International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-9 Issue-1, May 2020

Prediction of Rainfall using Machine Learning

S.Poornima, S.Devi, D.Oviya, A.Suhana Taj, G.Tamil Elakkiya

Abstract*: Rainfall is one of the major livelihood of this world. Each and every organism in this universe need some of water to order to survive in its own living conditions. As rainfall is the main source of water and its need to agriculture is inevitable, there arises a necessity to analyze the pattern of the rainfall. The main aim of our paper is to predict the rainfall considering various factors like temperature, pressure, cloud cover, wind speed, pollution and precipitation. There are various ideas and new methodologies proposed in order to predict rainfall. But our proposed concept is based on machine learning because of its wide range of development and preferability nowadays. Among the various technologies built in Machine Learning (ML), Feed Forward Neural Network (FFNN) which is the simplest form of Artificial Neural Network (ANN) is preferred because this model learns the complex relationships among the various input parameters and helps to model them easily. Rainfall in our proposed model is predicted using different parameters influencing the rainfall along with their combinations and patterns. The experimental results depicts that the proposed model based on FFNN indicates suitable accuracy.

Keywords* : Rainfall Prediction; Feed Forward Neural network; Humidity, precipitation

I. INTRODUCTION

Rainfall plays an important role in the economy of the country as agriculture is influenced mainly by rainfall. Rain is the main source for drinking water because it can become as a source of freshwater, which is important for the survival of human beings, plants and animals existing today.

The pattern of the rainfall should be regular as too much of variations in this pattern may be harmful which can destroy crops and even other living organisms. Prediction of rainfall in an accurate manner is a challenging task as it involves analysis of complex meteorological factors. Among various technologies and methodologies available for predicting the rainfall, the main focus of this paper is to provide a simple but effective method for predicting the rainfall.

This paper is structured in form of sections such that Section II describes Literature survey, Section III explains the methodology of the proposed work, Section IV portrays the results and Section V provides the conclusion of the paper respectively.

II. LITERATURE SURVEY

Tomoaki Kashiwao [3] used the concept of Multi-layer perceptron (MLP) with the combination of techniques such as Random Optimization (RO) method, Back -propagation(BP), and the least squares method (LSM) to predict the rainfall and the performance of prediction compared.

M.N. Ahuna [1] used rainfall data, to train back propagation neural network. Using the prediction, the statistics of attenuation of rain rates was calculated and compared with the ITU-R model. The prediction was flexible enough to different changes in climate even the time taken for sampling the data is comparatively low.

R.Arya [9] analyze the 30 years of rainfall data. He train the network using Feed Forward back propagation algorithm. Optimized result is obtained using Genetic algorithm (GA). BCNN-A gives better results compared to FF-BPNN alone.

E G Wahyuni [8] use the method of Backpropagation and System analysis using Recapitulating data. It uses Mamdani method which covers a wide field and enables to get input from the humans so that results obtained are quite accurate.

Reza Mohammadpour [4] implemented a sensitivity analysis on datasets to analyze the impact of various parameters causing rainfall. Artificial Neural Network along with the Learning Cellular Automation are employed to categorize the form of daily rainfall. This has the capability of eliminating the days in which rainfall does not occur.

III. PROPOSED SYSTEM

The model proposed in our work is based on prediction on rainfall using various parameters like temperature, pressure, cloud cover, wind speed, pollution and precipitation. The flow of the proposed work is described in Fig 1.

A. Collection of weather datasets

The data used for the implementation of this paper is the weather dataset which consists of rainfall predicting factors such as temperature, pressure, cloud cover, wind speed, pollution and precipitation are considered as an input.

B. Building model using ANN

The model built is this proposed work consists of three layers namely the input, the hidden and the output layer.

Manuscript received on April 02, 2020.
Revised Manuscript received on April 15, 2020.
Manuscript published on May 30, 2020.
* Correspondence Author
S. Poornima*, Department of Information Technology, Coimbatore Institute of Technology, Coimbatore, India. poornima.s@cit.edu.in
S. Devi, Department of Information Technology, Coimbatore Institute of Technology, Coimbatore, India. devi.s@cit.edu.in
D.Oviya, Department of Information Technology, Coimbatore Institute of Technology, Coimbatore, India. oviyavv@gmail.com
A.Suhana Taj, Department of Information Technology, Coimbatore Institute of Technology, Coimbatore, India. suhanaazam2610@gmail.com
G.Tamil Elakkiya, Department of Information Technology, Coimbatore Institute of Technology, Coimbatore, India. tamilillakiyag510@gmail.com

The FFNN model allows the flow of data in forward direction across the system.

The input parameters are passed through the input layer then the values are added with weights followed by activation function is being applied and is passed to hidden layers.

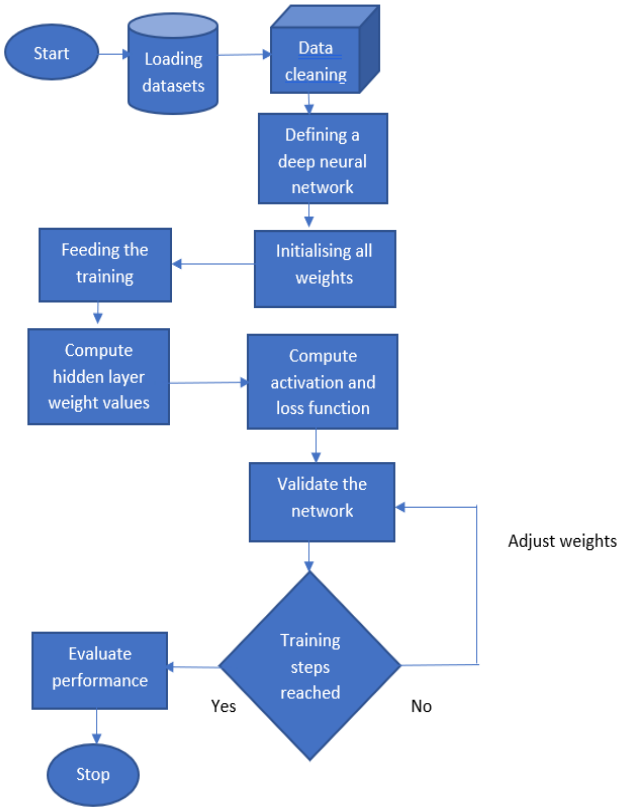

Fig. 1. Flow diagram

The accuracy will be based on the number of layers in the model. Then the values from hidden layer is again added with weights and applying activation function gives output to the output layer. Then the predicted values are found.

Feed Forward Neural Network can be visualized as a collection of neurons which act as processing units, constituting multiple layers as shown in Fig.2.

Data is fed into the input layer where the input is processed and the input is passed to the next consecutive layers . This process is repeated until the desired output is achieved in the output layer. The detailed algorithm of the proposed work is shown in Table I.

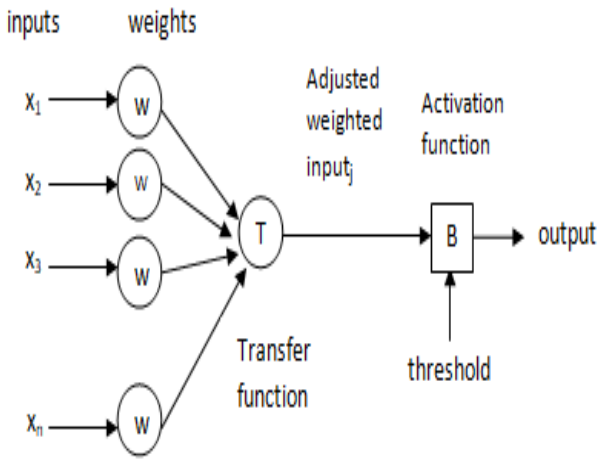

Fig. 2. Feed Forward Neural Network

Table – I: Algorithm for Rainfall Prediction

ALGORITHM FOR RAINFALL PREDICTION:
Rainfall Prediction (Input, Hidden, Output)
Begin
Import pandas and keras
Read Datasets
Perform Data cleaning
Initialize all network weights around 0-0.5 randomly
For each(x, t) in T
Begin
Propagate the input forward through network
Compile for every (p, q, r) by applying mean square error loss
Splitting datasets into training (80%) and testing (20%)
Print training and testing data
Fitting the model for every 1500 epochs
Predict the result and save the model
Calculate and plot mean square error
Plot the actual vs. predicted data
Print summary
End

Input layer

Input layer receives input in the form of patterns it communicates with one or more hidden layers. Our proposed system consists of single input layer with 20 nodes and 6 input shapes.

Hidden layer

Our proposed system consists of single hidden layer with 10 nodes. The values in the hidden layers are multiplied by the weights and are added with the bias function which can be otherwise called as activation function. The activation function used in the proposed work is sigmoid which is calculated as

$$f(x) = 1 / 1 + e^{-x} \quad (1)$$

The range of the above sigmoid function varies between 0 and 1.

Output layer

It is responsible for transforming predicted result to the outside world, which consists of single output layer with one node.

C. Training of datasets using FFNN model

Datasets are separated into training and testing datasets. The training data constitutes 80% of data that are used for training the proposed model.

The testing data comprises of remaining 20% data which are used for testing the model. The need for splitting the training set and testing is for better accuracy and evaluation.

D. Validation of datasets using FFNN model

The FFNN model is built using the training data. Then the proposed model is tested using the test data which is already designated as 20% of the actual data.

E. Evaluation of the model

The predicted values are found using above FFNN algorithm. Then based on the predicted values of the rainfall predicting factor, the result and summary are arrived.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The input parameters or factors that are used for this study includes temperature, pressure, wind speed, cloud cover, humidity and pollution which is selected so as to get accurate output. Then the null values in the datasets are dropped from the collected datasets using certain specified functions.
The comparison of rainfall predicting factors such as precipitation, pollution, wind speed, cloud cover and temperature is depicted in Fig. 3.

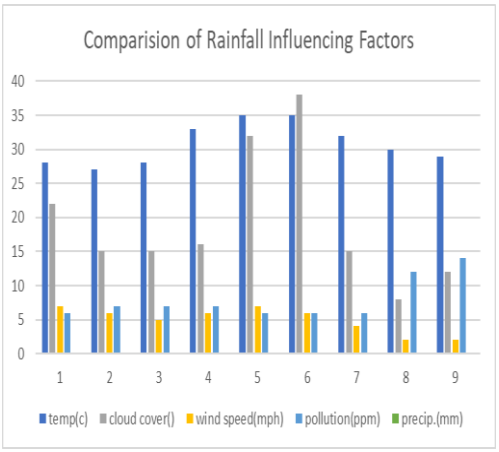

Fig. 3. Comparison Graph of input parameters

Fig. 4. shows the result of comparison of humidity and pressure. It can be seen that the factor pressure has a greater influence on rainfall when compared with the humidity factor.

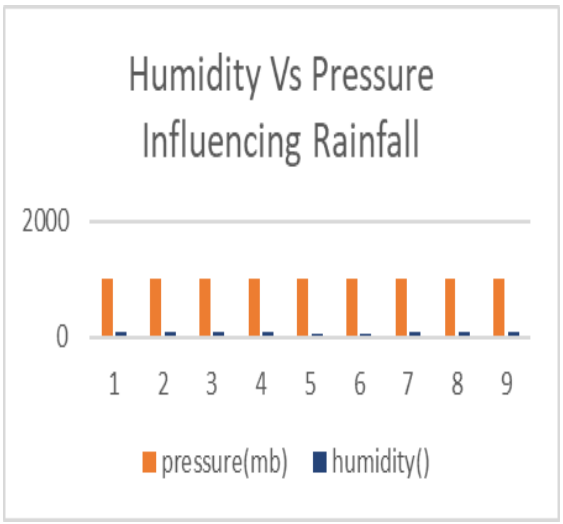

Fig. 4. Pressure Vs Humidity comparison graph

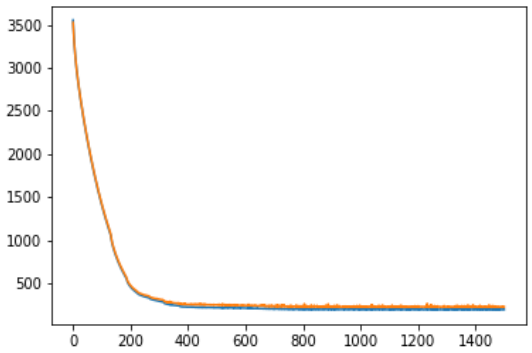

Fig. 5. Prediction of Mean Error

Root Mean Square Error (RMSE) is defined as the variance between the actual values and the predicted values. The value of RMSE is calculated as

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \quad (2)$$

$\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n$ are predicted values
$y_1, y_2, \ldots, y_n$ are observed values
n is the number of observations

The proposed work calculates RMSE value using the conventional formula and the results are shown in Fig. 5. Thus the proposed model effectively predicts the rainfall with the mentioned input parameters.

V. CONCLUSION AND FUTURE WORK

The proposed model employs a Feed Forward Network in order to predict the rainfall using various input parameters with almost effective accuracy. In future certain complex algorithms with more input parameters can be considered for achieving the high accuracy rate of predicting the rainfall.

REFERENCES

1. M.N. Ahuna ,T.J. Afullo, A.A. Alonge, Rain Attenuation Prediction Using Artificial Neural Network for Dynamic Rain Fade Mitigation.SAIEE Africa Research Journal Volume:110 issue:1, 2019
2. Aswin.S, et al., "Deep Learning Models for the Prediction of Rainfall", International Conference on Communication and Signal Processing, 2018.
3. T.Kashiwao, et al., "A neural network-based local rainfall prediction system using meteorological data on the internet: A case study using data from the Japan Meteorological Agency", Applied Soft Computing, vol.56, pp.317-330, ISSN 1568-4946, 2017.
4. Reza Mohammadpour, Zahra Asaie, Mohammad Reza Shojaeian, Mehdi Sadeghzadeh, "A hybrid of ANN and CLA to predict rainfall. Arabian Journal of Geosciences", 2018.
5. Devanshi Shukla, Vidhi Rajvir, Dr.Maulika S.Patel, "Rainfall Prediction using Neural Network", IEEE, 2018
6. Shilpa Hudnurkar, Neela Rayavarapu, "Prediction of Artificial Neural Network Nowcasting Summer Monsoon Rainfall: A case Study", IEEE Punecon, 2018.
7. E G Wahyuni, L M F Fauzan,F Abriyani, N F Muchlis,M Ulfa, "Rainfall Prediction with backpropagation method", International Conference on Mathematics, Science and Education (ICMSE), 2017.
8. R.Arya, et al., "Rainfall Prediction Using an Optimized Genetic-Artificial Neural Network Model", International Journal of Pure and Applied Mathematics, Volume 119, No. 10, 669- 678, 2018.
9. https://towardsdatascience.com/what-does-rmse-really-mean-806b65f2e48e

AUTHORS PROFILE

Poornima S received her Master degree in Information Technology from MIT Campus, Anna University, Chennai. She has 7 years of teaching experience. Her area of research includes Big data analytics, Information Retrieval and Speech processing.

Devi S received her B.E degree in Computer Science and Engineering and M.E degree in Computer Science Engineering from Anna University. She has 6 years of teaching experience. Her area of research includes Big Data Analytics, Security issues in Data Analysis, Natural Language Processing.

Oviya D Studying B.Tech Information Technology in Coimbatore Institute of Technology, Coimbatore. Interested in the research area of Data analytics and Machine learning

A. Suhana Taj Studying B.Tech Information Technology in Coimbatore Institute of Technology, Coimbatore. Interested in the research area of Data analytics and Machine learning

G. Tamil Elakkiya Studying B.Tech Information Technology in Coimbatore Institute of Technology, Coimbatore. Interested in the research area of Data analytics and Machine learning