Prediction of Rainfall and Wind Speed Using ANFIS

N.Suganthi
M.Phil Research Scholar, PG & Research Department of Computer Science,
Chikkanna Government Arts College, Tirupur, Tamil Nadu, India.
Email: suganca@gmail.com
Dr.A.Geetha
Assistant Professor, PG & Research Department of Computer Science,
Chikkanna Government Arts College, Tirupur, Tamil Nadu, India.
Email: gee_sam@yahoo.com

Abstract— Climate prediction is an ever tough place of research for scientists. The Adaptive Neuro-Fuzzy Inference System (ANFIS) has been widely used for modeling one-of-a-kind forms of nonlinear systems including rainfall forecasting. Adaptive Neuro-Fuzzy Inference systems (ANFIS) combines the talents of Artificial Neural Networks (ANN) and Fuzzy Inference Systems (FIS) to remedy different sorts of troubles, particularly green in rainfall prediction and also wind speed prediction. In this paper the utility of synthetic neural networks to predict the climate of Delhi town has been proposed the use of information base inside the Neuro-Fuzzy Inference gadget. The weather parameters like minimum temperature, maximum temperature, relative humidity, sea stage strain, rainfall, wind speed, wind route and sun shine etc. has been used for prediction. When appearing weather predictive model the key standards are usually accuracy. We're seeking to be expecting future weather circumstance based upon above parameters by way of artificial neural community. The version performance is contrasted with multi layered perceptron community. The proposed network train with actual records of the 5 years (2010 to 2015) of South station, Coimbatore and tested which comes from meteorological department. The Multilayer Perceptron (MLP) used with Fuzzy good judgment. The Extreme Learning Machine (ELM) as an emerging mastering approach affords green unified answers to generalized feed-ahead net-works including but not limited to (both single- and multi-hidden-layer) neural networks, radial basis characteristic (RBF) networks, and kernel mastering.

Keywords-ANFIS, NN, Fuzzy Logic, Neuro Fuzzy, BPA.

1. INTRODUCTION

Climate in reality refer to the situation of air on earth at a given region and time .The application of technological know-how and era are to predict the kingdom of the atmosphere in destiny time for a given location is so crucial because of its effectiveness in human existence. These

days, weather forecasts are made by means of accumulating quantitative records about the present day country of the environment and using scientific expertise of atmospheric strategies to task how the atmosphere will evolve. The chaotic nature of the atmosphere implies the need of large computational strength required to clear up the equations that describe the atmospheric conditions. This is resulted from incomplete understanding of atmospheric procedures which mean that forecasts turn out to be much less correct as the difference in time between the prevailing second and the time for which the forecast is being made will increase. Climate is a non-stop, information-intensive, multidimensional, dynamic and chaotic process and these homes make climate prediction a huge project. Commonly, two methods are used for climate forecasting (a) The empirical approach and (b) The dynamical technique. the first method is based on the occurrence of analogs and is regularly referred via meteorologists as analog forecasting. This technique is beneficial for predicting nearby scale weather if recorded information's are considerable. the second one technique is based totally on equations and forward simulations of the ecosystem and is often called pc modeling. The dynamical approach is simplest beneficial for modeling large-scale climate phenomena and won't forecast brief-time period weather correctly. most climate prediction systems use a mixture of empirical and dynamical strategies synthetic Neural community (ANN) provides a methodology for fixing many types of nonlinear issues which are tough to be solved through conventional strategies .most meteorological strategies often exhibit temporal and spatial variability. They are suffered by means of issues of nonlinearity of bodily techniques, conflicting spatial and temporal scale and uncertainty in parameter estimates [7].

2. RELATED WORKS

Pankaj Kumar was proposed ANFIS with 4- Bell-shaped Gauss types of membership functions and hybrid learning algorithms method was used for the optimization of Minimum Weekly Temperature Forecasting using [1]. In this paper three inputs are used for minimum temperature forecast and mean weekly value used as input data set. Kumar Abhishek applied multilayered artificial neural network with learning by back propagation algorithm configuration. There are two tools for implementing the algorithms in Matlab [2]. They area. Nntool – open network/data manager. The single layer and the multi layer algorithms are implemented in the nntool- open network/ data manager. b. Nftool – Neural network fitting tool. Only back propagation algorithm is implemented in this Matlab tool. Back Propagation Algorithm (BPA) was implemented in the Nftool. A minimum MSE was obtained and a graph was plotted between the predicted values and the target values. The following are the values recorded using the Nftool MSE=3.6456. The implementation of multi-layer architecture was done using NNTOOL in MTALAB. Three algorithms were tested in multi-layer architecture: a. Back Propagation Algorithm (BPA) b. Layer Recurrent Network (LRN) c. Cascaded Back-Propagation (CBP) A.C. Subhajini at all made comparisons among Radial Basis Function, Back Propagation Neural, Network, Regression Neural Networks, Fuzzy

ARTMAP (Neurofuzzy Hybrid with Recurrent Network as the host architecture). Find ARTMAP is best among all these method of forecasting [3]. B. Putra, at all applied Fusion of Fuzzy- Artificial Neural Network for Short Term Weather Forecasting.

Arti R. at all was proposed Back Propagation Feed Forward Neural Network for Weather Forecasting using Weather parameters temperature, pressure, humidity, wind direction [5]. Muhammad Buhari, Member, IAENG and Sanusi Sani Adamu, applied Levenberg-Marquardt back propagation n algorithm for Short term load forecasting. The input consists of daily 24 hour load data for 12 months of the year 2005 and daily average maximum temperature altogether making 25 inputs rows by 365 days. The output layer will be a day's 24 hours load forecast for the utility company. The Target data is the same as the input's daily 24 hours load data [6]. Ch. Jyosthna Devi, B. Syam Prasad Reddy, K. Vagdhan Kumar, B. Musala Reddy and N. Raja Naya, applied Back Propagation for weather forecast. The aim is to gather dataset consisting weather parameters like temperature, humidity, dew point, visibility ,atmospheric pressure ,sea level, wind speed, wind direction etc. The Work Done, How neural networks are useful in forecasting the weather and the working of most powerful prediction algorithm called back propagation algorithm was explained. A 3-layered neural network is designed and trained with the existing dataset and obtained a relationship between the existing non-linear parameters of weather. So many parameters are taken and their relationships are taken into consideration those factors for the temperature forecasting. Like temperature, humidity, dew point, visibility, atmospheric pressure, sea level, wind speed, wind direction etc. The data is normalized using min-max normalization to scale the dataset into the range of 0 to 1. Basically the work is done to check two different ANN architecture which is better. These are Back Propagation (BPN) feed forward network and Radial basis function network (RBN). BPN is found the best and taken for further development for prediction of temperature but there is a drawback that time consuming process. The research was focused on proper initialization of weights and bias of weather forecasting system [7]. After going through the above detailed study we find the following key results that play a major role in any forecasting model building first is the methodology ANFIS which is used for accuracy and convergence and second is combination of parameters is an important fact of weather forecasting because the parameters of weather are correlated each other. This reason selection of parameters is also an important work in forecasting.

3. METHODOLOGY

3.1 ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM

An adaptive neuro-fuzzy inference gadget or adaptive network-based totally fuzzy inference gadget (ANFIS) is a sort of artificial neural community this is primarily based on Takagi–Sugeno fuzzy inference device. The method was evolved inside the early 1990s. since it integrates both neural networks and fuzzy logic ideas, it has capacity to capture the blessings of both in a unmarried framework. Its inference gadget corresponds to a fixed of fuzzy IF–THEN policies that have mastering functionality to approximate nonlinear features. Therefore, ANFIS is taken into consideration to be a usual estimator for the usage of the ANFIS in a greater green and optimum manner, you may use the exceptional parameters obtained with the aid of genetic algorithm.

An Adaptive neuro-fuzzy inference gadget (ANFIS) is a mixture of ANN and Fuzzy Inference machine (FIS) in one of these manner that neural network gaining knowledge of algorithms are used to decide the parameters of FIS. an excellent more essential aspect is that the device ought to constantly be interpretable in phrases of fuzzy if-then guidelines, due to the fact it's miles based at the fuzzy gadget reflecting indistinct information. we have used first–order Sugeno fuzzy model amongst many FIS Models. The Sugeno fuzzy model is maximum widely implemented one for its high interpretability and computational efficiency and built-in most efficient and adaptive strategies. The Sugeno fuzzy model provides a systematic approach to generate fuzzy policies from a hard and fast of input–output statistics pairs. in addition the top-rated values of the ensuing parameters (parameters in adaptive neuro-fuzzy inference device) can be discovered by way of using the least rectangular approach (LSM). Whilst the idea parameters are not constant, the quest space becomes large and the convergence of education becomes slower. The hybrid gaining knowledge of (HL) algorithm combining LSM and BP algorithms may be used to clear up this trouble. It was shown in that the HL algorithm is pretty green in education the ANFIS. This set of rules converges tons quicker since it reduces the dimension of the hunt space of the BP set of rules. for the duration of the mastering system, the basis parameters and the ensuing parameters are tuned till the desired response of the FIS is achieved. The HL set of rules has a two-step manner. First, the ensuing parameters are diagnosed using LSM while the values of the basis parameters are constant. Then, the consequent parameters are held constant at the same time as the mistake is propagated from the output give up to the input cease, and the idea parameters are up to date through the BP algorithm [18].

3.2 Neuro-Fuzzy Method

The neural-fuzzy version is an powerful approach for modeling nonlinear systems which include climate statistics because of the aggregate of blessings of neural systems and fuzzy common sense structures. The Neuro-adaptive studying strategies provide a way for the bushy modeling manner to advantage records approximately a dataset, so that you can compute the membership function parameters which permit the associated fuzzy inference device to track the given enter/output facts (Jang 1993)[26].Every fuzzy device carries three essential elements: fuzzification, inference and defuzzification.

3.3 Structure of ANFIS

The ANFIS approach learns the regulations and club functions from records. The ANFIS structure is supplied in discern 2. The circular nodes represent nodes which might be constant whereas the rectangular nodes are nodes which have parameters to be learnt.

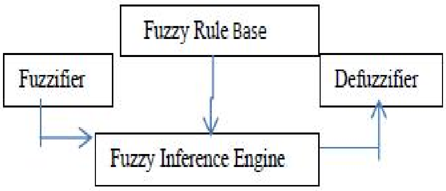

Fig. 1 Fuzzy interface system

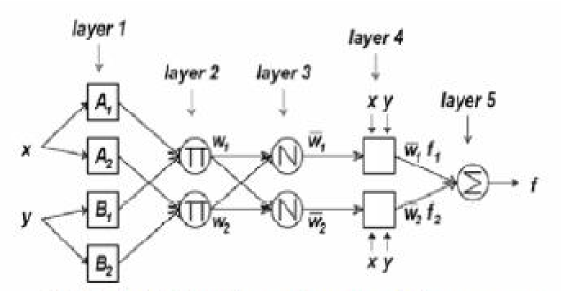

Fig. 2 An ANFIS architecture for a two rule Sugeno system

A two Rule Sugeno ANFIS has rules of the form:

A two Rule Sugeno ANFIS has rules of the form:

If x *is* A_1 *and* y *is* B_1 *THEN* $f_1 = p_1 x + q_1 y + r_1$.....................1

If x *is* A_2 *and y is* B_2 *THEN* $f_2 = p_2 x + q_2 y + r_2$..........................2

When Training the Network there is a forward pass and a backward pass. The forward pass propagates the input vector through the network layer by laye[19]r. In the backward pass, the error is sent back through the network in a way similar to back propagation [1].

In Layer 1, the output of each node is:

$$O_{1,i} = \mu_{B_{i-2}}(y) \qquad for\ i = 3,4$$

An is membership functions grade for x and y.The membership functions could be any shape. Using the Gaussian membership function given by:

$$\mu_{\bar{A}}(x) = \frac{1}{(1+x)^2}\text{.........}$$

In layer 2, every node in this layer is fixed. Here used 'AND', 'OR' membership grades for example the product:

$$O_{2,i} = w_i = \mu_{A_i}(x)\mu_{B_i}(y), \qquad i = 1,2\ldots$$

Layer3 contains the fixed node which calculates the ratio of the firing strengths of the rules:

$$O_{3,i} = \overline{w_i} = \frac{w_i}{w_1+w_2}$$

The nodes in layer 4 are adaptive and perform the consequent of the rules:

$$O_{4,i} = \overline{w_i} f_i = \overline{w_i}(p_i x + q_i y + r_i) \ldots\ldots$$

The parameters in this layer are to be determined and are referred to as the consequent parameters,

In layer 5 there is a single node that computes the overall output:

$$O_{5,i} = \sum \overline{w_i} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \ldots\ldots\ldots$$

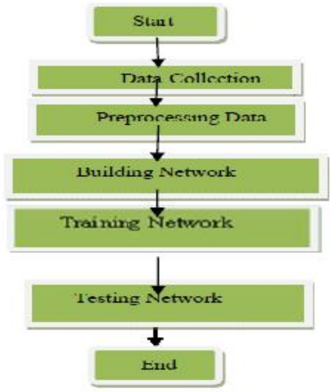

Fig.3 Neuro-Fuzzy Method

3.4 Machine Learning

Frequently harassed with artificial intelligence, machine learning truely takes the manner one step further by means of imparting the facts important for a machine to study and adapt whilst uncovered to new records. think about it as schooling a gadget: It relies upon on the other two strategies with the aid of reading mined data, creating a new algorithm via AI, and then updating contemporary algorithms accordingly to "study" a new assignment. Device getting to know is capable of generalizing information from big facts sets, and then detects and extrapolates patterns as a way to follow that facts to new answers and movements. Manifestly, positive parameters should be installation at the beginning of the device mastering system so that the device is able to locate, examine, and act upon new facts.

Machine mastering and artificial intelligence have reputedly never been as crucial and crucial to real-life programs as they may be in today's self reliant, huge facts generation. The achievement of machine gaining knowledge of and synthetic intelligence is based at the coexistence of three vital conditions: effective computing environments, rich and/or huge information, and efficient gaining knowledge of techniques (algorithms). the acute gaining knowledge of gadget (ELM) as an rising getting to know approach offers green unified answers to generalized feed-forward networks including but no longer restricted to (both unmarried- and multi-hidden-layer) neural networks, radial foundation characteristic (RBF) networks, and kernel gaining knowledge of. ELM theories1–4 display that hidden neurons are crucial however can be randomly generated and independent from packages, and those ELMs have each familiar approximation and classification competencies; in addition they construct an immediate link among a couple of theories (specifically, ridge regression, optimization, neural community generalization performance, linear device balance, and matrix principle). Consequently, ELMs, which may be biologically stimulated, offer characterize cant benefits including fast getting to know speed, ease of implementation, and minimal human intervention[21]. They for this reason have strong capacity as a feasible opportunity method for big-scale computing and system studying. This special version of traits & Controversies includes 8 original works that detail the similarly tendencies of ELMs in theories, programs, and hardware implementation. In "Representational studying with ELMs for big statistics," the authors recommend using the ELM as an auto-encoder for getting to know function representations the usage of singular values. In “A comfy and sensible Mechanism for Outsourcing ELMs in Cloud Computing,” the authors advise a method for dealing with massive statistics programs via outsourcing to the cloud that might dramatically reduce ELM schooling time. In “ELM-Guided Memetic Computation for vehicle Routing,” the authors bear in mind the ELM as an engine for automating the encapsulation of information memes from beyond trouble fixing stories. In “ELMVIS: A Nonlinear Visualization method the usage of Random permutations and ELMs,” the authors endorse an ELM technique for statistics visualization based totally on random permutations to map unique statistics and their corresponding visualization points. In “Combining ELMs with Random Projections,” the authors examine the relationships between ELM characteristic mapping schemas and the paradigm of random projections. In “decreased ELMs for Causal Relation Extraction from Unstructured text,” the authors recommend combining ELMs with neuron choice to optimize the neural network architecture and improve the ELM ensemble’s computational efficiency[22]. In “A device for Signature Verification based totally on Horizontal and Vertical additives in Hand Gestures,” the authors suggest a unique paradigm for hand signature biometry for contact much less applications without the need for hand-held devices. Eventually, in “An Adaptive and Iterative on-line Sequential ELM-based Multi-diploma-of-Freedom Gesture popularity machine,” the authors endorse an online sequential ELM-based green gesture popularity set of rules for contact less human system interaction.

A gadget gaining knowledge of algorithm’s generalization functionality relies upon at the dataset, which is why engineering a dataset’s functions to symbolize the statistics salient shape is critical. But, function engineering requires domain knowledge and human ingenuity to generate suitable functions. Geoffrey Hinton1 and Pascal Vincent2 showed that a limited Boltzmann machine (RBM) and automobile-encoders may be used for function engineering. These engineered features then could be used to teach a couple of-layer neural networks, or deep networks. types of deep networks based on RBM exist: the deep perception network (DBN)1 and the deep Boltzmann system (DBM).3 the 2 sorts of automobile-encoder-based totally deep networks are the stacked car-encoder (SAE)2 and the stacked denoising automobile encoder (SDAE).3 DBNs and DBMs are created by stacking RBMs, while SAEs and SDAEs are created by way of stacking automobile-encoders. Deep networks outperform traditional multilayer neural networks, single-layer feed-ahead neural networks (SLFNs), and aid vector machines (SVMs) for massive information, however are tainted by using sluggish mastering speeds[23]. Guang-Bin Huang and colleagues4 brought the acute studying system (ELM) as an SLFN with a quick gaining knowledge of velocity and precise generalization capability. similar to deep networks, our proposed multilayer ELM (ML-ELM) plays layer-by way of-layer unsupervised mastering. This newsletter also introduces the ELM vehicle-encoder (ELM-AE), which represents functions primarily based on singular values. Reminiscent of deep networks, ML-ELM stacks on pinnacle of ELM-AE to create a multilayer neural community. It learns considerably faster than current deep networks, outperforming DBNs, SAEs, and SDAEs and performing on par with DBMs on the MNIST5 dataset.

The ELM for SLFNs shows that hidden nodes can be randomly generated. The enter statistics is mapped to L-dimensional ELM random feature area, and the network output

$$f_L(\mathbf{x}) = \sum_{i=1}^{L} \beta_i h_i(\mathbf{x}) = \mathbf{h}(\mathbf{x})\beta,$$

Where b = [b1… bL] T is the output weight matrix between the hidden nodes and the output nodes, h(x) = [g1(x)… gL(x)] are the hidden node outputs (random hidden features) for input x, and gi(x) is the output of the ith hidden node. Given N training samples {(x t,)} i i i N =1 , the ELM can resolve the following learning problem:

$$\mathbf{H}\beta = \mathbf{T},$$

Where T = [t1… tN] T are target labels, and H = [hT(x1)... hT(xN)]T. We can calculate the output weights b from,

$$\beta = \mathbf{H}^{\dagger}\mathbf{T},$$

Where H† is the Moore-Penrose generalized inverse of matrix H. To improve generalization performance and make the solution more robust, we can add a regularization term as shown elsewhere:

$$\beta = \left(\frac{\mathbf{I}}{C} + \mathbf{H}^T\mathbf{H}\right)^{-1} \mathbf{H}^T\mathbf{T}.$$

ELM-AE’s main objective to represent the input features meaningfully in three different representations: • Compressed. Represent features from a higher dimensional input data space to a lower dimensional feature space. • Sparse. Represent features from a lower dimensional input data space to a higher dimensional feature space. • Equal. Represent features from an input data space dimension equal to feature space dimension[24]. The ELM is modified as follows to perform unsupervised learning: input data is used as output data t = x, and random weights and biases of the hidden nodes are chosen to be orthogonal. Bernard Windrow and colleagues7 introduced a least mean square (LMS) implementation for the ELM and a corresponding ELM-based auto encoder that uses no orthogonal random hidden parameters (weights and biases). Orthogonalization of these

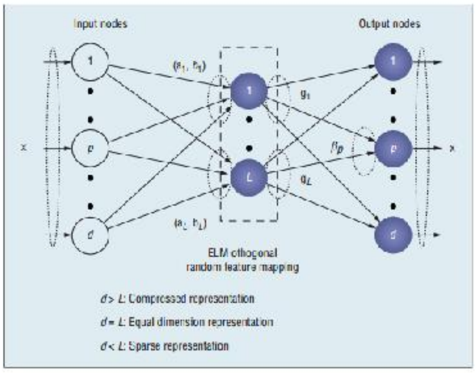

Fig.3 ELM-AE’s Network

Randomly generated hidden parameters tend to improve ELM-AE’s generalization performance. According to ELM theory, ELMs are universal approximates, 8 hence ELM-AE is as well. Fig.4 shows ELM-AE’s network structure for compressed, sparse, and equal dimension representation. In ELM-AE, the orthogonal random weights and biases of the hidden nodes project the input data to a different or equal dimension space as shown by the Johnson-Lindenstrauss lemma9 and calculated as,

$$\mathbf{h} = g(\mathbf{a} \cdot \mathbf{x} + \mathbf{b})$$
$$\mathbf{a}^T\mathbf{a} = \mathbf{I},\ \mathbf{b}^T\mathbf{b} = 1,$$

Where a = [a1, aL] are the orthogonal random weights, and b = [b1… bL] are the orthogonal random biases between the input and hidden nodes. ELM-AE's output weight b is responsible for learning the transformation from the feature space to input data. For sparse and compressed ELM-AE representations, we calculate output weights b as follows:

$$\beta = \left(\frac{I}{C} + \mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{X},$$

Where H = [h1… hN] are ELM-AE's hidden layer outputs, and X = [x1… xN] are its input and output data. For equal dimension ELM-AE representations, we calculate output weights b as follows:

$$\beta = \mathbf{H}^{-1}\mathbf{T}$$
$$\beta^T \beta = \mathbf{I}.$$

Singular value decomposition (SVD) is a commonly used method for feature representation. Hence we believe that ELM-AE performs feature representation similar to SVD. Equation 6's singular value decomposition (SVD) is

$$\mathbf{H}\beta = \sum_{i=1}^{N} \mathbf{u}_i \frac{\mathbf{d}_i^2}{\mathbf{d}_i^2 + C} \mathbf{u}_i^T \mathbf{X},$$

Where u is eigenvectors of HHT, and d are singular values of H, related to the SVD of input data X. Because H is the projected feature space of X squashed via a sigmoid function, we hypothesize that ELM-AE's output weight b will learn to represent the features of the input data via singular values. To test if our hypothesis is correct, we created 10 mini datasets containing digits 0 to 9 from the MNIST dataset. Then we sent each mini dataset through an ELM-AE (network structure: 784-20-784) and compared the contents of the output weights b (Figure 2a) with the manually calculated rank 20 SVD (Figure 2b) for each mini dataset. As Figure 2 shows, ELM-AE output weight b and the manually calculated SVD basis. Multilayer neural networks perform poorly when trained with back propagation (BP) only, so we initialize hidden layer weights in a deep network by using layer-wise unsupervised training and fine-tune the whole neural network with BP[27]. Similar to deep networks, ML-ELM hidden layer weights are initialized with ELM-AE, which performs layer-wise unsupervised training. However, in contrast to deep networks, ML-ELM doesn't require fine tuning. ML-ELM hidden layer activation functions can be either linear or nonlinear piecewise. If the number of nodes Lk in the kth hidden layer is equal to the number of nodes Lk−1 in the (k − 1) Th hidden layer, g is chosen as linear; otherwise, g is chosen as nonlinear piecewise, such as a sigmoidal function:

$$\mathbf{H}^k = g((\beta^k)^T \mathbf{H}^{k-1}),$$

Where Hk is the kth hidden layer output matrix. The input layer x can be considered as the 0th hidden layer, where k = 0. The output of the connections between the last hidden layer and the output node t is analytically calculated using regularized least squares.

4. Experimental Results

Algorithm	Dimension Reduction
Learning Vector Quantization (LVQ)	90.2
Linear Discriminant Analysis (LDA)	94.6
Principal Component Analysis (PCA)	97.4

Table 4.1 Dimension reduced of algorithm comparison method

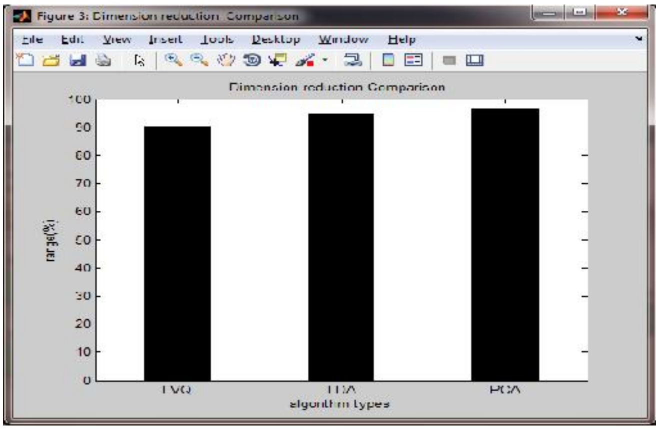

Fig.4.1 Algorithm Comparison Method

Algorithm	Accuracy	Time Period(Time seconds)
Bayesian	86.2	3.5
Principal Component Analysis (PCA)	92.3	2.8
Neuro Fuzzy	96.4	1.8
Elm_Nfuzzy	98.52	1.2

Table 4.2 Accuracy Comparison Elm_Nfuzzy

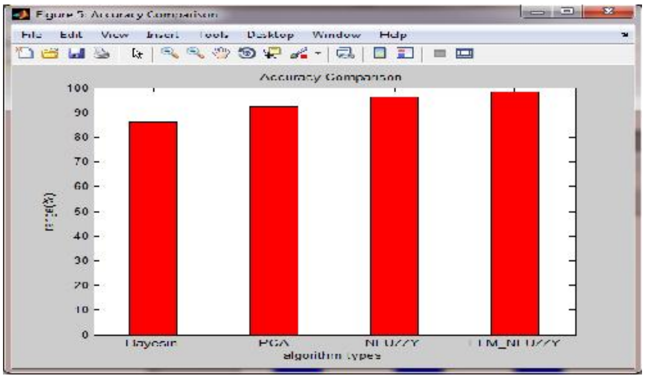

Fig.4.2 Accuracy Comparison Elm_Nfuzzy

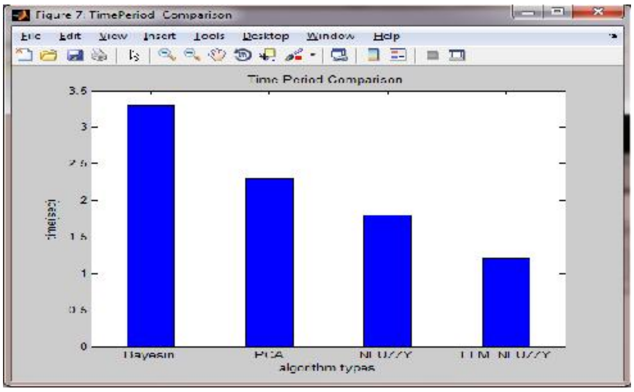

Fig.4.3 Time Period Comparison

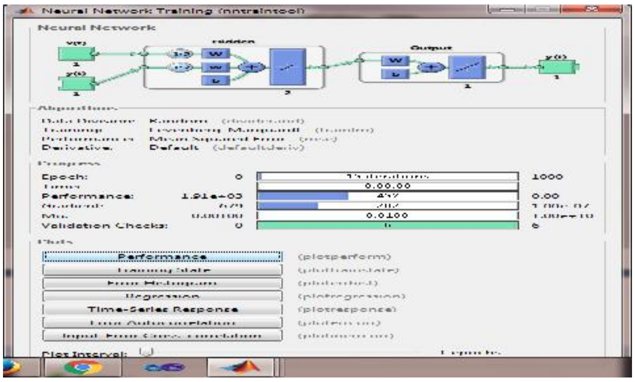

Fig.4.4 Neural network layer

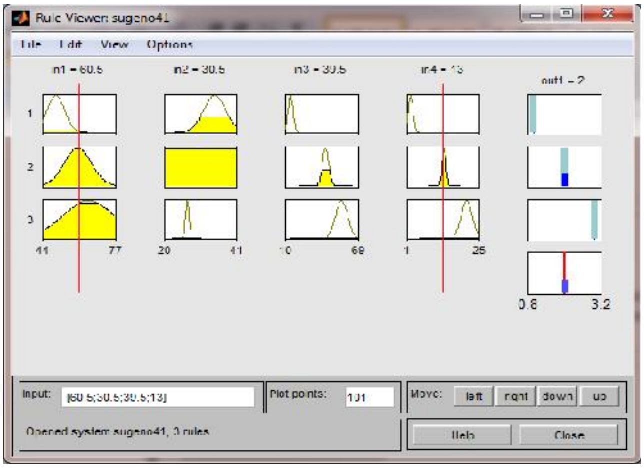

Fig.4.5 Neurofuzzy classification

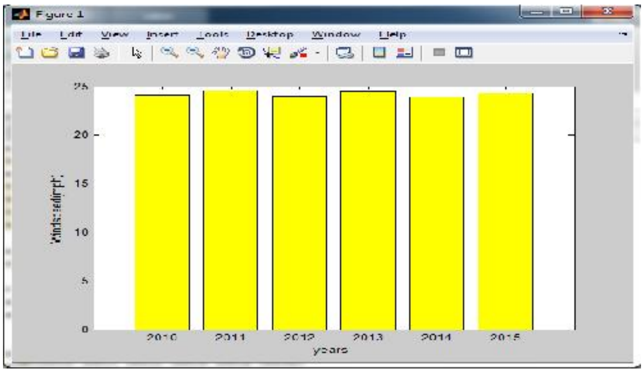

Fig.4.6 Year based Wind Speed

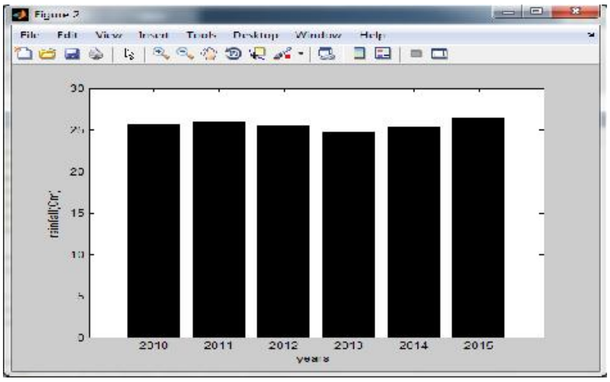

Fig.4.7 Year Based Rain Fall Prediction

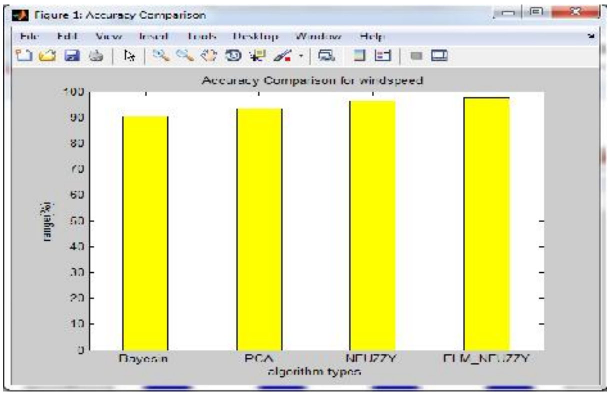

Fig.4.8 Accuracy Comparison Wind Data

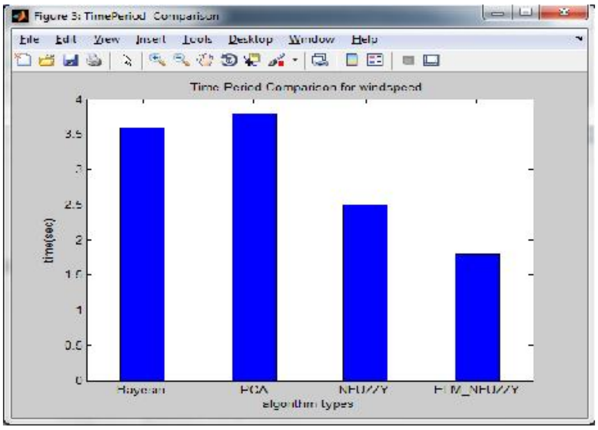

Fig.4.9 Time Period Comparison Wind Speed Data

6. CONCLUSION

A neural fuzzy network version changed into proposed for relative humidity prediction on the idea of exceptional aggregate of parameters. Short term prediction is an crucial part of the today's manipulate technology for operation of constructing systems. This Work discusses the possibility of using meteorological information with local commentary information for quick-term prediction. Neural network has gained outstanding popularity in climate prediction due to their simplicity and robustness. In this Look at the performance of different parameters of weather is as compared through using ANFIS (adaptive Neuro-fuzzy inference gadget) for relative humidity forecasting. They have a look at also says that Neural network with Fuzzy good judgment is the great aggregate for weather forecasting. The dataset choice, input variable selection, the connection and interdependencies many of the information, the right training set and the proper structure are maximum essential for great prediction consequences. Proposed strategies unearths high efficient and accuracy of facts prediction in future use a few changed algorithms may additionally b produced high accuracy.

7. REFERENCES

[1] Pankaj Kumar, 2012, "Minimum Weekly Temperature forecasting using ANFIS".

[2] Kumar Abhishek, Abhay Kumar, Rajeev Ranjan, Sarthak Kumar,2012 IEEE, "A Rainfall Prediction Model using Artificial Neural Network".

[3] A.C. Subhajini and T.Santhanam, 2011, "fuzzy artmap neural network architecture for weather forecasting", Journal of Theoretical and Applied Information Technology.

[4] B.Putra, B.T.Atmaja, S.Hidayat 2011 "Short–Term Weather Forecasting Using Fusion of Fuzzy-Artificial Neural Network",International Conference on Informatics for development.

[5] ArtiR.Naik, S.K.Pathan,2012,"Weather Classification and Forecasting using Back Propagation Feed forward Neural Network", International Journal of Scientific and Research Publications.

[6] Muhammad Buhari, Member, IAENG and Sanusi Sani Adamu,2012,"Short-Term Load Forecasting Using Artificial Neural Network".

[7] Ch. Jyosthna Devi, B. Syam Prasad Reddy, K. Vagdhan Kumar, B.Musala Reddy, N. RajaNayak, 2012,"ANN Approach for Weather Prediction using Back Propagation", International Journal of Engineering Trends and Technology.

[8] A.Geetha and G.M.Nasira, "Artificial Neural network"Application in weather Forcasting-using Rapidminer",published in Internatiolal journal of computational Intelligence and informatics, vol.4:no.2,October-Decamber 2014.

[9] A.Geetha and G.M.Nasira,"Rainfall Prediction using Logistic Regression Techniques",ciit Internatiolal journal of Artificial Intelligence system & machine learning,,vol.6,no.07 pp.246-250,aug-2014.

[10] Ratna Nayak, P. S. Patheja, Akhilesh A Waoo, 2012 "An Artificial Neural Network Model for Weather Forecasting in Bhopal", IEEEInternational Conference on Advances in Engineering, Science And Management.

[11] Gyanesh Shrivastava, Dr. C.V. Raman University, Bilaspur,Chhattisgarh, India, 2012,"Application of Artificial Neural Networks in Weather Forecasting: A Comprehensive Literature Review" International Journal of Computer Applications.

[12] Holger R. Maier , Graeme C. Dandy (ELSEVIER-2000)" Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications".

[13] M. Nasseri, K. Asghari, M.J. Abedini (ELSEVIER-2008) "Optimized scenario for rainfall forecasting using genetic algorithm coupled with artificial neural network".

[14] Mohammad Monfared, Hasan Rastegar, Hossein Madadi Kojabadi (ELSEVIER-2009) "A new strategy for wind speed forecasting using artificial intelligent methods".

[15] Brian A. Smitha, Gerrit Hoogenbooma, Ronald W. Mc. Clendona,(ELSEVEIR-2009) "Artificial neural networks for automated yearround temperature prediction".

[16] M. B. Abdul Hamid, T. K. Abdul Rahman (ICCMS-2010) Short Term Load Forecasting Using an "Artificial Neural Network Trained by Artificial Immune System Learning Algorithm".

[17] Rajasekaran S.Vijayalakshmi G.A. "Neural Networks, Fuzzy Logic and Genetic Algorithms", PHI, 2003.

[18] S.N.Sivanandam,,S.Sumathi,S.N.Deepa, "Introduction to Fuzzy Logic using MATLAB".

[19] G. E. Hinton and R. R. Salakhutdinov,"Reducing the Dimensionality of Data with Neural Networks," Science,vol. 313, no. 5786, 2006, pp. 504–507.

[20] P. Vincent et al., "Stacked Denoising Auto encoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion," J. Machine Learning Research, vol. 11, 2010, pp. 3371–3408.

[21] R. Salakhutdinov and H. Larochelle "Efficient Learning of Deep Boltzmann Machines," J. Machine Learning Research,vol. 9, 2010, pp. 693–700.

[22] G.-B. Huang, Q.-Y. Zhu and C.-K. Siew, "Extreme Learning Machine: Theory and Applications," Neurocomputing,vol. 70, 2006, pp. 489–501.

[23] Y. LeCun et al., "Gradient-Based Learning Applied to Document Recognition,"Proc. IEEE, vol. 86, no. 11, 1998,pp. 2278–2324.

[24] G.B. Huang et al.,"Extreme Learning Machine for Regression and Multiclass Classification," IEEE Trans. Systems, Man, and Cybernetics, vol. 42, no. 2, 2012, pp. 513–529.

[25] B. Widrow et al., "The No-Prop Algorithm: A New Learning Algorithm for Multilayer Neural Networks," Neural Networks, vol. 37, 2013,pp. 182–188.

[26] Gwo-Ching Liao, Ta-Peng Tsao (ELSEVIER-2004), "Application of fuzzy neural networks and artificial intelligence for load forecasting".

[27] Rahat Hossaina, Amanullah MaungThanOoa, A B M Shawkat Alia (ELSEVIER- 2012) "Historical Weather Data Supported Hybrid Renewable Energy Forecasting using Artificial Neural Network (ANN)".