International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 04 | April - 2023 Impact Factor: 8.176 ISSN: 2582-3930

RAINFALL PREDICTION USING MACHINE LEARNING

Dr.R.Prema
AP/CSE
SCSVMV University
Kanchipuram

SATTU SAI BALAJI
B.E(CSE)
SCSVMV University
Kanchipuram

P.CHAITHANYA REDDY
B.E(CSE)
SCSVMV University
Kanchipuram

ABSTRACT:

Predicting rainfall is one of the key methods for predicting the weather in any nation. Using our dataset, this research suggests a rainfall prediction model utilizing Multiple Linear Regression (MLR) methods. To more accurately forecast the frequency of rainfall, the input data includes several locations, months, years, and rainfall frequency factors. The parameters used to validate the suggested model are the Mean Square Error (MSE), accuracy, and correlation. The proposed machine learning model outperforms other algorithms in the literature, according to the results.

Keywords: Accuracy, Multiple Linear Regression, Machine learning, Prediction, Rainfall.

INTRODUCTION:

Rainfall forecasting is significant to Indian culture and has a significant impact on people's daily lives. Predicting rainfall frequency with uncertainty is a difficult task for the meteorological service. With the climatic conditions changing, it is challenging to forecast the rainfall with accuracy. Predicting the amount of rain during the summer and the wet season is difficult. Researchers from all over the world have created a number of models that are similar to climate data and use random numbers to forecast how much rain will fall. Multiple linear regressions are used in the development of the suggested model. The suggested approach predicts when it will shower using Indian meteorological data. Machine learning methods are typically divided into two groups:autonomous learning I guided learning (ii). All clustering techniques fall under the category of guided machine learning. The various classifications of machine learning methods are shown in Figure 1. The rainfall prediction study using a neural network for an Indian scenario is shown in Figure 2. Even though there are many models available, using machine learning algorithms in research is essential to obtain precise predictions. Better planning is made possible in agriculture and other sectors

by the error-free prediction. The structure of this essay is as follows: The literature's different related methods are discussed in Section II, and the MLR-based Rain Fall Prediction method is explained in Section III. Section IV elaborates on the findings, and Section V wraps up the report.

LITERATURE SURVEY:

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

In the study described in this paper, a growing hierarchical self-organizing map (GHSOM) was used to visualize the clustering behavior of annual rainfall as a result of variations in monthly rainfall for each year. The dataset for the study consisted of 142 years of data on Indian rainfall, and it allowed for the separation of the annual rainfall into smaller groups. Also, it has been found that the formation of clusters has a positive impact on the forecasting of the Indian summer monsoon rainfall using support vector machines (SVM). Analyses that were statistical and graphical were used to present the results.

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

Due to its simplicity in training, the multilayered artificial neural network with learning by back-propagation algorithm configuration is the most widely used. Back-propagation is thought to be used in more than 80% of all neural network projects currently under development. The backpropagation algorithm has two stages in its learning cycle: the first is to spread the input patterns throughout the network, and the second is to modify the output by altering the network weights. Numerous applications of the back-propagation-feed-forward neural network exist, including face detection, weather and financial forecasting, character recognition, and others. By creating training and test data sets and determining the optimal number of hidden neurons in these layers, the article executes one of these applications. The potential for forecasting average rainfall across the Karnataka district of Udupi has been examined in the current study using artificial neural network models. Three layered networks were built to create 12 artificial neural network-based predictive models. The amount of hidden neurons varies amongst the models being studied.

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

A major issue in the realm of meteorological service is the forecast of precipitation, particularly short-term rainfall. Practically speaking, the majority of current research focuses on using radar data or satellite imagery to develop forecasts. The collection of a set of meteorological features by diverse sensors at numerous observation sites constitutes a different scenario, though. Even though a site's observations are occasionally insufficient, they nonetheless offer crucial hints for predicting the weather at surrounding sites that have not yet been completely utilized in previous research. We suggest a multi-task convolutional neural network model to address this issue by automatically extracting features from time series obtained at observation locations and utilizing the correlation between the various sites for multi-tasking weather prediction. To the best of our knowledge, this is the first attempt to estimate short-term rainfall amount based on multi-site variables using multi-task learning and deep learning approaches. In particular, we characterize the correlations between various sites and structure the learning problem as an end-to-end multi-site neural network model that enables the use of gained knowledge from one site to other linked sites. Many tests demonstrate the value of the learnt site correlations, and the suggested model greatly beats a wide range of baseline models, such as the European Centre model for the system of medium-range weather forecasts (ECMWF).

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

One of the main sources of freshwater for all living things worldwide is rainfall. The information on numerous climatological variables that affect the amount of rainfall is provided by rainfall prediction models. The self-learning 13 data labels that allow for the creation of a data-driven model for a time series dataset were recently made possible by deep learning. It enables the detection of anomalies and changes in time series data and also forecasts future event data based on historical event data. The purpose of this research is to develop models of rainfall using deep learning. Deep Learning models will be effective and efficient for the modelers for any time series dataset, according to this analysis of the architectures (LSTM and ConvNet) and determination of the best design, with RMSE of LSTM as 2.55 and RMSE of ConvNet as 2.44.

[5] The Study Of Rainfall Prediction Models Based On Matlab Neural Network by Xianggen Gan, Lihong Chen, Dongbao Yang, and Guang Liu was published in the IEEE CCIS2011 Proceedings, pp. 45–48.

One of the most important aspects of agriculture is rainfall, and anticipating it is the most difficult undertaking. In general, weather and precipitation are very complicated, non-linear phenomena that necessitate sophisticated computer modeling and simulation for accurate prediction. To predict the rainfall, a variety of machine learning models, including Multiple Linear Regression, Neural Networks, K-means, Naive Bayes, and more, are used. In this study, a conventional neural network (CNN)-based rainfall prediction model for an Indian dataset is proposed. Several meteorological parameters are present in the input data, allowing for more accurate rainfall prediction. The parameters used to validate the suggested model are the Mean Square Error (MSE), accuracy, and correlation. The suggested machine learning model outperforms existing techniques in the literature, according to the results.

EXISTING SYSTEM:

The current system feeds computerised NWS numerical forecast models with observational data gathered by sensors, Doppler radar, radiosondes, weather satellites, buoys, and other devices. Random forest and Decision Tree are less accurate than other models in the current system, and their mean absolute error is likewise larger. In the suggested approach, we are substituting some high performance linear models for the less effective ones in order to accurately and precisely estimate the frequency of rainfall.

PROBLEM STATEMENT:

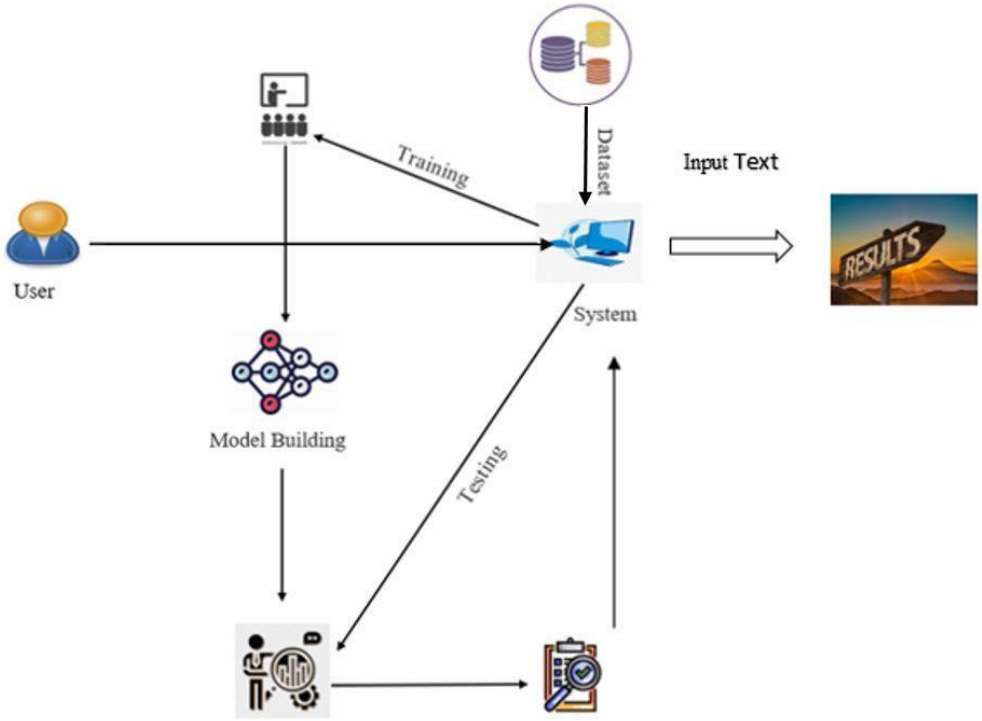
The main issue with numerical rainfall prediction models is how long it takes for the findings to be generated. Forecasts and predictions are seldom 100 percent accurate. It is nearly hard to foretell the future with accuracy, and forecasts are never 100% accurate. The accessibility, timeliness, and calibre of observational data are issues.

PROPOSED SYSTEM:

Multiple linear regressions are used in the development of the suggested model. The suggested approach predicts when it will rain using Indian meteorological data. This approach, which we suggest using, makes use of machine learning algorithms like linear regression, ridge regression, lasso regression, decision tree,

and random forest. These methods help forecast the likelihood of precipitation. Only linear regression, ridge regression, and lasso regression exhibit great performance in our model.

ARCHITECTURE:



ALGORITHMS:

Linear Regression:

A data analysis technique called linear regression uses another related and known data value to estimate the value of unknown data. It uses a linear equation to quantitatively model the relationship between the unknown or dependent variable and the known or independent variable. Predictive analysis and modelling frequently use linear regression.

Decision Tree Classifier:

The non-parametric supervised learning approach used for classification and regression applications is the decision tree. It is organised hierarchically and has a root node, branches, internal nodes, and leaf nodes. In order to structure the algorithm, decision trees are employed in machine learning. The dataset features will be divided using a cost function by a decision tree method. Before being optimised, the decision tree is pruned to remove any branches that might employ unnecessary features.

Random Forest Classifier:

Leo Breiman and Adele Cutler are the creators of the widely used machine learning technique known as random forest, which mixes the output of various decision trees to produce a single outcome. Its widespread use is motivated by its adaptability and usability because it can solve classification and regression issues.

Lasso Regression:

A regularisation method is lasso regression. For a more accurate forecast, it is preferred over regression techniques. Shrinkage is used in this model. When data values shrink towards the mean, this is referred to as shrinkage.

Ridge Regression:

Any data that exhibits multicollinearity can be analysed using the model tuning technique known as ridge regression. This technique carries out L2 regularisation. Predicted values differ much from real values when the problem of multicollinearity arises, least-squares are unbiased, and variances are significant.

RESULTS:

Fig:1 Home Page

Fig:2 Load Dataset

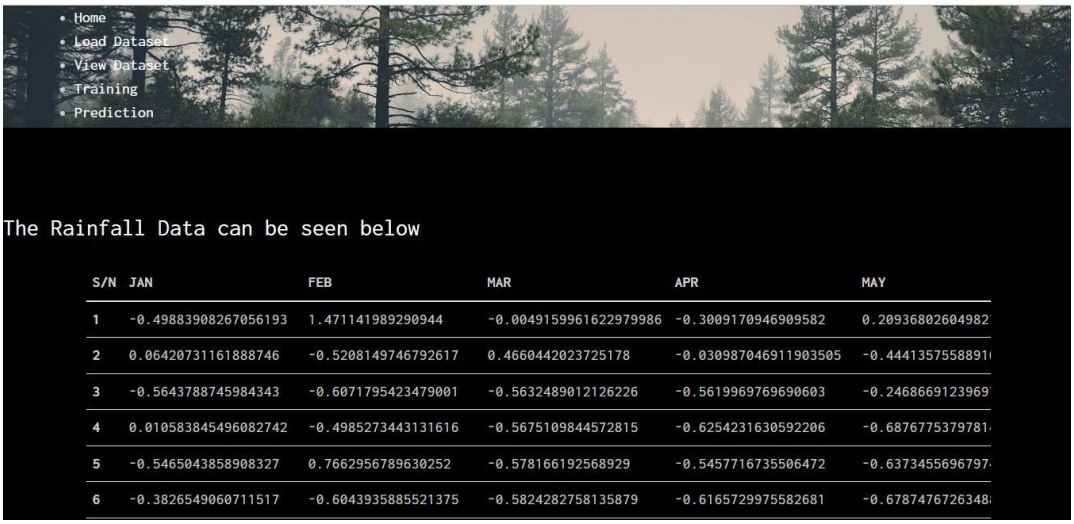


Fig:3 View Dataset

Fig4: Training Models

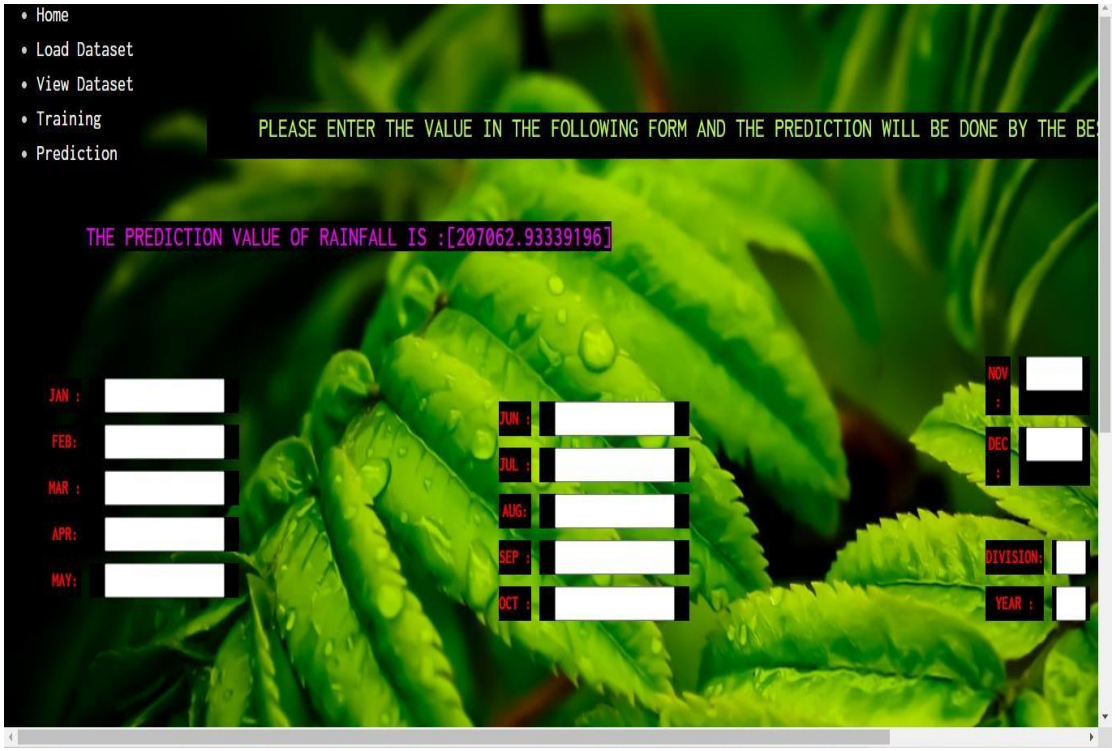


Fig5: Prediction Page

CONCLUSION:

The ability to predict when it will rain is crucial for agricultural production. Based on the amount of rainfall, agricultural items grow. So, it is essential to forecast rainfall in order to support farmers in their agricultural activities. The suggested method makes better predictions for accuracy, MSE, and correlation for the Indian dataset's rainfall using multiple linear regression.

REFERENCES:

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

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

[3] Minghui Qiu, Peilin Zhao, Ke Zhang, Jun Huang, Xing Shi, Xiaoguang Wang, and Wei Chu, "A Short-Term Rainfall Prediction Model Using Multi-Task Convolutional Neural Networks", IEEE International Conference on Data Mining, 2017, pp. 395–400, 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, India, April 3-5, 2018, pp. 0657-0661.

[5] Dongbao Yang, Guang Liu, Xianggen Gan, and Lihong Chen, "The Study Of Rainfall Prediction Models Based On Matlab Neural Network", IEEE CCIS2011 Proceedings, pp. 45–48.

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

[7] "Rainfall Forecasting Using Neural Network: A Survey," International Conference on Advances in Computer Engineering and Applications (ICACEA), 2015, IMS Engineering College, Ghaziabad, India, pp. 706-713.

[8] Sandeep Kumar Mohapatra, Anamika Upadhyay, and 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] Rainfall Prediction Using Hybrid Neural Network Technique, 2nd International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom), 2018, pp. 67-72. Soumya Sen, Nilanjan Dey, Sankhadeep Chatterjee, and Bimal Datta.

[10] "Weather Prediction: A Novel Approach for Measuring and Analyzing Weather Data," International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (ISMAC 2017), pp. 414 - 417.