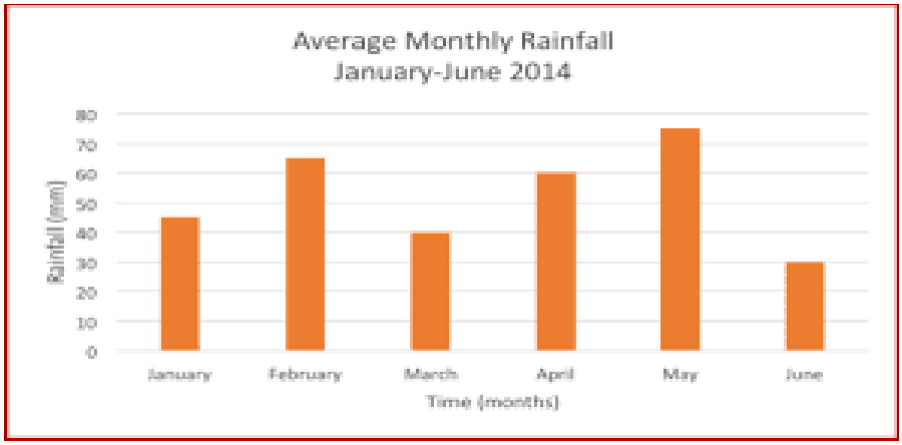
RAINFALL PREDICTION USING REGRESSION AND MULTIPLE ALGORITHMS

Jhalaka Hota
Department Of Computer Science & Engineering,
Aryan Institute Of Engineering & Technology, Bhubaneswar
Prasanna Kumar Chhotaray
Department Of Computer Science & Engineering,
Nm Institute Of Engineering & Technology, Bhubaneswar
Ashish Singh
Department Of Computer Science & Engineering,
Capital Engineering College, Bhubaneswar
Premananda Sahu
Department Of Computer Science & Engineering,
Raajdhani Engineering College, Bhubaneswar

Abstract: Rainfall is very important aspect for farmers in day-to-day agriculture and also very important in prospect to Indian economical growth. As now a day we see there is daily climate change in our India as there no stable season any month the rainfall occurs in India which also cause damage to the farming industry as the crop seeds gets damaged. In the proposed study the future rainfall is predicted using Linear Regression Technique as we also compared various model such as SVM, Random Forest Regression, Neural Network model so we obtained better results in Linear Regression so we further consider this model for future prediction of rainfall.
Keywords: Machine Learning, Rainfall Prediction, Data Analysis, Linear Regression.

INTRODUCTION

Rainfall Prediction is the application of science and technology to predict the annual amount of rainfall over a region. It is important exactly determine the rainfall for effective use of water resources, crop productivity and pre-planning of water structures. The need for rainfall prediction using a regression model is to predict the rainfall with the higher accuracy of prediction result and to provide better prediction results and to improve the prediction results over machine learning algorithms. Rainfall prediction is an important factor in various areas like when we were traveling outside such as in Maharashtra there is lot of water filling problem in roadside, farming aspects, sports, and while doing other outdoor activities for taking a decision.



For eg. Graph shows bar representation of Average Monthly Rainfall

As we know that any season the rainfall occurs which cause huge loss for agriculture work as farmer struggling daily to grow seeds work also damaged. So they need to start again from first step. It is also important in disaster management as it may saves lives of people and also in village sides places where the landslides occurs.

The main objective of the research is to predict annual rainfall of next year rainfall using machine learning model. The rainfall data sets are collected from kaggle where it consists of rainfall data from year 1901-2017 on basis of month and also consists of states. The datasets are divided into training and testing data sets model and also various models implemented and the model with higher accuracy and better results we considered for project and future annual rainfall value is predicted.

LITERATURE SURVEY

We considered here various research papers on rainfall prediction.

In paper [1] A. Kala, Dr.S.Ganesh Vaidyanathan (ICIRCA 2018) explained in paper that the rainfall is predicted using Artificial Neural Network (ANN) such as Feed Forward Neural Network and used for predicting rainfall. The model obtained accuracy of 93.55%. [2] Prediction Rainfall based Learning Machine by R. Kingsy Grace; B. Suganya 2020 in 6th International Conference on Advanced Computing and Communication Systems (ICACCS). The proposed method is based on the multiple linear regressions. The data for the prediction is collected from the publically available sources and the 70 percentage of the data is for training and the 30 percentage of the data is for testing, the accuracy of MLR gives results of 0.99.

In this paper. [3] When using a neural network for rainfall prediction some problems occurs explained by Mohini P. Darji (2015), The proposed approach uses Genetic Algorithm. The GA-TDNN (Time Delay Neural Network) model achieved monthly rainfall prediction accuracy of 89.83% and 81.38% for Anand and Navsari regions, respectively. In yearly prediction, the model achieved accuracy of 87.87%, and 87.86% for Anand and Navsari respectively. [4] Mohini P. Darji, Vipul K. Dabhi, Harshadkumar B. Prajapati. In this paper authors tried to comment on the enhancement of results by using RBM and DB for the time series network dataset, the designed model in this paper gives results with high accuracy in terms of numbers 70% to 80% etc.

METHODOLOGY

In this project we collected datasets from kaggle where the data sets consists of attributes such as subdivision, year, months, states and annual here we analyzed data and applied linear regression for prediction of future annual rainfall.

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	JF	MAM	JJAS	OND
4183	Lakshadweep	2013	26.2	34.4	37.5	5.3	88.3	426.2	296.4	154.4	180.0	72.8	78.1	26.7	1426.3	60.6	131.1	1057.0	177.6
4184	Lakshadweep	2014	53.2	16.1	4.4	14.9	57.4	244.1	116.1	466.1	132.2	169.2	59.0	62.3	1395.0	69.3	76.7	958.5	290.5
4185	Lakshadweep	2015	2.2	0.5	3.7	87.1	133.1	296.6	257.5	146.4	160.4	165.4	231.0	159.0	1642.9	2.7	223.9	860.9	555.4
4186	Lakshadweep	2016	59.6	12.1	3.2	2.6	77.4	321.1	262.6	86.2	75.6	58.6	32.0	74.7	1065.7	71.7	83.2	745.4	165.4
4187	Lakshadweep	2017	21.3	0.9	100.2	1.8	145.7	521.9	164.2	206.2	216.0	137.1	63.5	160.1	1738.9	22.2	247.8	1108.3	360.7

In this dataset the 16 attributes contain i.e. JAN, FEB, MAR, APR, MAY, JUN, JUL, AUG, SEP, OCT, NOV, DEC and JF, MAM, JJAS, OND through we will analyze data and also through ML model we will calculate accuracy, and annual future prediction rainfall.

Steps of working from flow chart

1. For predicting future rainfall value the collected datasets is read and then the null values are removed.
2. Then the data analysis is represented using the bar graph so we analyzed that which month the rainfall is high and also the comparison between the year and month visualized.
3. We consider various models to perform the mean square error, mean absolute error and r square and from that we consider the best model for predicting the future results depend on dataset.
4. First we performed the SVM model it was not ideal we manage to get the MSE of 661427.
5. Then we performed the linear regression as best result got minimum MSE compared to other model i.e. 7976.
6. Also then we implemented the random forest and managed to get MSE of 28484.
7. Neural Network model is implemented and the MSE result got is 32209.
8. After this entire model implemented we calculated the max and min where we came to know which month the rainfall was high and which month it was low.
9. Using the linear regression technique we found accuracy of our model.
10. At last we calculated the future year annual rainfall prediction using the linear regression technique.

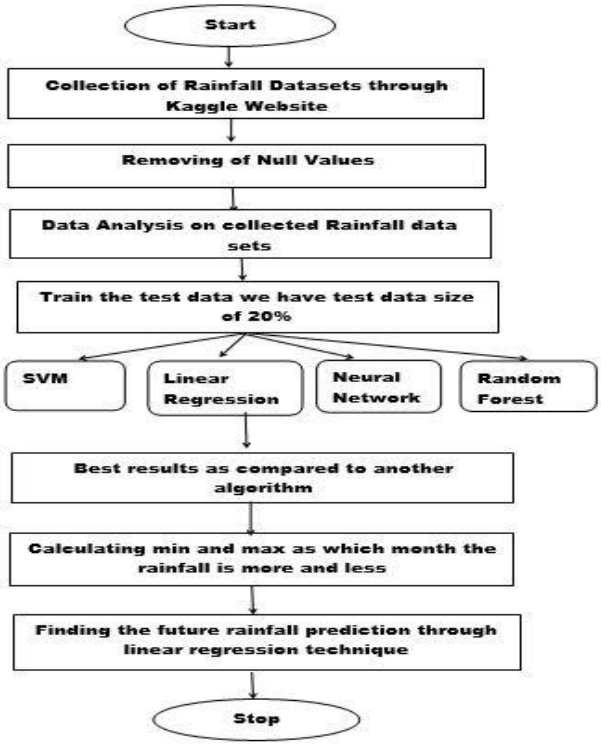


Figure shows the Flow Chart of Rainfall Prediction.

Table 1.1 Shows Comparison Results of Multiple Models Calculated.

Model	Mean Absolute Error	Mean Square Error	R^2 (R Square)
Linear Regression	-16.70 (14.83)	7976	0.985 (0.056)
Random Forest	-91.53 (23.23)	28484	0.964 (0.040)
Neural Network	-21.27 (16.668)	34846	0.985 (0.047)
SVM	-605792.12 (219447.00)	661427	0.244 (0.059)

Also calculated the min and max on the basis of which month the rainfall was more and which month it was less and we got result as month of JJAS (June-July-August-September) as more rainfall and month of March the rainfall is less.

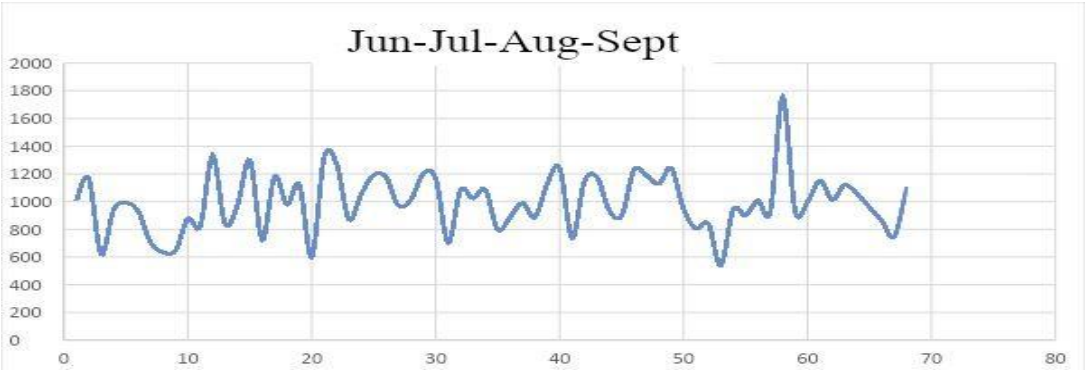


Figure represents the graph where we got maximum monthly rainfall.

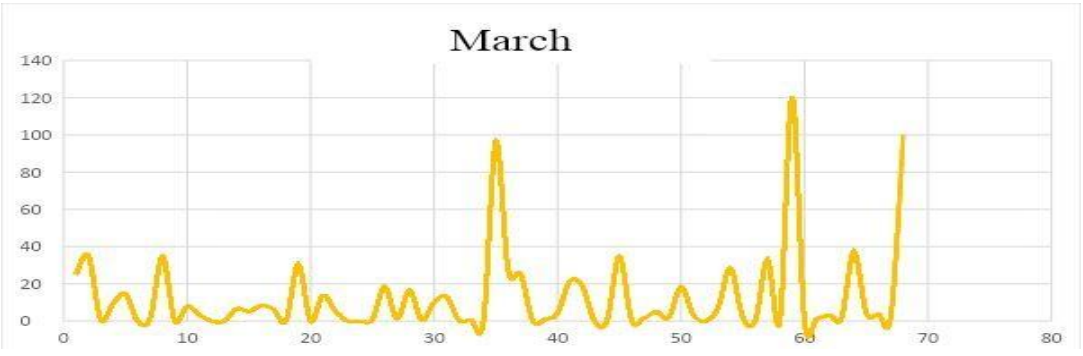


Figure represents the graph where we got minimum monthly rainfall.

Table 2.1 shows result of proposed method

	Method	Accuracy
A. Kala, Dr.S.Ganesh Vaidyanathan [1]	FFNN	93.55
Mohini P. Darji [3]	GA-TDNN	87.86
Proposed Work	Linear Regression	85

CONCLUSION

The proposed work is based on annual rainfall prediction as in future it is useful for farmers for crop fertility. The linear regression model is used as it provides the best accuracy of 0.85. In this we also compared various model from which we got best result through LR model and also the future annual rainfall prediction is calculated on basis of datasets. The rainfall prediction is important as it will be helpful in agriculture work.

REFERENCES

1. R. Kingsy Grace; B. Suganya "Machine Learning based Rainfall prediction" 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS).IEEE DOI: 10.1109/ICACCS48705.2020.9074233
2. Chandrasegar Thirumalai; K Sri Harsha; M Lakshmi Deepak; K Chaitanya Krishna "Heuristic prediction of rainfall using machine learning techniques" 2017 International Conference on Trends in Electronics and Informatics (ICEI).IEEE https://doi:/10.1109/ICACCS48705.2020.9074233
3. A. Kala; S.Ganesh Vaidyanathan "Prediction of Rainfall Using Artificial Neural Network" 2018 International Conference on Inventive Research in Computing Applications (ICIRCA).
4. Cmak Zeelan Basha; Nagulla Bhavana; Ponduru Bhavya; Sowmya V "Rainfall Prediction using Machine Learning & Deep Learning Techniques" 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC).IEEE https://doi:/10.1109/ICESC48915.2020
5. Zhengzhou Zhi-liang Wang; Hui-hua Sheng "Rainfall Prediction Using Generalized Regression Neural Network: Case Study", 2010 International Conference on Computational and Information Sciences. IEEE https://doi:/10.1109/ICCIS16966.2010
6. Emilcy Hernandez, Victor Sanchez-Anguix, Vicente Julian, Javier Palanca and N'estor Duque, "Rainfall Prediction: A Deep Learning Approach". CrossRef Google Scholar
7. Aakash Parmar, Kinjal Mistree and Mithila Sompura, "Machine Learning Techniques for Rainfall prediction: A Review", International Conference on Innovations in information Embedded and Communication Systems (ICIIECS).
8. Deepak Ranjan Nayak, Amitav Mahapatra and Pranati Mishra, "A Survey on Rainfall Prediction using Artificial Neural Network", International Journal of Computer Applications, vol. 72, no. 16, pp. 0975-8887, June 2013. Google Scholar
9. 9. N.K. James, Y.H. Liu, J. J. You and P. W. Chan, "Deep Neural Network Based Feature Representation for Weather Forecasting", pp. 261-267.Google Scholar
10. A. Kapoor, Z. Horvitz, S. Laube and E. Horvitz, "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.View Article Full Text: PDF (794KB) Google Scholar
11. N.D. Duque-Mndez, M. Orozco-Alzate and J.J. Vlez, "Hydro-meteorological data analysis using OLAP techniques", DYNA, vol. 81, no. 185, pp. 160. CrossRef Google Scholar
12. G. Bala Sai Tarun, J.V. Sriram, K. Sairam, K. Teja Sreenivas and M.V.B.T. Santhi, "Rainfall prediction using Machine Learning Techniques", Blue Eyes Intelligence Engineering & Sciences Publication. Google Scholar
13. M. Lngkvist, L. Karlsson and A. Loutfi, "A review of unsupervised feature learning and deep learning for time-series modeling", Pattern Recogn. Lett. vol. 42, pp. 11-24, 2014. CrossRef Google Scholar
14. C.I. Christodoulou, S.C. Michaelides, M. Gabella and C. S. Pattichis, "Prediction of rainfall rate based on weather radar measurements", IEEE, vol. 2, pp. 1393-1396, 2004.View Article Full Text: PDF (238KB) Google Scholar
15. 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 CrossRef Google Scholar
16. J. Wichard,"Forecasting the NN5 time series with hybrid models", International Journal of Forecasting, vol. 27, no. 3, pp. 700-707, 2011.CrossRef Google Scholar
17. 17.Valmik B Nikam and B.B. Meshram, "Modeling Rainfall Prediction using Data Mining Method", Fifth International Conference on Computational Intelligence Modeling and Simulation, no. 2166-8531, pp. 132-136, 2013.