ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

Clustering Based Crop Diversification Model Using K-Means Algorithm in Indian Punjab Region

[1] RadhikaSethi, [2] Sikander Singh Cheema
[1][2] Department of Computer Engineering, Punjabi University, Patiala, India

Abstract: **- A major area of research in Indian Punjab is in crop diversification. Crop Diversification means cropping system to agricultural production on same farm to increase biological stability of the farm. Then present studies and attempt to focus on the systematic review of the literature about the crop diversification and role of clustering in crop diversification. The finding showed that clustering provide better crop diversification. We will also consider possible future work to provide different clustering techniques to gain most of the knowledge in crop diversification.**

Keywords**: Diversification, Clustering, Cultivation, Model, Crop, Soil.**

I. INTRODUCTION

In Punjab the occupation of most of the people is agriculture. Agriculture is one of the most important land uses and its management practices have strong impact on the environment [4]. Their Income greatly affects with the production of crops. For sustainability in agriculture, it is necessary that particular crop or number of crops should not be grown in same soil for long run because various crops tap have many kind of soil layer to meet their water and nutrient requirement. So for this there is a need of diversification in cropping system [1]. Different algorithms have been created in different areas to diversify the crops. So, in this way we are trying to develop an algorithm which can help to provide a suitable crop according to soil specification. Crop diversification in different seasons (kharif, summer, Rabi) enhances the employment level and regular return to farmers through the year and maintains crop productivity without harming the soil health after year [2]. It was also observed that crops sown during different seasons have a good effect on soil rather than opposite effect. Pulses cultivation is good for soil health and increase the nutrient status.

II. OBJECTIVES

In view of above data, we come to think about expansion and need of diversification. Presently we have chosen to create clusters for Malwa, Majha, and Doaba area. The clusters will be based on the crops suitability. Each cluster will be made for different region. The main objectives of this study are defined as:-
Study of soil specification and corresponding crops. Study of different clustering methods and implementation using k-means algorithm.

III. BRIEF INTRODUCTION ABOUT DIFFERENT CROPS

There are various crops that are cultivated in Punjab but as wheat and rice are the two major crops. A brief description about the 10 major crops is provided below:-
1.Wheat: Wheat is main crop of Punjab. 20 percent of wheat in India is grown in Punjab. Wheat was native to the soil of Punjab [2]. It is developed in assortment of soils in India. It is rabbi season crop. The underlying climate necessity for wheat ought to be warm and clammy and ought to be radiant and dry toward the end. Temperature ought to be between 160C to 200C. Rained loam and clay loam are good for wheat. Sandy loam and black soil is also best for wheat production. It is the 1st major crop in all regions Malwa, Majha except region Doaba.
2.Rice: It is the most important crop of Punjab. It is a Kharif season crop. It covers around 33% of aggregate developed zone of the nation and India's12 percent rice is developed in Punjab [5]. It is likewise developed in Rabi season with the utilization of water system. It develops well in hot and moist conditions.Temprature ought to be between 220C to 320C and rainfall ranging between 150 cm to 300 cm is suitable for growth. Rice can be developed in different soil conditions however deep clayey and loamy soil gives better outcomes.

3.Cotton: It is 3rd major crop of state. This crop is sown mainly in Malwa region .It is kharif season crop and yield span is around 175 to 225 days. It requires constant temperatures between 18 and 30°. Ample sunshine and fairly dry conditions are good for this crop. All kinds of soil except coarse sandy, saline are suitable for cotton.

All Rights Reserved © 2018 IJERCSE

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

4. Maize: It is 4th major crop of state. It is kharif season crop and duration of this crop is about 175 to 225 days. It requires constant temperatures between 21° and 27°. It is grown well in alluvial soil. It is developed generally in areas having yearly precipitation between 60 cm to 110 cm.

5. Oil Seeds: It is Rabi season crop and cool climate is required for this crop, but this is more delicate to climatic variables than different yields developed in this season. It requires constant temperatures between 150° to 25°. At aging stage as germination high temperature is required. Cold temperature, bright sun shine and enough soil moisture increases oil content of seed. It was the 5th major crop . Oil Seeds had involved the third to fifth rank in all regions.

6. Potato: It is kharif season crop and cool climate is required for this crop, but it is adjusted to extensive variety of atmosphere conditions. Crop duration is about 70 to 80 days. It requires constant temperatures between 150° to 250° [5]. it is developed generally in areas having yearly rainfall 12mm. for this crop, well drained loose friable, non-saline, non-albaline, loamy sand to sandy loam soils are suitable. PH value should be 5.0 to 8.0.

7. Sugarcane: It has 7th rank in the state. It is kharif season crop and considered as tropic plant. Crop span is around 480 to 5400 days. Temperature for this crop shifts between 32° to 38° [5]. It needs a long warm developing season. it is developed generally in areas having yearly rainfall 1100mm to 1500mm.

8. Pulses: ranked 8th in the state. It is a Rabi season crop which is sown between September and November and is grown between February and April. It is grown well in cool and dry climate with constant temperatures between 20° and 25°.it is grown well in loamy soil.

9. Tomato: ranked 9th in the state. It is a warm season crop it grows well in area that are free from frost. Temperatures between 15° and 27°. It is grown well in sandy loam with well drained clay. A warm, sunny weather is most suited for proper ripening, shading quality & high return.

10. Onion: It is rabi season crop. It grows well under mild climate.. It requires constant temperatures between 15.5° to 21°. It is grown well in sandy loam, slit loam and heavy loam. PH value should be 5.8 to 6.5.

IV. DATABASE OF CROPS AND CORRESPONDING REQUIRED PARAMETERS

The given below table provides a description about 10 major crops their weather condition, temperature and soil type specifications:

S.no	Name of Crop	Weather	Temperature	Soil type
1.	Wheat	Warm and moist	16°- 20°	Heavy Loam
2.	Rice	Hot and Humid	22°-32°	Slit Clay
3.	Cotton	Sunshine and Fairley dry	18°-30°	All types except coarse sandy, saline
4.	Maize	Rainfall	21°-27°	Slit Loam
5.	Oilseed	Cool and dry	15°-25°	Sandy loam
6.	Potatoes	Hot	45°-80°	Clay Loam
7.	Sugarcane	Warm and humid	32°-38°	Light Clay
8.	Pulses	Mild cool and dry climate	20°-25°	Loamy soil
9	Tomato	Warm sunny weather	15°-27°	Sandy loam
10	Onion	Mild	15.5° - 21°	Sandy loam

After doing the study we come to know that weather, temperature and soil type plays an important role in good production of crops. The impacts of environmental change are as take after:-

- Due to increment or decline in temperature of air or soil, the growth of crop is affected.
- The changes in climate also affect the crop's biophysical processes.
- The effects of changes in soil properties are as follows:-

a) Poor growth of crops.
b) For one kind of crop this type of soil can be useless.
c) Result in low yield of crops.

There are many ways to make soil perfect for crops that are as follow:-

a) From outside by adding nutrition elements like phosphorous, carbon, nitrogen etc.
b) Another way is crop diversification which is to grow one kind of crop for one time and for next time growing different crop on same land. When same type of crop is grown for year's soil loses its nutrition values. For example if wheat and rice are grown in same land for many years, it can be circumventing by using crop diversification.

V. CROP DIVERSIFICATION

Diversification means adding a new crop or cropping system to agricultural production on same farm to increase biological stability of the farm, to fulfill the

All Rights Reserved © 2018 IJERCSE

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

market demand, to get proper rates of crops and for higher market opportunities [2]. There is a need to meet food grain requirement of growing population as well as sustaining reasonably high productivity level. So there is high need to diversify into new area. It is an efficient strategy for achieving the objectives of and maximum income growth and food security [6].

The various benefits of crop diversification are listed below:

The better use of land, water, manpower and other resources.

Realizing quicker returns for the farmer.

Reducing the risk factor of crop failure, market failure loss of yield.

Less damage to equipment.

Increase productivity.

To improve soil fertility.[3]

5.1 Diversification in Punjab

As we know Punjab is divided into three regions: Majha, Malwa and Doaba. Each region is divided on the bases of soil type, temperature and weather conditions. Now we will select the crop according to soil and weather conditions of particular area. For example-wheat requires an area which is warm and moist and temperature must be between 160C to 200C. Rained loam and clay loam are good for wheat. Sandy loam and black soil is also best for wheat production. So for wheat production we will choose the area which has comparative properties as said above and similarly for other crops. Now subdivision is necessary as it helps to classify region on basis of soil and weather conditions. For subdivision, we need to make clusters. For this purpose, various clustering techniques can be employed.

VI. CLUSTERING

Clustering means Partitioning of data into sub classes or grouping of similar objects or partitioning of data based on similarities. Clustering helps user to understand data set. In clustering inter group similarity of data should be low and intra group similarities should be high.

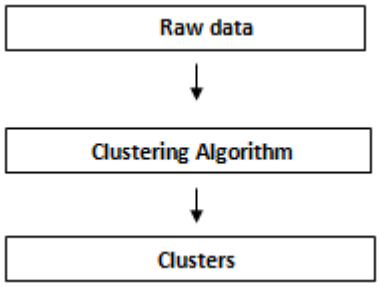


Fig 1: Steps of Clustering

6.1 Clustering Methods

1. Partitioning method: clustering based on data partition.
2. Hierarchal method: finding a new clusters using previously found area.

a) Agglomerative method: It is a bottom up approach. Starts from one point and add two or more clusters. Stops when all groups are merged or k number of clusters is formed.

b) Devise method: It is a bottom up approach. Recursively separate the cluster into sub clusters. I stop when k number of cluster is achieved.

3. Density based method: based upon density region with sufficiency high density point.
4. Model based method: it is formulated for each cluster and try to find best fit model to each other.
5. Constraint based method: this model is based on some specific constraints..

VII. IMPLEMENTATION OF PROPOSED METHOD WITH K-MEANS CLUSTERING ALGORITHM

K Means clustering is an unsupervised learning estimation that tries to group information in light of their closeness. Unsupervised learning implies that there is no result to be anticipated, what's more, the calculation just tries to discover designs in the information. In k implies clustering, we have to indicate the quantity of clusters we need the information to be gathered into[11]. The calculation haphazardly appoints every perception to a cluster, and finds the centroid of each group. At that point, the calculation repeats through two stages:

- Reassign information focuses to the cluster whose centroid is nearest.
- Calculate new centroid of each cluster.

These two steps are rehashed till the inside cluster variety can't be diminished any further. The inside cluster variety is figured as the aggregate of the Euclidean separation between the information focuses and their particular cluster centroids

$$\text{EuclideanDistance} = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}$$

In our problem, the total area of Punjab will be divided into 3 clusters, depending upon the Weather, Temperature, and Soil type. Using K-mean algorithm, minimum distance will be treated as calculated value of these parameters (Weather, Temperature, and Soil type), if values of these parameters will be optimum for the production of a particular crop then that crop will be listed in that particular cluster, in this case Weather,

All Rights Reserved © 2018 IJERCSE

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 5, Issue 4, April 2018

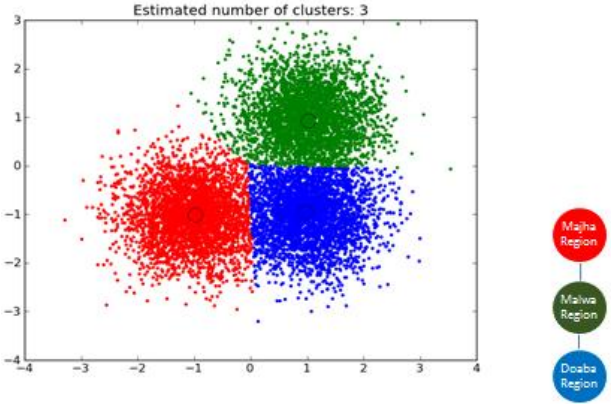
Temperature, and Soil type will be input to the K-Means algorithm and output will be the name of the crop. So values of these parameters will be assigned as a training set x (1),..., x(m), and want to group the data into 3 cohesive "clusters." Here, we are given feature vectors for each data point x (i)∈Rn as usual; our goal is to predict k crops.

For each data point. The k-means clustering algorithm is as follows:

1. Initialize Cluster centroids $\mu 1, \mu 2, \ldots \mu_1, \mu_2, \ldots\ldots\ldots, \mu_k \in R^n$ randomly.
2. Repeat until convergence: {
   For every i, set
   $$c^{(i)} := \arg\min_j \| x^{(i)} - \mu_j \|^2 .$$
   For every j, set
   $$\mu_j := \frac{\sum_{i=1}^{m} 1\{c^{(i)} = j\} x^{(i)}}{\sum_{i=1}^{m} 1\{c^{(i)} = j\}}$$
   }

As described in above algorithm μ is number of parameters(soil,weather,temperature), i is number of crops, C is number of clusters.

Here the value of C will be 3 because we have taken data of 3 regions that are Majha, Malwa and Doaba. After giving these inputs (soil,weather,and temperature) k-means algorithm will provide the suitable results. K-means algorithm helps to generate the clusters of the region on the basis of their soil and climate conditions. The k-means algorithm will help to identify suitable crops for diversification for each region.



We have taken data from three regions Majha, Malwa and Doaba, Total number of instances were taken 3000, which means 1000 from each region. As I have taken 10 Crops and computed on 100 variations for each crop. On the bases of above mentioned parameters(Temprature,soil and weather) different values are given to the algorithm. After processing these given inputs using the k-means algorithm the value of the suitable crop will the output.

VIII. CONCLUSION

After the review process we came to know about that crop diversification can provide benefits to farmers as well as for soil wellness. Proper crop diversification is possible with the process of clustering. K-means algorithm helps cluster the region on the bases of their soil and climate conditions. After clustering, the diversification results reveals that these crops should be in these regions according to the soil and weather requirement of crops. Wheat can be replaced with potato, Rice can be replaced with maize/pulses/sugarcane/cotton. By doing this farmers can get maximum profit and other benefits too. Farmers can get proper rates of their crops.

S.no	Name of Crop	Regions		
		Majha	Malwa	Doaba
1.	Wheat	Yes	Yes	Yes
2.	Rice	Yes	Yes	No
3.	Cotton	Yes	No	No
4.	Maize	No	No	Yes
5.	Oilseed	Yes	Yes	Yes
6.	Potatoes	Yes	Yes	Yes
7.	Sugarcane	Yes	No	Yes
8.	Pulses	Yes	Yes	No
9	Tomato	No	Yes	Yes
10	Onion	Yes	Yes	No

IX. FUTURE SCOPE

There is high need to diversify into new area. Crop Diversification is an efficient strategy for achieving the objectives of maximum income growth and food security. The various clustering and classification techniques can be used for diversification purpose. Also Feature Selection at various levels can help to improve the process of diversification.

REFERENCES

[1] R. D. Singh, Shivani, A. R. Khan & N. Chandra(2012) , Sustainable productivity & profitability of diversified rice-based cropping system in an irrigated ecosystem, Archives of Agronomy and Soil Science, Vol. 58, No. 8, 859-869.

All Rights Reserved © 2018 IJERCSE

[2] Nivedita Sharma, S.P Singh (2014), Agriculture Diversification in Indian Punjab: An Assessment of Government Intervention through Contract Farming, Journal of Agriculture & Food Information, 191-213.

[3] J.G. Lauren, R. Shreshta, M.A Sattar and R.L. Yadav (2008), Legumes and Diversification of the Rice-Wheat Cropping System, Journal of Crop Production, 3:2, 67-102.

[4] Louis Mahy, B.E.T.I. Dupeux, G.V. Huylenbroeck, JeroenBuysse (2015), Simulating farm level response to crop diversification policy, Land Use policy, 36-42.

[5] Parneetkaur, Sikander Singh Cheema (2017), Crop Simulation Model for Malwa Region Punjab, International Journal of Advanced Research in Computer Science, Vol.8, No.7, 172-175.

[6] B. Khoshnevisan, E. Bolandnazar, S. Barak, S. Samshirband ,H. Maghsoudlou, Torki A. Altameen, A. Gani (2015), A clustering model based on anevolutionary algorithm for better energy use in crop production, 1921-1935.

[7] Rahul Sarkar, Tapabrata Ray(2009), An improved evolutionary algorithm for solving multi-objective crop planning models, Computers and Electronics in Agriculture 68, 191-199.

[8] R. k Singh, J. S bohra, T. Nath, Yeshwant Singh and Kalyansingh(2011) Integrated assessment of diversification of rice-wheat cropping system in indo-Gangetic plain, Archives of Agronomy and Soil Science, 57:5, 489-506.

[9] S.K Sharma, S.N. Sharma (2008) Effect of Crop Diversification of Rice-Wheat Cropping System on Productivity and Profitability, Journal of Sustainable Agriculuture,26:1,39-48.

[10] M. R. Herman, A. P. Nejadhashemi, F. Daneshvar, M. Abouali, D. M. Ross, S. A. Woznicki, Z. Zhang(2016), Optimization of bioenergy crop selection and placement based on a stream health indicator using an evolutionary algorithm, Journal of Environment Management, 413-424.

[11] G Sehgal, ,Dr. K Garg (2014), Comparison Of Various Clustering Algorithms, Journal of computer Science and Information Technologies,vol.5(3), 3074-3076

All Rights Reserved © 2018 IJERCSE