



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 7, July 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.542



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

Prediction of Rainfall using Machine Learning

D Sivakumar¹, M Chandan Agrahar², Meghana A³, Kirti Singh⁴, Likith S⁵

Department of Computer Science and Engineering, Rajarajeswari College of Engineering, Bangalore,
Karnataka, India^{1,2,3,4,5}

ABSTRACT: One of the most significant approaches for predicting meteorological conditions in any country is rainfall prediction. Rainfall is caused by a variety of natural elements such as temperature, humidity, cloudiness, wind speed, and so on. Rainfall forecasting is a crucial problem for meteorological departments since it is so intimately linked to the economy and human survival. Four machine learning algorithms have been applied on a selected set of features from the dataset to achieve a higher accuracy. This work proposes a rainfall prediction model using Multiple Linear Regression (MLR) for Indian dataset. The parameters considered for training the model includes the daily recorded temperature, humidity, cloud speed, wind speed and wind direction. The input data is having multiple meteorological parameters and to predict the rainfall in more precise. The suggested model is validated using the Mean Square Error (MSE), accuracy, and correlation metrics. This system is a modified framework that uses data analysis and machine learning methods to blend observed and estimated information to predict rainfall. From the results, the proposed machine learning model which consists of four algorithms for the analysis provides better results than the other algorithms used in the existing system.

KEYWORDS: Machine Learning, Mean Square error, Rainfall Prediction.

I. INTRODUCTION

Prediction of Rainfall is vital role in India. It performs indispensable way in living being life to a wonderful extent. It is worrying duty of climate department to analyse and predict the Measurement of rainfall. It is complex to predict precisely with altering conditions. It is hard to forecast the rainfall for every summer time season and wet weather. The researchers have developed vary of fashions to predict the rainfall frequently the utilization of random values and these are comparable to the nearby weather datasets. The proposed model is developed the use of extra than one linear regression. The Rainfall prediction system using machine learning is a system that gives information using datasets from different sources like temperature, humidity, wind, wind speed, day, month, precipitation, etc. to predict the occurrence of rainfall so that preventive measures can be taken.

One of the most important concerns in hydrological study has been reliable rainfall forecasting, because early warnings of severe weather can help reduce casualties and damages caused by natural catastrophes if projected in a timely and precise manner. Forecasting is one of the most difficult difficulties facing researchers from a variety of domains, including meteorological data mining, environmental machine learning, operational hydrology, and statistical forecasting. How can one analyse the past and apply future prediction is a prevalent concern in these problems.

The parameters that are required to predict rainfall are enormously complex and subtle even for a short term period. Different classification of studying algorithms. In particular based on neural community for India. Many fashions have developed, however it is imperative for doing research the usage of laptop analysing algorithms to get accurate prediction. This lookup appears into the raining prediction hassle using deep studying algorithms. Predicting rainfall is a difficult task, and the findings must be correct. Many hardware devices are available for predicting rainfall based on weather parameters such as temperature, humidity, and pressure. Because traditional approaches are inefficient, we can achieve accurate results by employing machine learning algorithms. We can apply many techniques like classification, regression according to the requirements and also we can calculate the error between the actual and prediction and also the accuracy.

Different techniques produce different accuracies so it is important to choose the right algorithm and model it according to the requirements. Machine is getting to know the potential to answer questions. One of the objectives is a prediction on sickness data. Rainfall forecasting is very necessary due to the fact of heavy rainfall that can have many affects like damage of plants, damage of property. By the early information regarding the rainfall can help the farmers to cultivate and prepare for the upcoming rainfall with respective measures.

Various existing data mining processes and its applications were observed. Use of algorithms was connected in many data sets. The traditional machine learning techniques resulted less outcome and results additionally these results have been used for the prediction. The framework considers just the temperature and precipitation as the main factors that can impact the ascent of the level of rainfall and can lead to danger due to rainfall. Various sensors were used in this system but the main drawback of the research is that it that it won't consider the criteria like rainfall and cannot be utilized in highly populated places.

There are many units for predicting rainfall via the use of the climate conditions like temperature, humidity, and wind speed. It can be used to provide forecasting in some regions where there are no meteorological stations. This system is also enhanced in adopting to environmental parameters, evolution and climate changes. It mainly comprises of factors such as Humidity, Max. Temperature, Min. Temperature, Wind and Precipitation with possibly giving the best accuracy.

The main purpose of this system is to provide an information system which can be used to predict the rainfall occurrences and take necessary actions by alerting the authorities. It mainly focuses on the data analysis to predict the rainfall especially for the developing countries.

1. Gather data from online repository or climatic agencies and format it appropriately.
2. To study all four algorithms and put them into practise in order to determine their accuracy and precision.
3. Using database to discover the many outcomes that represent data
4. Calculating all of the algorithms' accuracy and precision values.
5. Comparing algorithms to determine which algorithm is the best.
6. Predicting the rainfall with accuracy.

After training our models on Humidity, Max. Temperature, Min. Temperature, Wind and Precipitation and many more independent features to predict the rainfall amount. The findings of the presented project can be used in the coming times to improve accuracy in predicting under various climatic conditions.

II. LITERATURE SURVEY

[1] The MSE is excessive in BPNN. Model have added rainfall prediction the use of quick term approach due to the fact of its challenges in the prediction. The SVM is used for data accumulation. The major merit is, This approach gives less error in MSE and has good performance. The main demerit is that it uses convolutional techniques which leads to have lesser accuracy. This paper represents the Approach of Deep Learning for predicting the rainfall by using the Multilayer Perceptron. In this paper, a simple method which uses features extracted from dataset and uses an SVM classifier.

[2] The goal of this project is to provide non-experts with easy access to the techniques and approaches used in the field of precipitation prediction, as well as a comparison of different machine learning algorithms. The Multilayer Perceptron is used in this research to depict a Deep Learning Approach for predicting rainfall. In this paper, a simple method which uses features extracted from dataset and uses an SVM classifier

[3] The planned analysis work provides the performance comparison of assorted machine learning algorithms like simple regression, back propagation neural network (BPNN), support vector regression (SVR) and long memory network (LSTM) wont to forecast precipitation. The planned algorithms use preprocessing techniques followed by knowledge standardisation to extend the accuracy of prognostication models. This study involves the development comparison of various machine learning algorithms to predict monthly rainfall intensity for finding the optimal technique for making early warnings. The developed models are able to predict lower as well as medium intensity rainfalls efficiently, whereas inefficient in mapping high-intensity rainfalls accurately. The demerit was the performance was less in quality as compared to other models.

[4] Rainfall prediction is one of the challenging tasks in weather forecasting. Accurate and timely rainfall prediction can be very helpful to take effective security measures in advance regarding: ongoing construction projects, transportation activities, agricultural tasks, flight operations and flood situation, etc. Data mining techniques can effectively predict the rainfall by extracting the hidden patterns among available features of past weather data. This research contributes by providing a critical analysis and review of latest data mining techniques, used for rainfall prediction. This study involves the development and performance comparison of different machine learning algorithms for predicting monthly rainfall intensity to find the optimal technique for making early warnings of landslide occurrence. Published papers from year 2013 to 2017 from renowned online search libraries are considered for this research. This review will serve the researchers to analyze the latest work on rainfall prediction with the focus on data mining techniques and also will provide a baseline for future directions and comparisons. The advantage observed here is the faster training in the reading of data sets. The limitation is it consumes more memory space.

III. SYSTEM DESIGN AND IMPLEMENTATION

Researchers are particularly concerned about the prediction of rainfall. Because of the high number of damages caused by heavy rainfall, predicting the rainfall with accuracy is challenging.

The Complexity to predict the accurate results was too much, it was the main reason for us to choose the particular topic and develop in such a way that it would give accurate and efficient results. Rainfall is responsible for damages that cost a great deal of cash. It is also responsible for lives also. Rainfall prediction can also affect the agriculture as our farmers are mostly dependent on the rainfall. The Project aims of using different Machine Learning Algorithms like Support Vector Machine, Decision Tree, Linear Regression, Random forest Algorithm for predicting the rainfall. These analytics motivated the researchers to incorporate and develop a predicting system for rainfall with an aim of lowering the damages as much as possible.

Tools and Technologies

Python is a high-level programming language that may be used in a variety of situations (human understandable languages are High level programming languages) Python Guido Van Rossum developed it for the National Research Institute in 1989. (NRI) Python was first made available to the general public in 1991 in the Netherlands: FEBRUARY 20TH, 1991 Guido van Rossum of the Centrum Wiskunde en Informatica (CWI) in the Netherlands was the first to adopt Python in December 1989 as a replacement for the ABC dialect (which was preceded by SETL) for dealing with and interacting with the Amoeba working system. Van Rossum, the creator of Python, is still alive and well.

Python includes libraries which allow programmers for employing more efficient algorithms. Which makes use of well-known machine learning technique like classification, recommendation and clustering. As a result, before proceeding, a basic introduction to machine learning is required.

Machine learning is a subfield of computer science concerned with programming systems to learn and improve on their own. Identifying and digesting the input data, as well as making informed judgments based on the information presented, are all examples of learning. When making judgments, it is difficult to consider all available inputs. To overcome this problem, algorithms are being created that use statistical science, probability, logic, mathematical optimization, reinforcement learning, and control theory to construct knowledge from specific facts and prior experience.

Machine learning is a type of learning that is automated and requires less human being support. This includes programming the computers in such a way that they may learn from the dataset that they are given. The main objective of machine learning is to research and develop algorithms which can learn from previous data and make the predictions based on new data. A learning algorithm's input is training data, which represents experience, and its output is any expertise, which is usually in the form of another algorithm capable of executing the task. A machine learning system's input data can be numerical, textual, auditory, visual, or multimodal. The system's output data can be a floating-point number, like the rocket's velocity, or an integer, like a pigeon or a sunflower from image recognition.

Unsupervised learning is used to find anomalies, outliers, such as fraud or malfunctioning equipment, or to group clients who behave similarly for a sales campaign. It is diametrically opposed to guided instruction. There are no

data points that have been labelled. When learning data only offers a few hints and no descriptions or labels, it is up to the coder or algorithm to figure out the underlying data's structure, uncover hidden patterns, or decide how to characterize the data. Unlabelled data is a term used to describe this type of learning data.

The process of establishing the architecture, components, modules, interfaces, and data for a system in order to satisfy specific requirements is known as system design. It may be thought of as a systems theory application in product development. Systems analysis, systems architecture, and systems engineering all have some overlap. Design is the process of obtaining marketing data and translating it into a design for a produced product, if product development as a whole "combined the views of marketing, design, and manufacturing into a unified approach to product creation."

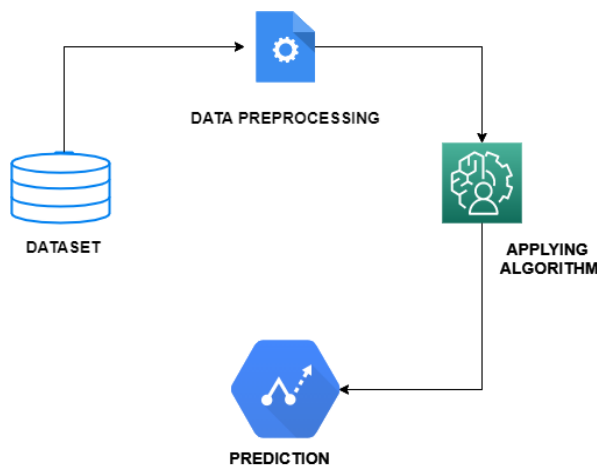


Fig. 1 System Architecture

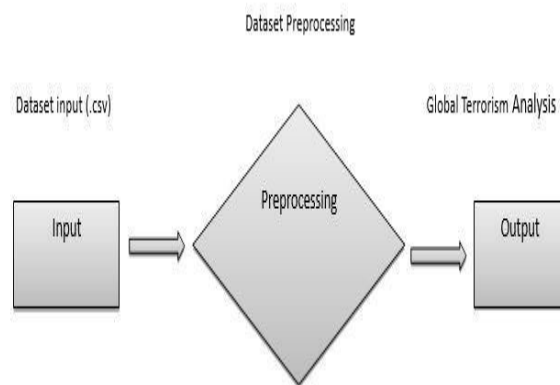


Fig. 2 IO Design

Flow Diagram

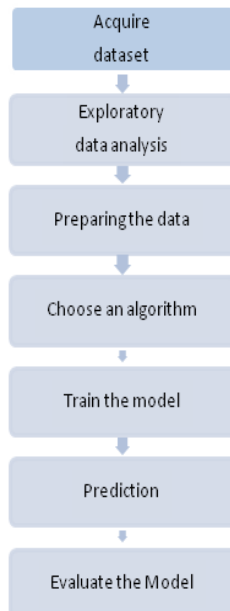


Fig. 3 Data Flow

In the process of analyzing the success we have considered machine learning algorithms to predict the output. Some of the models are taken into consideration to check the efficiency of the machine learning algorithm.

Gathering Data

The datasets used in our prediction was obtained from online repository. Such repository has n number of datasets of different places across the globe. Every sample of dataset had important attributes which was used to predict the rainfall for particular place. The key attributes were considered in order to have an accurate and efficient prediction of the rainfall with the given inputs. The dataset obtained was verified to check whether the datasets were correct or not by climate agency person.

Exploratory Data Analysis

At the beginning stage of creating the model, we undertook some exploratory data analysis to gain a high-level understanding of the dataset's features.

Preparing the Data

The process of cleaning and altering raw data prior to processing and analysis is known as data preparation. Reformatting data, making data changes, and merging data sets to enhance data are all part of this crucial stage before processing. For data professionals or business users, data preparation might be time consuming, but it is necessary to put data into context in order to transform it into insights and minimize bias caused by poor data quality. Standardizing data formats, enhancing source data, and/or eliminating NAN values are all common steps in the data preparation process. Which, in any case, cannot be used in a machine learning manner.

Algorithm

1. Support Vector Machines

It's a supervised learning algorithm that uses the largest margin to reduce misclassification during the training phase. SVM is a high-dimensional space classification function that establishes the decision border between two classes. The data points that define gutters on either side of the hyperplane are called support vectors. It is included under a class of models that is capable of analysing data sets used for classification. There is a collection of weights w to predict y . The margin is optimized which will minimize the overall number of non-zero vector weights correlated with the critical features.

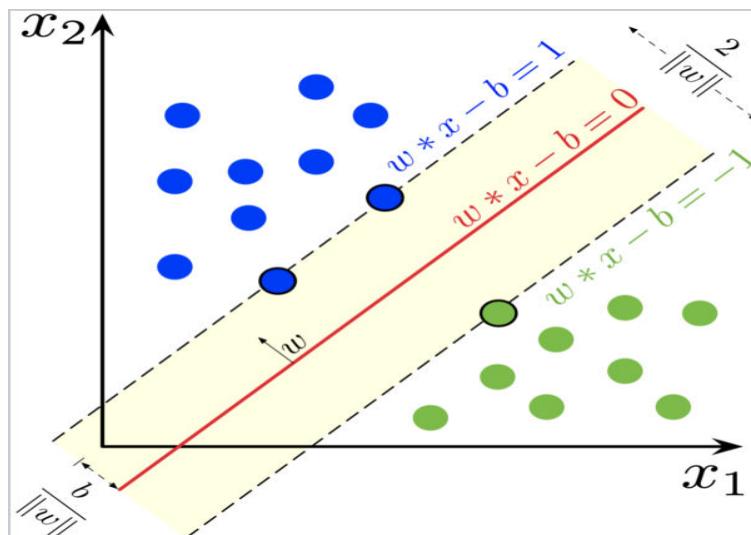


Fig. 4 SVM Mode

2. Decision tree

The Decision Tree algorithm is part of the supervised learning algorithms family. The decision tree approach, unlike other supervised learning algorithms, may also be utilised to solve regression and classification issues.. By

learning simple decision rules inferred from past data, the purpose of employing a Decision Tree is to develop a training model that can be used to predict the class or value of the target variable (training data).

The trees for training samples of data D are built using high entropy inputs. In a top down recursive divide and conquer (DAC) strategy, these trees are simple and quick to build. On D, tree pruning is done to get rid of the samples that aren't relevant.

$$Entropy = - \sum_{j=1}^m p_{ij} \log_2 p_{ij} \quad \text{eq(1)}$$

3. Random forest classifier

Random Forest Classifier is a high-level machine learning approach that represents a powerful classifier by combining a few significantly weaker decision tree learners with a function name. The independent decision tree learners formed as a result of this algorithm process are trained individually, making them irrelevant. RFC has been the most used ensemble learning method in recent decades. The built classifier is an environment friendly mannequin that gives a predictive framework for the resolution of more than one getting to know tasks.

IV. RESULT AND DISCUSSION

Prediction of rainfall is used to display the rainfall with the help of past data and facts. It is determined by various meteorological factors which is taken from meteorological and climate stations of any country. The system can be applicable in any part of the country with the database of the climates. In this process, analysis is made to display the rainfall. The user is meant to give the input values for all the necessary attributes in the text box. Once the data is entered, the prediction phase starts as per the inputs given by the user. The result is displayed on the screen to the user. The user can log out or can predict with different data as per the user's requirement.

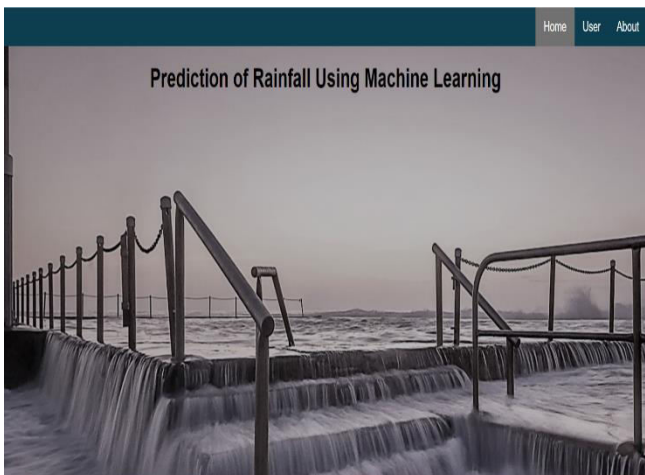


Fig. 5 Home Page

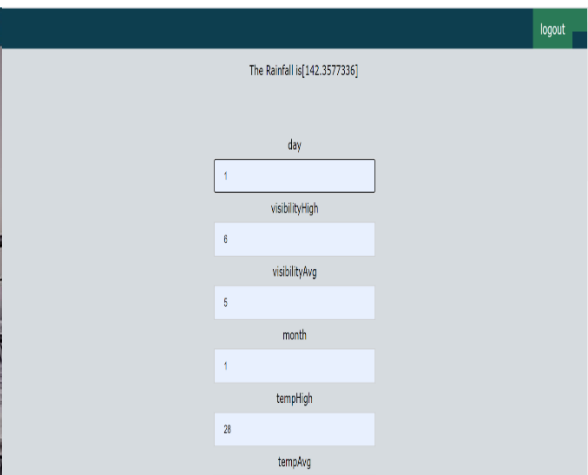


Fig. 6 Prediction Page

Once the user is redirected to the home page, the user can sign up or log in with the credentials. Once the user logs in he/she can give the inputs and get the result. Fig 6 is the result page where the user gets the result which is displayed on the screen.

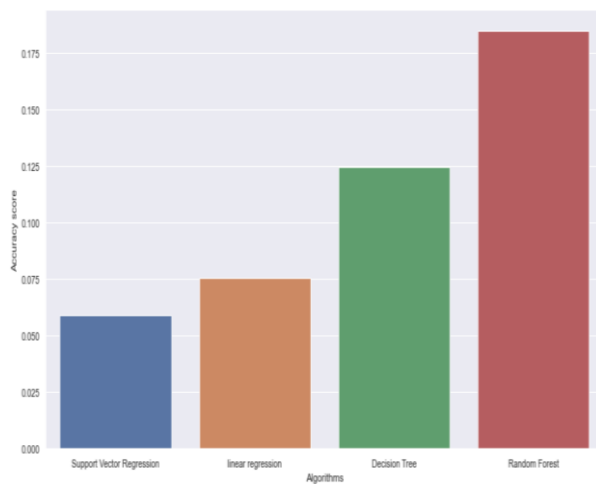


Fig. 7 Accuracy of algorithms

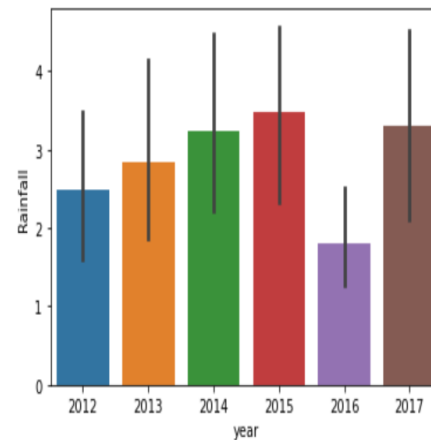


Fig. 8 Bar graph

The prediction is done with the inputs given by the user. Based on the analysis and process, the Random forest is observed to have more accuracy than other algorithms. The result is displayed on the screen with the measurement unit.

V. CONCLUSION

Prediction of rainfall is used to display the rainfall with the help of past data and facts. It is determined by various meteorological factors which is taken from meteorological and climate stations of any country. In this process, analysis is made to display the rainfall. These algorithms are used which depicts different accuracy measures. In the final accuracy score value, the Random Forest has the higher accuracy compared to Decision Tree, Linear Regression and Support Vector Machine.

Future experiment with the IoT sensors will lead to improvement in predicting rainfall with more accurate results. Sensory networks can be considered and can be implemented in the project for improved results in collection of data. The future enhancement can be done with the help of sure forms of desktop mastering strategies. However, the lookup can be multiplied success by way of the use of hydraulic models. Time collection structure can be executed by laptop gaining knowledge of algorithms for higher outcomes.

Apart from that, we want to utilize web scraping technologies and emotional analysis in the future to examine various posts and comments on social media sites for hate speech and text, filter them, and develop a classifier to combine the present project with the social media texts.

REFERENCES

1. Forecasting Raining Using Multilayer Perceptron Artificial Neural Network Model Author: Ali, Z., Hussain, I., Faisal, M., Nazir, H. M., Hussain, T., Shad, M. Y., & Hussain Gani, S, 2017.
2. Operational raining monitoring in Kenya using MODIS NDVI time series. Author: Anja Klisch and Clement Atzberger; Published In: Remote Sens, 20164
3. Prediction Of Rainfall Using Machine Learning Techniques Authors: Moulana Mohammed, Roshitha Kolapalli, Niharika Golla, Siva Sai Maturi, issue 01, January 2020
4. Machine Learning based Rainfall Prediction Authors: R. Kingsy Grace 1 Published in: ICACCS 2020
5. Rainfall Prediction using Data Mining Techniques: A Systematic Literature Review Authors: Shabib Aftab, Munir Ahmad, Noureen Hameed
6. Basha, Cmak Zeelan; Bhavana, Nagulla; Bhavya, Ponduru; V, Sowmya (2020), "Rainfall Prediction using Machine Learning & Deep Learning Techniques"
7. Srivastava, Shikha; Anand, Nishchay; Sharma, Sumit; Dhar, Sunil; Sinha, Lokesh Kumar (2020) "Monthly Rainfall Prediction Using Various Machine Learning Algorithms for Early Warning of Landslide Occurrence"

8. Tharun, V.P; Prakash, Ramya; Devi, S. Renuga (2018) "Prediction of Rainfall Using Data Mining Techniques" pg 1507- 1512, doi:10.1109/ICICCT.2018.8473177.
9. Chatterjee, Sankhadeep; Datta, Bimal; Sen, Soumya; Dey, Nilanjan; Debnath, Narayan C. (2018), "Rainfall prediction using hybrid neural network approach
10. 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
11. 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.
12. 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.
13. 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.
14. 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.
15. BhaskarPratap Singh, Pravendra Kumar, Tripti Srivastava, Vijay Kumar Singh. Estimation of Monsoon Season Rainfall and Sensitivity Analysis Using Artificial Neural Networks. Indian Journal of Ecology 44, 2017.
16. Government of Kenya. (2012). Kenya Post-Disaster Needs Assessment: 2008-2011 Raining.
17. Cody, B. A. (2010). California raining: Hydrological and regulatory water supply issues. DIANE Publishing
18. Ding, Y., Hayes, M. J., & Widhalm, M. (2011). Measuring economic impacts of raining: a review and discussion. Disaster Prevention and Management: An International Journal, 20(4), 434-446.
19. 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.
20. BhaskarPratap Singh, Pravendra Kumar, Tripti Srivastava, Vijay Kumar Singh. Estimation of Monsoon Season Rainfall and Sensitivity Analysis Using Artificial Neural Networks. Indian Journal of Ecology 44, 2017.
21. Emilcy Hernandez¹(B), Victor Sanchez- Anguix², Vicente Julian³, Javier Palanca³, and Néstor Duque⁴ Rainfall Prediction: A Deep Learning Approach .
22. 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).
23. 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.
24. James, N.K., Liu, Y.H., You, J .J., Chan, P.W.: Deep Neural Network Based Feature Representation for Weather Forecasting, 261– 267.
24. 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.



INNO  **SPACE**
SJIF Scientific Journal Impact Factor
Impact Factor: 7.542



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 **9940 572 462**  **6381 907 438**  **ijircce@gmail.com**



www.ijircce.com

Scan to save the contact details