RESEARCH ARTICLE OPEN ACCESS

Estimation of the Influence of Rainfall on the Groundnut yield in India - a Data Mining Approach

Ananthoju Vijay Kumar, Dr. T. V. Rajini Kanth
Faculty Member, Department of Computer Science, Satavahana University, Karimngar, Professor in CSE, SNIST, Hyderabad.

ABSTRACT
In Indian agriculture, particularly, rainfall shows a vivid effect on agriculture. Rain is essential to agriculture because without water no plants will be survived. A regular rainfall is essential for healthy plants, too much or very little rainfall can harm plants and agriculture. Drought can destroy crops and increase corrosion, while excessively wet weather can help harmful fungus grow. Need of rainfall is varying from plant to plant. Since Indian agriculture is rainfall dependent, in the present paper an attempt is made to predict the influence of rainfall on yield of the groundnut in India. To carry the present research work a dataset is prepared with the annual measurements of the groundnut yield and rainfall for a period of 62 years. The groundnut yield and rainfall data is collected from secondary sources like, the Department of Agriculture and Cooperation, India, and the Directorate of Economics and Statistics, India. Over the dataset prepared, an experiment is conducted to know the fact. The experiment results revealed that the groundnut yield is negligible negatively influenced by the rainfall.
Keywords: Data mining, prediction, correlation, regression, correlation analysis, regression analysis.

I. Introduction

India is the second largest country in the world for the agriculture output with at most ten thousand years history. Today, agriculture's share in the GDP of India is about 38.2 per cent, providing employment for 73 percent [1][5][6][9]. Agriculture plays a major role in the development of the nation. India stands first in the production of castor oil seed, pulses, fruits, coriander, jute, pulses, spices, millets, lemons, limes, milk, chilly and peppers, chick peas, ginger, cashew, turmeric, mangoes and meat and cattle production[2][5][10]. India stands second in the production of cabbages, cotton seed and lint, fresh vegetables, garlic, eggs, silk, nutmeg, cardamom, onions, wheat, rice, sugarcane, groundnut, tea, green peas, cauliflowers, potatoes, mace, pumpkins, squashes, gourds and inland fish[2][5][10]. Similarly, India is the third largest producer of tomatoes, rapeseed, coconut, tobacco and sorghum [2][5][10].

Rain makes, rain breaks [7] is the motivation behind doing this experiment. In the experiment, to estimate the influence [3][4][8] of rainfall on the groundnut yield in India, a dataset of groundnut yield per hectare and rainfall is prepared. To assess the impact of rainfall on yield of the groundnut, an experiment is conducted over the prepared dataset with a data-mining tool. The results of the experiment [3][4][8] proved that the groundnut yield is negligible negatively influenced by the rainfall.

II. Growth rates of the groundnut yield during 1950-2011

The groundnut production [10][11][12] in India during the study period was almost declined due to some of the reasons except, a few years, in which there is a drastic increase in yield. During the years 1955-56, 1960-61, 1965-66, 1980-81, 1985-86, 2000-01 and 2010-11, there is a decrease in yield and during the years, 1970-71, 1990-91 and 2005-06 there is a drastic increase in yield.

Year	Yield of the Groundnut	Growth rate of the groundnut
1950-51	775	0.00
1955-56	752	-2.97
1960-61	745	-0.93
1965-66	554	-25.64
1970-71	834	50.54
1975-76	935	12.11
1980-81	736	-21.28
1985-86	719	-2.31
1990-91	904	25.73
1995-96	1007	11.39
2000-01	977	-2.98
2005-06	1187	21.49
2010-11	1144	-3.62

Table 1: Growth rates of the groundnut yield

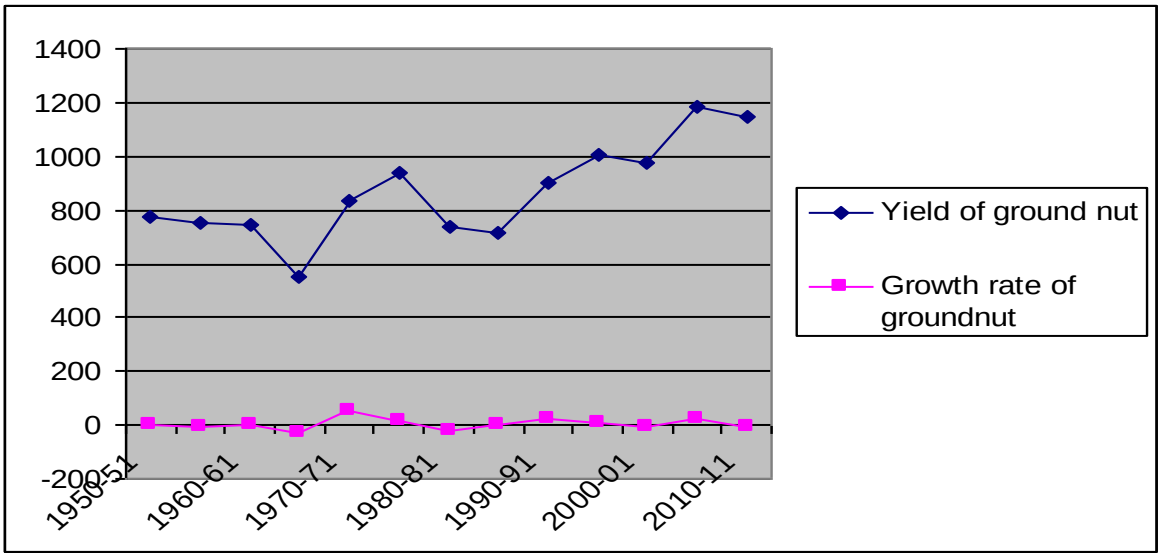

Figure 1: A chart view of growth rates of the groundnut during 1950 - 2011

III. The effect of rainfall on yield of the groundnut in India

To estimate the effect of rainfall on yield of the groundnut in India, a statistical analysis is performed. The following table shows the dataset of rainfall and yield of the groundnut year wise from 1950 -2011.

Year	Rainfall	Groundnut-Yield
1950	1174.2	775
1951	1060.6	649
1952	1110.1	611
1953	1222.1	811
1954	1181.4	766
1955	1275.4	752
1956	1362.6	783
1957	1131.9	734
1958	1312.3	828

1959	1376.9	708
1960	1154.8	745
1961	1399.2	725
1962	1198	695
1963	1220.9	769
1964	1244.4	814
1965	947.4	554
1966	1058	604
1967	1154	759
1968	1059.3	653
1969	1147.8	720
1970	1255	834
1971	1216.9	823
1972	947.1	585
1973	1219.5	845
1974	1055.3	724
1975	1294.8	935
1976	1131.6	747
1977	1269.7	866
1978	1237.2	835
1979	1030.2	805
1980	1182.3	736
1981	1170.7	972
1982	1084.4	732
1983	1320.9	940
1984	1160.8	898
1985	1144.9	719
1986	1137.6	841
1987	1088.9	855
1988	1342.1	1132
1989	1127.4	930
1990	1401.4	904
1991	1170.2	818
1992	1102.7	1049
1993	1207.8	941
1994	1295.3	1027
1995	1242.4	1007
1996	1182.9	1138
1997	1183.1	1040
1998	1208.8	1214

1999	1116.6	766
2000	1035.4	977
2001	1100.7	1127
2002	935.9	694
2003	1187.3	1357
2004	1106.5	1020
2005	1208.3	1187
2006	1161.6	866
2007	1179.3	1459
2008	1118	1459
2009	953.7	1163
2010	1215.5	0
2011	1116.3	1144

Table 2: Rainfall and the groundnut yield dataset

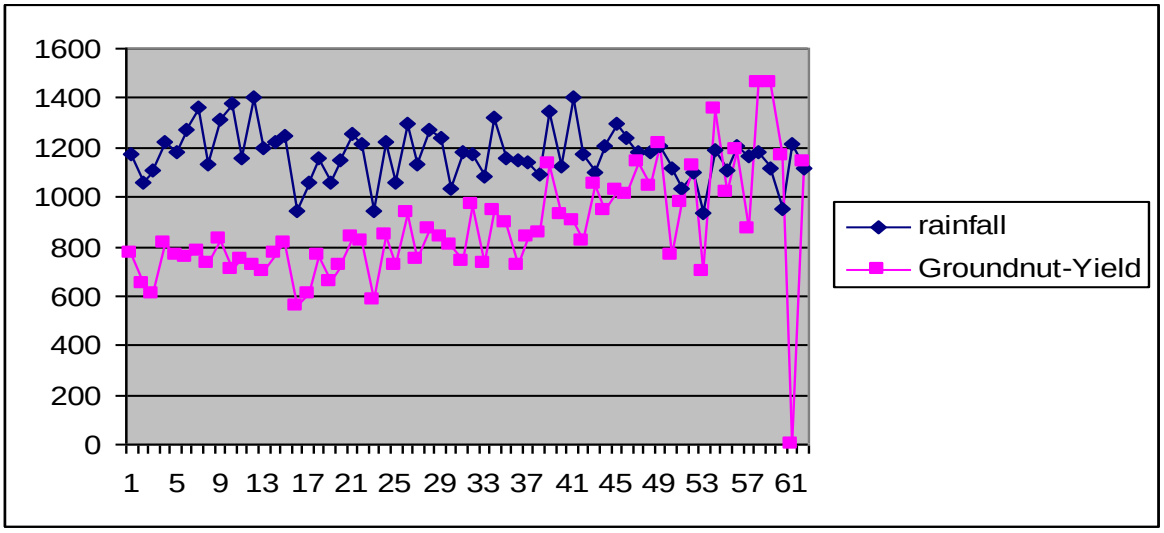

Figure 2: A chart view of rainfall and the groundnut yield

IV. Correlation analysis of the rainfall and yield of groundnut

To know the correlation between rainfall and yield of the groundnut in India during the study period, the correlation analysis is performed over the dataset of rainfall and the groundnut yield. The correlation between yield and rainfall is -0.44, this indicates that there is a very low negative impact on the yield of groundnut. The following table shows the results of the correlation analysis of groundnut yield and rainfall.

Correlations

		GROUNDNU	RAINFALL
Pearson Correlation	GROUNDNU	1.000	-.044
	RAINFALL	-.044	1.000
Sig. (2-tailed)	GROUNDNU	.	.803
	RAINFALL	.803	.
N	GROUNDNU	34	34
	RAINFALL	34	62

Table 3: Correlations

V. Regression analysis of the rainfall and yield of groundnut

To measure the dependency between the groundnut yield and rainfall in India, regression analysis is performed over the dataset of the groundnut yield and rainfall. The results of regression analysis also proved that there is low negligible negative dependency between yield of the groundnut and rainfall during the study period. The following table shows, results of the regression analysis.

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	RAINFALL[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: GROUNDNU

Table 4: Variables Entered/ removed

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.044[a]	.002	-.029	218.7990

a. Predictors: (Constant), RAINFALL

Table 5: Model Summary

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3017.696	1	3017.696	.063	.803[a]
	Residual	1531936	32	47873.002		
	Total	1534954	33			

a. Predictors: (Constant), RAINFALL

b. Dependent Variable: GROUNDNU

Table 6: ANNOVA

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1076.058	401.691		2.679	.012
	RAINFALL	-8.49E-02	.338	-.044	-.251	.803

a. Dependent Variable: GROUNDNU

Table 7: Coefficcients

Output of the Linear Regression of rainfall and groundnut yield

Independent variable:	RAINFALL
Dependent variable:	Groundnut yield
Method:	Linear regression
R squire	0.010
Degree of freedom:	60
F	0.60
Significance level:	0.442
b0	611.989
b1	0.2151

The Linear Regression Equation of the groundnut yield and rainfall is as follows

$$Y=611.989+0.2151X \quad \text{--- (1)}$$

Where,
X: Independent variable (Rainfall)
Y: Dependent Variable (Groundnut yield)
b0: 611.989
b1: 0.2151

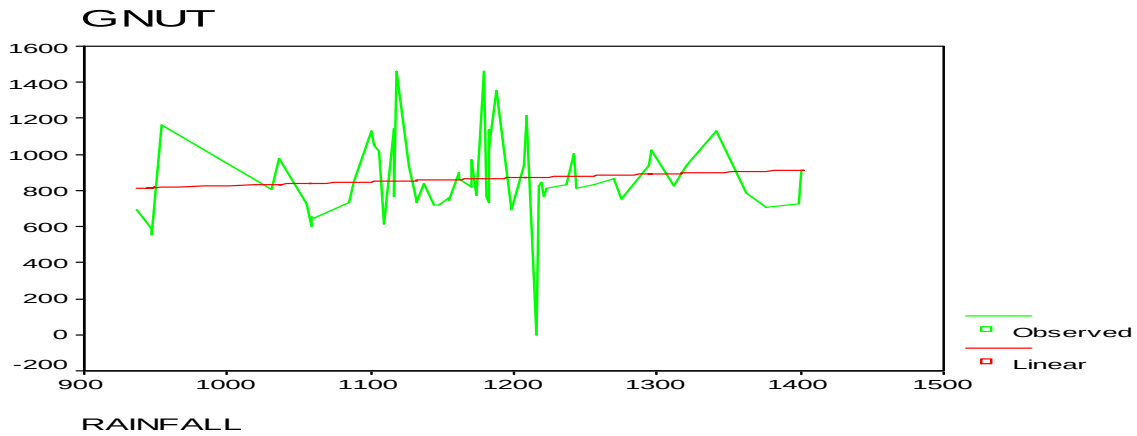

Figure 3: The Linear Regression Curve of the groundnut yield and rainfall

VI. Clustering with K-means

To classify, yield of the groundnut in India during the study period, K-means quick clustering technique is used. Initially, the K-means divided yield data into two clusters with center points, 353, and 1021. Finally, K-means grouped yield data again into two clusters only, with center points, 516.56, and 818.87. The variation in the first and second cluster center points is 163.563and 202.133 respectively. The first cluster has 32 cases and the second cluster has 30 cases only. The total valid cases are 62 and invalid cases are null. The following tables shows, the results of k-means quick clustering.

Initial Cluster Centers

	Cluster	
	1	2
JOWARYEI	353.00	1021.00

Table 8: Initial Cluster Centers

Iteration History[a]

Iteration	Change in Cluster Centers	
	1	2
1	163.563	202.133
2	.000	.000

a. Convergence achieved due to no or small distance change. The maximum distance by which any center has changed is .000. The current iteration is 2. The minimum distance between initial centers is 668.000.

Table 9: Iteration History

Final Cluster Centers

	Cluster	
	1	2
JOWARYEI	516.56	818.87

Table 10: Final Cluster Centers

Number of Cases in each Cluster

Cluster	1	32.000
	2	30.000
Valid		62.000
Missing		.000

Table 11: Number of Cases in each Cluster

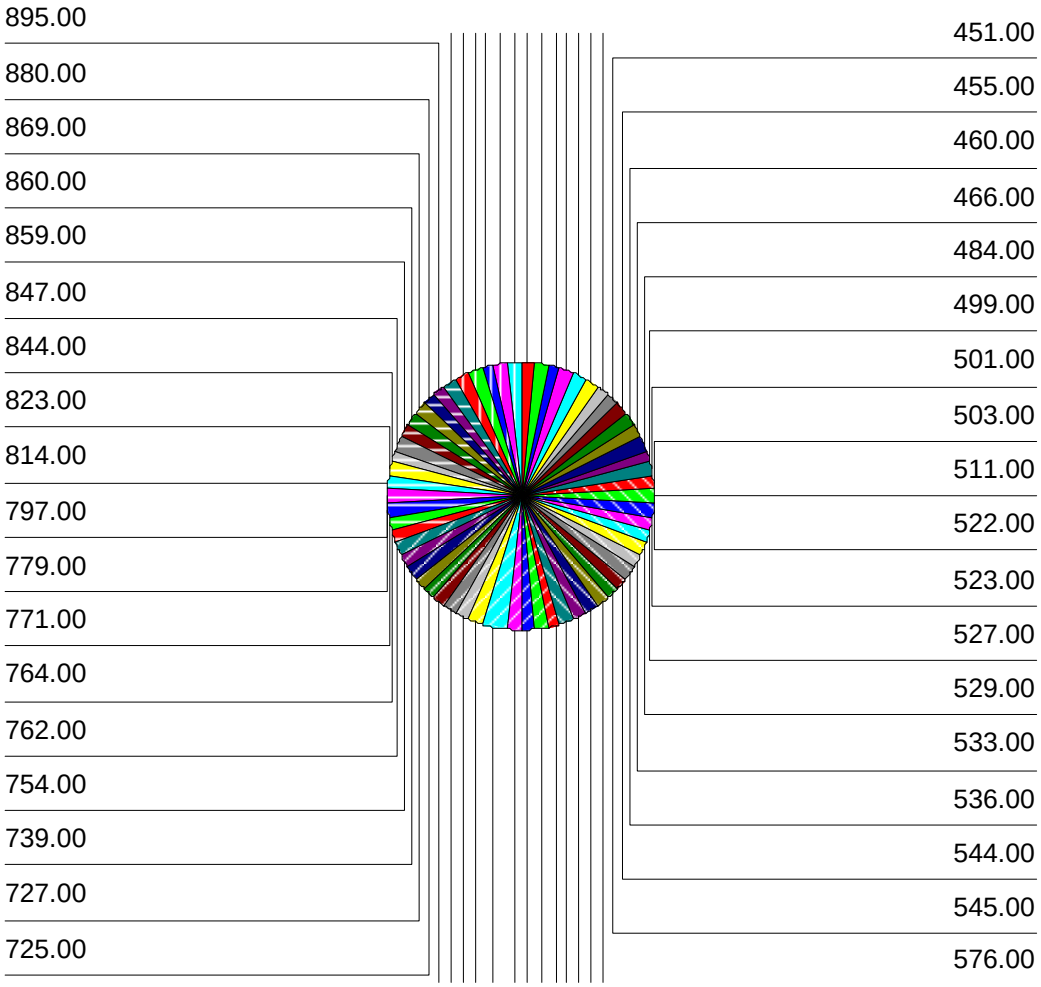

Figure 4: A pie chart view of the groundnut yield using K-means clustering technique

VII. CONCLUSION

In India, the groundnut yield has some ups and downs, but on the whole, yield is increased from 1950 – 2011. The groundnut yield is very high in the years, 2007, and 2008 and very low in the year, 1952.The present research work is mainly devoted for the estimation of rainfall impact on the groundnut yield. In the experiment, it is proved that the groundnut yield is extremely little negatively influenced by the rainfall. Finally, on the basis of the results of the experiment we can conclude that to have high yields, India needs a moderate rainfall.

REFERENCES

[1] Economic survey 2010.
[2] Hand book of statistics on Indian Economy 2013.
[3] Ananthoju Vijay Kumar, Dr.T.V.Rajini Kanth, "A Data mining approach for the estimation of climate change on the Jowar crop yield in India," International Journal of Emerging Science and Engineering, Dec 25, 2013, Vol/Issue: 2/2, Pages: 16-20.
[4] Ananthoju Vijay Kumar, Dr.T.V. Rajini Kanth, "Estimation of the Influence of the fertilizer nutrients consumption on the Wheat crop yield in India – a data mining approach", Vol/Issue:3/2, Pages: 316-320.
[5] http://wikipidea.org/wiki/Agriculture_in_India.
[6] State of Indian Agriculture, www.agricoop.nic.in/SIA111213312.pdf
[7] http://kvgktrailblazers.weebly.com/rain-and-agriculture.html.
[8] Ananthoju Vijay Kumar, Dr.T.V.Rajini Kanth, "Estimation of the influential factors of the rice crop yield in India, ICACM-2013, 2nd International Conference, ISBN NO: 9789351071495.
[9] http://indiastat.com.
[10] http://apy.dacnet.nic.in.

[11] http://mospi.nic.in.
[12] http://isas.org.in