Mahatma Jyotiba Phule Rohilkhand University, Bareilly

**Institute of Engineering and Technology**

**2021-2022**



A Project Report on

**Time Series Analysis- Rainfall Prediction Using ANN**

**CS-305P Advanced Programming Lab**

Submitted By:- Under the Guidance:- **Anurodh Kumar (19CS13) Dr. Iram Naim Aradhana Yadav (19CS14) Dr. Anil Kumar Bisht Arpit Maurya (19CS15)**

# INDEX

Certification of the Guide………………………………………………………………………………………(i)

Declaration of the project…………………………………………………………………………………....(ii)

Acknowledgement……………………………………………………………………………………………….(iii)

Abstract……………………………………………………………………………………………………………….(iv)

1. [Introduction](#_TOC_250004) 7
   1. Objective………………………………………………………………………….8
2. [Literature Review](#_TOC_250003) 9
3. [Methodology](#_TOC_250002) 11
   1. Technique Used 11
   2. Diagram Related To ANN 12
   3. Algorithm… 14
   4. [Flow chart](#_TOC_250001) 16
4. Result Analysis… 17
   1. About Dataset 17
   2. Performance… 18
   3. Accuracy of the Model 20
5. Conclusion 22

[References…](#_TOC_250000) 23

# ABSTRACT

Artificial neural networks (ANNs) are relatively new computer tools that have gained widespread use in solving many problems of the real world. The attraction of ANNs comes from their amazing information processing features that are highly correlated with ambiguity, high similarity, error tolerance and noise, and learning and generality skills. This paper aims to familiarize the reader with ANN-based computing (neurocomputing) and to serve as a useful tool an effective guide and ANN model tool kit during the development of the ANN project.

History of the emergence of neurocomputing and its relation to the field of neurobiology is briefly discussed. ANNs are compared to both professional systems and declining statistics and their benefits and limitations are presented. A variety of bird eye reviews for a variety types of ANNs and related learning rules are introduced, with particular emphasis on the theory of backpropagation (BP) ANNs and design. A common way to develop successful ANN projects from design, to design, to launch, explained.

The most common problems BPAN developers face during training are summarized in them in line with possible problems and solutions. The improved model was reasonably accurate in imitating both training and testing growth curves are time-dependent as they are affected by temperature and pH.

## **1.INTRODUCTION**

* Rainfall plays an important role in managing water resources and is one of the most important variables of the rainfall forecasting model. However, rainfall is one of the most complex and complex aspects of the hydrological cycle to understand and model due to the wide variety of space and time. The complexity of atmospheric rainfall systems makes predicting rainfall extremely difficult. Therefore, predicting rainfall is one of the major challenges in hydrology, despite many advances in climate forecasting. However, these models require a high level of data and calculation methods, one of the worst things about numerical models is the unavailability of input data in general.
* The artificial neural network (ANN) model used rainfall prediction. The model used the neural network to predict two-dimensional rain at the same time in advance, and used only current rainfall as an input of the training data set.
* The ANN and LSTM model strategies were used to develop the model for predicting future rainfall using previous rainfall. Here we have used vanilla LSTM methodology for predicting the rainfall. A vanilla LSTM is an LSTM model that has a single hidden layer of LSTM units, and output layer used to make a prediction.
* We are working with a univariate series, so the number of features is one, for one variable. After the model is fit, we can use it to make a prediction. We can predict the next value in the sequence by providing the input.
* In ANN there is no assurance of proper network structure i.e., there is no specific rule for determining the structure of artificial neural network.
* Rainfall forecasting is important for many catchment management applications, in particular for flood warning systems. This Rainfall forecasting is important to agriculture sector , alarming them and might provide better rainfall prediction for entire catchment. The reliable and true rainfall prediction helps to provide beneficiaries with information on water supply control, farm planning and their integrated crop insurance applications on monthly and seasonal time scales.

1.1 OBJECTIVE OF THE PROJECT

* To come up with strong and intuitive model for Predicting the Rainfall and Alarming the Heavy Rainfall Areas.
* Predicting the rainfall to establish a landslide early warning for investigated area’s or for area’s of heavy rainfall**.**

# 2.Literature Review

**[2.1]** Chadwick et al. (2011) **[1]** applied a synthetic neural network approach to lower GCM temperatures and rain fields to regional model scales throughout Europe.

**[2.2]** A Support Vector Machine (SVM) novel presented by Lu and Wang **[2]** (2011) to predict the monthly rainfall in Guangxi, China. Finally, they showed that predicting using the SVM composite model was generally better than that obtained using other models submitted according to the same test ratings. The authors strongly believe that it can be used as another forecast tool for the Weather app to achieve greater predictive accuracy and improve the quality of continuous prediction.

**[2.3]** Model based on the sensory networks network (ANNs) and wave decomposition proposed by Charaniya et al. (2011) **[3]** predicting monthly rainfall in accounts of previous rainfall data events. A wavelet conversion was used to extract the coefficient and details of a series of rain data. These coefficients have been used in conjunction with ANN in the processes of reading and extracting information. They tested the rainfall data model in various parts of India and across the country and found that the proposed model was able to predict monthly rainfall in less than one month.

**[2.4]** El-Shafie et al. (2011) **[4]** compared and studied two different neural networks and one dynamic neural network namely; Multi-Layer Perceptron Neural Network (MLP-NN), Radial Basis Function Neural Network (RBFNN) and Input Delay Neural Network (IDNN), respectively. Those models were developed twice on the horizon in predicting the monthly and weekly rainfall in Klang River, Malaysia. Finally, they concluded that IDNN may be suitable for modeling the temporal size of the rain pattern, thus providing better predictability.

**[2.5]** Geetha and Selvaraj (2011) **[5]** developed a network model of back propagation neural rainfall prediction in Chennai, India. The average monthly rainfall was predicted by them using that model. The model can do well both in training and in independent periods.

**[2.6]** In another study an artificial neural network-fuzzy logic-wavelet model was used to predict long-term rainfall by Afshin et al. (2011) **[6]**. The results of the integrated model showed significantly higher outcomes compared to two-year forecasts for six months and

seasons. As a result of the mean square root, the forecast for the two annual and annual seasons is 6.22 and 7.11, respectively. However, the six-month forecast shows 13.15

**[2.7]** Abhishek et al. (2012) **[7]** developed an ANN model for predicting the average monthly rainfall in the Udupi region of Karnataka, India. In Udupi, the months of April to November were identified as the rainy season in May, June, July, August, and October as the main rainy seasons. They have therefore reviewed the data for these 8 months from 1960 to 2010. Finally, three algorithms were tested in multi-layer construction: Back Propagation Algorithm (BPA), Layer Recurrent Network (LRN) and Cascaded Back-Propagation (CBP) and authors found that BPA was the best of the three tested algorithms.

**[2.8]** The Multilayer Perceptron Neural Network was proposed by Deshpande (2012) **[8]** as a clever tool for predicting the Rainy Season Series. Rain samples were collected at the Government Rainfall Monitoring Center in Yavatmal, Maharashtra Province, India. Predictions of the many further steps (1, 5, 10, 20) of this Rain Data Series were made using the Multilayer Perceptron Neural Network. They found that performance measurements such as the mean square error, as well as the standard measurement error in the test and the training data set for short-term guessing were correct compared to other networks such as the Jordon Elmann Neural Network, SOFM (Personal Feature Map.), RNN (Network of normalneural)

# 3.Methodology

Here, The Technique use to predict the future of rainfall is going to happen or not is

#### ANN (Artificial Neural Network).

We are using RNN (Recurrent Neural Network) which is a type of artificial neural network. The specific version of ANN we are dealing with is **Vanila LSTM**. A Vanilla LSTM model have **model single hidden layer of LSTM units, and an output layer used to make predictions**.

#### ANN (Artificial Neural Network)

"**Artificial Neural Network**" is derived from Biological neural networks that develop the structure similar to the human brain. It has neurons interconnected to one another, artificial neural networks also have neurons that are interconnected to one another in various layers of the networks. These neurons in ANN called as nodes.

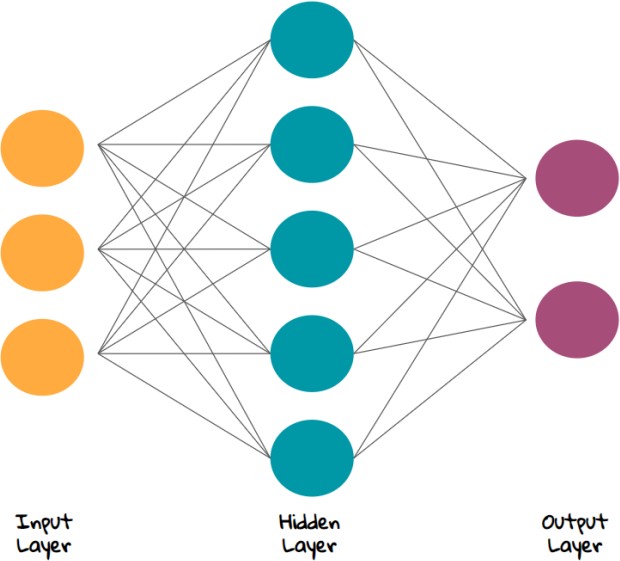


Figure 1. types of Neural Network layers

### Input Layer:

As from the name, It simply accepts the inputs from the programmer.

### Hidden Layer:

It performs all the calculations to get hidden features and patterns.

### Output Layer:

The input goes through a many transformations by the hidden layer, which finally results in output that is conveyed using this layer.

### Typical diagram of Biological Neural Network.

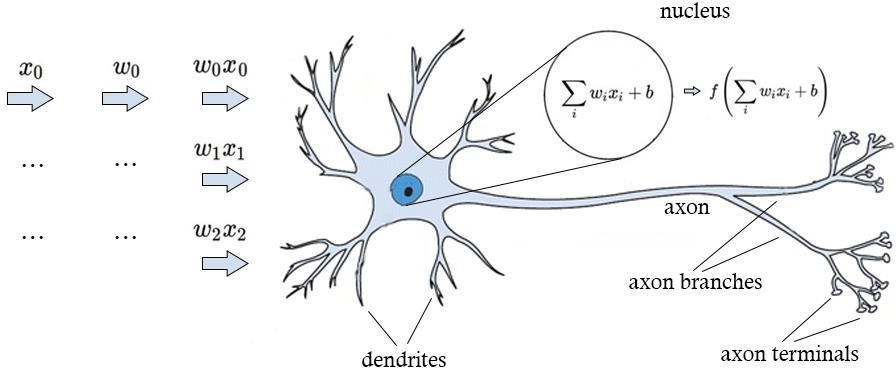


Figure 2. biological neural network

#### Typical Artificial Neural Network .

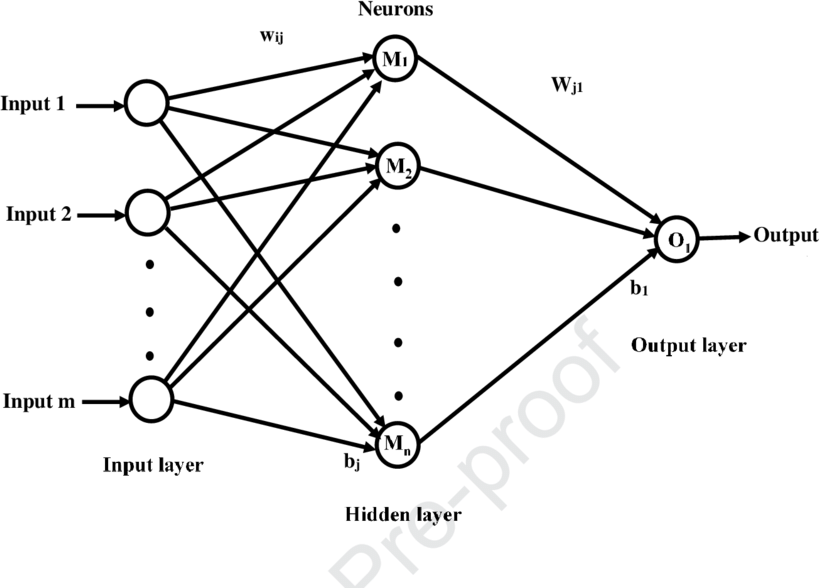
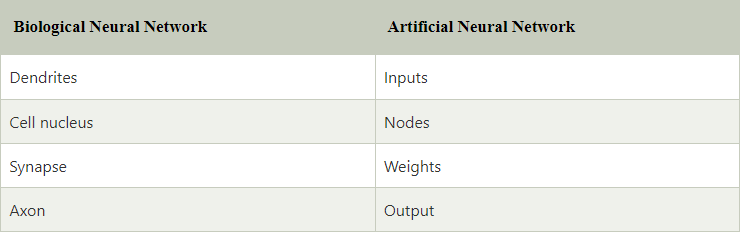


Figure 3. ANN layers

Dendrites represent inputs in Artificial Neural Networks, cell nucleus represents Nodes, synapse represents Weights and Axon represents Output.

#### Biological neural network and artificial neural network:

Table 1. Comparision between biological and artificial neural network



An **Artificial Neural Network** in the field of **Artificial intelligence** where it attempts to completely mimic the network of neurons makes up a human brain so that computers will have an choice to understand conditions and make decisions in a human like manner. The artificial neural network is perfectly designed by programming computers to behave simply like interconnected brain cells.

There are more then 1000 billion neurons in the human brain. Each neuron has one association point somewhere in the range of 1,000 and 100,000. In the case of human brain, data is stored in such a manner as to be distributed, and we can extract more than one piece of this data when necessary from our memory parallelly. We can say that the human brain is made up of incredibly amazing parallel processors that works synchronously.

What is Artificial Neural Network

Figure 4. Activation function

Let’s understand the ANN, consider one example of a digital logic gate that takes input and gives output simply. "OR" gate, which takes 2 inputs. If one or both the inputs are On that means high, then we get On in output. If both the inputs are Off or down , then we get Off in output. Here the output depends upon input. Our brain doesn’t perform the same task. The outputs to inputs keep changing because of the neurons in our brain, which are "learning."

”The ANN takes inputs in the form of data and computes the weighted sum. This computation comes from this transfer function.”

### Activation function:

The role of Activation function is to decide, whether a neuron should be activated or not by calculating weighted sum and further adding bias with it. The purpose of the activation function is to **include non-linearity** into the output of a neuron.

“We know, In neural network, it has nodes that work in correspondence of *weight, bias* and their respective activation function. In a neural network, we would update the weights and biases of the nodes on the basis of the error at the output after every cycle. This process is known as *back-propagation*. Activation functions make the back-propagation possible since the gradients are supplied along with the error to update the weights and biases, and make output more accurate, that means it comes with feedback features”

There are many type of activation function present to give a good result and accuracy. Here we are using the **RELU** Activation Function, which stands for ***Rectified linear unit.***

“It **converts the node's input into some output**. The rectified linear activation function (called ReLU) is a very high-performance network. This function takes a single number as an input, returns 0 if the input is negative, and the input if the input is positive.”

* **Equation :- *A(x) = max(0,x)***. It gives output x if x is positive and other it gives 0.

#### Value Range :- the range of value lies in [0, inf).

* **Nature :-** non-linear, which means we can easily back propagate the errors and have multiple layers of neurons being activated by the ReLU function, that means feedbacks.
* **Uses :-** ReLu is less computationally expensive than many usual activation functions because it involves simpler mathematical operations so easy.In simple words, RELU learns *much faster* than sigmoid ,Tanh function and many others too.

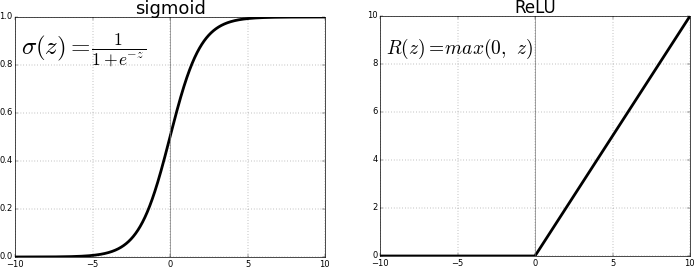
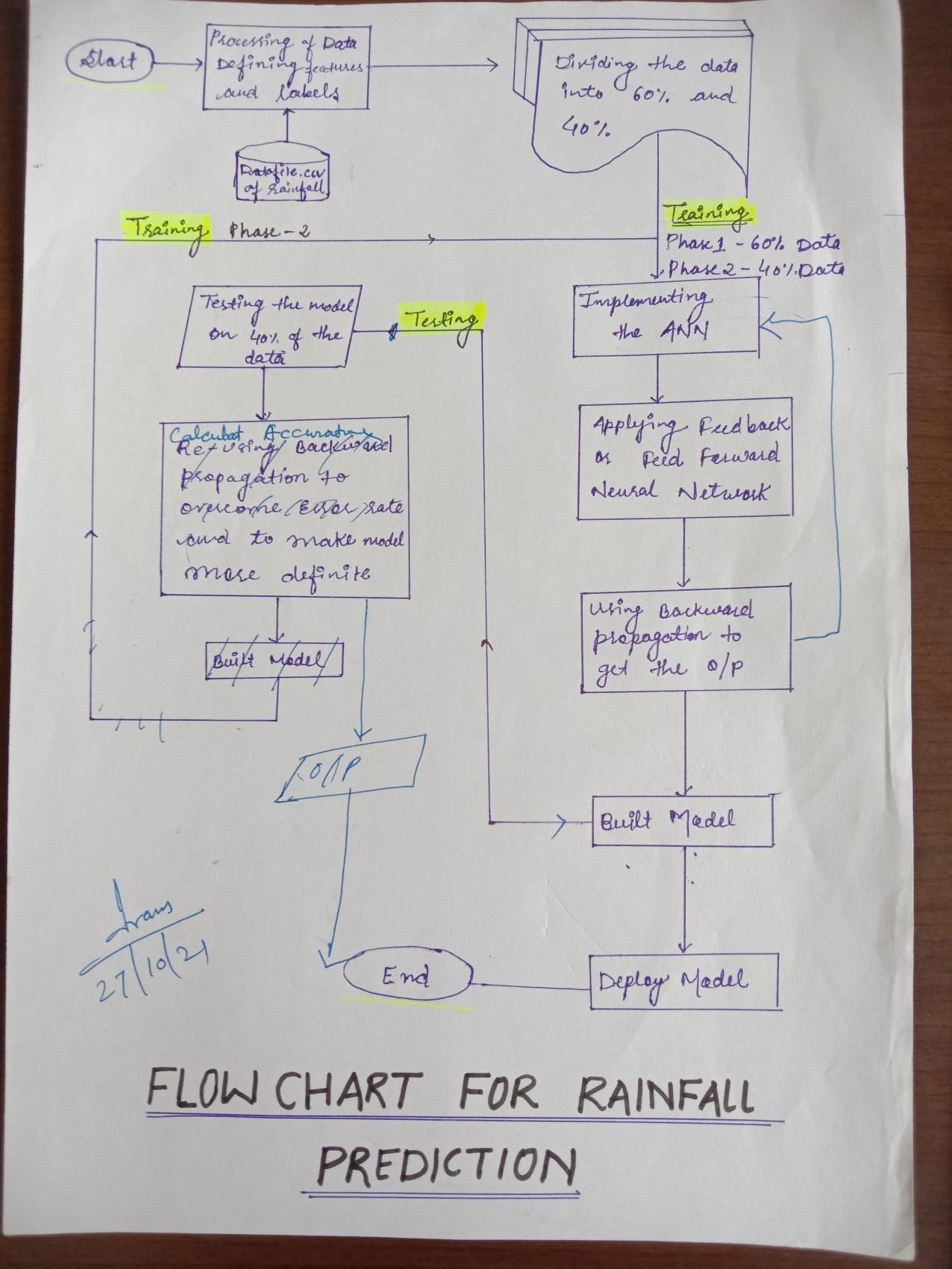


Figure 5. Equation comparision

# Flowchart



**4.Result and Analysis**

### About The Dataset

* + 1. **Source**:

We have taken our dataset from **Kaggle.com** .

Kaggle is a subsidiary of Google. It is an online platform or community which enables data scientists and other developers to engage in running machine learning contests.

#### Founder : Anthony Goldbloom Founded : April 2010

Details About Various Fields:

Since our dataset is univariate i.e a type of data which consist of observation on only a single characteristic or attribute.

We have one attribute i.e months value which describes the annual rainfall per month in year,therefore we have data month-wise as well as year-wise.

We have collected 110 years data on which we have trained our model. The rainfall data differ from month to month and seasonality also have great impact on the forecast. We have taken 80% of total data for training part while 20% data for testing purpose.

Summary Of the Dataset: No of Rows :1416

No of columns:1

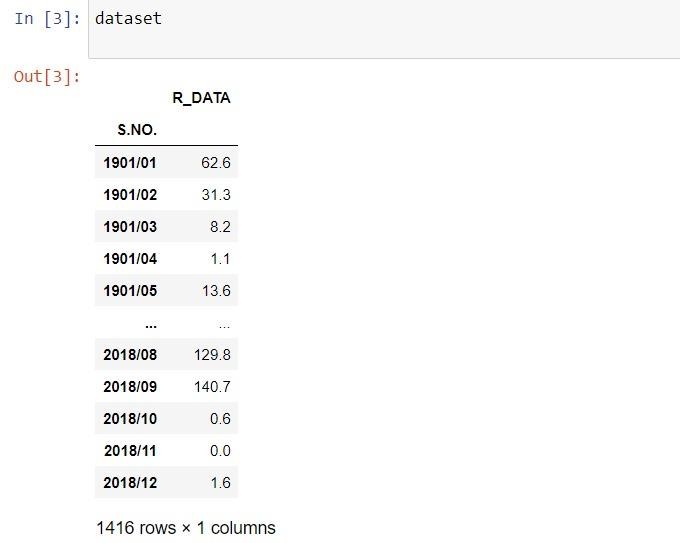
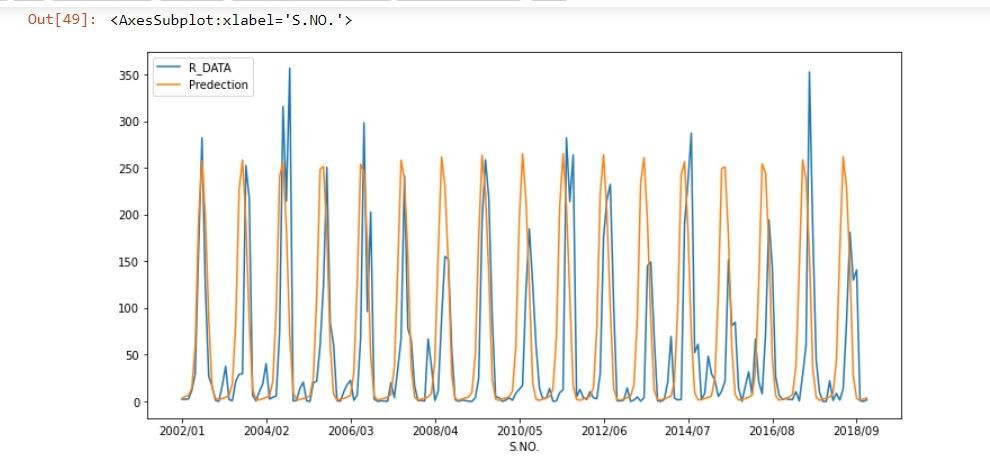


Figure 6. Dataset description

Here ( S.NO.) describes the month-year ,where as (R\_DATA) describes the annual rainfall per month on particular year.

Result Obtained:



.

Figure 7. Rainfall Prediction graph

The blue line curves defines the actual data we have taken to train our model.

While, orange coloured line defines the prediction values i.e., the prediction which has been done by our model .

### Performance :

The model we have developed were evaluated using the performance criteria, root mean square error (RMSE).

The root mean square error offer a general picture of the error involved in the prediction.

Hence lower the value of RMSE, the better is the prediction.

The ANN and LSTM model strategies were used to develop the model for predicting future rainfall using previous rainfall.

It is fair to say that the performance of the models during the training will be provided only if their performance during the test is better.

In this case, the results is consistent throughout the world, but compared to the time measure used for rainfall predicting, the results are much better. So work during training as well.

During training and testing sessions respectively with RMSE value ( ).

It is also found to be non-existent improvements in performance or with increasing nodes.

It should also be mentioned that all ANN models provide only positive and predictable rainfall forecasts differs from the observed data, especially during the testing period.

The Vanila LSTM network has led to slightly improved performance Although the development is small, the network shows a significant reduction of nodes is in the hidden layer.

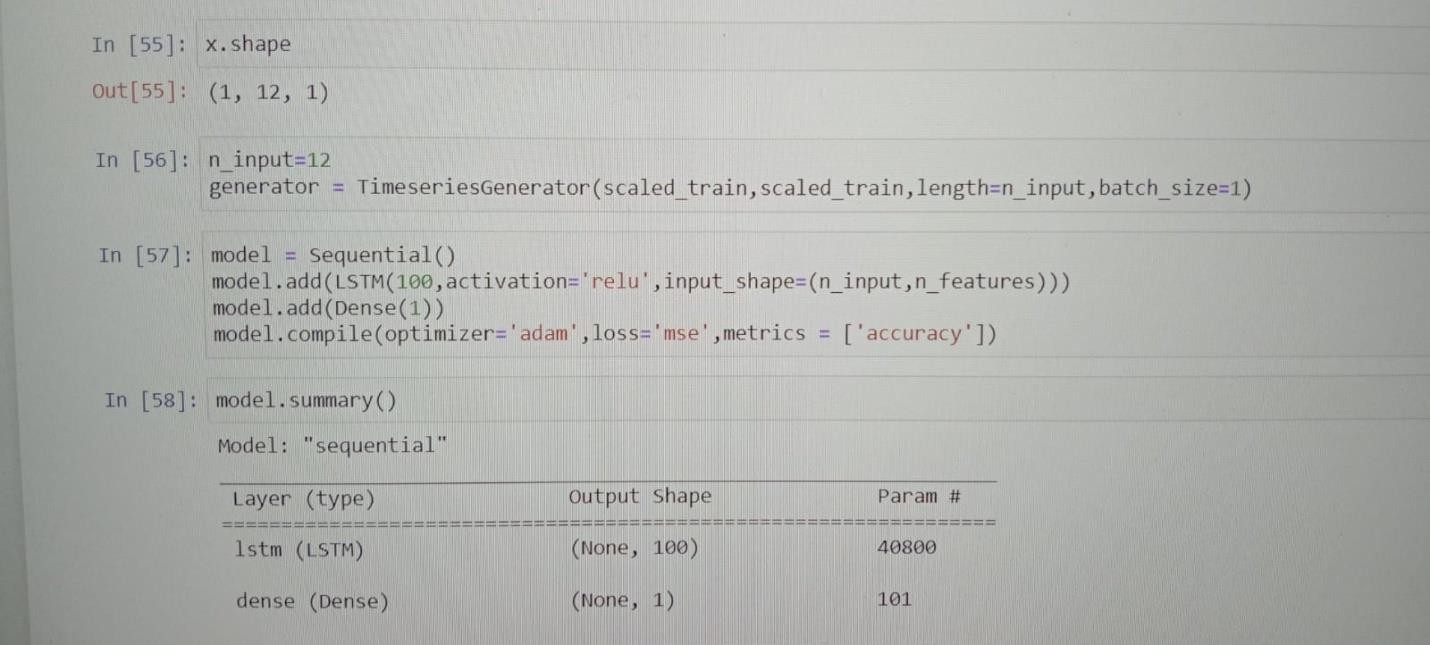


Figure 8. Model summary

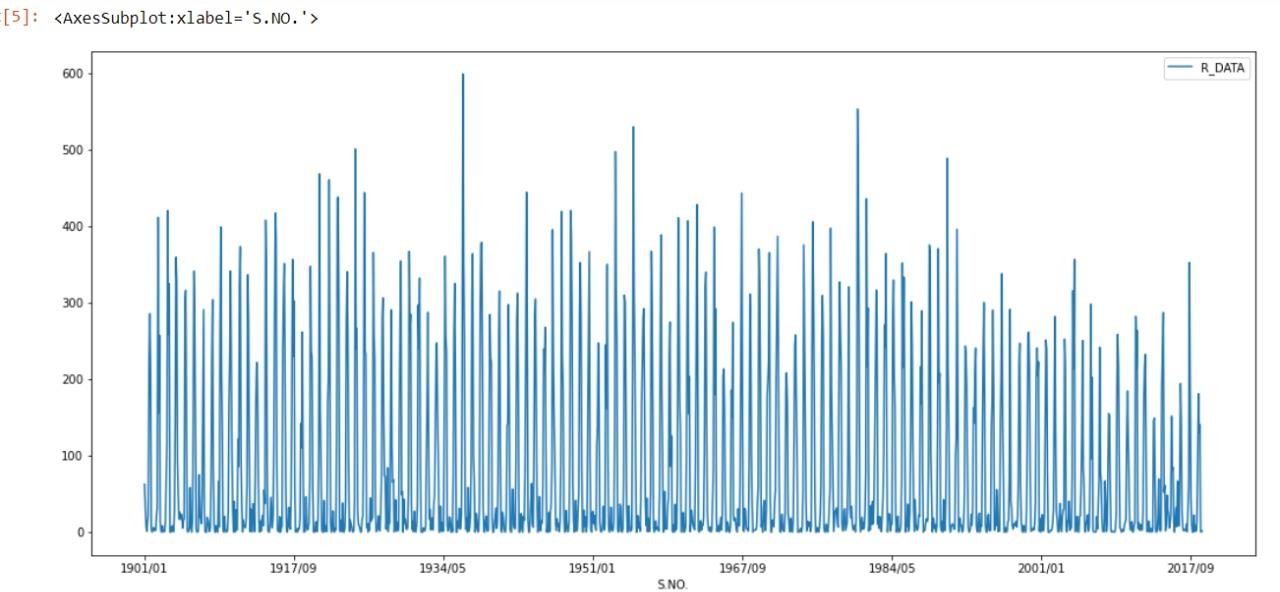


Figure 9. Graphical Representation of year-wise data

**x-axis** : defines the value of annual rainfall.

**y-axis** : defines the month-year-wise data

### Accuracy of the model

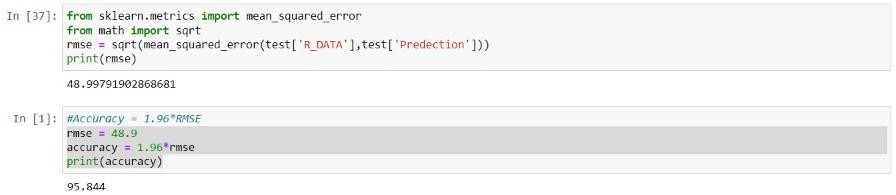


Figure 10. Accuracy

An epoch indicates the number of passes of the entire training dataset the machine learning algorithm has completed

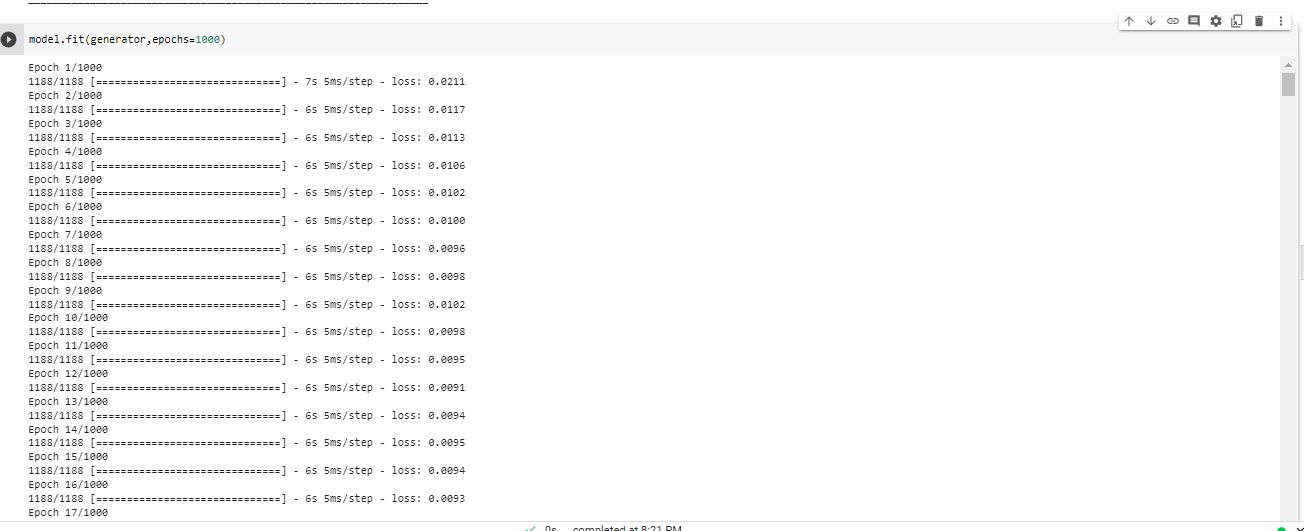


Figure 11. epoch

# 5.Conclusion

This paper report is a detailed rainfall prediction using the RNN (Recurrent Neural Network) Architecture which is a type of Artificial Neural Network for one hundred and eighteen years.

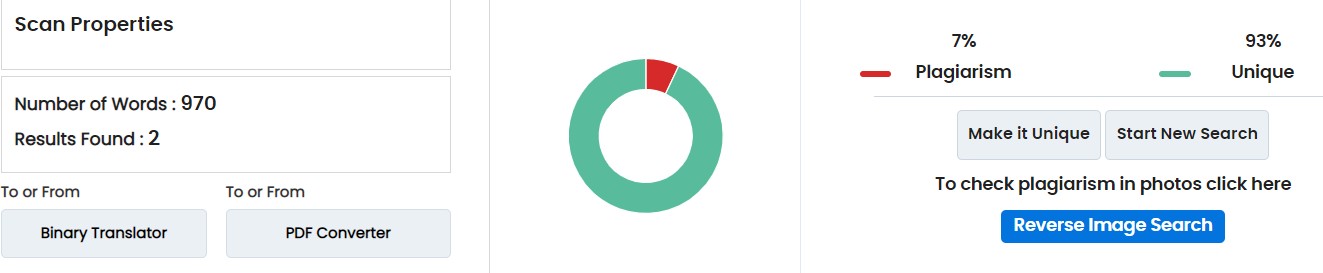
The study found that most researchers used RNN to predict rainfall and obtained significant results. The study also concluded that predictive strategies using ANN are more appropriate for predicting rainfall than other predictors such as numbers and numbers. However, some limitations have been identified. Extensive indicators that support the unique development of ANN research provided by the paper should be of great help to ANN researchers in accurately predicting future rainfall.

# References

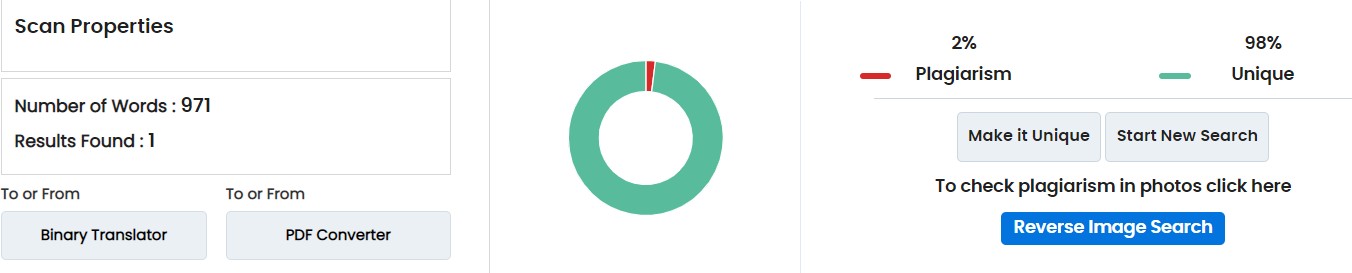
1. S. Haykin, 1999: Neural Networks- A Comprehensive Foundation. Addison Wesley Longman
2. D. Rumelhart, G. Hinton and R. Williams, (1986) Reading presentations in support of broadcast errors, Environment, 323, 533-536.
3. M. J. D. Powell, (1987) Radial basis functions for multivariable interpolation: a review. In J.C. Mason, and M.G. Cox (Eds.), Algorithms for Approximation, pp. 143-167. Oxford: Clarendon Press.
4. J Park, and I Sandberg, (1991) The measurement of the universe using radial basis activity networks. Neural Computation, 3 (2), 246-257.
5. T. Kohonen, (1990), Editing map. IEEE Procedures, 78 (9), 1464-1478.
6. W. Patterson Dan, (1996), Artificial Neural Networks: Theory and Applications, Prentice Hall.
7. G.E.P. Box & G.M. Jenkins, 1970. Timeline Analysis, Predictability and Control, Holden- Day, CA, San Francisco.
8. M.J.C., Hu, Use of the ADALINE system in weather forecasting, Technical Report, Stanford Electron, 1964.
9. D.F. Cook and M. L. Wolfe, “Neural network broadcast back to predict moderate air temperature”, AI Applications, 5, 40-46, 1991.
10. D.W. McCann, “Short-Term Neural Network Forecast Forecast”, Weather and Forecast, 7, 525-534, doi: 10.1175 / 1520-0434, 1992.
11. M. N. French, W. F. Krajewski, and R. R. Cuykendall: Predicting rainfall in space and time using the neural network, J. Hydrol., 137, 1–31, 1992.
12. Tao Chen, M Takagi, “rain forecast for local satellite imagery using artificial neural network”, GARSS, 1993, vol.3, 1247 - 1249.
13. M. Zhang, and A. R. Scofield, “Artificial Neural Network Techniques for Measuring Rain and Detecting Cloud Connectivity from Satellite Data”, Remote Sensing International Journal, 16,3241-3262, 1994

# Plagiarism Report

* + **Total words- 2812 970 Words**



**971 words**



**871 words**

