International journal of science, engineering, and information technology
Volume 02, Issue 01, December 2017
Journal homepage: https://journal.trunojoyo.ac.id/ijseit

Application of Land Grouping Using Self Organizing Map (Som) in Sumenep Regency

Aditami Yoanta[a], Rika Yunitarini[a]

[a]Informatics Engineering Study Program, Trunojoyo University Madura, Bangkalan, Indonesia

ABSTRACT

Farm is one of the biggest main sector that give the highest economic contribution towards sumenep regency. One of the procedure is aimed to sharpening the farming quality and gradeness, Service agriculture sumenep regency structuring a program for farmer empowerment and farmer organization in increasing productivity, prosperity and farmer income through farmig ability increasement accebility in palawija comodity including palawija commodity equalization potention. For the realisation, goverment need to mapping the area potential characteristic in each region based on potential farm list, it will give a farming organization needs an improvement.Self organizing map (som) is used as a methodology in this final assignment as a unsupervised clustering method that appropriate to have a result potential characteristic area in in each cluster that get with area potential indicator area which make goverment decrit clarification easier in spreading potential palawija farming.

Keywords: Clustering Aplication, Farming, Clustering, Palawija Commodity, Palawija, Self Organizing Map, SOM.

Article History
Received 10 August 17 Received in revised form 20 August 17 Accepted 10 November 17

1. Introduction

Sumenep Regency has Agriculture is one of the main sectors that provides the largest contribution to the economic structure. Where agriculture becomes the main point of regional development in realizing the mission of the Sumenep Regency Government in improving the quality of the environment and optimism in the utilization of natural resources. From the main agricultural products of the various Madurese communities, the commodity palawija as a sub-sector of the needs of the community which has a very important role in meeting food needs, but the agricultural product of the palawija commodity needs to develop the potential of plant quality so that it can be used as a regional superior product.

The difficulty of the Department of Agriculture of Sumenep Regency in determining the potential of each region can prevent local governments from providing policies on agricultural needs, data that can change at any time are also a supporting factor and there is no mapping of land potential in accordance with local government criteria. To overcome this, the government needs to map the characteristics of the potential of land in each of its regions so that it can be seen the similarity of the characteristics of each district to be given further policies in improving regional agricultural progress.

In this study, Clustering Self Organizing Map is an effective data mining method as a reference to map and classify the characteristics of the potential of plants in each sub-district, especially in crops. This method is one of the artificial neural network models that uses unsupervised methods that classify classes based on the same characteristics. By carrying out the clustering process the agriculture department of Sumenep Regency can understand the potential segmentation of crops commodities in each sub-district. The existence of this research is expected to be able to show sub-district and problem groups - issues that need attention in efforts to equalize the potential of secondary crops, as well as being used as a reference for local governments to take agricultural policy.

2. Literatur Review

2.1. Agriculture

Fatah [7] Definition of Agriculture is divided into 2, namely agriculture in the broad sense is a business field that includes the fields of plants, livestock, and fisheries. While in the narrow sense is a business only in the field of plants, without expressing the factors that support. Broadly speaking, agriculture is a business or cultural activity which covers the fields of plants, animal husbandry and fisheries, as well as environmental factors that affect its production such as climate, soil, pests and diseases and the technology used, as well as processing and marketing of products. All these activities are carried out in order to meet the needs and obtain income and not damage the environment.

* *Corresponding author*
E-mail address: adit.amiy@gmail.com, bainkk@gmail.com, rika_yunitarini@yahoo.com.

E-ISSN 2548-4214

2.2. Palawija Commodity

In the Indonesian Big Dictionary [8] Palawija is a plant other than rice; planted in rice fields or in fields (such as beans, corn, sweet potatoes). [9] Palawija (Sanskrit: phaladwija) literally means the second plant. Based on the meaning of Sanskrit, palawija means the second yield, and is the second crop besides rice. The term palawija developed among farmers in Java to refer to other types of agricultural crops besides rice.

Suparman [10] The term Palawija is the origin of the term given by the farming community in rural Java, especially farmers in rural Central Java, for types of plants other than rice. From here on, the term spread widely and has become a general term in the field of agricultural products on dry land.

In the farming community in rural Java, the types of secondary crops are grouped into several types, namely:

1. Nutmeg Nutmeg, for the mention of palawija plants whose fruit is hanging in the trunk of a tree and away from the ground surface. These types for example: Corn, Soybeans, Green Beans, and the like.
2. Nutmeg Kependhem, namely for the mention of plants whose fruit is in the ground. These types such as: Sweet Potatoes, Cassava etc.

2.3. Data Mining

In simple terms data mining is mining or finding new information by looking for certain patterns or rules from a very large amount of data. Data mining is also referred to as a series of processes to explore the added value of knowledge that has not been known manually from a data set. Data mining, often also referred to as knowledge discovery in database (KDD). KDD is an activity that includes the collection, use of data, historical to find regularities, patterns or relationships in large data sets. Data mining is the activity of finding interesting patterns of large amounts of data, data can be stored in a database, data warehouse, or other information storage.

Data mining is related to other fields of science, such as database systems, data warehousing, statistics, machine learning, information retrieval, and high-level computing. In addition, data mining is supported by other sciences such as neural networks, pattern recognition, spatial data analysis, image databases, signal processing [11].

Data mining is defined as the process of finding patterns in data. This process is automatic or often semi-automatic. The pattern found must be meaningful and the pattern provides benefits, usually economic benefits. Data needed in large quantities.

2.4. Cluster Analysis

Cluster analysis is organizing a collection of patterns into clusters (groups) based on their similarity. The patterns in a cluster will have similar characteristics / characteristics than the patterns in other clusters [12]

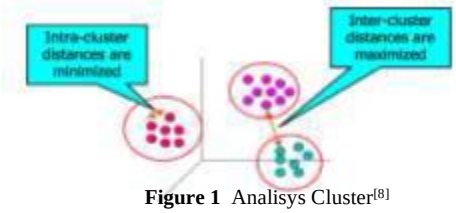

Figure 1 Analisys Cluster[8]

Clustering is different from classification, in that there are no target variables for clustering. Clustering does not classify, predict, or predict the value of a target variable. Clustering algorithms are used to determine the segment of the entire set of data into relatively equal subgroups or clusters, with the similarity of records in the cluster maximized and the similarity of records outside the cluster drunk [7].

2.5. Neural Network

The neural network training aims to find the weights found in each layer. There are two types of training in artificial neural network systems, namely:

1. Supervised Learning. In this training process, the network is trained by providing data called training data which consists of expected input-output pairs and is called associative memory. After the network is trained, associative memory can remember a pattern. If the network is given new input, the network can produce outputs as expected based on existing patterns.
2. Unsupervised Learning. In this training process, the network is trained only by being given input data that has similar characteristics without being accompanied by outputs.

Unsupervised Learning is a method that does not require target output. In this method the results cannot be determined as expected during the learning process. During the learning process, the weight values are arranged in a certain range depending on the value of the input provided. The purpose of this learning is to group almost the same units in a certain area. Learning like this is usually very suitable for grouping (classification) patterns. In the process of learning without supervision, ANN will classify examples of input patterns available into different groups. The network used for the training process without supervision is a feedback network.

Recurrent structure (feedback) is if a network is repeated (has a connection back from output to input) will cause instability and will produce very complex dynamics. Repeated networks are very interesting to study in Artificial Neural Networks, but so far the feedforward structure is very useful for solving problems. Included in the recurrent (feedback) structure:

a. Competitive Network
b. Self-Organizing Map
c. Hopfield networks
d. Adaptive-resonansetheory models.[13]

2.6. Data Normalization

Normalization is a preprocessing process carried out to reshape existing data with a range of 0 (zero) to 1 (one). This is done with the intention that the data processed has a narrower range, but still accommodates the differences and values of the data.

Although there is a nominal change in the value of the data to be processed, it should not have an impact on the change in the value of the data. This is done by comparing with data that has maximum and minimum for each variable. Data with a maximum value will be worth 1 (one) while data with a minimum value will be given a value of 0 (zero). Thus, data that has a value between maximum and minimum will have values greater than 0 (X> 0) and smaller than 1 (X <1) or with mathematical notation, these conditions can be expressed with {X | 0 <X> 1, X R}. Thus, there are

equations that can be used in the process of data normalization as explained above.

$$\sum_{i=1}^{n} f(x) = \frac{X_i - X_{min}}{X_{maks} - X_{min}} \quad (1)$$

Information:
Xn : normal data
Xi : actual data
Xmin : data with a minimum value Xmax: data with a maximum value [4]

2.7. Clustering

Muhammad[14] Clustering is unsupervised classification of certain patterns in certain clusters. The problem of clustering has been addressed in many contexts and by many researchers in many branches of science, this illustrates the broad appeal and usefulness as one way in exploratory data analysis. In general, clustering is a method of dividing data into groups based on their similarity. One significant research area in data mining is to develop methods to modernize knowledge using existing knowledge, because in general it can improve the efficiency of extracting data, especially for very large databases.
Clustering has many methods, one of which is SOM. However, of the many methods possessed by the clustering can be grouped into two major groups, namely clustering with a hierarchical approach (Hierarchical Approach Clustering) and clustering with a partitioning approach (Partial Approach Clustering).

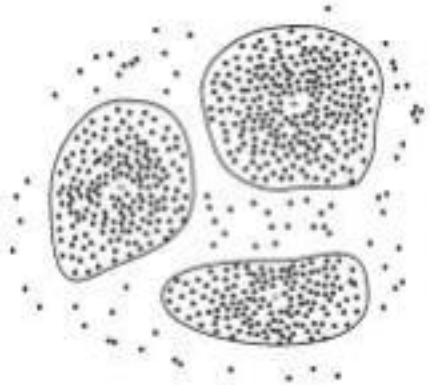

Figure 2. Illustration of Clustering [14]

2.8. Self-Organizing Map (SOM)

Fahlisian [3] The SOM (Self-Organizing Map) or Kohonen network is one of the neural network models that uses unsupervised learning methods. The Kohonen SOM network consists of two layers, namely the input layer and the output layer. Each neuron in the input layer is connected to each neuron in the output layer. Each neuron in the output layer represents the class of the given input.

3. System Planning

3.1. System flowchart

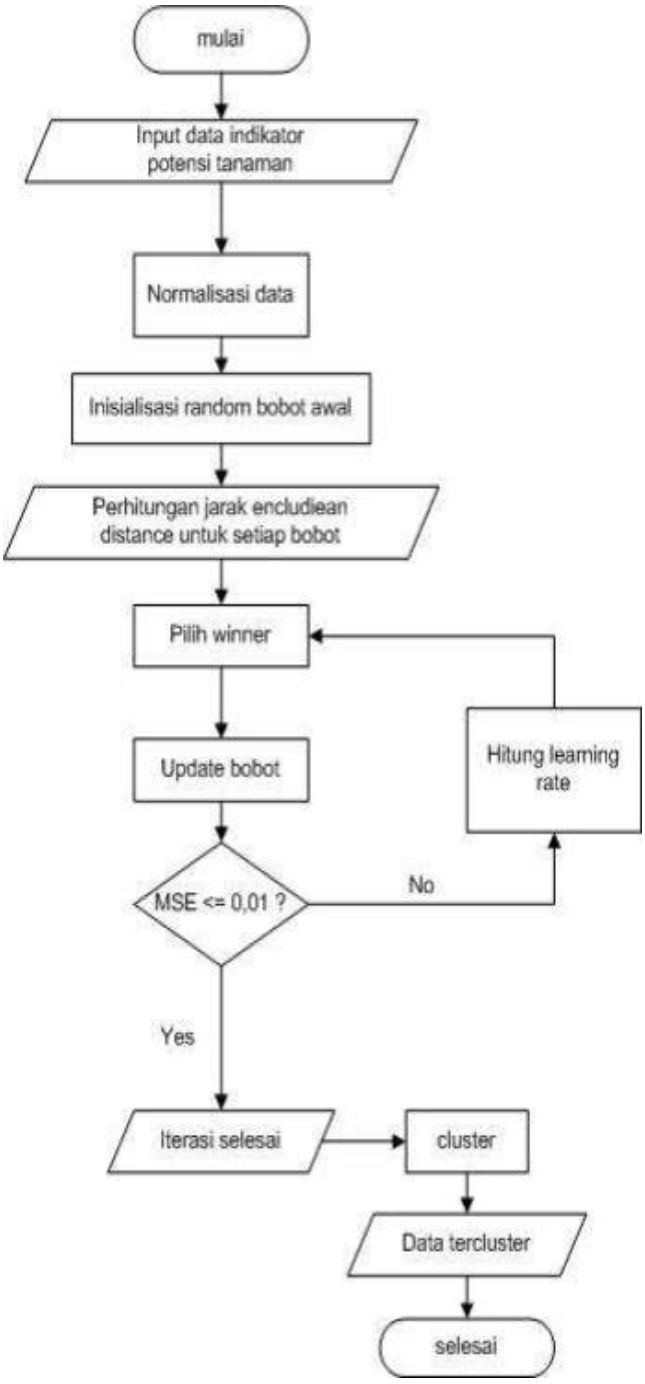

Figure 3. Flowchart calculation Som

SOM Calculation Flowchart can be explained as follows:
1 Input data used are educational data in 2014, then the clustering process is calculated using the SOM method.
2 Before carrying out the SOM calculation process, normalize the data to reshape the existing data so that it has a narrower range of values, namely

with a range of values from 0 to 1, but still accommodates the differences and values of the data.
3 Next, the calculation uses the SOM method, beginning with initial initial weights at random.
4 Establish the initial learning rate (α) for the calculation process in the first iteration
5 For each data the weight is calculated using the Euclidean Distance formula.
6 After calculating the weight, find and select the weight that has the smallest value (winner).
7 The data that has the smallest value from step 5 is used for the weight update process for the next Xn value.
8 Check the requirements for stopping iteration or checking the accuracy of data processed using MSE. If the MSE value indicates a value close to or equal to 0.01, then the iteration stops and the data that has been processed has reached the optimum level of accuracy.
9 If the MSE value is still greater or equal to the specified MSE value of 0.01, then the SOM calculation process returns to the Xn iteration calculation step.
10 Before carrying out an iteration calculation, first calculate the learning rate according to Equation 2.4.
11 Repeat step 5 through step 9. After calculating the data in the iteration, check the MSE value. If the MSE value meets the specified MSE value, the iteration stops.
12 Iteration is completed with MSE fulfillment requirements. By the time the iteration is finished, the process will be carried out by displaying a sub-district calculation table of all the data in the iteration along with the distance and cluster position values.
13 The final result of this process is in the form of data clusters by displaying the name of the sub-district and the position of the sub-district cluster.

3.2. Use Case System

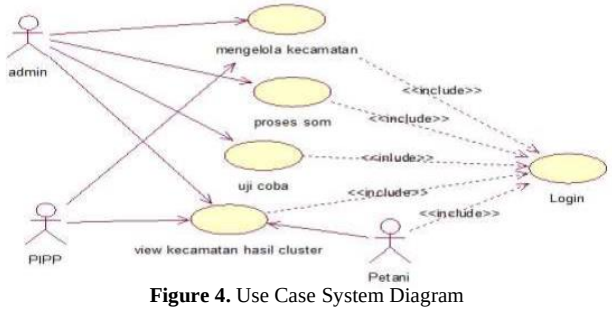

Figure 4. Use Case System Diagram

In Figure 4, use case clustering Self Organizing Map Admin is a system user who has full access rights to the system, which are as follows;
1. Admin can add, change, delete data from the value of each district. Subdistrict value data contains the value of 4 indicators of the planting potential of each crop commodity variant, namely corn, soybeans, peanuts, green peas, cassava, and sweet potatoes.
2. Admin can do the SOM cluster calculation process. The cluster calculation process is the process of processing raw data into grouping calculations using the SOM method.
3. Admin can test the system. System testing is a testing process that will produce subdistrict data that has been grouped according to the similarity of group characteristics.
4. Admin can see the results of the subdistrict cluster. While the agricultural service employee actors and user farmers can only see the results of the grouping of districts from the SOM clustering process.

3.3. Diagram aktivity Som

The diagram in Figure 5 explains each user's access flow on the menu page som. Each user has different interests, in this system only admin users can manage the full menu som.

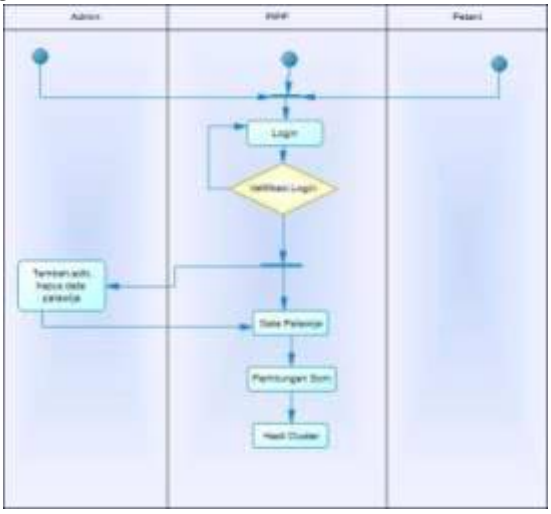

Figure 5. Activity diagram SOM

4. Results and Discussion

4.1. SOM network

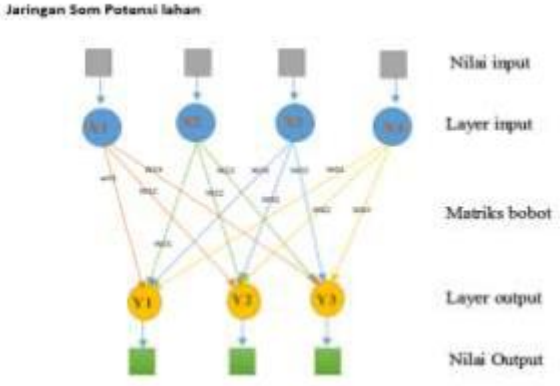

Figure 6. SOM Network

Figure 6 is a network of land potentials that uses a single layer with four inputs (x1, x2, x3, .. x4) that correspond to the number of land potential indicators needed in the clustering process. Whereas for output units there are outputs (y1, y2, ..y3) which are in accordance with the initial

assumptions of land areas in Sumenep Regency which have three categories namely mountains, plains, and beaches.

4.2. Analysis of Results

From the overall results of experiments on scenarios that have been carried out with the number of iterations of 20, 50 and 100 there are differences in the results of MSE on learning rates ranging from 0.6, 0.8 and 0.9. Table 4.4 below shows the results of the maximum iteration value and the Optimal MSE value.

5. Conclusion

After conducting this research of various theories and implementation in this application, also by considering the analysis of the results of the implementation of the Potential Land Grouping Application Using the Self Organizing Map Method in Sumenep Regency, several conclusions can be obtained, namely:

1. From the results of the trial using the Self Organizing Map (SOM) method, sub-district groups which have the same characteristics in the potential of commodity crops land are produced. In this trial, 3 clusters of sub-district groupings were produced. In the first cluster c1 obtained as many as 2 subdistricts, namely Prasional and Saronggi sub-districts. Whereas c2 cluster has 10 members, namely Talango, Kalianget, Sumenep City, Pasongsongan, Ambunten, Manding, Batuputih, Gapura, Batang-Batang, and Sapeken districts. And the last one is c3 has 15 members, namely the districts of Bluto, Giligenting, Batuan, Lenteng, Ganding, Guluk-guluk, Rubaru, Dasuk, Dungkek, Nonggunong, Gayam, Ra'as, Arjasa, Kangayan, Masalembu.

2. The results of the trial analysis on cluster C1 have similar patterns with LT and LP indicators with medium values and have low PR and PV indicator values with high values. C2 has the similarity of patterns with LT and LP indicators with low values and has medium PR values and PV values with high values. C3 has a similar pattern to the LT and LP indicators with High values and has a low PV and PR indicators with medium values. The similarity of the pattern aims to make it easier for the Sumenep Regency Agriculture Office to provide policies in managing the potential of agricultural crops in each region, to maintain high value indicators and improve low value indicators.

REFERENCES

[1] Afiffaturrohmah, M.Pengelompokkan Kecamtanan Berdasarkan Indikator Pendidikan Menggunakan Metode Self Organizing Map (SOM) di Kabupaten Lamongan.Tugas Akhir Program Studi Teknik Informatika. Universitas Trunojoyo Madura.2014.

[2] Ariana,G.B. Costumer Segmentation Dengan Metode Self Organizing Map (Studi Kasus :UD.Fenny).Lontar Komputer.2:1-7.2011.

[3] Fahlisian,R. Klasifikasi Mutu Lembaga Pendidikan Menggunakan Self Organizing Maps (SOM).Tugas Akhir Program Studi Teknik Informatika. Universitas Trunojoyo Madura.2012.

[4] Kurniaty,V. Klasterisasi Data Kependudukan Menggunakan Metode Self Organizing Maps (SOM). Tugas Akhir Program Studi Teknik Informatika.Universitas Trunojoyo Madura. 2012.

[5] Purba,G.S., Daulay,S.B., Rindang, A., dan Sigalingging, R. Sistem Informasi Komoditas Tanaman Pangan (Padi dan Palawija) di Kabupaten Karo. J.Rekayasa Pangan dan Pert.1:77-80.2012

[6] Santosa,A.D.,Priyandari,Y.,Suletr a,I,W. Perancangan Sistem Pendukung Keputusan Pemetaan Profil Kecamatan di Kabupaten Sragen Berdasarkan Indeks Pembangunan Manusia Menggunakan Algoritma Self Oraganizing Maps (SOM).Seminar Nasional IENACO.2337_4349.1-8.2013.

[7] Fatah,Luthfi. Dinamika Pembangunan Pertanian dan Pedesaan. Banjar Baru : KDT.2006.

[8] Alwi,Hasan.Kamus Besar Bahasa Indonesia Edisi Ketiga.Jakarta : Balai Pustaka.2005.

[9] Wikipedia Indonesia, Ensiklopedia Bebas. Palawija. URL: http://id.wikipedia.org/wiki/Pala wija, diakses tanggal 29 November 2014 pukul 08.46.

[10] Suparman. Bercocok Tanam Ubi Jalar.Jakarta : Azka Press. 2007.

[11] Kusrini. Dan Luthfi, Taufiq, Emha. "Algoritma Data Mining". ANDI, Yogyakarta, 2009.

[12] Siong, A.W. dan Resmana. Juli 1999 "Pengenalan Citra Objek Sederhana Dengan Menggunakan Metode Jaringan Saraf Tiruan SOM" Seminar Nasional I Kecerdasan Komputasional Universitas Indonesia

[13] Mulyadi, Y., 1) Abdullah, A.G. dan Nurjaman, R. 2009. "Estimasi Beban Puncak Harian Berbasis Algoritma Self Organizing Map (SOM)" Seminar Nasional Electrical, Informatics and Its Education, FT UNM

[14] Muhammad,A. Pengelompokan Tingkat Kesehatan Masyarakat Jawa Timur Menggunakan SOM dengan Cluster Validation IDB dan I-DUNN. Tugas Akhir Program Studi Teknik Informatika. Universitas Trunojoyo Madura. 2014.