Journal of Cardiovascular Disease Research

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

Drought Prediction Using Ai Based Multiple Linear Regression Technique

[1]Dr. J. Angel Ida Chellam , [2]Dr.M.Siva Sangari , [3]Dr.R.Rajendra Kumar

[1]Assistant Professor(Sr.G), Dept of IT, Sri Ramakrishna Engineering College, Coimbatore

[2]Assistant Professor(Sl.G) , Dept of CSE, KPR Institute of Engineering and Technology, Coimbatore

[3]Frugal Expert & CEO, Dr.R.R.Consultancy, Coimbatore

*Corresponding author: angel.anbuseelan@srec.ac.in

Abstract

Drought Prediction is to be focused of great importance for India. Drought can be Predicted using the amount of rainfall data. Two types of predictions can be done, they are – Long term Predictions: Predict rainfall over few weeks/months in advance. Short term Predictions: Predict rainfall a few days in advance in specific locations. Indian meteorological department provides forecasting data required for the project. In this Paper we are planning to work on long term predictions of rainfall. The main motive of the paper is to predict the amount of rainfall in Particular division and state whether the drought will occur or not in advance. We predict the amount of rainfall using Multiple Linear Regression(MLR) which is a branch of AI by using the past dataset.

Keywords*:* AI(Artificial Intelligence),MLR(Multiple Linear Regression),Data set

INTRODUCTION

Forecasting future events in a region is an significant first step for funding sustainable solutions to water management and risk assessment of drought occurrences. However, it remains a multifaceted task, because of the random character of precipitation field, which is the elementary variable commonly used for drought assessment. Drought, in fact, is a usual recurrent feature of climate occurring in all climatic zones and it originates from a deficiency of precipitation over a given period of time. In Short time, characterize meteorological drought, while longer time, hydrological drought. The methods involved can be further analyzed and processed by using machine learning approaches.[2][4][5][10][12]

The system uses the MLR (Multiple Linear Regression).It is used to predict the value based on the value of two or more other variables.[3][6][11][13][12]The variable we want to predict is called the dependent variable. The variable we are using to predict the value of dependent variable are called the independent variables. [14][15]

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

II.LITERATURE REVIEW

Various researches have been done for drought prediction using different data sets and methodologies. This research is done prior to taking up the project and understanding the various methods that were used previously. This study helped to identify the benefits and drawbacks of the existing system.

A. **Norbert A Agana and Abdouah Hamifar in"A Deep Learning based approach for long term drought prediction"** This paper observed into the drought prediction problem using a deep learning based approach. They proposed a deep belief network (DBN) for long-term drought prediction and compared its performance to that of standard MLP and SVR models[16][17][18]. The DBN model was found to provide better prediction results, recording lower prediction errors compared to MLP and therefore can be more reliable andefficient[19][20][21].

B. **Daniel Hong and kee An Hong in"Drought forecasting using MLP neural networks"** This paper proposes an idea to develop the SPI with longer duration time scale can be predicted more accurately than shorter series data. In short, the 9 month SPI values are predicted more accurately than those 3 and 6 months using the MLP neural networkmodel[22][23].

C. **Tin Huang,Hong-ling Li and Lin Qiu "Reasearch on Predicting drought bassed fuzzy sets and R/S analysis model"** This paper proposes an idea to develop the fuzzy set and R/S analysis model predicting the year of occurrence of agricultural drought has good accuracy. Then, the years of 2006 and 2008 also used as new information and added into the new time line series. Then they predicted the next drought year is 2011.

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

III. PROPOSED SYSTEM

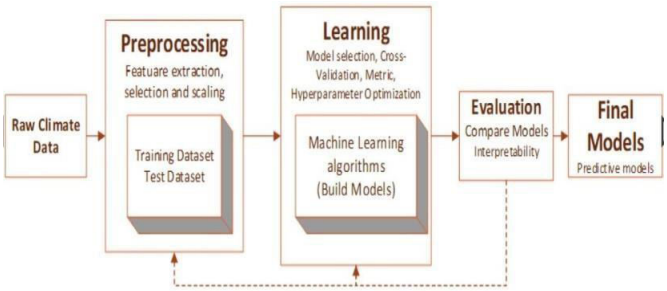
The proposed system is designed for the purpose: The new proposed system is to predict the drought all over India because there is only very less predictions available for Indian data set and make further processing for approximate desired results. We are going to use AI(Artificial Intelligence) for predicting drought. Machine learning is the subset of AI where we teach machine to learn by itself without the help of any external source. In machine learning we teach the mission to learn from previous data and try to improve its result in future by taking lessons from its previous decision. The drought is predicted using LR (Linear Regression) which comes under supervised is one of the machine learning approaches.

IV. WORKING PRINCIPLE

A data is converted to a correct format to conduct experiments. Good analysis of data and observe variation in the patterns of rainfall are done. We try to predict the average rainfall by untying data into training and testing. For prediction we formulated data in the way, the rainfall in the last three months. We try to forecast the rainfall in the next consecutive month and state whether the drought will occur or not. For experiments we used 80:20 training and test ratio. Testing metrics used is Mean Absolute Error to train the models. It also shown the amount of rainfall actually and predicted with the histogram plots. The training is done on complete data set and observed mean and standard deviation written, first one represents ground truth, and second one representspredictions

V. SYSTEM ARCHITECTURE

Fig.1 Steps for Prediction



VI. LINEAR REGRESSION ALGORITHM

Linear Regression is a machine learning algorithm built on supervised learning. It performs a regression task. Regression models a object prediction value based on independent variables. It is frequently used for finding out the relationship between variable (x). So, this regression technique discoveries out a linear relationship.

A. Hypothesis Function of Linear Regression:

$$y = \theta_1 + \theta_2.x$$

While training the model we are given:
x: input training data (univariate – one input variable(parameter))
y: labels to data (supervised learning)
When training the model, it turns the best line to predict the value of y for a given value of x. The model develops the best regression fit line by finding the best θ_1 and θ_2 values.
θ_1: intercept
θ_2: coefficient of x

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

B. **Linear Regression with Python**:

Before moving on, we encapsulate 2 basic steps of Machine Learning as per below:

1. Training
2. Predict

We used 4 libraries such as **numpy** and **pandas** to work with data set, **sklearn** to implement machine learning functions, and **matplotlib** to envision our plots for viewing.

C. Data Set:

The rainfall data of India is obtained from kaggle dataset.

D. Code:

data = pd.read_csv("/content/drive/My Drive/rainfall in India 19 01-2015.csv",sep=",")

E. Code Explanation:

dataset: the table contains all values in our csv file.
X: the dependent Variable
Y: the independent Variable

F. Testing and Training:

Next, we have to split our dataset (total 36 observations) into 2 sets: training set which used for training and test set which used for testing.

After train set and test set, now we have to build the Regression Model

Regressor = LinearRegression():Our training model which will implement the Linear Regression.

Fig.2 Graph of input dataset

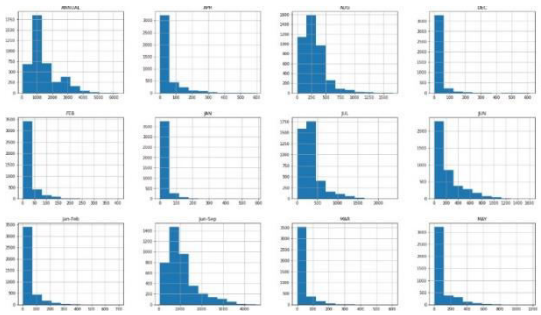


Fig.3Testing and training of the year 2005

```
#test 2005
temp = data[['SUBDIVISION','JAN', 'FEB', 'MAR', 'APR', 'MAY', 'JUN', 'JUL',
       'AUG', 'SEP', 'OCT', 'NOV', 'DEC']].loc[data['YEAR'] == 2005]

data_2005 = np.asarray(temp[['JAN', 'FEB', 'MAR', 'APR', 'MAY', 'JUN', 'JUL',
       'AUG', 'SEP', 'OCT', 'NOV', 'DEC']].loc[temp['SUBDIVISION'] == 'TELANGANA'])

X_year_2005 = None; y_year_2005 = None
for i in range(data_2005.shape[1]-3):
    if X_year_2005 is None:
        X_year_2005 = data_2005[:, i:i+3]
        y_year_2005 = data_2005[:, i+3]
    else:
        X_year_2005 = np.concatenate((X_year_2005, data_2005[:, i:i+3]), axis=0)
        y_year_2005 = np.concatenate((y_year_2005, data_2005[:, i+3]), axis=0)
```

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

Fig.4 Testing and training of the year 2010

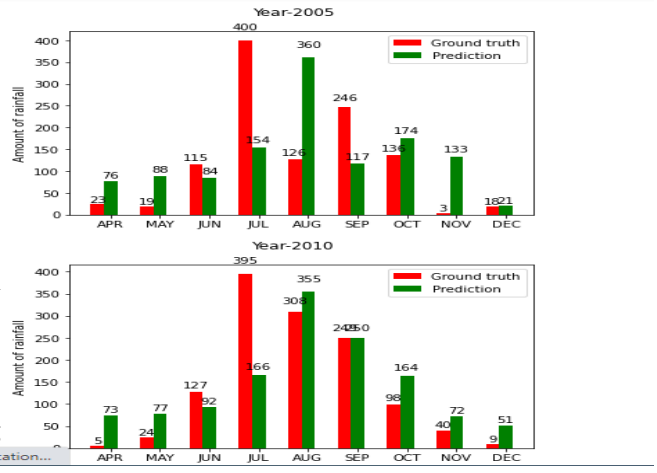
```
#test 2010
temp = data[['SUBDIVISION','JAN', 'FEB', 'MAR', 'APR', 'MAY', 'JUN', 'JUL',
       'AUG', 'SEP', 'OCT', 'NOV', 'DEC']].loc[data['YEAR'] == 2010]

data_2010 = np.asarray(temp[['JAN', 'FEB', 'MAR', 'APR', 'MAY', 'JUN', 'JUL',
       'AUG', 'SEP', 'OCT', 'NOV', 'DEC']].loc[temp['SUBDIVISION'] == 'TELANGANA'])

X_year_2010 = None; y_year_2010 = None
for i in range(data_2010.shape[1]-3):
    if X_year_2010 is None:
        X_year_2010 = data_2010[:, i:i+3]
        y_year_2010 = data_2010[:, i+3]
    else:
        X_year_2010 = np.concatenate((X_year_2010, data_2010[:, i:i+3]), axis=0)
        y_year_2010 = np.concatenate((y_year_2010, data_2010[:, i+3]), axis=0)
```

In conclusion, with Simple Linear Regression, we have to perform 5 steps as per below:

1. Importing thedataset.
2. Splitting dataset into two sets: training set and testing set (2 dimensions of X and Y per each set). Normally,the testing set should be 5% to 30% ofdataset.
3. Visualize the training set and testing set todouble check (you can bypass this step if youwant).
4. Initializing the regression model and suitable it using training set (both X and Y)
5. Predict theoutput.

VII. EXPERIMENTAL RESULTS ANDDISCUSSIONS

As we have concentrated in long term predictions. The rainfall can be predicted few weeks/months in advance. So it helps farmers in taking decision about cultivation of crops few months before. It also helps government to be planned and prepared for future dry events.



VIII. CONCLUSION

The progress in technology is a nonstop process. New technology is being invented and updated on a daily basis. Hence this system can predict drought using one of the machine learning approaches which is the subset of AI

IX. FUTURE SCOPE

● We can predict drought more accurately using the upcoming technologies.

● We can create an app which helps the farmers for deciding the cultivation of crops during drought

● In future we can be more specific by determining the results of drought prediction over districts in each state.

X. REFERENCES

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

[1].A Belayneh, J Adamowski, B Khalil, and J Quilty. "Coupling machine learning methods with wavelet transforms and the bootstrap and boosting ensemble approaches for drought prediction" Atmospheric Research, 172:37–47, 2016.

[2]. Antonino Cancelliere, G Di Mauro, B Bonaccorso, and G Rossi. "Drought forecasting using the standardized precipitation index" Water resources management, 21(5):801–819, 2007.

[3]. Anne F Van Loon. "Hydrological drought explained" Wiley Interdisciplinary Reviews: Water, 2(4):359–392, 2015.

[4]. Donald A Wilhite and Michael J Hayes. "Drought planning in the United States: status and future directions" In The arid frontier, pages 33–54. Springer, 1998.

[5]. Stanley A Changnon, Roger A Pielke Jr, David Changnon, Richard T Sylves, and Roger Pulwarty. "Human factors explain the increased losses from weather and climate extremes" Bulletin of the American Meteorological Society, 81(3):437, 2000.

[6]. A. Norbert A Agana and Abdouah Hamifar "A Deep Learning based approach for long term drought prediction" IEEE Explore Conference, 2017

[7]. Daniel Hong and kee An Hong in"Drought forecasting using MLP neural networks", 8th International Conference on u- and e-Service, Science and Technology (UNESST), 2015

[8]. C. Tin Huang,Hong-ling Li and Lin Qiu "Reasearch on Predicting drought bassed fuzzy sets and R/S analysis model"

[9]. Anurag Malik, Anil Kumar1, Sinan Q. Salih,Sungwon Kim, Nam Won Kim, "Drought index prediction using advancedfuzzy logic model: Regional case study overKumaon in India" Plos One, 2020

[10]Rajendra kumar R(2018) Components of Consumer Factor and Its Influence on Attitude of the Student Online Shopping Community: Consumer Factor and Its Influence on Attitude .Vol 8(3). 36-51.

[11]M.Siva Sangari, Dr.K.Baskaran, 2016, Secured Efficient Fast Handover Multihoming Based NEMO+ (SEFMNEMO+) for Vanets, Circuits and Systems, Vol. 7, pp.239-254.

[12] Dr.M.SivaSangari , Dr.R. Rajendra Kumar,, Dr.P.Thangaraj, Dr.A.Devipriya, Ms.M.SalomiSamsudeen, 2020, "Technology Factors Of Online Shopping And Its Impact On Attitude With Special Reference to Student Community", European Journal of Molecular & Clinical Medicine, SSN 2515-8260 Volume 07, Issue 08.

[13] Dr.M.Siva Sangari, Dr.K.Baskaran,2021, " Comparative Analysis of Multihoming based VANET Protocols", Turkish Journal of Physiotherapy and Rehabilitation, Vol 32,Issue 2.

[14] Dr.M.Siva Sangari, E.Dhivakar, .Gowtham.K,2021, "Secret Knock Detecting Door Lock", Annals of the Romanian Society for Cell Biology,Volume 25,Issue 5,pg.no 406-410.

[15] Dr.M.Siva Sangari,Kaviya L, Meenakshi K.L,Nikildharan R.,2021, "Intelligent Smart Shopping and Stock Notification Using IoT", Annals of the Romanian Society for Cell Biology,Volume 25,Issue 5,pg.no 306-311.

[16]Amin Salih Mohammed, Saravana Balaji B, Saleem Basha M S, Asha P N, Venkatachalam K(2020),FCO — Fuzzy constraints applied Cluster Optimization technique for Wireless AdHoc Networks,Computer Communications, Volume 154,Pages 501-508.

[17]Ponmagal, R.S., Karthick, S., Dhiyanesh, B. et al. Optimized virtual network function provisioning technique for mobile edge cloud computing. J Ambient Intell Human Comput (2020).

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

[18]Ramamoorthy, S., Ravikumar, G., Saravana Balaji, B. et al. MCAMO: multi constraint aware multi-objective resource scheduling optimization technique for cloud infrastructure services. J Ambient Intell Human Comput (2020). **IF 4.5**

[19]R. Ali, Y. B. Zikria, A. K. Bashir, S. Garg and H. S. Kim, "URLLC for 5G and Beyond: Requirements, Enabling Incumbent Technologies and Network Intelligence," in IEEE Access, vol. 9, pp. 67064-67095, 2021, doi: 10.1109/ACCESS.2021.3073806.

[20]M. Abdel-Basset, R. Mohamed, M. Elhoseny, A. K. Bashir, A. Jolfaei and N. Kumar, "Energy-Aware Marine Predators Algorithm for Task Scheduling in IoT-Based Fog Computing Applications," in *IEEE Transactions on Industrial Informatics*, vol. 17, no. 7, pp. 5068-5076, July 2021, doi: 10.1109/TII.2020.3001067.

[21] Shehzad Ashraf Chaudhry, Khalid Yahya, Marimuthu Karuppiah, Rupak Kharel, Ali Kashif Bashir, Yousaf Bin Zikria,GCACS-IoD: A certificate based generic access control scheme for Internet of drones,Computer Networks,Vol.191,2021.

[22] S. Anbalagan *et al*., "Machine Learning-based Efficient and Secure RSU Placement Mechanism for Software Defined-IoV," in *IEEE Internet of Things Journal*, doi: 10.1109/JIOT.2021.3069642.

[23] Banupriya, S., Kottursamy, K. & Bashir, A.K. Privacy-preserving hierarchical deterministic key generation based on a lattice of rings in public blockchain. *Peer-to-Peer Netw. Appl.* (2021)