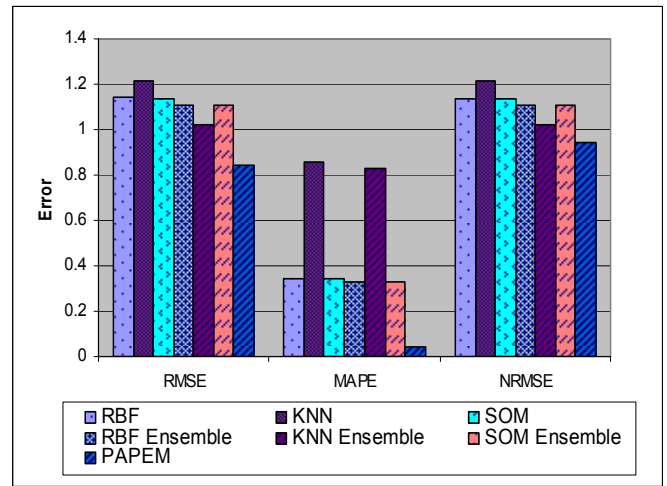


Figure 5. Performance Comparison of Different Classifiers with PAPEM for Mackey Dataset

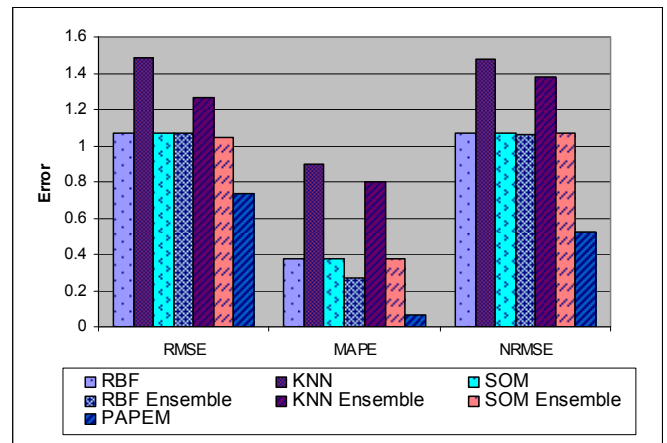


Figure 6. Performance Comparison of Different Classifiers with PAPEM for Sunspot Dataset

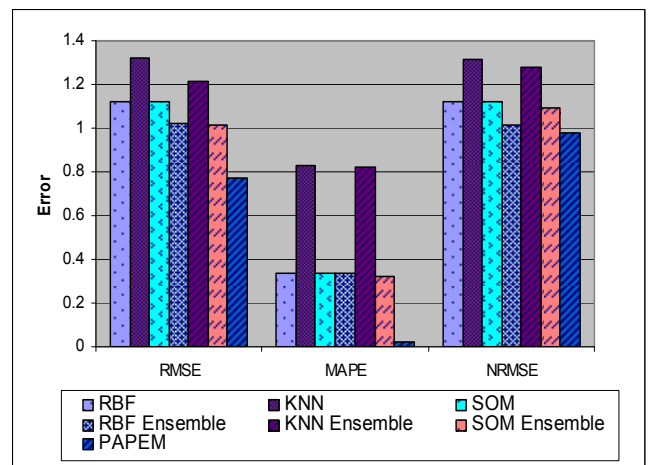


Figure 7. Performance Comparison of Different Classifiers with PAPEM for Stock Dataset

Sunspots are useful in the prediction of conditions of short-wave propagation or satellite communications that benefits a larger population. A larger mass of people are actively involved in stock trading. Prediction of stock movements increases the profits even in a chaotic market. Thus, the increase in prediction accuracy of the PAPEM model benefits a larger population. The PAPEM model can

be used for the pattern prediction in all kinds of non-linear time series data.

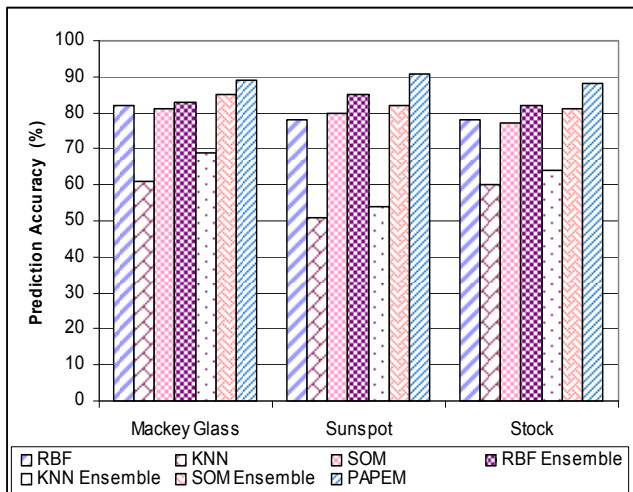


Figure 8. Comparison of prediction accuracy of different classifiers with PAPEM

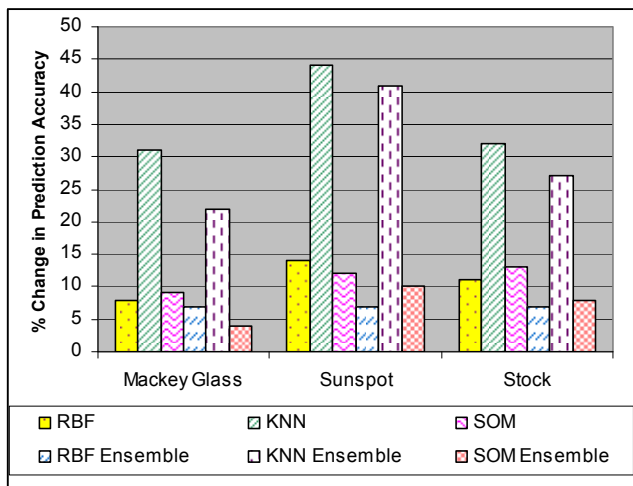


Figure 9. Comparison of percentage change in prediction accuracy of different classifiers with PAPEM

VI. CONCLUSION

The proposed model, PAPEM is designed for pattern prediction for non-linear time series applications like financial forecasting, Sunspot, Mackey-Glass, electricity demand forecasting, commodity and forex trading. Experimentation of the model is done for pattern prediction of Mackey Glass, Sunspot and a Stock data. The comparison of the generalization ability of the various classifiers is done using root mean square error, mean absolute percentage error, normalized mean square error and prediction accuracy. The results show that the PAPEM model is better than the other standalone classifiers like RBF, KNN, SOM and single level ensembled classifiers of RBF, KNN, SOM. The prediction accuracy of PAPEM is 91%. It is experimentally studied that it outperforms the other models.

REFERENCES

- [1] Kin Keung Lai, Lean Yu, Shouyang Wang, and Huang Wei, "A Novel Nonlinear Neural Network Ensemble Model for Financial Time Series Forecasting", International Conference on Computational Science (1) 790-793, 2006.
- [2] LeanYua B, ShouyangWanga, Laic K.K, "A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange

- rates", www.sciencedirect.com, Elsevier . Computers & OR 32 2523-2541, 2005.
- [3] Zhuo Zheng, "Boosting and Bagging of Neural Networks with Applications to Financial Time Series", 2006.
- [4]
- [5] Hansen L. K and Salamon P, "Neural network ensembles," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, no. 10, pp. 993-1001, 1990.
- [6] Bo Qian, Khaled Rasheed, "Stock market prediction with multiple classifiers", Applied Intelligence, Vol. 26, No. 1, pp. 25-33, February 2007.
- [7] Amir Atiya, Noha Talaat and Samir Shaheen, "An Efficient Stock Market Forecasting Model using Neural Networks", IEEE, 1997.
- [8] Antonio Galaria-Bengochea, Claudio Ordonz-Ureta, Manuel Marchant-Saavedra, Nelson Opazo-Medina, "Stock Market Indices in Santiago-de Chile : Forecasting using neural Networks", IEEE, 1996.
- [9] Atsalakis George, Skiadas Christos, Nezis Dimitrios, "Forecasting Chaotic time series by a Neural Network", Chaotic Modelling and Simulation International Conference, June 3-6, Chania Crete, Greece, 2008.
- [10] Carl G. Looney, "Pattern Recognition Using Neural Networks", Oxford University Press, 1997.
- [11] Venu Gopala Rao K, Prem Chand Pand, Ramana Murthy M. V, "Soft Computing-Neural Networks Ensembles", Journal of Theoretical and Applied Information Technology, JATIT 2005 - 2007.
- [12] LeanYua B, ShouyangWanga, Laic K.K, "A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates", www.sciencedirect.com, Elsevier, Computers & OR 32: 2523-2541, 2005
- [13] Ding Tao , Xiao Hongfei, "Chaotic Time Series Prediction Based on Radial Basis Function Network", Proceedings of the Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, p.595-599, July 30-August 01, 2007
- [14] Chng, E. S., Chen, S. and Mulgrew, B, "Gradient radial basis function networks for nonlinear and nonstationary time series prediction", IEEE Transactions on Neural Networks, 7 (1). pp. 190-194. ISSN 1045-9227
- [15] Habib Dhahri, and Adel. M. Alimi, "The Modified Differential Evolution and the RBF (MDE-RBF) Neural Network for Time series Prediction," In2006 Int. Joint Conference on Neural Networks, Vancouver, BC, Canada, July 16-21, pp. 2938-2943, 2006.
- [16] Ales Leonardis, Horst Bischof, "An efficient MDL-based construction of RBF networks", Neural Networks, Vol. 11, No.5, pp. 963-973, 1998
- [17] Padraig Cunningham and Sarah Jane Delany, "K-Nearest Neighbour Classifiers", Technical Report UCD-CSI-2007-4, March 27, 2007
- [18] J.-H. Lee, G.-H. Cha, and C.-W. Chung, "A Model for k-Nearest Neighbor Query Processing Cost in Multidimensional Data Space", Information Processing Letters, Vol. 69, pp. 69-76, 1999.
- [19] A. Lendasse, J. Lee, V. Wertz, M. Verleysen, "Forecasting electricity consumption using nonlinear projection and self-organizing maps", Neurocomputing 48 (2002) 299-311
- [20] Barreto, G.A. Araujo, A.F.R., " Identification and control of dynamical systems using the self-organizing map" IEEE Transactions on Neural Networks, Sept. 2004, Volume: 15, Issue: 5, 1244- 1259
- [21] Yudong Chen, Yi Zhang, and Jianming Hu, "Multi-Dimensional Traffic Flow Time Series Analysis with Self-Organizing Maps ", Tsinghua University Press Published by Elsevier B.V, 2008
- [22] <http://en.wikipedia.org/wiki/Sunspot>

Dr. A. Chitra is Professor, Department of Computer Science and Engineering, PSG College of Technology, Coimbatore, Tamil Nadu, INDIA. She holds a Ph.D., in Computer Science and Engineering and is actively involved in research and teaching. She has published prolifically and her books include programming with C and Internet and Java Programming. She has published around 75 papers in national / international journals / conferences. She is a Fellow member of Institution of Engineers, Life member of ISTE, ACS and CSI. She is a recipient of Tamil Nadu Young women Scientist Award, ISTE National Award for Outstanding Academician and IETE Award for the best paper published in IETE journals.

S.Uma is Assistant Professor, Department of Information Technology, Hindusthan College of Engineering and Technology, Coimbatore. She is pursuing Doctoral degree under the guidance of Dr. A. Chitra in Anna University, Coimbatore, Tamil Nadu, INDIA. She is a life member of Indian Society for Technical Education and a member of Computer Society of India.