

FUZZY ARTMAP NEURAL NETWORK ARCHITECTURE FOR WEATHER FORECASTING

A.C. SUBHAJINI¹ AND T. SANTHANAM²

¹Department of Computer Applications, Noorul Islam University,
Kumaracoil, Kanyakumari, India

²Department of Computer Science, D.G. Vaishnav College, Chennai, India
E-mail: acsubhajini@yahoo.com

ABSTRACT

Accurate weather forecasting plays a vital role for planning day to day activities. Neural network has been used in numerous meteorological applications including weather forecasting. A neural network model has been developed for weather forecasting, based on various factors obtained from meteorological experts. The back propagation neural network, radial basis function neural network, Generalized Regression Neural Networks and Fuzzy ARTMAP architecture are used in this study to test the performance in order to investigate which technique for weather forecasting is most effective. The results indicate that proposed Fuzzy ARTMAP Neural Networks is better than back propagation neural network, Radial Basis Function and Regression neural network. The prediction accuracy of Fuzzy ARTMAP is 97.95%.

Keywords: *Weather forecasting, Radial Basis Function, Back Propagation Neural, Network, Regression Neural Networks, Fuzzy ARTMAP*

1. INTRODUCTION

Weather simply refers to the condition of air on earth at a given place and time. The application of science and technology to predict the state of the atmosphere in future time for a given location is so important due to its effectiveness in human life [4]. Today, weather forecasts are made by collecting quantitative data about the current state of the atmosphere and using scientific understanding of atmospheric processes to project how the atmosphere will evolve. The chaotic nature of the atmosphere implies the need of massive computational power required to solve the equations that describe the atmospheric conditions. This is resulted from incomplete understanding of atmospheric processes which mean that forecasts become less accurate as the difference in time between the present moment and the time for which the forecast is being made increases. Weather is a continuous, data-intensive, multidimensional, dynamic and chaotic process and these properties make weather forecasting a big challenge. Generally, two methods are used for weather forecasting (a) the empirical approach and (b) the dynamical approach. The first approach is based on the occurrence of analogs and is often referred to by meteorologists as analog forecasting. This approach is useful for predicting local-scale weather if recorded data are plentiful. The second approach is based on

equations and forward simulations of the atmosphere and is often referred to as computer modeling. The dynamical approach is only useful for modeling large-scale weather phenomena and may not predict short-term weather efficiently. Most weather prediction systems use a combination of empirical and dynamical techniques [10].

Artificial Neural Network (ANN) provides a methodology for solving many types of nonlinear problems that are difficult to be solved by traditional techniques (Shereef 2010). Most meteorological processes often exhibit temporal and spatial variability. They are suffered by issues of nonlinearity of physical processes, conflicting spatial and temporal scale and uncertainty in parameter estimates. The ANN has capability to extract the relationship between the inputs and outputs of a process, without the physics being explicitly provided [13]. Thus, these properties of ANN are well suited to the problem of weather forecasting [1]. The main purpose is to develop the most suitable ANN architecture and its associated training technique for weather prediction. Back propagation feed forward network, radial basis function network, regression neural network and Fuzzy ARTMAP which were trained by differential algorithm are the selected architectures in this study. The basic architecture of the both neural networks are given. These neural network

architectures are used as a prediction tools for the weather forecasting.

2. WEATHER FORECASTING SYSTEM STRUCTURE

Components of a modern weather forecasting system include the following modules: data collection, data assimilation and numerical weather prediction.

2.1 Data collection: Observations of atmospheric pressure, temperature, wind speed, wind direction, humidity and precipitation are made near the earth's surface by trained observers, automatic weather stations. The World Meteorological Organization acts to standardize the instrumentation, observing practices and timing of these observations worldwide.

2.2 Data assimilation: During the data assimilation process, information gained from the observations is used in conjunction with a numerical model most recent forecast for the time that observations were made to produce the meteorological analysis. This is the best estimate of the current state of the atmosphere. The features considered in this study are bar temperature, bar reading, sea level pressure, mean sea level pressure, dry bulb temperature, wet bulb temperature, dew point temperature, vapor pressure, wind speed, humidity, cloudiness, precipitation, wind direction, wind speed and for prediction of rain. It is easy to implement and produces desirable forecasting result by training the given data set.

2.3 Numerical weather prediction: Numerical Weather Prediction (NWP) uses the power of computers to make a forecast. Complex computer programs, also known as forecast models, run on supercomputers and provide predictions on many atmospheric variables such as temperature, pressure, wind and rainfall. A forecaster examines how the features predicted by the computer will interact to produce the day's weather.

3. MATERIALS AND METHODS

ANN is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) to solve specific problems. Learning in biological

systems involves adjustments to the synaptic connections that exist between the neurons. Neural networks have broad applicability to real world problems. In fact, they have already been successfully applied in many industries [5].

3.1 Back propagation neural network

A typical neural network consists of layers. In a single layered network there is an input layer of source nodes and an output layer of neurons. A multi-layer network has in addition one or more hidden layers. A multi layer neural network is displayed in Fig. 1. Extra hidden neurons raise the network's ability to extract higher-order statistics from (input) data. Furthermore a network is said to be fully connected if every node in each layer of the network is connected to every other node in the adjacent forward layer [8]. The network "learns" by adjusting the interconnections (called weights) between layers. When the network is adequately trained, it is able to generalize relevant output for a set of input data. A valuable property of neural networks is that of generalization, whereby a trained neural network is able to provide a correct matching in the form of output data for a set of previously unseen input data.

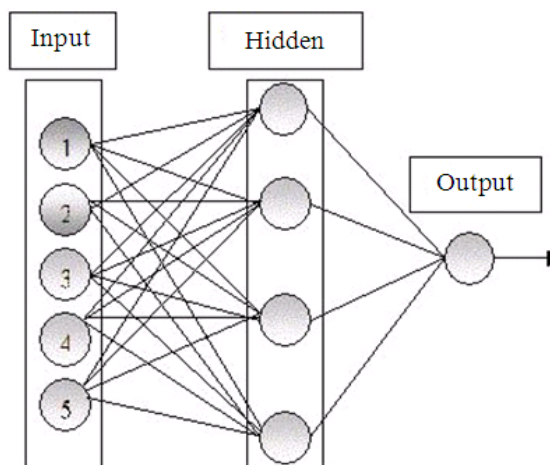


Fig. 1: Neural network architecture

Learning typically occurs by example through training, where the training algorithm iteratively adjusts the connection weights (synapses). Back propagation is one of the most famous training algorithms for multilayer perceptions [3]. BP is a gradient descent technique to minimize the error E for a particular training pattern. For adjusting the weight w_{ij} from the i -th input unit to the j -th output, in the batched mode variant the descent is based on

the gradient $\nabla E \left(\frac{\delta E}{\delta w_{ij}} \right)$ for the total training set:

$$\Delta w_{ij} = -\Sigma^* + \alpha * \Delta w_{ij} (n-1) \quad (1)$$

The gradient gives the direction of error E. The parameters ϵ and α are the learning rate and momentum respectively [3]. Normally, the learning rate is held constant throughout the training. If the learning rate is too high, the algorithm may oscillate and become unstable. If the learning rate is too small, the algorithm is slow to converge. The performance of the steepest descent algorithm can be improved by using an adaptive learning rate, which will keep the learning step size as large as possible while keeping learning stable. The learning rate is made adaptive to the complexity of the local error surface. If the new error exceeds the old error by more than a predefined ratio, the new weights are discarded. In addition, the learning rate is decreased. Otherwise the new weights are kept. If the new error is less than the old error, the learning rate is increased. This architecture was used in the field of forecasting [9].

3.2 Basic architecture of RBF

A Radial Basis Function (RBF) network is a special type of neural network that uses a radial basis function as its activation function. RBF networks are very popular for function approximation, curve fitting, time series prediction, control and classification problems [11]. In RBF networks, determination of the number of neurons in the hidden layer is very important because it affects the network complexity and the generalizing capability of the network. In the hidden layer, each neuron has an activation function. The Gaussian function, which has a spread parameter that controls the behavior of the function, is the most preferred activation function [2]. The training procedure of RBF networks also includes the optimization of spread parameters of each neuron. Afterwards, the weights between the hidden layer and the output layer must be selected appropriately. Finally, the bias values which are added with each output are determined in the RBF network training procedure. RBF network is a type of feed forward neural network composed of three layers, namely the input layer, the hidden layer and the output layer. A general block diagram of an RBF network is illustrated in Fig. 2.

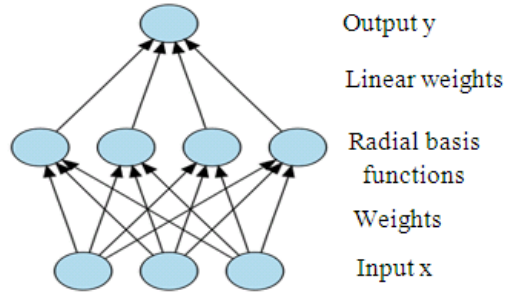


Fig. 2: Block diagram of a RBF network

A Radial basis function network with m outputs and hidden nodes can be expressed as:

$$y_i(t) = w_{io} + \sum_{j=1}^{n_h} w_{ij} \Phi(\|v(t) - c_j(t)\|) \quad ; i = 1, \dots, m \quad (2)$$

Considering this argument, the RBF network with additional linear input connections is used. This network allows the inputs to be connected directly to the output node via weighted connections to form a linear model in parallel with the non-linear standard RBF model [7]. The new RBF network with m outputs, n inputs, hidden nodes and n_1 linear input connections can be expressed as:

$$y_i(t) = w_{io} + \sum_{j=1}^{n_h} \lambda_{ij} v_l(t) + \sum_{j=1}^{n_h} w_{ij} \Phi(\|v(t) - c_j(t)\|) \quad (3)$$

$i = 1, 2, \dots, m$

where the λ 's and v_l 's are the weights and the input vector for the linear connections may consist of past inputs and outputs [7].

3.3 General Regression Neural Network

Robert Hecht-Nielsen developed the Generalized regression neural networks. The General Regression Neural Network (GRNN) is one of the most popular neural networks. The GRNN works quite accurately with prediction of weather forecasting. The GRNN infrastructure consists of four layers: input, hidden, summation and output layer. The basic GRNN comprising of an input, hidden, summation and output layer. A general block diagram of an GRNN network is illustrated in Fig. 3.

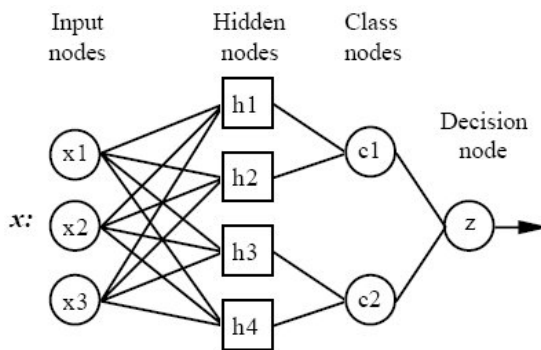


Fig. 3: Block diagram of a GRNN network

-> GRNN is a non iterative neural network that is data flows from input layer to the output layer so it consumes less memory than the iterative neural networks like back propagation neural networks etc.

-> These qualities make GRNN very feasible for real time applications such as weather forecasting.

Steps involved in working of GRNN

Step 1: Input layer — There is one neuron in the input layer for each predictor variable. In the case of categorical variables, $N-1$ neurons are used where N is the number of categories. The input neurons (or processing before the input layer) standardizes the range of the values by subtracting the median and dividing by the interquartile range. The input neurons then feed the values to each of the neurons in the hidden layer.

Step 2: Hidden layer — This layer has one neuron for each case in the training data set. The neuron stores the values of the predictor variables for the case along with the target value. When presented with the x vector of input values from the input layer, a hidden neuron computes the Euclidean distance of the test case from the neuron's center point and then applies sigma value(s). The resulting value is passed to the neurons in the pattern layer..

Step 3: Pattern layer / Summation layer — For GRNN networks, there are only two neurons in the pattern layer. One neuron is the denominator summation unit the other is the numerator summation unit. The denominator summation unit adds up the weight values coming from each of the hidden neurons. The numerator summation unit

adds up the weight values multiplied by the actual target value for each hidden neuron.

Step 4: the decision layer divides the value accumulated in the numerator summation unit by the value in the denominator summation unit and uses the result as the predicted target value.

4. FUZZY ARTMAP

Fuzzy ARTMAP is a supervised neural network learning algorithm that has many desirable properties. Fuzzy ARTMAP can be used both for classification as for function approximation.

Fuzzy ARTMAP, called Fast amplified Fuzzy ARTMAP [16], that reduce some of the redundancies of Simplified Fuzzy ARTMAP and speed up its convergence to a solution. A general block diagram of Fuzzy ARTMAP Neural Network is illustrated in Fig. 4. Fuzzy ARTMAP Neural Network uses the similarity of fuzzy subset calculation and adaptive resonance dynamic type selection method through feature space search and learning in form, and effectively integrates the merits of fuzzy logic and Adaptive Resonance Theory. The basic structure of fuzzy ARTMAP neural network should include two modules-ARTa and ARTb. Each module can be divided into two subsystems in function, namely, attention and orientation subsystem. Attention subsystem is used to deal with the mode that has been studied, and establish precise internal representation and fine tuning for familiar mode. Orientation subsystem is used to deal with the new emerging mode. When unfamiliar events emerge, it calls back attention subsystem and interiorly establishes new internal codes to represent the unfamiliar events. This is substantially the same as the principle of ART-1 and ART-2, while the difference is the addition of fuzzy processing capability.

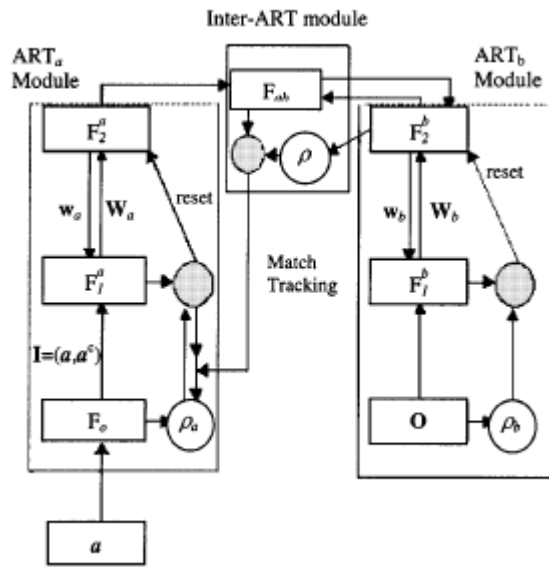


Fig. 4: Block diagram of a Fuzzy ARTMAP network

4.1The Fuzzy ARTMAP learning algorithm

- (1) Initialize all weights.
- (2) Present input at ARTa and target output at ARTb. Allow category formation in both modules. The F2 layer nodes in ARTb encode the class information.
- (3) Get predicted output class information from ARTa by using the map field weights w_{ab} .
- (4) Compare predicted class with actual class information. If they are the same, go to step 6.
- (5) Reset current representation in ARTa and search for a better representation. Repeat steps 4 and 5 until a correct match is found.
- (6) Update all weights in both modules.
- (7) Go to step 2.

5. RESULTS and DISCUSSION

To experiment the proposed system a sample dataset is taken from metrological department. These data sets contain real time observation of the weather for a particular period of time. For this experiment, an observation of the complete previous ten year of data is collected from meteorological department, Kanyakumari District. The data set contain many attribute. The basic

input data for classification needs preprocessing and the above attributes are processed for weather forecasting using BPN,RBF,GRNN and Fuzzy ARTMAP. Improvement of classification accuracy in weather forecasting is an important issue. The factors temperature, air pressure, humidity, cloudiness, precipitation, wind direction wind speed of weather forecasting are consolidated from meteorological experts. Fuzzy ARTMAP are trained with samples and outputs are namely no rain and rain.

BPN,RBF,GRNN and Fuzzy ARTMAP are trained with sample of thousand patterns. The performance of the Fuzzy ARTMAP is compared with performance of the BPN , RBF and GRNN for weather forecasting. The training of Fuzzy ARTMAP is faster compared that of BPN, RBF and GRNN. The classification to predict rainfall is enhanced by GRNN which is depicted in Table 1. Performance of classification of rainfall prediction by BPN , RBF, GRNN and Fuzzy ARTMAP is shown in Fig. 5 and 6.

Table 1:Percentage of classification of rainfall by BPN, RBF ,GRNN and Fuzzy ARTMAP

		Percentage of correct classification (%)		
-----		-----		
Weather Forecasting	No of patterns	BPN	RBF	Fuzzy GRNN
ARTMAP				
Rain	1000	80.30	86.60	97.40
97.50				
No rain	1000	83.60	90.30	96.20
98.40				
Overall Percentage (%)		81.90	88.40	96.80
		97.95		

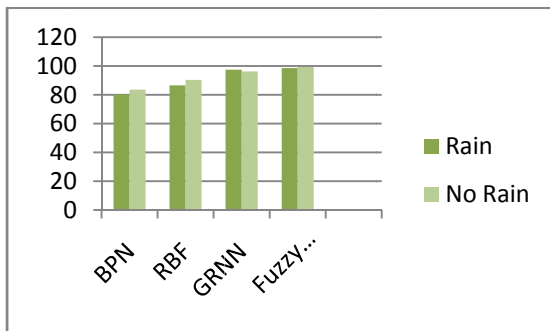


Fig. 5: Performance classification of rainfall prediction by BPN ,RBF, GRNN and Fuzzy ARTMAP

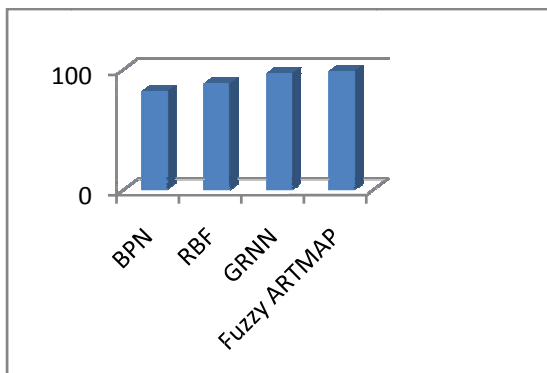


Fig. 6: Overall percentage of classification by BPN,RBF, GRNN and Fuzzy ARTMAP

6.CONCLUSION

Neural network has gained great popularity in weather prediction because of their simplicity and robustness. In this study, the performance of back propagation neural network (BPN), radical basis functioned neural network (RBF), generalized regression neural network (GRNN) and Fuzzy ARTMAP is compared. Back propagation algorithm, radical basis functioned neural network and regression neural network are too time consuming and the performance is heavily dependent on the network parameters. Compared to BPN, RBF and GRNN, Fuzzy ARTMAP gives the overall best results in terms of accuracy and fastest training time. Fuzzy ARTMAP are much faster and more reliable for the weather forecasting. These proportions make it more effective for fast real time weather forecasting.

REFERENCES

- [1] Abd, M.K., 2009. Electricity load forecasting based on frameset neural network technique. *Am. J. Applied Sci.*, 6: 970-973. DOI: 0.3844/ajassp.2009.970.973
- [2] Ajeel, S.A., 2010. A novel carbon steel pipe protection based on radial basis function neural network. *Am. J. Applied Sci.*, 7: 248-251.
- [3] AL-Allaf, O.N.A., 2010. Improving the performance of back propagation neural network algorithm for image compression/decompression system. *J. Comput. Sci.*, 6: 1347-1354.
- [4] Cheng, C.S., G.L., Q. Li and H. Auld, 2010. A synoptic weather typing approach to simulate daily rainfall and extremes in Ontario, Canada: Potential for climate change projections. *J. Applied Meteorol. Climatol.*, 49: 5: 845- 866.
- [5] Dastorani, M.T., H. Afkhami, H. Sharifidar and M. Dastorani, 2010. Application of ANN and ANFIS models on dryland precipitation prediction. *J. Applied Sci.*, 10: 2387-2394.
- [6] Ghods, L. and M. Kalantar, 2010. Long-term peak demand forecasting by using radial basis function neural networks. *Iranian J. Elect. Eng.*, 6.
- [7] Goriz, M. and T. Punithori, 2011. A new model for time-series forecasting using radial basis functions. *Int. J.*, 181: 903-916.
- [8] Khan, M.S. and P. Coulibaly, 2010, M.S. and P. Coulibaly, 2010. Assessing hydrologic impact of climate change with uncertainty estimates: Back propagation neural network approach. *J. Hydrometeorol.*, 11: 482-495.
- [9] Luo, L. and L. Zhou, 2010. Application of radial basis function neural network in modeling wastewater sludge recycle system. *Commun. Comput. Inform. Sci.*, 98: 117-122.
- [10] Moro, Q.I., L. Alonso and C.E. Vivaracho, 2011. Application of neural networks to weather forecasting with local data. *Proceedings of the 12 IASTED International Conf Applied Informatics, Liverpool, United Kingdom*, pp: 68-70.
- [11] Park, J. and I. Sandberg, 2008. Universal approximation using radial basis function networks. *Neural Comput.*, 3: 246-257.
- [12] Shereef, I.K. and S. Baboo, 2010. An efficient weather forecasting system using artificial neural network. *Int. J. Environ. Sci. Dev.*, 1: 248-251



[13]Veisi, H. and M. Jamzad, 2009. A complexity-based approach in image compression using Neural networks. *Int. J. Sign. Process.*, 5: 82-92.

[14]LeszekRutkowski,, Generalized Regression Neural Networks inaweather forecasting, *IEEE Transaction Neural Network*VOL. 15, NO. 3,MAY 2008

[15]ReceipDiizgiion,Generalized Regression Neural networks for InflationForecasting, *International Research Journal of Finance and Economics* ISSN 1450-2887 Issue 51 (2010)

[16]G. Carpenter, S. Grossberg, N. Markuzon, J. Reynolds, and D. Rosen, Fuzzy ARTMAP: a neural networkarchitecture for incremental supervised learning of analog multidimensional maps, *IEEE Trans. Neural Network.*,vol. 3, no. 5, pp. 698–713, Sep2009