ISSN: 2277-9655
Impact Factor: 1.852

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

# Rainfall Forecasting Using Artificial Neural Network: A Data Mining Approach

**Amruta A.Taksande[*1], Dr.S.P.Khandait[2], Prof.Manish Katkar[3]**
Department of Computer Science and Engineering,RTMNU, India
ttejswini71@gmail.com

## Abstract

Rainfall forecasting or Weather forecasting has been one of the most challenging problems around the world because it consists of multidimensional and nonlinear data such as in the field of agriculture to determine initial growing season. Recently, climate change causes much trouble in rainfall forecasting. This paper describes five data mining algorithms namely neural network (NN), random forest, classification and regression tree (CRT), support vector machine (SVM) and k-nearest neighbour. Generally these algorithms are used for the prediction. In this paper rainfall forecasting using Artificial Neural Network (ANN) and Genetic Algorithm ( GA) is made. In genetic algorithm we use Hidden Markov Model (HMM) for records the previous data. The data used within this research is taken from Yahoo Weather API is the type of Interface. Those data include temperature, air pressure, rainfall, relative humidity , and wind speed. Based on experiment result, it can be concluded that the combination of GA and HMM weather data can gives prediction graph with higher than 90% accuracy with several population size and crossover probability.

**Keywords**: Data Mining Algorithms, Prediction, Artificial Neural Network, Genetic Algorithms, and Weather Forecasting.

## Introduction

Weather forecasting is the application of science and technology to predict the state of the atmosphere for a given location. Human beings have attempted to predict the weather informally for millennia, and formally since the nineteenth century. Weather forecasts are made by collecting quantitative data about the current state of the atmosphere on a given place and using scientific understanding of atmospheric processes to project how the atmosphere will evolve on that place. Weather forecasting is one of the most imperative and demanding operational responsibilities carried out by meteorological services all over the world. The term "monsoon" seems to have been derived either from the Arabic mausin or from the Malayan monsin..As first used it was applied to southern Asia and the adjacent waters, where it referred to the seasonal surface air streams which reverse their directions between winter and summer, southwest in summer and northeast in winter in this area. In 1686 Halley explained the Asiatic monsoon as resulting from thermal contrasts between the continent and oceans. Rainfall is one of several important factors affecting watershed water quality. The downstream flux of nitrogen and phosphorus originating in the watershed basin depends on the amount of rainfall. Once an all-human endeavour based mainly upon changes in barometric pressure, current weather conditions, and sky condition, weather forecasting now relies on computer-based models that take many atmospheric factors into account. Present paper endeavours to develop an Artificial Neural Network (ANN) model to forecast average rainfall during summer-monsoon in India. Human input is still required to pick the best possible forecast model to base the forecast upon, which involves pattern recognition skills, tele connections, knowledge of model performance, and knowledge of model biases.The chaotic nature of the atmosphere, the massive computational power required to solve the equations that describe the atmosphere, error involved in measuring the initial conditions, and an incomplete understanding of atmospheric processes mean that forecasts become less accurate as the difference in current time and the time for which the forecast is being made (the *range* of the forecast) increases. There are a variety of end uses to weather forecasts. Weather warnings are important forecasts because they are used to protect life and property. Forecasts based on temperature and precipitation are important to agriculture, and therefore to traders within commodity markets. Temperature forecasts are used by utility companies to estimate demand over coming days. On

ISSN: 2277-9655
Impact Factor: 1.852

an everyday basis, people use weather forecasts to determine what to wear on a given day. Since outdoor activities are severely curtailed by heavy rain, snow and the wind chill, forecasts can be used to plan activities around these events, and to plan ahead and survive them. Thus, rainfall forecasting can warn of happening flood or drought so that peoples can save their lives and properties. Rainfall forecasting is also important for engineering applications, mainly for the design of hydroelectric power projects, because this system requires prior information about average rainfall, maximum/minimum rainfall, maximum intensity, duration etc. for a year/each month. Thus, we believe that precise rainfall prediction is important for practitioners who are interested to make wise policies related to this event. In this paper we implement the Neural network, Genetic algorithm and Hidden Markov Model for the future prediction and optimization.

## A Brief Literature Surveys

In paper [1], five data-mining algorithms, neural network, random forest, classification and regression tree, support vector machine, and *k*-nearest neighbour were used to build the prediction models. In paper [2] ,rainfall forecasting system using fuzzy system based on genetic algorithm (GA) is made. This paper [3] investigates the use of weather ensemble predictions in the application of ANNs to load forecasting for lead times from one to ten days ahead. This article [4] , presents a comparison of two sub sampling nonparametric methods for designing algorithms to forecast time series from the cumulative monthly rainfall. Both approaches are based on artificial feed-forward neural networks (ANNs).

## Data Mining Algorithms

Five data-mining algorithms, neural network (NN), random forest, classification and regression tree (C&RT), support vector machine (SVM), and *k*-nearest neighbour (*k*-NN) were used to build the prediction models. NN consists of a group of interconnected neurons, making it an adaptive system that can change its structure based on external or internal information flowing through the network during the learning phase. NNs are usually used to model complex relationships between input and output variables.. Random forest combines decision tree predictors in a way that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. It integrates a bagging idea and a random selection of features in constructing a collection of decision trees. C&RT, popularized by Breiman, is a nonparametric technique producing logical if–then rules that are easy to interpret. An SVM is a supervised learning method used for classification and regression analysis. SVM constructs one or a set of hyper planes in a high or infinite dimensional space. The key advantage of SVM is the use of kernel functions making SVM suitable for modelling in complex nonlinear domains. *k*-NN is an instance-based learning method accounting for contributions of the neighbours.. It offers good performance for some classes of applications.

## Data and Methodology

In this paper, we use Artificial Neural Network (ANN) and Genetic Algorithm (GA) for weather forecasting. Neural networks are widely applied in areas such as prediction system modelling, and control .

A neural network is good at learning using some algorithms such as a genetic algorithm (GA) and back propagation. Traditionally neural network has three layers (input, hidden, and output layers) of nodes connected in a layer-to-layer manner. The universal approximation theorem for neural network states that a two-layer feed forward perceptron network with non constant, bounded, and monotone-increasing continuous activation function can perform arbitrary nonlinear input-output relationship mapping . A neural network learns the input–output relationship through the training process. The learning process in a neural network is an interactive procedure in which its connection weights are adapted through the presentation of a set of input–output training example pairs. A novel neural network model is proposed in this paper. Two activation functions are used in the neuron and a node-to-node relationship is proposed in the hidden layer. This network model is found to be able to give better performance than the traditional feed forward neural network. A GA with arithmetic crossover and non uniform mutation can help in tuning the parameters of the proposed network. Numerical examples (three-inputs XOR problem and sunspot number forecasting) are used to test the proposed network and good results are obtained. Two applications are also given, which are short-term daily load forecasting and pattern recognition. The GA is widely applied in optimization problems where the number of parameters is large and the analytical global solutions are difficult to obtain. It has been applied in different areas such as fuzzy control, path planning, greenhouse climate control, modelling and classification etc. This paper introduces proposed Neural Network. Training of the neural network with the GA is presented in this paper.

A data is collected from Yahoo Weather API. That data included temperature, humidity, pressure, wind speed etc. Yahoo Weather API in nothing but the type of interface. In this interface there is special unique code for the particular city that included percentage of input parameters that are responsible for rainfall prediction. These input parameters can be saved as a database server for the future prediction.from the percentage of these input parameters we can predict there is a rainfall or not.

## Conclusion

In this paper we have presented a framework to develop neural network estimates of rainfall. Among the five data-mining algorithms tested in this paper, the MLP(multilayer perceptron) has performed best ANNs are being used increasingly for the prediction and forecasting of a number of water resources variables, including rainfall, flow, water level and various water quality parameters. In most papers, a good description of basic ANN theory, the case study considered and the results obtained is given. This paper modelled the complex multi-dimensional behaviours of rainfall for a number of stations using several soft computing techniques. In this way we are collecting the weather information from yahoo weather API, the information such as pressure, temperature, humidity etc. These information will be stored as a database server for the future prediction. Our future research plan is to model the daily monsoon rainfall data, which have significance in the field of agriculture, transportation, sports, tourism activities and others by using Genetic algorithm (GA) and Hidden Markov Model (HMM).

## *References*

[1] *Andrew Kusiak,"Modelling and Prediction of Rainfall Using Radar Reflectivity Data: A Data-Mining Approach" IEEE trans on geo science and remote sensing, VOL.51, NO.4,APRIL 2013, pp.2238-2239.*

[2] *Fhira Nhita , Adiwijaya ,"A Rainfall Forecasting using Fuzzy System Based on Genetic Algorithm" 2013 International Conference of Information and Communication Technology (ICoICT).*

[3] *James W. Taylor and Roberto Buizza, "Neural Network Load Forecasting With Weather Ensemble Predictions" IEEE TRANSACTIONS ON POWER SYSTEMS, VOL. 17, NO. 3, AUGUST 2002*

[4] *J. Pucheta, C. Rodriguez Rivero, Member, IEEE, M. Herrera, C. Salas and V. Sauchelli "Rainfall Forecasting Using Sub sampling Nonparametric Methods" IEEE LATIN AMERICA TRANSACTIONS, VOL. 11, NO. 1, FEB. 2013.*

[5] *Stephen Dunne;Bidisha Ghosh, "Weather Adaptive Traffic Prediction Using Neuro wavelet Models" , IEEE trans on intelligent transportation system, VOL. 14, NO. 1, MARCH 2013, pp.370.*

[6] *Dezhi Li;Wilson Wang; Fathy Ismail,"Fuzzy Neural Network Technique for System State Forecasting" IEEE TRANSACTIONS ON CYBERNETICS, VOL. 43,OCTOBER 2013, NO. 5.[7] Carlos Domenech;Tobias Wehr, "Use of Artificial Neural Networks to Retrieve TOA SW Radiative Fluxes for the Earth CARE Mission" IEEE trans on geo science and remote sensing,VOL.49,NO.6,JUNE 2011,pp.1841-1843.*