A Project Based Seminar Report

on

"Artificial Neural Network in Rainfall Prediction"

Submitted to the

Savitribai Phule Pune University

In partial fulfillment for the award of the Degree of

Bachelor of Engineering

in

Information Technology

by

Pankaj Satyawan Bagal

Roll No: T150058511

Class: TE-9

Under the guidance of

Kavita A. Sultanpure

Department of Information Technology

Pune Institute of Computer Technology

Sr. No 27, Pune-Satara Road, Dhankawadi, Pune - 411 043

2018-2019

CERTIFICATE

This is to certify that the project based seminar report entitled "**Artificial Neural Network in Rainfall Prediction**" being submitted by **Pankaj Satyawan Bagal (Exam No. T150058511 / Roll No. 3906 & Div. TE-09)** is a record of bonafide work carried out by him/her under the supervision and guidance of **Kavita A. Sultanpure** in partial fulfillment of the requirement for **TE (Information Technology Engineering) – 2015 course** of Savitribai Phule Pune University, Pune in the academic year 2018-2019.

Date: 05/04/2019

Place: Pune.

Kavita A. Sultanpure   Dr. B. A. Sonkamble

Guide   Head of the Department

Dr. P. T. Kulkarni

Principal

This Project Based Seminar report has been examined by us as per the Savitribai Phule Pune University, Pune requirements at Pune Institute of Computer Technology, Pune – 411043 on

. . . .

Internal Examiner   External Examiner

ACKNOWLEDGEMENT

I am extremely grateful to Dr. P. T. Kulkarni, Principal, Pune Institute of Computer Technology and Dr. B. A. Sonkamble, Head of Department, Department of Information Technology, for providing all the required resources for the successful completion of my seminar. My heartfelt gratitude to my seminar guide Kavita. A. Sultanpure, for his valuable suggestions and guidance in the preparation of the seminar report. A special thanks to my team who made this seminar a great success by giving their valuable suggestions.

Furthermore, my sincere thanks to all academic staff members and the technical staff and friends for all the help and co-ordination extended in bringing out this seminar successfully in time. I am very much thankful to my parents who guided me in every step I took. At last I want to thank everyone who gave their moral support and enabled me to complete this seminar successfully.

PANKAJ SATYAWAN BAGAL

(Students Name & Signature)

ABSTRACT

The country like India, where most of the farmers depend on monsoon for their crop, it is required to be having advance knowledge of actual rainfall. Many states of India also suffered from the flood whereas some state is facing the problem of drought. Rainfall is recognizing as one of the main important element of the hydrological process. Accurate rainfall forecasting will help in evaluating drought and flooding situations in advance. Therefore, it is significant to have a perfect model for rainfall forecasting. Data mining is a set of techniques used to extract unknown pieces of information from the large database repository. There are various data mining techniques available to extract valuable and useful information from spatial, temporal, sequencing and time series data.

Statistical techniques fails to provide good accuracy for rainfall prediction due to change in atmospheric data dynamically. Artificial Neural Network handles non-linearity of data very neatly. So, in this seminar comparison will be done between various techniques for rainfall prediction and advantage of using ANN in rainfall prediction over other techniques.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Introduction to Project

Rainfall is major component of the hydrological process. It takes significant part in evaluating drought and flooding events, transportation activities, agricultural tasks, flight operations ,ongoing construction projects etc. Data mining techniques and neural network methods gives a better approach to handle large and dynamic changing weather data. We discuss advantages, obstructions and challenges over these techniques that might come during the implementation.

1.2 Motivation behind the project topic

India is an agricultural country and its economy is largely based upon crop productivity and rainfall. For analyzing the crop productivity, rainfall prediction is require and necessary to all farmers. Untimely heavy rainfall causes large scale floods such as Mumbai floods of 2006 Uttarakhand crisis and Kerala floods. Many states of India also suffer from the flood whereas some states are facing the problem drought. This might have prevented if there was prior information about such situations.

1.3 Aim and Objective(s) of the work

The aim of this project is to study some techniques such as Decision Tree algorithm, Support Vector Machine, Artificial Neural Network and Time series data analysis that can achieve efficient solution for untimely rainfall and subsequent consequences.

Objectives are :

   - To identify flaws in existing models for rainfall prediction

   - Detecting reasons behind the failures of previously implemented models

   - Finding solutions on these failures

   - To identify specific regions that are continuously being affected

- To compare and select best solution among found solutions

- To compare results of proposed solution with previously existing model's results

1.4 Introduction to seminar topic

Artificial Neural Network (ANN) is a computing system based on biological neural networks that constitute human brains. The neural network is not itself algorithm rather it is a framework for many different machine learning algorithm to work together and process complex data inputs. Generally weather dataset consist of several characteristics and large number of records. ANN is capable of handling such dynamically changing dataset and has ability to converge to accurate output.

CHAPTER 2

LITERATURE SURVEY

Kumar Abhishek and et al (2017) [1] implemented multi-layered backpropagation neural network forecasting weather. It showed that as number of neurons increases in ANN, the MSE decreases. In the paper they have compared various functions and their ease to use it in accuracy model. Model building process is illustrated well i.e collection of data, cleansing , model construction, training and testing in this paper.

Michaelides et al (1995) [2] compared the performance of ANN with multiple linear regressions in estimating missing rainfall data over Cyprus.

Mohini P. Darji and et al (2015) [3] discussed the issues while applying different neural networks for yearly/monthly/daily rainfall forecasting. They have presented different accuracy measures used by researchers for evaluating performance of ANN. It shows that Neurals Networks such as RNN, FFNN and TDNN are more suitable to predict rainfall than other forecasting techniques such as statistical and numerical methods.

Akash Parmar and et al (2017) [4] made a review on various machine learning techniques that are used in predicting rainfall prediction. The paper showed effects of non linear relationships in rainfall data.

P. Goswami and Srividya (1996) [5] have combined RNN and TDNN features and conclusion of their work was that composite models gives better accuracy than the single model.

SR. NO.	Title	Author	Description
1.	A Rainfall prediction model using Artificial Neural Network	Kumar Abhishek, Rajiv Ranjan, Sarthak Kumar	Model building process and multi layered BPNN
2.	Rainfall forecasting using neural network : A survey	Mohini P. Darji, Vipul K.Dabhi, Harshadkumar B. Prajapati	Issues while applying neural network techniques
3.	Machine Learning Techniques For Rainfall Prediction	Aakash Parmar, Kinjal Mistree, Mithila Sompura	Categorization of Approaches of Rainfall Prediction
4.	A novel neural network design for long range prediction of rainfall pattern	P. Goswami and Srividya	Composite models for long range prediction

Table 2.1: Summary of Literature Survey

CHAPTER 3

ANN IN RAINFALL PREDICTION

Artificial Neural Network are computing systems vaguely inspired by the biological neural networks that constitute animal brains. The neural network itself is not an algorithm, but rather a framework for many different machine learning algorithms to work together and process complex data inputs.

3.1 Architecture of ANN

Architecture of ANN mainly consist of three components i.e input layer, hidden layer and output layer.

3.1.1 Input Layer

At input stage number of input nodes are decided depending upon the dataset available. This nodes are also called neurons of layer. An *input neuron* has no predecessor but serves as input interface for the whole network. Similarly an *output neuron* has no successor and thus serves as an output interface of the whole network.

Formally neuron can be defined as a set of input values X_i with given weights W_i .

These neurons are given to the input function f () as summation over whole set of input values and corresponding weights.

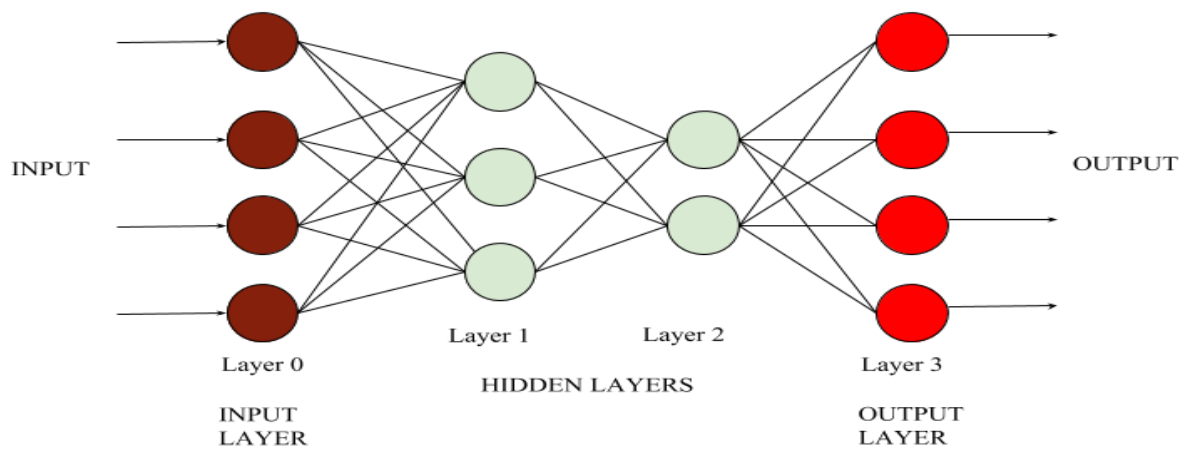$$y_{in} \qquad = \qquad \sum_{i=0}^{n} x_i . y_i$$

Figure 3.1: Architecture of ANN

3.1.2 Hidden Layer

Hidden layers are the layers in between output layer and input layer. Neurons in hidden layer are activated by activation function. This activation function helps to maps the input data based on the given weights to it.

3.1.3 Output Layer

Output of a given algorithm is collected at output layer. Based on activation function and applied algorithm some of the neurons in output layer are finally activated. These neurons are the final output the algorithm.

3.2 Background Theory

There are two widely used methods for rainfall forecasting are :

1) Statistical methods

2) Numerical Weather Prediction (NWP) model

Nature of rainfall data is nonlinear and dynamic in nature. Frequency, intensity and amount of rainfall are main characteristics of rainfall dataset. Statistical model have limitation over computational power on large scale dataset. Non linear relationships in rainfall data is not handled by traditional approaches.

Many engineers are trying to find solution on non linear rainfall data for accurately predicting the amount of rain or probability of rain. ANN has showed a very good capability of handling non linear data.

3.3 Types of Neural Network

There are several variations of neural network. Every neural network has its own advantage and disadvantages.

Types of neural networks are as follows :

3.3.1 Feed Forward Neural Network (FFNN)

A feedforward neural network is an artificial neural network wherein connections between the nodes do not form a cycle. As such, it is different from recurrent neural networks.

The feedforward neural network was the first and simplest type of artificial neural network devised. In this network, the information moves in only one direction, forward, from the input nodes, through the hidden nodes (if any) and to the output nodes. There are no cycles or loops in the network.

Figure 3.2: Feed Forward Neural Network (FFNN)

3.3.2 Back Propagation Network (BPNN)

Backpropagation is a method used in artificial neural networks to calculate a gradient that is needed in the calculation of the weights to be used in the network. Backpropagation is shorthand for "the backward propagation of errors," since an error is computed at the output and distributed backwards throughout the network's layers. It is commonly used to train deep neural networks. The output and error are feed to input layer recursively to input layer.

Figure 3.3: Back Propagation Neural Network (BPNN)

3.4 Model Proposal

3.3.1 Flow Of algorithm

Following is flowchart of model with ANN for rainfall prediction :

There are five steps for model building:

1) Network definition

2) Compiling Network

3) Network fitting with training data

4) Evaluating Network based on testing data
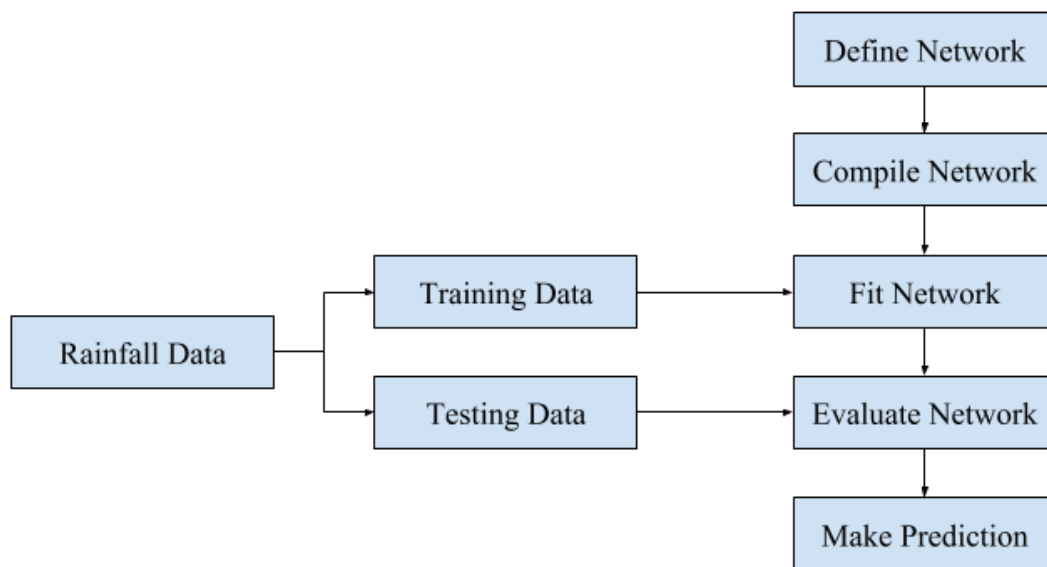
5) Prediction on test data

Figure 3.4: FlowChart of Prediction Algorithm

3.4.2 Data collection and Preprocessing

Rainfall data is collected from Indian Meteorological website. Data consists of classes of data as average annual rainfall data, minimum average temperature, maximum average temperature, mean temperature year wise. Data is collected over time period of 115 years(1901-2015).

All records with "NA" values are removed from dataset. Mean rainfall data was not present so, mean of annual rainfall is calculated for each year.

YEAR	MIN TEM-PERATURE	MAX TEM-PERATURE	MEAN TEM-PERATURE	MEAN ANNUAL RAINFALL

Table 3.1: Characteristics in dataset

3.4.3 Implementation and Result

Simple rainfall prediction model is implemented using ANN model in python using keras library.

In the model input layer is designed with 100 nodes. Two hidden layers are designed with 50 nodes and 20 nodes and one output layer is designed with one output node. "relu" activation function is used to activate nodes at input layer and in hidden layers. After designing model, model was trained on input training dataset with 100 epochs. After training of model, model is ready for prediction. With the help of pre compiled model, predictions were made on testing data. The accuracy of prediction is measured on the basis of mean squared error.

| SR. NO.	Parameter	Description
1.	Model	ANN with no feedback
2.	Dataset	Rainfall Data of IN-DIA from 1901-2015
3.	Input Parameter	Temperature classes
4.	Output Parameter	Mean Annual Rainfall
5.	Accuracy Measure	Mean Squared Error
6.	Python libraries	Keras, sklearn, pan-das, numpy, matplotlib
7.	Layers	3 (100 -> 50 -> 20 -> 1) nodes/layer
8.	Mean squared error	91.54650502838729
9.	Mean Absolute Percentage Error	6.586825506065352 %

Table 3.2: Model Overview
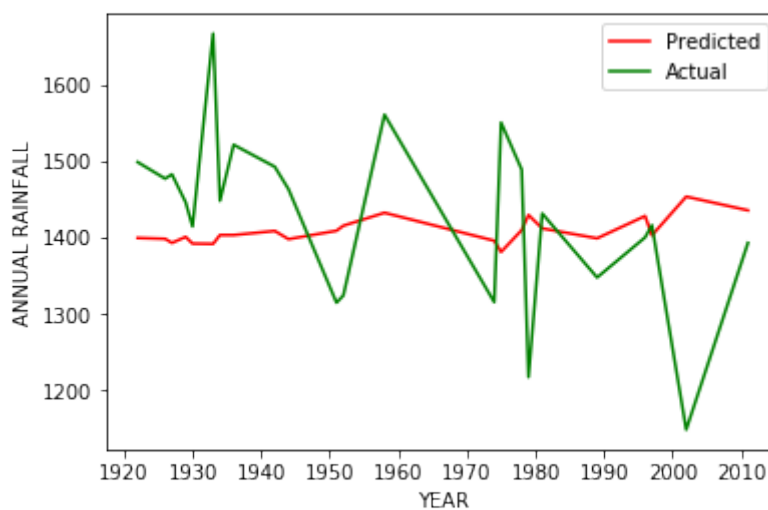
Figure 3.5: Actual VS Predicted

3.5 Advantages and Disadvantages of using ANN for Prediction model

- Advantages of using ANN in rainfall prediction :

1) ANNs have the ability to learn and model non-linear and complex relationships, which is really important because in real-life, many of the relationships between inputs and outputs are non-linear as well as complex.

2) ANNs can generalize—After learning from the initial inputs and their relationships, it can infer unseen relationships on unseen data as well, thus making the model generalize and predict on unseen data.

3) Unlike many other prediction techniques, ANN does not impose any restrictions on the input variables (like how they should be distributed).

4) Many studies have shown that ANNs can better model heteroskedasticity i.e. data with high volatility and non-constant variance, given its ability to learn hidden relationships in the data without imposing any fixed relationships in the data.

- Disadvantage of ANN in rainfall prediction :

1) It is non trivial to extract what is the process going underneath with regards to the decisions the ANN is making (it can be done but it is not as straightforward as other algorithms.

2) Speed of convergence to particular result is slow in Artificial Neural Network.

3) Also it has disadvantage such as less generalizing performance, arriving at local minimum and over fitting problems.

CHAPTER 4

ENHANCEMENT

In the above ANN model more characteristics like amount of moisture, humidity, wind speed etc. can be added to obtain better accuracy. Characteristics such as moisture,wind speed, humidity etc. gives more information about rainfall predictability. Such characteristics will make model more effective and logical. A particular characteristics can be made as dominant characteristics by assigning more weight to it.

Error occurred in above model can be used as input to first layer to improve accuracy. It simply implies that "Learn from mistakes" . Thus back propagation will make model more effective.

CHAPTER 5

CONCLUSION

ANN is very effective technique in the prediction model also it has ability to "learn from past mistakes" . Rainfall prediction model with appropriate number of hidden layer(s) and node(s) per layer makes model best fit for prediction. By using keras and other python libraries, ANN model for rainfall prediction is partially implemented to predict Mean Annual Rainfall in (mm). Thus, by this seminar architecture of ANN and it's partial implementation in rainfall prediction is studied very effectively.

CHAPTER 6

BIBLIOGRAPHY

[1] Kumar Abhishek, Abhay Kumar, Rajeev Ranjan, Sarthak Kumar, "A rainfall prediction model using artificial neural network," *2012 IEEE Control and System Graduate Research Colloquium, Shah Alam, Selangor, Malaysia, 2012*. Doi: 10.1109/ICCV.2015.55

[2] Michaelides,S.C, Neocleous C.C, and Schizas C.N, "Artificial Neural networks and multiple linear regression in estimating missing rainfall data" . In proceedings of the DSP95 International Conference on Digital Signal processing, Limassol Cyprus .pp.668-673 1995

[3] Mohini P. Darji, Vipul K. Dabhi, Harshadkumar B. Prajapati, "Rainfall forecasting using neural network: A survey" , 2015 International Conference on Advances in Computer Engineering and Applications, Ghaziabad, India. Doi : 10.1109/ICACEA.2015.7164782

[4] Akash Parmar, Kinjal Mistree, Mithila Sompura "Machine Learning Techniques For Rainfall Prediction: A Review" , 2017 International Conference on Innovations in information Embedded and Communication Systems (ICIIECS), Coimbatore, India

[5] Goswami, Prashant and Srividya, "A novel neural network design for long range prediction of rainfall pattern" , (1996) Current Science (Bangalore), 70 (6). pp. 447-457