ISSN No:-2456-2165

Rainfall Prediction using Machine Learning Algorithms: A Comparative Analysis Approach

Yeluri Kiranmai[1], Thota Kedhara Varshini[2] , Tarun Kumar Kodali[3], Venigalla Sai Teja[4], N. Md. Jubair Basha[5]
[1,2,3,4] B.Tech Student, [5] Associate Professor,
CSE Department, Kallam Haranadhareddy Institute of Technology, Guntur, AP

Abstract:- Rainfall prognosis is one of the most important technique to anticipate the climatic conditions across globe. Heavy Rainfall is a major issue for meteorological department as it is closely connected to the economy and life of human beings. It is a source for natural disasters like flood and drought which are encountered by people across the globe every year. Prompt and explicit predictions can assist proactively reduce human and financial loss.

The paper introduces a Rainfall prediction model using Multi Linear Regression, Random Forest Regression and KNN Regression. The given input data is having numerous inputs and to anticipate the rainfall in more specific. The Mean Squared Error (MSE), accuracy score of training and testing, R2_Score for the parameters used to verify the proposed model. The comparative outcome among various algorithms is also presented in this paper.

Keywords:- *Rainfall, Prediction, Machine Learning.*

I. INTODUCTION

Rainfall prediction plays vital role in Indian civilization and human life in a great range. It is challenging duty of meteorological department to anticipate the rainfall. It is complicated to anticipate the rainfall accurately with changing climatic conditions. It is challenging to predict the rainfall for both summer and rainy seasons. It is important to exactly calculate the rainfall for productive use of water resources, crop productivity and pre-planning of water structures.

Researchers across the globe have flourished various models to anticipate the rain fall mostly using random numbers and they are homogenous to the climate data. Even though many models have succeeded, but it is imperative for doing research using machine learning algorithms to get accurate prediction. In this project, we used Linear Regression, Random Forest and KNN to predict the annual density of rainfall for Indian dataset.

II. SYSTEM ANALYSIS

A. EXISTING SYSTEM

Rainfall is one of the extreme phenomena within an atmosphere system. It has a level impact on water resource management, agriculture, and ecosystems. The heavy rainfall mostly comes up with landslides, mudslides, floods, and other natural disasters. The rainfall-induced disasters in each year lead to consequential destruction and mislaying of both life and infrastructure. Therefore, rainfall prediction is a obstacle that needs to be solved to elude or minimize losses of both life and possessions. LSTM and ConvNet Architectures are used to anticipate the monthly average rainfall. There are so many models are there to predict the rain fall but the models will take more time and they not provide satisfactory predictions.

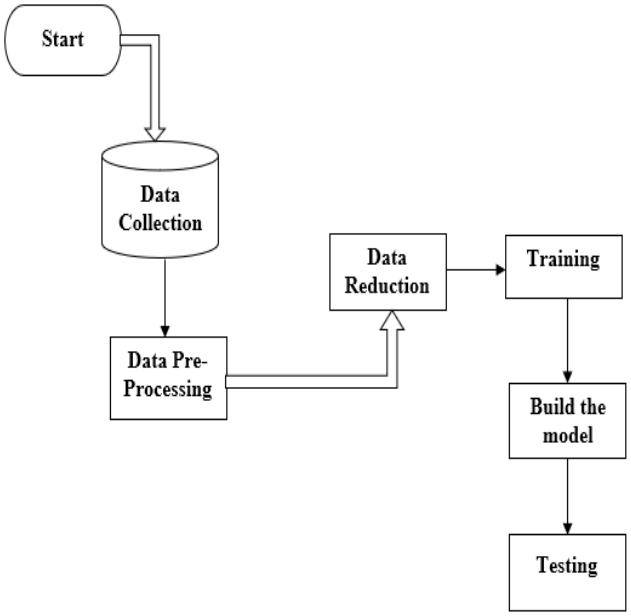


Fig. 1: Existing System Architecture for Rainfall Prediction

III. PROPOSED SYSTEM

Knowing rainfall density plays a vital role because heavy and irregular rainfall can have many influences like wrecking of crops and farms, damage of infrastructure. so a proper prediction model is crucial for an early warning that can diminished the risks to life and infrastructure and also managing the agricultural farms in proper way. These conventional methods cannot work in an efficient way so by using machine learning techniques we can produce faultless results. We can just do it by having the previous data analysis of rainfall and can predict the rainfall for the future . We can apply many techniques according to the collected data and also, we can calculate the accuracy. Different techniques produce different results, so it is important to choose the right algorithm and model it according to the requirements here we are using MLR (multi linear regression), Random Forest regressor and KNN regression. Feeding multiple variables and predicting the result.

ISSN No:-2456-2165

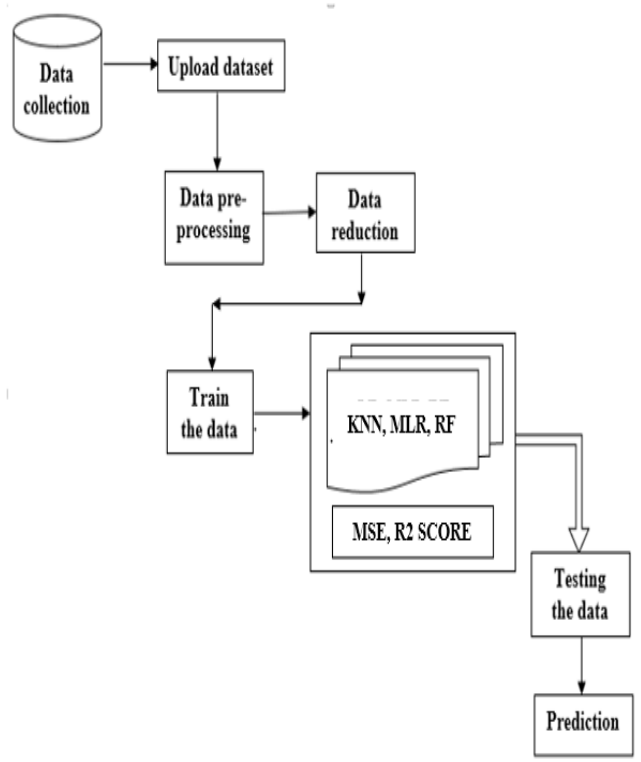


Fig. 2: Proposed System Architecture for Rainfall Prediction

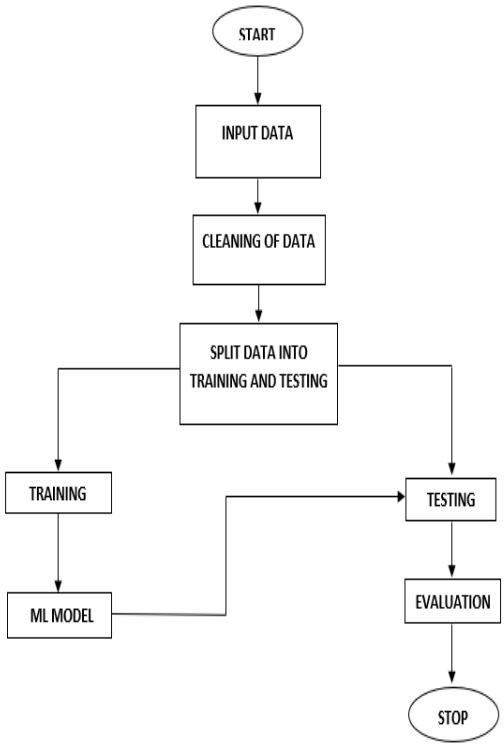


Fig. 3: Algorithm for Rainfall Prediction

IV. METHODOLOGY

In this, the study examines the use of machine learning in the prediction of rainfall to specifications status based on data produced. The model achieved best accuracy of using Linear Regression. The training and testing dataset were performed by the system from Scikit-learn module.

Precipitation is predicted using a predictive model. A good examination of the data and observation of variations in the patterns of rainfall are the initial steps in this process. Datasets are separated into training and testing sets before different machine learning and statistical techniques are applied to estimate the amount of rain. Innumerable techniques are used to reduce the mistake.

Techniques utilized in the production of this work include: By applying an equation to collected information, multiple linear regression attempts to describe the relationship between two or more parameters and a given answer. As far as I can see, it's nothing new than an extension of a simple regression toward the mean.

The general formula of an multivariable linear regression model is: $y=\alpha+\beta 1x1+ \beta 2x2+\ldots+ \beta kxk+\varepsilon$ where y is an dependent variable and x1, x2… xk are independent variables α, β are co-efficients.

When one definite variable is not clear enough to map the relationship between the independent and also the variable quantity, multiple Regression can be used to describe the relationship.

The data which we have been collected will undergo for the cleaning process. We perform splitting of the input data into training and testing datasets with a desired ratio by using a powerful python module named Scikit-learn.

Machine learning algorithms are evaluated using the train-test split process when they are used to generate predictions on data that was not used to train the algorithm. You may differentiate the performance of different machine learning algorithms for the predictive modeling issue. If you have a tiny dataset or if further setup is necessary (such as if the dataset is not balanced), you should not utilize the technique.

V. OUTCOMES AND DISCUSSIONS

This paper shows the study about the results that are obtained by performing the above method and how they are transformed by using the AI architecture.

ISSN No:-2456-2165

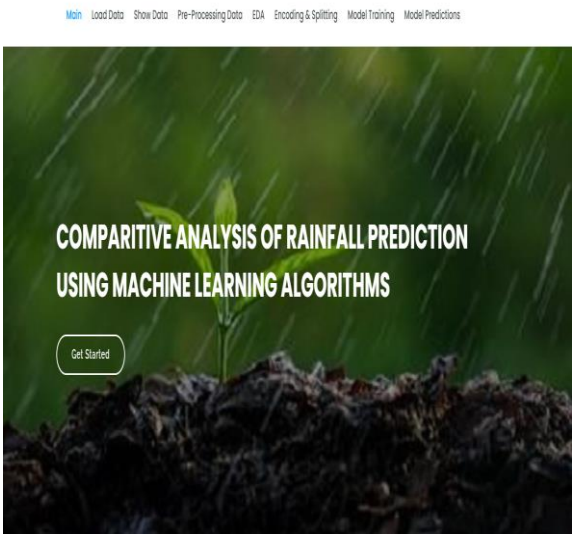


Fig. 4: Main page

The main consist Title of the project Rainfall Prediction using Machine Learning Algorithms:a Comparative Analysis Approach and its an image format

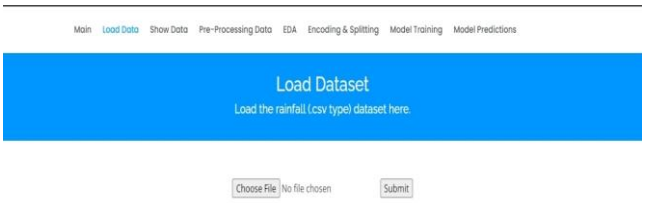
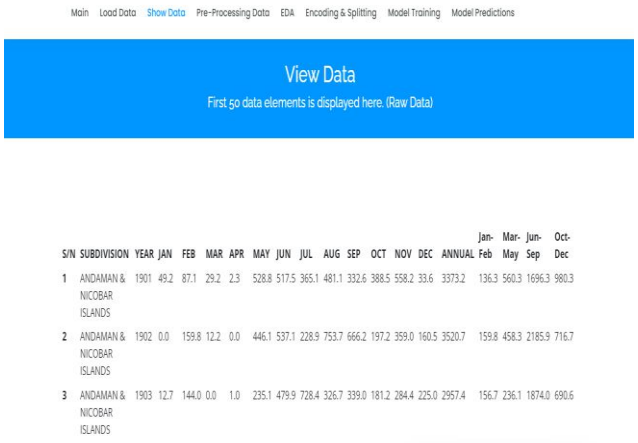


Fig. 5: Load Data

To preprocess and predict the result, primarily it is required to need to Upload the data set in csv format.



S/N	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	Jun-Sep	Oct-Dec
1	ANDAMAN & NICOBAR ISLANDS	1901	49.2	87.1	29.2	2.3	528.8	517.5	365.1	481.1	332.6	388.5	558.2	33.6	3373.2	136.3	560.3	1696.3	980.3
2	ANDAMAN & NICOBAR ISLANDS	1902	0.0	159.8	12.2	0.0	446.1	537.1	228.9	753.7	666.2	197.2	359.0	160.5	3520.7	159.8	458.3	2185.9	716.7
3	ANDAMAN & NICOBAR ISLANDS	1903	12.7	144.0	0.0	1.0	235.1	479.9	728.4	326.7	339.0	181.2	284.4	225.0	2957.4	156.7	236.1	1874.0	690.6

Fig. 6: View Data

After successfully uploading the dataset we can observe the visual format of the Dataset. The Data set consist of States, Year, Months, Annual Rainfall.

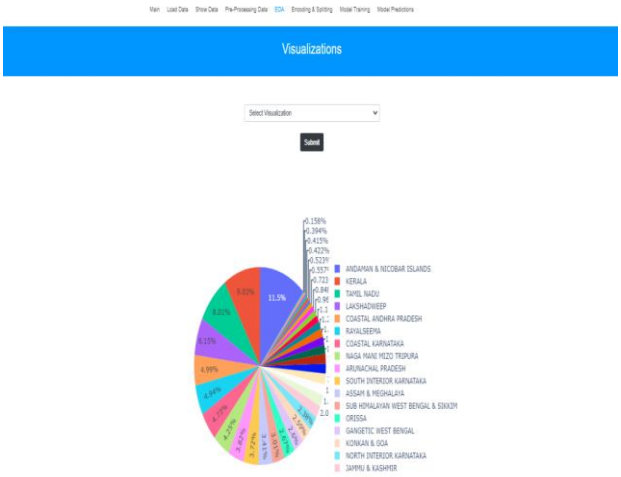


Fig. 7: Pie chart for Rainfall Prediction

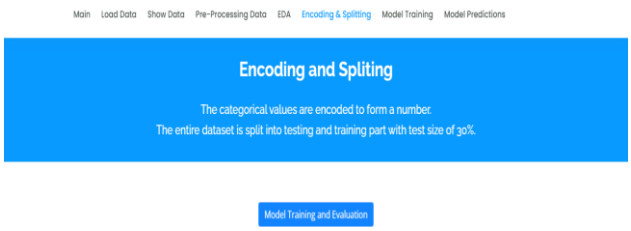


Fig. 8: Encoding & Splitting

The above shown figure is performing encodings and splitting of dataset using Scikit-learn. As the encoding and splitting will be performed by using a powerful Scikit-learn module.

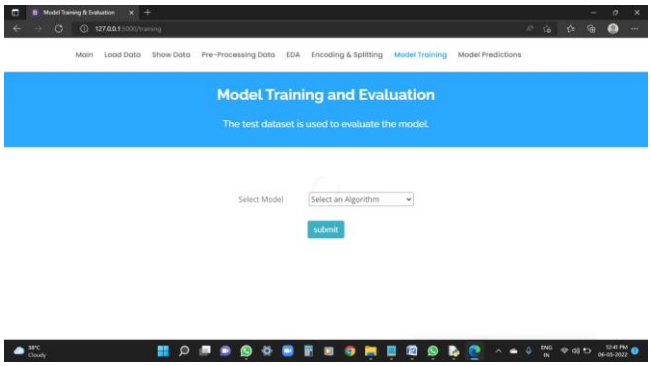


Fig. 9: Model Training

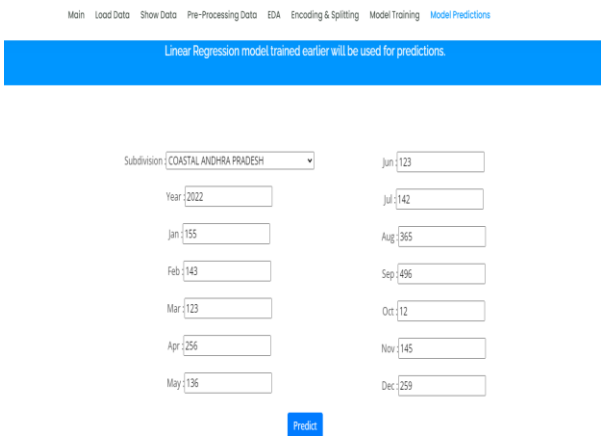


Fig. 10: Model Predictions

The above shown figure is for performing predictions to get the required output from the given inputs.

VI. CONCLUSION

The estimation of rainfall is foremost for water resource management, human life and environment. It also plays vital role in agriculture production. The development of the agriculture products is based on the amount of the rainfall So, it is required to anticipate the rainfall to assist farmers in agriculture field.

The proposed methodology presented in this paper predicts the annual density of rainfall for the Indian dataset with regarding to multiple linear regression, random forest regression, KNN regression and provides upgraded results in terms of accuracy score in training and testing wise, MSE and correlation. The R2 Score for Linear Regression, Random Forest and KNN Algorithms are 0.98, 0.96 and 0.97 respectively. The Mean Squared Error (MSE) for Linear Regression, Random Forest and KNN Algorithms are 8000.45, 39383.4 and 33456.5 respectively.

The Linear Regression algorithm showing the best result in this model. The predicted values for the linear regression are R2 : 0.98, MSE : 12083.66.

REFERENCES

[1.] Manojit Chattopadhyay, Surajit Chattopadhyay, "Elucidating the role of topological pattern discovery and support vector machine in generating predictive models for Indian summer monsoon rainfall", Theoretical and Applied Climatology, pp. 1-12, July 2015, DOI: 10.1007/s00704-015-1544-5

[2.] Kumar Abhishek, Abhay Kumar, Rajeev Ranjan, Sarthak Kumar," A Rainfall Prediction Model using Artificial Neural Network", 2012 IEEE Control and System Graduate Research Colloquium (ICSGRC 2012), pp. 82-87, 2012.

[3.] Minghui Qiu, Peilin Zhao, Ke Zhang, Jun Huang, Xing Shi, Xiaoguang Wang, Wei Chu, "A Short-Term Rainfall Prediction Model using Multi-Task Convolutional Neural Networks", IEEE International Conference on Data Mining, pp. 395-400, 2017, DOI 10.1109/ICDM.2017.49.

[4.] Aswin S, Geetha P and Vinayakumar R, "Deep Learning Models for the Prediction of Rainfall", International Conference on Communication and Signal Processing, April 3-5, 2018, India, pp. 0657-0661.

[5.] Xianggen Gan, Lihong Chen, Dongbao Yang, Guang Liu, "The Research of Rainfall Prediction Models Based on Matlab Neural Network", Proceedings of IEEE CCIS2011, pp. 45- 48.

[6.] Sam Cramer, Michael Kampouridis, Alex A. Freitas and Antonis Alexandridis, "Predicting Rainfall in the Context of Rainfall Derivatives Using Genetic Programming", 2015 IEEE Symposium Series on Computational Intelligence, pp. 711 – 718.

[7.] 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 (ICACEA) IMS Engineering College, Ghaziabad, India, pp.706 – 713

[8.] Sandeep Kumar Mohapatra, Anamika Upadhyay, Channabasava Gola, "Rainfall Prediction based on 100 years of Meteorological Data",2017 International Conference on Computing and Communication Technologies for smart Nation, pp.162 – 166.

[9.] Sankhadeep Chatterjee, Bimal Datta, Soumya Sen, Nilanjan Dey, "Rainfall Prediction using Hybrid Neural Network Approach", 2018 2nd International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom), pp. 67 – 72.

[10.] KangWoon Hong, Taegyu Kang "A study on Rainfall Prediction Based on meterological TimeSeries", 2021 twelfth International Conference on ubiquitous and Future Networks (IUFN)