

# A Novel Machine Learning Based Approach for Rainfall Prediction

Niharika Solanki and Gaurang Panchal<sup>(✉)</sup>

U & P U. Patel Department of Computer Engineering,  
Chandubhai S Patel Institute of Technology, Changa, India  
15pgce035@charusat.edu.in, gaurangpanchal.ce@charusat.ac.in

**Abstract.** The climate changes effortlessly nowadays, prediction of climate is very hard. However, the forecasting mechanism is the vital process. It is also a valuable thing as it is the important part of the human life. Accordingly to the research, the weather forecast of rainfall intensity conducted. The remarkable commitment of this proposal is in the implementation of a hybrid intelligent system data mining technique for solving novel practical problems, Hybrid Intelligent system data mining consists of the combination of Artificial Neural Network and the proper usage of Genetic Algorithm. In this research, Genetic algorithm is utilized the type of inputs, the connection structure between the inputs and the output layers and make the training of neural network more efficient. In ANN, Multi-layer Perceptron (MLP) serves as the center data mining (DM) engine in performing forecast tasks. Back Propagation algorithm used for the trained the neural network. During the training phase of the proposed approach, it gains the optimal values of the connection weights which, in fact, utilized as the part of the testing phase of the MLP. Here, the testing phase is used to bring about the rainfall prediction accuracy. It may be noted that the information/data is used to cover the information from the variables namely temperature, cloud fraction, wind, humidity, and rainfall.

**Keywords:** Rainfall prediction · Genetic algorithm · Artificial neural network

## 1 Introduction

Machine learning is an advanced mechanism for analysis is various data (i.e. data sets) that computerises explanatory which are use to build the models. The knowledge/information gain by the various calculations which are performed based on the information, the Machine Learning (ML) allows the systems to find out the hidden bits (i.e. information) of knowledge without applying alternative solutions [1–10].

There are various utilizations of machine learning. It's quite to acknowledge how much machine learning has accomplished in real world applications.

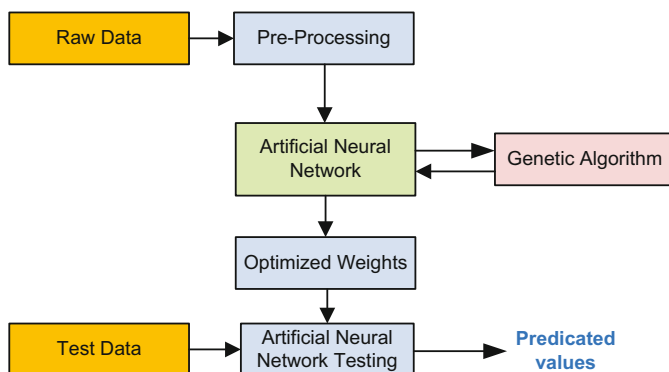
Machine learning is regularly connected in the disconnected training phase. Thus machine learning is utilized to enhance the applications such as face recognition, face detection, speech recognition, genetics, image classification, weather forecast etc. [3–15]. Machine learning is connected in weather forecasting programming to enhance the nature of the estimate. Machine learning makes it moderately less demanding to create complex programming frameworks without much exertion on the human side [10–18].

Weather prediction is a confused methodology that incorporates various particular fields of mastery [6]. Weather prediction has been a standout amongst the most deductively and mechanically troublesome issues the world over in the latest century [18–28].

Since its conception Weather Prediction has made tremendous progress [6]. In any case, numerous new issues have developed. Therefore, critical topics in the community as a whole are in fact considered as a lack of timely. There are few issues to solve weather prediction which described as below;

1. The most difficult issue is enhancing accuracy.
2. High speed of the streaming data and scale it for the high dimensions.
3. Reducing the computational complexity and the memory requirement.
4. To tackle the over fitting issue, a standout amongst the most widely recognized assignment is to fit a model to a set of training data.

To solve this problem we propose a hybrid genetic algorithm based neural network rainfall prediction system which will be given in Fig. 1.



**Fig. 1.** Proposed approach.

## 2 Related Work

The work by Kumar Abhisek *et al.* [7] used Layer Recurrent Network (LRN), Artificial Neural Network in which Back-Propagation Algorithm (BPA) and Cascaded back Propagation (CBP) algorithm utilized for rainfall prediction. For the

best performance in this paper, executes rainfall prediction by utilizing data set and train it. The test data sets and setting up the number of neurons in the hidden layer is done in this phase. In the present research, ANN models analyzed possibility of predicting average rainfall. In formulating, ANN constructed predictive models three layered network. BP algorithm used for the train the normalized data. Finally compare the predicted output with desired output.

The work by Ahmet Erdil *et al.* [2] utilised the concept of Neural Network along with the addition process of Back propagation Algorithm (BPA) for Prediction of meteorological factors that is radiation of secondly solar, minimum and maximum atmospheric pressures. In the Real-world scenarios, the issues and the problems are solved effectively using the artificial neural network by straightforward use. The objective of the proposed algorithm the minimization of the level of the global error. For example, the mean error and root mean square error.

The work by Kumar Abhishek *et al.* [8] used Artificial Neural Network as a part of which Feedforward with back propagation algorithm used for forecast maximum temperature. In these paper, investigates the pertinence of ANN approach by creating viable and reliable nonlinear predictive models for climate examination also additionally assess and look at the execution of the developed models utilizing different transfer functions, hidden layers and neurons to forecast maximum, temperature for 365 days of the year.

The work by Abbas Ahmadi *et al.* [1] utilized ensemble of neural network using Multi-Layer Perceptron (MLP), Radial basis function (RBF), and General Regression Neural Network (that is, GRNN) and Time Delay Neural Network (that is, TDNN) for Temperature Prediction (Dew Point, Humidity, Sea Level Pressure and Visibility.). For weather forecasting, propose a new hybrid model in this research paper, which is depends on an ensemble of neural networks. Likewise, introduce a mutual information (MI) concepts that handle the various challenges that uses the results of the different networks' combinations and try to minimise the redudancy in the combined hybrid model.

The work by Morten Gill Wollsen *et al.* [9] used Artificial Neural Network using Back-Propagation Algorithm for Solar irradiate and temperature forecasts. In this paper, (Nonlinear Auto Regressive with eXternal input) NARX network is used which is the artificial neural network which has been shown to be appropriate for the displaying of non-linear systems. The back-propagation training algorithm used for trained the neural network. The network is trained until accomplishing an error of 0.05%.

### 3 Proposed Approach

The proposed approach includes different process namely collection of Rrw data, pre-processing of the raw data, training GA-MLP, testing GA-MLP and the final rainfall prediction. In pre-processing first remove the missing value by the taking the centre measure value like mean of the given variable or if missing value is small in size, then remove the missing value by erasing. After the removing missing data the rainfall data will be normalised between the range of  $[0 \cdots 1]$  which minimise the error.

After the pre-processing, there will be a training of GA-MLP. In this stage, genetic algorithm [5] and multi-layer perceptron [4] is used. Genetic algorithm is utilized the type of inputs, the connection structure between the inputs and the output layers. This, in fact, represents the simplification of Multilayer Perceptron structure in the Model, and makes the training process more efficient. The evaluation of the fitness function for the different combination of the variables are resolved by the MSE (means square error) when such an input combination is utilised as a part of a MLP to perform the classification/prediction task.

Also, in pre-processing missing value and normalization is carried out to improve the accuracy of the algorithm. Missing value in the data set can be given as “missing” data. To solve this missing value by, if missing value is small in size then remove the missing value by erasing or remove the missing value by the taking the centre measure value like mean of the given variable. After the removing the procedure of missing data the rainfall data will be normalised between the various ranges (i.e.,  $[0 \cdots 1]$ ) which, in fact, used to minimize the overall error and produces the better performance. The equation of normalization given as follow;

$$X_{norm} = \frac{X_t - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Here,  $X_{norm}$  is normalized data,  $X_t$  is actual data,  $X_{max}$  is maximum value of  $x$  variable and  $X_{min}$  is minimum value of  $x$  variable.

After the pre-processing, there will be a training of GA-MLP. In this stage, genetic algorithm [5] and multi-layer perceptron [4] is used. The proposed architecture is trained using the multi-layer perceptron in combination of the genetic algorithm. Genetic algorithm is utilized the type of inputs, the connection structure between the inputs and the output layers.

This simplifies the MLP structure in Model, and makes the training process more efficient. In Genetic algorithm, initial population of the chromosomes is randomly generated. Then from each chromosome, the weights are extracted based on the number of genes a chromosomes having. The error and the fitness value are calculated. After the weights are extracted successfully, the initial weights are combined together and again given as an input to the network for the calculation of the actual output. These outputs when compared with the desired output, error values are obtained.

The assessment of the fitness function for every input variable combination is resolved by mean squared error (MSE) when such an input combination is utilised as a part of a MLP to perform the classification/prediction task. Here, the fitness function is examined by the calculating the RMS error. Then generate a new population, the crossover operator is applied. These steps are calculated in repetition until the stopping criteria has been satisfied. The chromosomes archives same value of fitness then the algorithm get stopped. The weights which are obtained at the end of the training phase are used to predicts the temperature, wind speed, humidity.

## 4 Conclusion

The use of the Genetic Algorithm to optimise the structure parameters, the connection weights between the inputs and the output layers. This, in fact, represents that the MLP structure in Model, and enables the classifier training phase more efficient and computationally less expensive. The evaluation of the problem fitness functions for the different input combination is calculated by the mean squared error (MSE) when such an input combination is exploited as a part of a multilayer perceptron to perform the classification/prediction task. The training must be repeated a few circumstances keeping in mind the end goal to get satisfactory results.

## References

1. Ahmadi, A., Zargaran, Z., Mohebi, A., Taghavi, F.: Hybrid model for weather forecasting using ensemble of neural networks and mutual information. In: 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, pp. 3774–3777 (2014)
2. Erdil, A., Arcaklioglu, E.: The prediction of meteorological variables using artificial neural network, pp. 1677–1683 (2012)
3. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, New York (2006)
4. Du, K.L., Swamy, M.N.S.: Neural Networks in a Soft Computing Framework. Springer, London (2006)
5. Goldberg, D.E.: Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley Longman Publishing Co., Inc., Boston (1989)
6. Upadhaya, J.: Assam University, Climate Change and its impact on Rice productivity in Assam
7. Abhishek, K., Kumar, A., Ranjan, R., Kumar, S.: A rainfall prediction model using artificial neural network. In: 2012 IEEE Control and System Graduate Research Colloquium (ICSGRC), Shah Alam, Selangor, pp. 82–87 (2012)
8. Abhishek, K., Singh, M.P., Ghosh, S., Anand, A.: Weather forecasting model using Artificial Neural Network. *Procedia Technol.* **4**, 311–318 (2012)
9. Wollsen, M.G., Jørgensen, B.N.: Improved local weather forecasts using artificial neural networks, pp. 75–86 (2015)
10. Ganatra, A., Panchal, G., Kosta, Y., Gajjar, C.: Initial classification through back propagation in a neural network following optimization through GA to evaluate the fitness of an algorithm. *Int. J. Comput. Sci. Inf. Technol.* **3**(1), 98–116 (2011)
11. Panchal, G., Ganatra, A., Kosta, Y., Panchal, D.: Forecasting employee retention probability using back propagation neural network algorithm. In: IEEE 2010 Second International Conference on Machine Learning and Computing (ICMLC), Bangalore, India, pp. 248–251 (2010)
12. Panchal, G., Ganatra, A., Shah, P., Panchal, D.: Determination of over-learning and over-fitting problem in back propagation neural network. *Int. J. Soft Comput.* **2**(2), 40–51 (2011)
13. Panchal, G., Ganatra, A., Kosta, Y., Panchal, D.: Behaviour analysis of multilayer perceptrons with multiple hidden neurons and hidden layers. *Int. J. Comput. Theory Eng.* **3**(2), 332–337 (2011)

14. Panchal, G., Panchal, D.: Solving np hard problems using genetic algorithm. *Int. J. Comput. Sci. Inf. Technol.* **6**(2), 1824–1827 (2015)
15. Panchal, G., Panchal, D.: Efficient attribute evaluation, extraction and selection techniques for data classification. *Int. J. Comput. Sci. Inf. Technol.* **6**(2), 1828–1831 (2015)
16. Panchal, G., Panchal, D.: Forecasting electrical load for home appliances using genetic algorithm based back propagation neural network. *Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET)* **4**(4), 1503–1506 (2015)
17. Panchal, G., Panchal, D.: Hybridization of genetic algorithm and neural network for optimization problem. *Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET)* **4**(4), 1507–1511 (2015)
18. Panchal, G., Samanta, D.: Comparable features and same cryptography key generation using biometric fingerprint image. In: 2nd IEEE International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), pp. 1–6 (2016)
19. Panchal, G., Samanta, D.: Directional area based minutiae selection and cryptographic key generation using biometric fingerprint. In: 1st International Conference on Computational Intelligence and Informatics, pp. 1–8. Springer (2016)
20. Panchal, G., Samanta, D., Barman, S.: Biometric-based cryptography for digital content protection without any key storage. In: *Multimedia Tools and Application*, pp. 1–18. Springer (2017)
21. Panchal, G., Kosta, Y., Ganatra, A., Panchal, D.: Electrical load forecasting using genetic algorithm based back propagation network. In: 1st International Conference on Data Management, IMT Ghaziabad. MacMillan Publication (2009)
22. Patel, G., Panchal, G.: A chaff-point based approach for cancelable template generation of fingerprint data. In: *International Conference on ICT for Intelligent Systems (ICTIS 2017)*, p. 6 (2017)
23. Patel, J., Panchal, G.: An IoT based portable smart meeting space with real-time room occupancy. In: *International Conference on Internet of Things for Technological Development (IoT4TD 2017)*, pp. 1–6 (2017)
24. Soni, K., Panchal, G.: Data security in recommendation system using homomorphic encryption. In: *International Conference on ICT for Intelligent Systems (ICTIS 2017)*, pp. 1–6 (2017)
25. Patel, N., Panchal, G.: An approach to analyze data corruption and identify misbehaving server. In: *International Conference on ICT for Intelligent Systems (ICTIS 2017)*, pp. 1–6 (2017)
26. Bhimani, P., Panchal, G.: Message delivery guarantee and status Up-date of clients based on IOT-AMQP. In: *International Conference on Internet of Things for Technological Development (IoT4TD 2017)*, pp. 1–6 (2017)
27. Mehta, S., Panchal, G.: File distribution preparation with file retrieval and error recovery in cloud environment. In: *International Conference on ICT for Intelligent Systems (ICTIS 2017)*, p. 6 (2017)
28. Kosta, Y., Panchal, D., Panchal, G., Ganatra, A.: Searching most efficient neural network architecture using Akaikes information criterion (AIC). *Int. J. Comput. Appl.* **1**(5), 41–44 (2010)