Journal of Electrical Engineering and Computer Sciences
Vol. 5, Issue 1, June 2020
P-ISSN: 2528-0260 E-ISSN: 2579-5392
p.773-778

DISTRICT AND CITY CLASTERIZATION BASED ON ADIPURA ASSESSMENT IN THE REGION OF INDONESIA USING THE FUZZY METHODC-MEANS

[1]Syariful Alim, [2]M. Mahaputra Hidayat, [3]R Dimas Adityo, [4] Elta Deka Wilinda

Student Department of Informatics, Faculty of Engineering, Bhayangkara University, Surabaya

Jl. Ahmad Yani 114 Surabaya 60231

Tel. 031-8285602, 8291055

E-mail: alim@ubhara.ac.id, mahaputra@ubhara.ac.id

ABSTRACK

Adipura is an award for Indonesian cities that succeed in cleanliness and management of the urban environment. Adipura is organized by the State Ministry of Environment. Adipura is actually used as a tool to encourage the motivation of government officials and the community to improve and improve environmental hygiene conditions in Indonesia. The problem is to determine the recipient of Adipura takes 2 weeks to select Regencies and Cities in all regions of Indonesia, in this study there were 374 Cities and Regencies collected. Clustering is a method of grouping based on the size of closeness (similarity). Clustering is different from a group, if a group means the group is in the same condition, if not, then definitely not the group. One method that can be used in clustering is fuzzy c-means. In the concept of Fuzzy C-Means first determine the center of the cluster, which will mark the average location of each cluster which is then carried out the process of repairing the cluster center and the degree of membership of the point to the minimization of objective functions that describe the distance from the data points given to the center of the weighted cluster by the degree of membership of that point. From the results of testing a regional clustering system that deserves Adipura with the result that each cluster in each process gets a value that is always changing, so it cannot determine the labeling of a patent. Data greatly influences the time for application processing. The more data it will take more time for the fastest time is in the Big City category while the longest application execution time is in small cities with a total data of 273.

Keywords : Adipura Award, Regency, City of Indonesia, Clustering, Fuzzy CMeans.

1 INTRODUCTION

Adipura is an award for Indonesian cities that succeed in cleanliness and management of the urban environment. Adipura is organized by the State Ministry of Environment. Understanding the city in Adipura's assessment is not an autonomous city, but it can also be part of a regency area that has characteristics as an urban area with certain regional boundaries. Adipura is actually used as a tool to encourage the motivation of government officials and the community to improve and improve environmental hygiene conditions in Indonesia. Some of the goals behind the Adipura award include, to reduce pollution levels from domestic waste, realize environmental health, and realize a culture of clean environment. The Adipura assessment process begins with monitoring in Adipura participating areas. The assessment conducted by the Adipura team covers the condition of the Adipura participant areaThe Clustering Process for Adipura is a grouping of objects based on information that is on the object. The way is by grouping the characteristics of the same information into the same class and an object with different information characteristics to another class.

DOI: https://doi.org/10.54732/jeecs.v5i1.102

2 Basic Theory

The determination of Adipura winner's assessment uses attributes which are special requirements set by the Minister of the Environment. These attributes are taken from the criteria that KLH has set to simplify and speed up the awarding system of Adipura winners. Clustering data uses attributes that have been determined by the Ministry of Environment. Attributes are taken from the data of each region that occurs every year in every region in the entire territory of Indonesia, but the data can not help find out in detail, areas that are entitled to get the Adipura Award in every district in all regions of Indonesia. It is expected that the system that will be designed can provide recommendations to the Recipient in determining the City that is entitled to get the Adipura Cup. In determining Adipura Award Winners, there are several attributes that must be fulfilled by the City.

2.1 Fuzzy C-Means

Fuzzy C-Means (FCM) is one of the clustering methods which is part of the Hard K-Means method. FCM uses a fuzzy grouping model so that data can be members of all classes or formed with degrees or levels of membership that differ between 0 to 1. The level of data presence in a class or determined by the degree of membership. This technique was first introduced by Jim Bezdek in 1981. The basic concept of FCM, first is to determine the center that will mark the average location for each data. In the initial condition, this center was still inaccurate. Each data has a degree of membership for each cluster. By improving the data center and the membership value of each data repeatedly, it can be seen that the center will go to the right location. This iteration is based on the minimization of objective functions.Model system development Fuzzy C-Means has several steps that must be done and can be seen in Figure 1, design Fuzzy C-Means in this study are as follows.

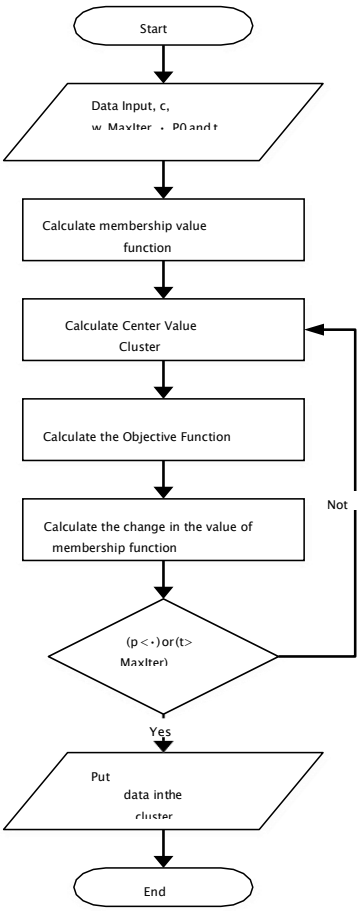

Figure 2.1 Process Flowchart Fuzzy C-Means

Algorithm Fuzzy C-Means is as follows :

a) Input data to be clustered X, in the form of nxm matrix (n = number of samples data, m = attribute of each data). • .. = i-i sample data (i = 1,2, ... n) j-attribute (j = 1,2, ... m).

b) Determine:

a. Number of clusters = c
b. Rank = w
c. Maximum iteration = MaxIter
d. The smallest expected error = •
e. Initial objective function = • 0 = 0
f. Initial iteration = t = 1

c) Generate random numbers μik,i=1,2,...,n; k=1,2,...,c as matrix elements initial partition U. Count the number of each column:

$$Qj = \sum_{k=1}^{c} \mu_{ik}$$

with j = 1,2,...,n.
Count:

$$\mu_{ik} = \frac{\mu_{ik}}{Q_i}$$

d). Calculate the center of the cluster to k: V-kj with K=1,2,..,C; and j=1,2,..,m

$$V_{kj} = \frac{\sum_{i=1}^{n}((\mu_{ik})^{w} * X_{ij})}{\sum_{i=1}^{n}(\mu_{ik})^{w}}$$

e). Calculate the objective function on the t-iteration, Pt :

$$P_1 = \sum_{i=1}^{n}\sum_{k=1}^{c}\left(\left[\sum_{j=1}^{m}(X_{ij} - V_{kj})^2\right](\mu_{ik})^{w}\right)$$

f). Calculate the matrix partition Change

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{m}(X_{ij} - V_{kj})^2\right]^{\frac{-1}{w-1}}}{\sum_{k=1}^{c}\left[\sum_{j=1}^{m}(X_{ij} - V_{kj})^2\right]^{\frac{-1}{w-1}}}$$

with : i = 1,2,…,n ; dan k = 1,2,…,c.

g) Check stop condition:

a. If: $(|P_t - P_{t-1}| < \xi)$ or (t> MaxIter) then stop;
b. If not: t = t + 1, repeat step 4.

3 Methodology

Data testing conducted on the system is to search the amount of data in the cluster, to prove whether the results of the program are the same as the manual results or even far different from the manual results, and therefore need to be tested and compared the results of the program with manual data. There are 4 clusters used in this system, namely cluster one as a group that receives Adipura Kencana Adipura Kencana, Adipura, and Plaques (Certificates).

3.1 2013-2014 Period Data Testing Results

Testing is done by giving cluster value 2, rank of weight 2, maximum iteration 100 and the smallest error value is 0.001. The criteria used are:

a) Waste Management:
- Housing
- Street, Market
- Department store
- Office space
- School
- Bus / Angkot Terminal
- Railway Station
- River / Sea Port
- Hospital
- City Forest
- City Park
- Open Water
- TPA, sorting
- Processing of Tourist Beach

b) River Water Pollution Treatment:
- River Water Quality Data
- Availability of Wastewater Treatment Facilities

c) Treatment of Air Pollution (Size of Air Pollution):
- 'Radio Side Monitoring'
- Urban Traffic Performance
- Vehicle 'Spotcheck' emission test on the highway.

The observations of each process are manually checked to determine the stability of the results of each process obtained. The test results on the program can be seen in Figure 2 which is a display program using the calculation results Fuzzy C-Means. The time required in this process is 1.5511 seconds. The following is the distribution of the first test shown in table 1

P-ISSN: 2528-0260 E-ISSN: 2579-5392
p.773-778

Nomor	Kota	Fungsi Objektif 1	Fungsi Objektif 2	Fungsi Objektif 3	Fungsi Objektif 4	MU 1	Mu 2	MU 3	MU 4	Kelompok
1	KOTA MALANG	687.185	1886.07	1102.7	1662.31	0.128728	0.353312	0.206566	0.311394	2
2	KOTA BALIKPAPAN	1429.67	1030.94	1428.61	1097.28	0.286708	0.206746	0.286496	0.22005	1
3	KOTA DENPASAR	512.354	1688.88	1088.39	929.052	0.121449	0.400333	0.257994	0.220224	2
4	KOTA PEKANBARU	1844.06	1381.81	1066.79	1253.46	0.332495	0.249149	0.192349	0.226007	1
5	KOTA MANADO	1086.36	527.063	619.895	1598.94	0.283478	0.137533	0.161757	0.417232	4
6	KOTA BATAM	1676.29	519.381	1062.8	1640.83	0.342149	0.106011	0.216929	0.334911	1
7	KOTA YOGYAKARTA	1273.23	633.424	899.616	1914.85	0.269689	0.134168	0.190552	0.405592	4
8	KOTA SAMARINDA	505.491	1526.06	1683.51	1284.6	0.101105	0.305232	0.336725	0.256938	3
9	KOTA PADANG	1119.55	741.218	1853.31	1640.32	0.209091	0.138432	0.346128	0.30635	3
10	KOTA BOGOR	888.389	918.416	1689.71	1945	0.163261	0.16878	0.310522	0.357437	4
11	KOTA BANJARMASIN	1238.75	1740.62	1466.03	526.75	0.249137	0.350074	0.294849	0.10594	2
12	KOTA SURAKARTA	1758.12	1333	1083.47	940.301	0.343726	0.260612	0.211826	0.183836	1
13	KOTA PONTIANAK	1729.31	1429.27	1576.04	1335.96	0.284867	0.235442	0.25962	0.220071	1
14	KOTA BANDAR LAMPUNG	1169.55	788.535	863.572	1752.13	0.255707	0.172403	0.188809	0.383081	4

Figure 3.1 Display Calculation Results

Table 1 Test Results of Small Town Data

Cluster	total
1	5
2	3
3	2
4	4

3.2 Result Discusion

The results from each cluster are never the same. The data obtained can be the same, but not by laying in each class. Each class never has a permanent labeling, and cluster labeling is determined after the iteration process stops, by comparing membership values. The labels used are cluster 1 as a group that gets Adipura Kencana and cluster 2 as a group that gets Adipura, Cluster 3 group that gets a Plaque (Certificate) and Cluster 4 Group that does not get the Adipura Award.The results of testing on the data in Small Cities, can be obtained information about the tendency of districts and cities to enter cluster 1, cluster 2, cluster 3 and cluster 4. The tendency of districts and cities is influenced by the value of the degree of membership in the last iteration. If the cluster data is compared with the original data, the correct results are obtained: 22, Incorrect results: 16, Accuracy: 57.9% and Error: 42.1%.

4. Conslusion

From the research conducted by the author, the conclusions obtained from this study are:

4.1 Each cluster in each process gets a value that is always changing, so it cannot determine the labeling of a patent.

4.2 Cluster labeling is largely determined by membership value / membership function on the last iteration after the error difference is met.

4.3 In 2012 data obtained an accuracy rate of 57.9% and an error rate of 42.1%. Whereas in Small Town data

5. REFERENCES

[1] Bahar (2011), Determination of High School Majors with Fuzzy C-Means Algorithm,
[2] Dian Nuswantoro University, Semarang. Sanmorino Ahmad, (2012), Clustering Batik Images using Fuzzy C-Means Algorithm Based on Log-Average Luminance. University of Indonesia, Depok.
[3] Charezita Sandhi Yudha, Suyanto, 2012. Clustering Indonesian Documents with Using Fuzzy C-Means, Telkom University, Bandung. Jaya Tri Sandika, 2012, Housing Selection System with Fuzzy C- Combination Method Means and Simple Additive Weighting, Diponegoro University, Semarang.
[4] Mirza Muhammad Faisal, 2013, Clustering Method with Fuzzy C-Means Algorithm for Recommendations for the Selection of Field Expertise in the Informatics Engineering Study Program. Diponegoro University, Semarang.
[5] .Farmadi Andi, Fatma Indriani, Nurjanah, 2014, Fuzzy C-Means Method Implementation on Rice Varientas Data Clustering System. Lambung Mangkurat University, South Kalimantan.
[6] Huda Miftakhul, 2015. Regional Clustering in East Java to Get Intensive Funds Regions Using Fuzzy C-Means, Bhayangkara University, Surabaya.
[7] Yudistira Novanto, S.Kom., M.Sc, Dinda Novitasari, W. Lisa Yunita, Nur Alfiyah, 2014, Image Segmentation on Maps with Fuzzy C-Means Method, Universitas Brawijaya (UB), Malang.
[8] Lestari Wiji, 2015 Mapping Student Learning Styles By Using Clustering Fuzzy C-Means, STMIK Duta Bangsa, Surakarta.
[9] .Astuti Femi Dwi, 2016, Implementation of Fuzzy C-Means for Population Clustering (Study Case: Bantul District), STMIK AKAKOM, Yogyakarta.
[10] Hadi Rosalia, I Ketut Gede Darma Putra, I Nyoman Satya Kumara, 2016, Determination Student Competencies with Genetic Algorithms and Fuzzy C-Means Method. Udayana University, Denpasar.
[11] Prihantyo Achmad Ika 2013, Data Mining Feasibility Take Final Projects With the Fuzzy C-Means Method, Bhayangkara University, Surabaya.