Climatic Assessment Of Rajasthan's Region For Drought With Concern Of Data Mining Techniques

NEHA KHANDELWAL
M-Tech Scholar (SGUV)

RUCHI DAVEY
SGUV

Abstract

Meteorological data base is the part of data mining which focused on the hidden patterns available in the meteorological data. From this data we retrieve the information which can be transformed into usable knowledge. Now Weather is one of the meteorological data that is rich by knowledge by including different climatic factors like- rainfall, sea-level, humidity etc. As we see that when the fluctuation occurs in the climate, it greater impact on the world's economy, social life of humans. So in this paper, We consider the one climatic factor for weather i.e. rainfall .rainfall is major issue for nature, food productions, water resource management etc. so in this paper first we consider weather daily historical data collected at JAIPUR city[1] for trying to extract useful knowledge . Then we apply data mining techniques regression analysis on the rainfall dataset because Rainfall forecasting has been one of the most scientifically and technologically challenging problems around the world. So we apply multiple regression analysis to predict rainfall of a year by using different 4 climatic factors temperature, humidity, pressure, sea level etc. for selecting these factors we apply correlation analysis on that .

Keywords: Data Mining, Meteorology, Weather, Rainfall, Regression analysis.

Introduction

As we know that India is an agricultural country and the success or failure of the harvest and water scarcity in the year is always considered with the greatest concern. These problems are closely related with the rainy seasons of India [2]. The monsoon plays the major role in India's rain fall.

As we know that climatic changes has vast problem of the world which effect on the social life, economics, food, crops etc. In today's scenario, water crises are big problem because of dry rainfall comes on the earth. Every year rainfall statics changes rapidly. This fluctuation depends on climatic factors like sea level rises, ice sheet melting, pressure, wind speed, population, temperature etc. in these factors, some are directly related to the climate but others are indirectly but the main thing is that these direct or indirect factors effect on the weather's climate. The fluctuation in the factors arise the problem in the world like-food

crisis, water crisis, flood, drought etc. Now as we see in today's scenario that water is major problem all over the world because of rainfall .In India rainfall seasons are divided into 4 parts- pre monsoon(winter) comes in Dec-Apr, summer comes in Apr-June, monsoon (rainy season) comes in June-Sep, post monsoon comes in Oct-Dec. The main rainfall comes in monsoon seasons all over the year. So the monsoon has greater impact on the climate.

So in this paper we have focused on the drought problem of the Rajasthan's one of the region Jaipur.This problem is arises due to less or dry rain. Here we have apply some data mining techniques on the rain fall data set[1] to predict the rain and then apply some statistical analysis on the predictable rainfall data for find drought condition which comes or not.

Proposed work:

Data Analysis And Approach Towards Climate: if we see today's scenario we can see that in climate dramatically changes has come due to climate reasons like global warming, sea level rising, ice sheet melting etc . From all of these reasons now a day's global warming is main issue due to heat up of the earth. As the earth heats up it can increase the temperature of weather due to green houses gases CO_2, methen etc.Resulting in eavesdropping of water from the surfaces and then rain comes.

Human resources also affect on the climate due to increase in population gradually and from these affected fields are forest, water resources, agriculture, and health etc.After analyze the strategy of the climate we can predict the rain by using these climate factors. So In this paper we have use Jaipur rainfall dataset which is collect from the India Meteorological Department [3] and Weather Forecast & Reports - Long Range & Local _ Wunderground _ Weather Underground [1] which contain different climate factors average temperature, gas emission, pressure, winds direction, wind speed . After doing whole study of the data we extract useful data for the research area because there are so many factors as we include are greater impact on the climate. So we extract some climatic factors for this paper.

Now for extracting these factors we apply some data mining techniques. So we apply correlation analysis on the dataset and calculate

correlation between these factors and select those factors whose correlation coefficient is positive which is shown below in tabular form.

S. NO	CLIMATE FACTORS	CORRELATION COEFFICIENTS
1	Humidity	0.6715
2	Max temp	0.216736
3	Pressure	0.380933
4	Winds speed/hr	0.022494

Table-1 correlation between the climate factors

In Data Mining, we use classification, clustering, regression or prediction, association rules etc for extract information. Using the classification we classify what is the reason for rainfall fall in the ground level. By Using clustering technique, we grouping the element that is particular area occupied by the rainfall region. In Prediction we have to predict the rainfall occurs in the particular region in the particular year.

Finally, we perform regression analysis in data mining because rainfall occupied in the region can be done by this approach. In regression analysis we use correlation Coefficient for finding dependency between the variables so that we can say that how many rainfalls fall in the particular region by using different climatic factors.

In the regression analysis there are different types of approach like- linear regression, log based, nonlinear regression for prediction. Here we have used multiple regression approach on the data set. From this approach we can predict rainfall in any one of the future's year by using climatic factors. Now for moving towards this approach first we select 4 climate factors with rain dataset of Jaipur. Then we apply multiple regression approach on that data set and find out predictable equation between rain and climate factors. So MLR is –

Y = 30.7516741 X1 + 37.86154806 X2 + 30.60160245 X3 + 4.528472701 X4 - 34083.26584
Here X1= max. Temp
X2=humidity
X3=max pressure
X4= winds speed
Y=predictable rain

From this equation we can predict rainfall for the future year with knowing max temp, humidity, pressure, wind speed.

After calculate predictable rain then we apply statistical analysis on that data for finding drought possibility. For finding drought possibility we used standard deviation, variance of coefficient, drought indices, drought perception etc.

Statistical Analysis

Regression is a statistical empirical technique that utilizes the relation between two or more quantitative variables.

Y = 30.7516741 X1 + 37.86154806 X2 + 30.60160245 X3 + 4.528472701 X4 - 34083.26584

After calculating MLR equation we can predict rain fall. Now we apply statistical analysis [11] on the resulting equation's data for drought condition. As we already know that drought can be observed by the normal rain which is compare by predictable rain. If rain comes below some level then drought can be occurs otherwise not or some other climate condition can be influenced. So for calculating drought condition we apply some statistical operations on the dataset after MLR which are following-

1) Firstly perform MLR and calculate predictable rain. Then select normal rain, observed rain and predictable rain.
2) Observe these rain data and calculate variation between the rain data.
   a) $D_1 = N_r - O_r$
   b) $D_2 = P_r - O_r$
   c) $D_3 = P_r - N_r$
   Here N_r=average of normal rain
   O_r= average of observed rain
   P_r= average of predict rain
3) Now calculate standard deviation **(σ)**
   $\sigma = \Sigma \, SQRT \, (x_i - x')^2 / SQRT \, (N)$
   X_I=predictable rain
   X'= average rain
   N=number of years (consider rain fall dataset i.e. 20 years) and
   Also N=number of factors i.e. 4
4) Then calculate difference between rain and climate factors by-
   $\Delta = \sigma_{20} - \sigma_4$
5) Then calculate variation between climatic factors i.e. CV-
   $CV = \sigma_4 / x'$
6) Then calculate total difference between dataset i.e. T_d
   $T_d = \Delta - CV$
7) Calculate drought perception Dp
   $Dp = P_r / D_3$
8) After that calculate drought index SPI[10]-
   SPI=X-X'/ Perception=D2/D1
   X= average rain
   X'= difference between the rain
9) Now we compare the parameter by the standard drought indices[12] which is shown below-

S.NO	INDEX	STAGE
1	-.99 to +.99	Normal
2	-1 to -1.49	Drought water
3	-1.5 to -1.99	Drought warning
4	-2 to (<-2)	Drought emergency
By Drought Perception Conditions (%)		
5	$5>D_P>=10$	Light Drought
6	$10>D_P>=20$	Medium drought
7	$D_P<20$	Heavy Drought

Result and conclusion

In this paper we have proposed a model after analysis of Jaipur rainfall dataset which is derived by some data mining techniques like firstly apply correlation analysis secondly regression analysis and last apply statistical analysis. So that we can predict rain in the future year by knowing climate factors and then use this dataset for calculating drought possibility. This is the only prediction regarding rain but not accurate because of climate factors. As we know that climate factors changes due to different reasons and here we have used some factors so other remaining factors can influenced the rain.

The topic climate is very highly complex due to climate conditions which are rapidly changes. So here we apply only prediction on the dataset.

Future Enhancement

Rainfall is essential part of our life. Here we consider Jaipur certain period of rain so that we can predict rain by using data mining techniques. For avoiding flood, global warming, drought etc. In future we have switch on the whole Rajasthan climate for my proposed model with some more climate factors and apply some other techniques artificial intelligence, neural network etc. so that we can give contribution for saving world.

References:

[1]   Weather Forecast & Reports - Long Range & Local _ Wunderground _ Weather Underground http://www.wunderground.com/history/airport/VIJP/2001/1/31/MonthlyHistory.html?req_city=NA&req_state=NA&req_statename=NA

[2]   The Prediction of Indian Monsoon Rainfall: A Regression Approach Goutami Bandyo padhyay.

[3]   **India Meteorological Department** http://www.imd.gov.in/section/nhac/mean/110_new.html

[4]   "Rainfall Forecasting Using Data Mining Technique "M.Kannan, S.Prabhakaran, P.Ramachand .

[5]   Klaush Juliseh, "Data mining for Intrusion Detection – A critical review", Applications of Data mining in computer security, Daniel Barbara, Sushil jajodia, Published by Springer.

[6]   Guhathakurta, P (2005) "Long-range monsoon rainfall prediction of 2005 for the districts and sub-division Kerala with artificial neural network", Current Science, 90, 773-779

[7]   "The Prediction of Indian Monsoon Rainfall: A Regression Approach" by Goutami Bandyopadhyay

[8]   Hastenrath, S (1988) "Prediction of Indian Monsoon Rainfall: Further Exploration", Journal of Climate, 1, 298-304

[9]   Implementation of Data Mining Techniques for Meteorological Data Analysis" (A case study for Gaza Strip) by Sarah N. Kohail, Alaa M. El-Halees , Faculty of Information Technology The Islamic University of Gaza.

[10]   Seasonal and annual regional drought prediction by using data-mining approach" K. YUREKLI,M. TAGHI SATTARI,A. S. ANLI, Department of Farm Structure and Irrigation, Agriculture Faculty, University of Ankara, Ankara-Turkey.

[11]   "Some aspects of daily rainfall distribution over INDIA during the south-west monsoon seasons" M.K.Soman and K. Krishna Kumar.

[12]   "DROUGHT INDICES"Andrej Ceglar Biotechnical faculty, University of Ljubljana