Proceedings of the International Conference on Electronics and Sustainable Communication Systems (ICESC 2020)
IEEE Xplore Part Number: CFP20V66-ART; ISBN: 978-1-7281-4108-4

Rainfall Prediction Using Machine Learning & Deep Learning Techniques

CMAK Zeelan Basha
Department of CSE
Koneru Lakshmaiah Education Foundation
Vaddeswaram,Guntur,India.
cmak.zeelan@gmail.com

Nagulla Bhavana
Department of CSE
Koneru Lakshmaiah Education Foundation
Vaddeswaram,Guntur,India.
bhavana07sweetie@gmail.com

Ponduru Bhavya
Department of CSE
Koneru Lakshmaiah Education Foundation
Vaddeswaram,Guntur,India
bhavyapondhuru@gmail.com

Sowmya V
Department of CSE
Koneru Laksmaiah Education Foundation
Vaddeswaram, Guntur, India
sowmyasri40008@gmail.

***Abstract*:- In India, Agriculture is the key point for survival. For agriculture, rainfall is most important. These days rainfall prediction has become a major problem. Prediction of rainfall gives awareness to people and know in advance about rainfall to take certain precautions to protect their crop from rainfall. Many techniques came into existence to predict rainfall. Machine Learning algorithms are mostly useful in predicting rainfall. Some of the major Machine Learning algorithms are ARIMA Model(Auto-Regressive Integrated Moving Average), Artificial Neural Network, Logistic Regression, Support Vector Machine and Self Organizing Map. Two commonly used models predict seasonal rainfall such as Linear and Non-Linear models. The first models are ARIMA Model. While using Artificial Neural Network(ANN) predicting rainfall can be done using Back Propagation NN, Cascade NN or Layer Recurrent Network. Artificial NN is same as Biological Neural Networks.**

Keywords:- Rainfall, Prediction, Artificial Neural Networks, Deep Learning

1) INTRODUCTION

India's welfare is agriculture. The achievement of agriculture is dependent on rainfall. It also helps with water resources. Rainfall information in the past helps farmers better manage their crops, leading to economic growth in the country. Prediction of precipitation is beneficial to prevent flooding that saves people's lives and property. Fluctuation in the timing of precipitation and its amount makes forecasting of rainfall a problem for meteorological scientists. Forecasting is one of the utmost challenges for researchers from a variety of fields, such as weather data mining, environmental machine learning, functional hydrology, and numerical forecasting, to create a predictive model for accurate rainfall. In these problems, a common question is how to infer the past predictions and make use of future predictions. A variety of sub-processes are typically composed of the substantial process in rainfall. It is at times not promising to predict the precipitation correctly by on its global system. Climate forecasting stands out for all countries around the globe in all the benefits and services provided by the meteorological department. The job is very complicated because it needs specific numbers and all signals are intimated without any assurance. Accurate precipitation forecasting has been an important issue in hydrological science as early notice of stern weather can help avoid natural disaster injuries and damage if prompt and accurate forecasts are made. The theory of the modular model and the integrati2on of different models has recently gained more interest in rainfall forecasting to address this challenge. A huge range of rainfall prediction methodologies is available in India. In India, there are two primary methods of forecasting rainfall. Regression, Artificial Neural Network (ANN), Decision Tree algorithm, Fuzzy logic and team process of data handling are the majority frequently used computational methods used for weather forecasting The basic goal is to follow information rules and relationships while gaining intangible and potentially expensive knowledge. Artificial NN is a promising part of this wide field.

II.METHODOLOGIES

2.1 ARIMA MODEL(Auto-Regressive Integrated Moving Average)

This model is used for time series prediction and analysis and forecasting. It contains four methods and is proposed by Box and Jenkins. The following are the four steps used in the ARIMA model.

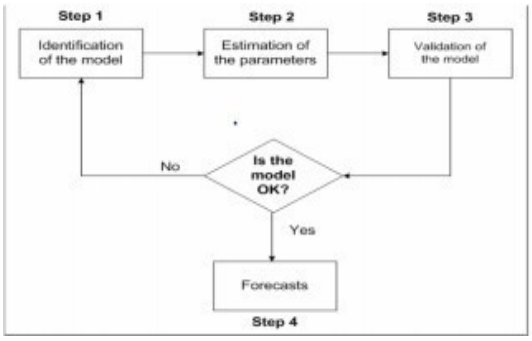
Stage-1: Identification of a series of responses is

done in the first stage which is used in calculating the time series and autocorrelations using statement IDENTIFY

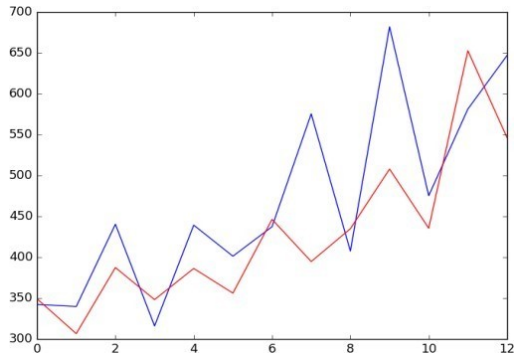
Stage-2: In this stage Estimation of the previously identified variables is done and also the parameters are estimated using the statement ESTIMATE.

Stage-3: Diagnostics checking of the above-collected variables and parameters is done in this stage.

Stage-4: In this stage the predicting values of time series are forecasted which are future values, using the ARIMA model using the statement FORECAST. The parameters used in this model are p,d,q which describes 'p' as the number of lag observations, 'q' as the degree of differencing and' as the moving average order.



Figure(1):



Figure(2)

2.2 ARTIFICIAL NEURAL NETWORK

These are considered on the base of human neural networks. It is comprised of specially organized with computing elements that they can learn and acquire the knowledge from the dataset. An artificial neural network is associated with nodes. Artificial NN activities mainly resemble human brain activities. ANN does the computations, pattern recognising and so on. After building the ANN they are trained with different datasets.

Artificial neural networks have been comprised of two types:

- Single layered neural networks.
- Multi-layered neural networks.

Here backpropagation NN is being used. Firstly, have to train the NN by a dataset which is having both input and outputs.

Firstly the input is given to the NN. The weights which are giving to neural networks are random numbers which are untrained. The value of weights which are giving to the NN must have values in the middle of -1 and +1. Weight training is used to decrease the error function in neural networks. By doing all This it gives the output. If the output is incorrect then adjusts the weights on the NN. By doing so on up to the resemblance output is given by the neural network. Then stop adding the weights and consider the weights and check it for another dataset. The weights consider in the next data in the data goes wrong then the process repeats.

Backpropagation NN consists of 3 layers:

- Input layer
- Hidden layer
- Output layer

In this backpropagation, NN comprises of 32 inputs and also having 185 neurons in the hidden layer. Hidden layered neural networks are calculated by the sigmoid function. It is having one neuron in the output layer. All these neurons are interrelated with each layer. In Backpropagation NN input signals are captured by input neurons and output signals are captured by output neuron.

So, it is called acyclic nature. When a neural network is constructed then the weightage of the input has to be adjusted according to the outputs. Among them, it is an effective approach which is driving from the input layer to the output layer.

Sigmoid function: This function is used to bring linearity in the provided data.

- Supervised learning
- Semi-supervised learning
- Unsupervised learning

Supervised learning: It includes a method of providing that the network having the expected output moreover by mathematically categorizing the network's production or by understanding the expected output with the inputs. Input and output information is specified to the NN to help in future data processing of the dataset.

Semi-supervised learning: semi-supervised learning makes use of unlabeled data for training. It is having less percentage of labeled data. It is having more percentage of unlabeled data. The network is trained with unlabeled data to define the limits.

Unsupervised learning: In unsupervised learning, the NN should decide by avoiding external help. The data is grouped by the NN by using unsupervised learning.

They have collected the data samples by which cancer is detected by using a sonar image technique. The data is distributed in three different kinds:

- Training set
- Validation set
- Testing set

2.3 Support Vector Machines:

The accumulation of experience and estimated memory of the job forecast model based on different weather types are still used in weather forecast work. The accumulation of forecasting expertise is a long-term process, and climate evolution's complexity and non-linearity make it difficult to define forecasting information. The evolution of any climate or meteorological variable is the result of a combination of certain meteorological elements ' conditional combinations, and the combination of these factors is varied and complex. Smart machine recognition skills have been well developed to communicate complex nonlinear relationships between meteorological elements in existent time and space with advances in computer technology and smart machines. SVM is commonly used for many problems in machine learning. SVM is the one which has the key multi-layer feed-forward network classification. Like multi-layer perceptrons and radial function networks, support vector machines are used for identification of patterns and non-linear regression. The SVM is used for performing the following functionalities

- Better potential for generalization when compare to the other NN models.
- The SVM solution is the same, efficient and not present in local minima.
- Used for non-vectorial data.

Only a few scientists have used this method for rainfall prediction and the results are accepted.

2.4 Self Organizing Map:

SOM is a type of competitive learning which is a part of a neural network. SOM is also known as Kohonen Map and is trained using unsupervised data which requires human interactions. In SOM only the weight of the winning neurons is updated. This method is also known as "Winner takes

everything".SOM reduces the size and shows similarities.

3. PROPOSED METHODOLOGY:

In this paper, the rainfall has been predicted using deep learning techniques. Two deep learning techniques which were used are Multilayer Perceptron and Auto-Encoders. Auto-Encoders are responsible in time series forecasting by performing the feature extraction as mentioned in

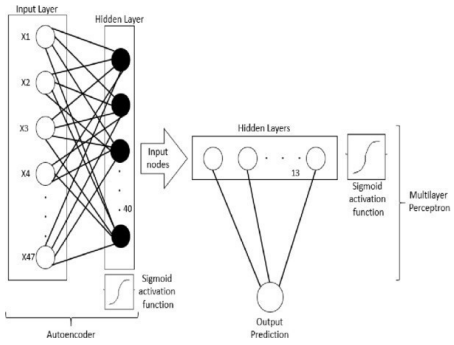
[8] and the MLP is used in prediction and classification tasks. Firstly the auto-encoders extract all the non-linear features from the data. This auto-encoder consists of three layers:

- Input Layer
- Hidden Layer
- Output Layer

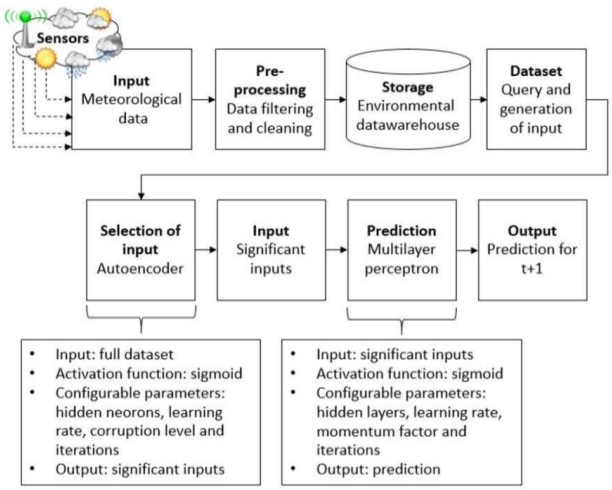
The auto-encoder extracts all the non-linear features and then send to MLP, this helps in making a prediction. The performance of methodology is also evaluated by using RMSE(Root MSE) and MSE(Mean Square Error). The architecture in the figure(3) shows how the inputs Of the autoencoder network are connected to the MLP using a sigmoid function

MSE=1/n

RMSE=



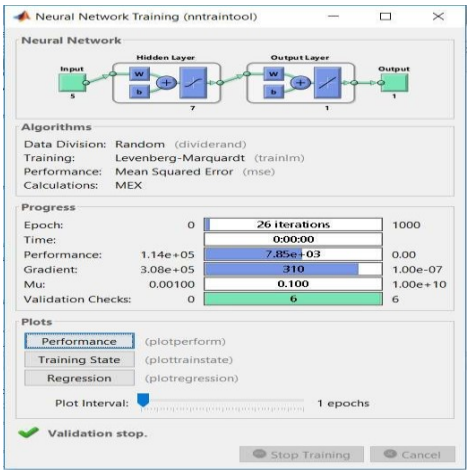
Figure(3): Auto-encoder Architecture



Figure(4): Flow of Architecture

RESULTS

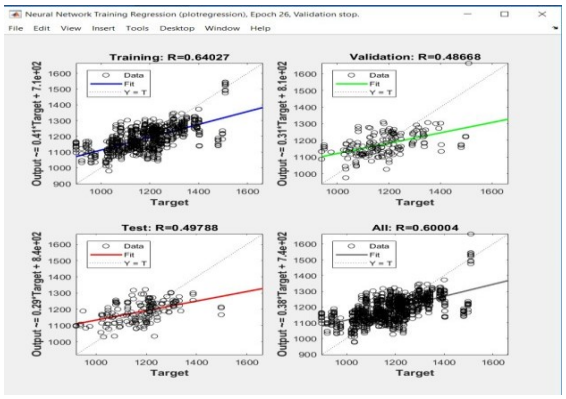
In Multilayer perceptron input is taken from multi-levels and predict the future data from the past data. Neural network training tool is used as shown in Fig(5). Generally, sensors are taken as input like wind sensor, light sensor etc.. But here using convolution neural networks to take the input from the past data. In this, the datais separated into training and testing and calculating Mean absolute error from this NN training. No of iterations are being calculated and checking the best Epoch value.



Figure(5): Neural Network Training

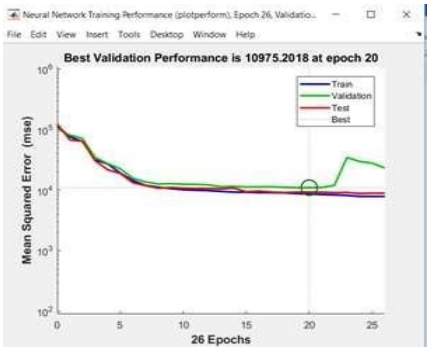
The best Epoch value results at 26th iteration as seen in Fig(6), got the best performance at 20. When the training and testing is performed there is a slight

variation. It indicates Mean square error is less but when performed validation got the best performance at Epoch 20 as shown in Fig(6).



Figure(6): Finding the best Epoch value

Now the training testing and validation set results are combined to get the best result at the targeted value as shown in Fig(7).



Figure(7):Targeted vs Prediction

3. *COMPARISON OF EXISTING METHODOLOGIES*:

Authors	Techniques	Accuracy Measure	Rainfall predicting Attribute
Victor Sanchez-Anguix2, Vicente Julian3, Javier Palanca3, and N´estor Duque4	Deep learning, MLP, Auto-encoder network	RMSE,MSE	temperature, relative humidity, barometric pressure, direction and speed of the wind
[2]	Artificial Neural Network (EBP)	RMSE	Min-Max temperature
[3]	Artificial Neural Network (EBP)	RMSE, a correlation coefficient	Min-Max temperature
[4]	ABFNN	RMSE	Latitude- Longitude, wind, sea surface pressure, wind, temperature
[5]	ABFNN	RMSE	Precipitation, Wind, temperature,
[6]	Artificial Neural Network (EBP)	MSE	Temperature
Deepak Ranjan Nayak, Amitav Mahapatra, Pranati Mishra	BPN,MLP RBFN,SVM,SOM	RMSE	Min-Max temperature

4. CONCLUSION

This paper represented the Deep Learning Approach for predicting the rainfall by using the Multilayer Perceptron and Auto-encoder Neural Network. Comparing the present architecture with other state approaches. The results intend that in terms of MSE and RMSE, our proposed architecture outperforms remaining approaches. The accuracy can be measured by the MSE and RMSE comparing with the other models. In circumstances of water resource and management, human being life and the climate they possess, precipitation prediction is of huge importance. Wrong or unfinished estimation issues can be faced because the measurement of precipitation is influenced by spatial and local change and property. This paper provided a study of different types of methodologies used to forecast and predict rainfall and issues that could be found when applying different approaches to forecasting rainfall. Because of nonlinear relationships in rainfall datasets and the ability to learn from the past, Artificial Neural Network makes a superior solution to all approaches available.

5.FUTURE WORK

The future work of the project would be the improvement of architecture for light and other weather scenarios. Also can develop a model for small changes in climate in future.

6.REFERENCES:

1. Emilcy Hern´andez1(B), Victor Sanchez- Anguix2, Vicente Julian3, Javier Palanca3, and N´estor Duque4 Rainfall Prediction: A Deep Learning Approach .
2. Aakash Parmar , Kinjal Mistree, Mithila Sompura Machine Learning Techniques for Rainfall prediction:A Review International Conference on Innovations in information Embedded and Communication Systems (ICIIECS).
3. Deepak Ranjan Nayak, Amitav Mahapatra, Pranati Mishra A Survey on Rainfall Prediction using Artificial Neural Network International Journal of Computer Applications (0975 – 8887) Volume 72– No.16, June 2013. James, N.K., Liu, Y.H., You, J .J., Chan, P.W.: Deep Neural Network Based Feature Representation for Weather Forecasting, 261 – 267.
4. Kapoor, A., Horvitz, Z., Laube, S., Horvitz, E.: Airplanes aloft as a sensor network for wind forecasting. In: Proceedings of the 13th International Symposium on Information Processing in Sensor Networks, IPSN 2014, pp. 25–34. Piscataway, NJ, USA, IEEE Press.
5. Duque-Mndez, N.D., Orozco-Alzate, M., Vlez, J.J.: Hydro-meteorological data analysis using OLAP techniques. DYNA 81(185), 160 G. Bala Sai Tarun, J.V. Sriram, K. Sairam,
6. K. Teja Sreenivas, M.V.B.T. Santhi Rainfall prediction using Machine Learning Techniques, Blue Eyes Intelligence Engineering & Sciences Publication.
7. Lngkvist, M., Karlsson, L., Loutfi, A.: A review of unsupervised feature learning and deep learning for time-series modeling. Pattern Recogn. Lett. 42, 11–24 (2014)
8. C.I. Christodoulou, S.C. Michaelides, M. Gabella, C.S. Pattichis, “Prediction of rainfall rate based on weather radar measurements”,
9. IEEE, 2004, vol. 2, 1393-1396
10. G. Zhang, B. E. Patuwo, and M. Y. Hu, “Forecasting with Artificial Neural Networks:: The state of the art,” Int J. of forecasting, vol 14, no. 1, pp. 35-62, 1998
11. Wichard, J. (2011). Forecasting the NN5 time series with hybrid models, International Journal of Forecasting 27(3): 700–707.
12. Valmik B Nikam and B.B. Meshram, “Modeling Rainfall Prediction using Data Mining Method”, Fifth International Conference on Computational Intelligence, Modeling and Simulation, Issue No: 2166- 8531, PP:132-136, 2013.
13. Basha, C. Z., Srinivasa Rao, S., Lahari, P. L., Navya, B., & Divya, S. V. S. (2019). An effective and robust computerized library management system. International Journal of Innovative Technology and Exploring Engineering, 9(1), 1647–1649.
14. Basha, C. Z., Simha, G. K. J., & Krishna, Y. V. (2019). An efficient and robust fracture detection in femur bones. International Journal of Innovative Technology and Exploring Engineering, 9(1), 1954–1957.
15. Basha, C. Z., Krishna, A., & Savarapu, P. R. (2019). Automatic detection of lung infection. International Journal of Recent Technology and Engineering, 8(3), 200–203.
16. Zeelan Basha, C. M. A. K., Shanmuka Sharan, T. V., Ravi Kishore Reddy, M., & Venkatesh, P. S. M. (2019). Computer based registration for skill development. International Journal of Engineering and Advanced Technology, 9(1), 5190–5193.