CLUSTER ANALYSIS TO CLASSIFY THE LEVEL OF SOCIAL WELFARE OF THE COMMUNITY IN DELI SERDANG REGENCY USING FUZZY C-MEAN CLUSTERING DURING THE COVID-19 PANDEMIC

Dwi Salasa Riana[1], Fibri Rakhmawati[2], Rina Widyasari[3],

Departement of Mathematics, Universitas Islam Negeri Sumatera Utara, Medan

ABSTRACT

The Covid-19 pandemic has greatly affected the social welfare of the affected communities in each region. Based on data sources that there are many people who have lost their jobs and lack of income, the government provides a number of assistances to help people affected by covid-19. This study uses data on the number of recipients of cash social assistance (BST) in Deli Serdang Regency in 2020 which aims to look at the problems of assistance received by the community towards the social welfare of the community. This study uses fuzzy c-mean clustering method because social welfare groupings can be grouped appropriately. Based on the results of fuzzy c-mean clustering analysis will produce three clusters that have different characteristics. The sub-district included in cluster 1, namely Hamparan Perak, Percut Sei Tuan, Sunggal, Tanjung Morawa are sub-districs that have a low level of welfare during the covid-19 pandemic because in these sub-districs the population is more than other sub-districs. Located in Deli Serdang regency.

Corresponding Author:

Dwi Salasa Riana,
Departement of Mathematics,
Universitas Islam Negeri Sumatera Utara, Medan, Indonesia
Email: dwisalasariana0109@gmail.com

1. INTRODUCTION

Social welfare is very important to evaluate because it is very influential on the economy and stability of government. Whether or not a development is good or not is not only seen from the physical side but also must be balanced non-physically, including the extent to which the government's efforts to improve social welfare. The government has made various efforts to overcome social welfare problems, including through the improvement of social facilities such as rehabilitation of nursing homes, provision of facilities for disabled people, and others. Information about the state of social welfare is very necessary, so an analysis is needed to see the extent to which each district has a level of social welfare. In several studies, for example: education, social, economic, and political fields, large data and many variables are often obtained. Of course, it is difficult to interpret directly, so it is necessary to do the data dimension preduction stage. In statistics, there is a data analysis that uses more than one variable and is analyzed simultaneously, namely multivariate analysis.

The world is currently being hit by a disaster that has resulted in a decline in the welfare of its people. The 2019-2020 coronavirus pandemic, also known as the covid-19 pandemic, is an event that spreads the 2019 coronavirus disease throughout the world. This disease is caused by a new type of coronavirus named SARS-

CoV-2. The covid-19 outbreak was first detected in Wuhan city, Hubei Province, China in December 2019 and was designated a pandemic by the world health organization (WHO) on march 11, 2020. Deli Serdang Regency is the second most affected Regency in North Sumatra to be exposed to the covid-19 virus with a total of 162 cases as of July 7, 2020, making most of the people in Deli Serdang Regency affected by layoffs by companies, in addition to workers in the company, traders also experience the impact of the virus, because of the reduced number of buyers resulting in losses. The government has made a lot of efforts to help residents affected by covid-19, by providing assistance in the form of money and basic necessities for 3 consecutive months from May-July 2020. The government's goal to provide social assistance (Bansos) to the community is to help improve the welfare of society.

2. RESEARCH METHOD
2.1 Cluster Analysis
2.1.1 Definition of Cluster Analysis
Cluster analysis is a technique used to generate objects or cases (respondents) into relatively homogeneous groups, called clusters. Objects/cases in each group tend to be similar to each other and very different (not the same) from objects from other clusters. Cluster analysis groups a number of n objects based on p variance which has relatively similar characteristics among these objects, so that the diversity within a group is smaller than the diversity between groups. Objects can be goods, services, plants, animals, and people (respondents, consumers, or others).

2.1.2 Cluster Analysis Procedure
Cluster analysis consist of several basic processes, namely: formulate the problem, choose a measure of distance or similarity, data standardization, choosing a clustering procedure, interpreting the clusters that have been formed, perform cluster validation and profiling.

2.1.3 Cluster Interpretation
The centroid function used in the fuzzy c-mean method is as follows:

$$v_{kj} = \frac{\sum_{k=1}^{N}(\mu_{ik})^{W} X_{ij}}{\sum_{k=1}^{N}(\mu_{ik})^{W}}$$

2.2 Fuzzy Logic
2.2.1 Definition of Fuzzy Logic
Fuzzy logic is a logic that has a value of ambiguity or ambiguity between true or false. However, how bg the precence and error of a product depends on the weight of its membership. Fuzzy logic shows the extent to which a value is true and the extent to which the value is false.

2.2.2 Fuzzy C-Mean Clustering
Fuzzy c-mean clustering (FCM) is a data grouping technique where the existence of each data in a cluster is determined from the membership value. This technique was first introduced by Jim Bezdek in 1981. The basic concept of FCM, the first time is to determine the center of the cluster which will mark the average for each cluster. The objective function used in fuzzy c-means is as follows:

$$J(U,V,X) = \sum_{i=1}^{c} \sum_{k=1}^{n} (u_{ik})^{w} \|x_k - v_i\|^2$$

2.3 Community Social Welfare
2.3.1 Definition of Social Welfare According to Experts
Gertrude Wilson (social welfare is a shared concern of everyone for everyone), Walter Friedlander (social welfare is an organized system of social institutions and services designed to assist individuals or groups to achieve better standards of living and health), Elizabeth Wickenden (social welfare, including laws, programs, benefits, and services that guarantee or strengthen services to meet the basic social needs of the people and maintain order in society).

2.3.2 The Meaning of Social Welfare According to Suharto (2006)
Social welfare as a condition of well-being, social welfare as a social service, social welfare as social benefits, social welfare as a planned process or effort carried out by individuals, social institutions, communities and government agencies to improve the quality of life and provide social services.

2.3.3   Social Welfare Goals

According to Leonard Schneiderman, the main objectives of the social welfare system are: system maintsenance, system control, system change.

2.3.4   Social Welfare Factors

Social problems can be caused by actors and environmental factors. Internal and external factors interact and are related, so social problems are usually complex and not easy to solve. Social problems have various dimensions, both economic, social, cultural, biological, psychological, spiritual, legal, and security, so that social problems can only be approached across sectors and various disciplines.

2.3.5   Community Welfare Indicators

The social welfare indicators are: (welfare indicators, 2020)

The amount and distribution of income, easy to reach education, improved and equitable health quality.

2..4   History of Covid-19

The covid-19 pandemic in Indonesia is part of the ongoing coronavirus disease 2019 (Covid-19) pandemic around the world. The disease is caused by severe acute respiratory syndrome coronavirus (SARS- Cov-2). The first positive case of Covid-19 in Indonesia was detected on March 2, 2020 when two people were confirmed to have contracted it from a Japanese citizen. On April 9, the pandemic had spread to 34 provinces with East Java, DKI Jakarta, and South Sulawesi as the most exposed provinces.

3.   RESULT AND ANALYSIS

In this study, the authors took data from 22 sub-districts in Deli Serdang Regency. The data will be grouped using the fuzzy c-mean clustering (FCM) method by determining the number of clusters as many as 3 clusters. As for the cluster 1st cluster, it is categorized with a less prosperous level of welfare, the 2nd cluster is categorized with a fairly prosperous, and in the 3rd cluster it is categorized with a prosperous level of welfare.

Appendix 1
Data on recipients of covid-19 assistance for Deli Serdang Regency in 2020

NO	NAMA KECAMATAN	PENERIMA BANTUAN COVID-19				JUMLAH
		X_1	X_2	X_3	X_4	
1	BANGUN PURBA	3080	997	662	448	5187
2	BATANG KUIS	4488	1464	1412	1337	8701
3	BERINGIN	4030	1315	973	887	7205
4	BIRU-BIRU	3312	1073	805	677	5867
5	DELI TUA	2620	850	628	629	4727
6	GALANG	5116	1671	1369	1062	9218
7	GUNUNG MERIAH	631	205	142	102	1080
8	HAMPARAN PERAK	16493	5361	3564	1674	27092
9	KUTALIMBARU	4367	1416	1014	430	7227
10	LABUHAN DELI	5170	1675	1105	487	8437
11	LUBUK PAKAM	4668	1555	1076	1281	8580
12	NAMORAMBE	3032	983	680	445	5140
13	PAGAR MERBAU	5011	1617	1205	807	8640

14	PANCUR BATU	6811	2203	1712	1319	12045
15	PANTAI LABU	5837	1903	1019	666	9425
16	PATUMBAK	5500	1791	1418	1377	10086
17	PERCUT SEI TUAN	15004	4909	3417	2248	25578
18	SIBOLANGIT	2651	859	616	387	4513
19	STM HILIR	4253	1347	916	329	6845
20	STM HULU	1695	549	539	520	3303
21	SUNGGAL	7806	2526	2113	2077	14522
22	TANJUNG MORAWA	11447	3729	2511	2372	20059

Data on the number of cash food aid recipients (BST) in appendix 1 is mapped or clustered using the fuzzy c-mean clustering algorithm as follows:
1. Set the initial partition matrix.
2. Determine the initial parameter value.
3. Generate random numbers.
   To determine the value of the initial parttion matrix by using a random value with a total value equal to 1.

$$
\begin{array}{ccc}
1 & 2 & 3
\end{array}
$$
$$
\begin{bmatrix}
0{,}3 & 0{,}43 & 0{,}27 \\
0{,}25 & 0{,}40 & 0{,}35 \\
\vdots & \vdots & \vdots \\
0{,}42 & 0{,}23 & 0{,}35
\end{bmatrix}
$$

4. Determine the center of the cluster (V)

$$
v_{kj} = \frac{\Sigma_{i=1}^{n}(\mu_{ik})^{w}x_{ij}}{\Sigma_{i=1}^{n}(\mu_{ik})^{w}}
$$

Table 4.1 the results of the calculation of the cluster center in the first iteration of 1^{a} cluster

| μ_{ik} | Cluster membership degree 1 | Data to be clustered | | | | $(\mu_{ik})^2$ | $(\mu_{ik})^2$ $* x_1$ | $(\mu_{ik})^2$ $* x_2$ | $(\mu_{ik})^2$ $* x_3$ | $(\mu_{ik})^2$ $* x_4$ |
		X_1	X_2	X_3	X_4					
μ_{11}	0,3	3080	997	662	448	0,09	277,2	89,73	59,58	40,32
μ_{21}	0,25	4488	1464	1412	1337	0,063	280,5	91,5	88,25	83,563
[1] Σ						2,33	14772,143	4804,868	3446,993	2586,581
$v_{kj} = \frac{\Sigma_{i=1}^{22}(\mu_{ik})^{2}*x_{ij}}{\Sigma_{i=1}^{22}(\mu_{ik})^{2}}$							6339,976	2062,175	1479,396	1110,121

$$
v_{11} = \frac{\Sigma_{i=1}^{22}(\mu_{11})\, 2 *_{X_{11}}}{\Sigma_{i=1}^{22}(\mu_{11})^{2}} = \frac{14.772{,}143}{2{,}33} = 6.339{,}976
$$

$$v_{12} = \frac{\sum_{i=1}^{22}(\mu_{11})2*X_{12}}{\sum_{i=1}^{22}(\mu_{11})^2} = \frac{4.804,868}{2,33} = 2.062,17$$

$$v_{13} = \frac{\sum_{i=1}^{22}(\mu_{11})2*X_{13}}{\sum_{i=1}^{22}(\mu_{11})^2} = \frac{3.446,993}{2,33} = 1.479,396$$

$$v_{14} = \frac{\sum_{i=1}^{22}(\mu_{11})2*X_{14}}{\sum_{i=1}^{22}(\mu_{11})^2} = \frac{2.586,581}{2,33} = 1.110,121$$

Do the same thing until the 3rd cluster in the first iteration.

The cluster V library formed in the first iteration is:

$$V = \begin{bmatrix} 6.339,978 & 2.062,175 & 1.479,396 & 1.110,120 \\ 5.265,059 & 1.710,661 & 1.242,344 & 915,717 \\ 5.253,676 & 1.707,932 & 1.233,299 & 913,963 \end{bmatrix}$$

5. Calculate objective function (Pt) The objective function in the first iteration P1 is calculated using the equation

$$P_1 = \sum_{k=1}^{3}\sum_{i=1}^{22}\left(\left[\sum_{j=1}^{4}(X_{ij} - V_{kj})^2\right]\mu_{ik}{}^2\right)$$

No	Square of degree of data membership to= i			L_1	L_2	L_3	L_T $=L_1+L_2+L_3$
	$(\mu_{i1})^2$	$(\mu_{i2})^2$	$(\mu_{i3})^2$				
1	0,090	0,185	0,073	1158172,119	1079696,079	420909,551	2658777,749
22	0,176	0,053	0,123	5559519,331	2434476,004	5659546,03	13653541,36
Objective function = Σ							137.140.017,4

So to find the value of $P_1 = L_{T1} + L_{T2} + \cdots + L_{Tn}$

$$P_1 = 2.658.777,749 + 484.516,046 + \cdots + 13.653.541,3 = 137.140.017,4$$

6. Calculating partition matrix Change (U): the partition matrix change (U) is calculated using the equation:

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

To find the value:

$$\mu_{11} = \frac{[\sum_{j=1}^{m}(x_{ij} - v_{kj})^2]^{\frac{-1}{w-1}}}{\sum_{k=1}^{c}[\sum_{j=1}^{m}(x_{ij} - v_{kj})^2]^{\frac{-1}{w-1}}}$$

$$\mu_{11} = \frac{L_1}{L_T} = \frac{0,0000000777}{0,0000004222} = 0,184 \quad \text{, etc}$$

7. Checking the stop condition:

If t > Maxiter then stop. If not t = t+1, repeat step 4.

8. Cluster interpretation

After doing 8 steps, you will get a function to the members of each cluster as below:

Table 4.7 Function Value to Membership

subdistrict	Membership function value		
	1	2	3
P_1	0,019	0,459	0,523
P_2	0,007	0,523	0,470
P_3	0,002	0,340	0,604
P_4	0,013	0,453	0,534
P_5	0,028	0,462	0,510
P_6	0,014	0,546	0,441
P_7	0,075	0,452	0,473
P_8	0,632	0,185	0,182
P_9	0,002	0,464	0,535
P_{10}	0,013	0,542	0,446
P_{11}	0,005	0,548	0,447
P_{12}	0,019	0,459	0,521
P_{13}	0,007	0,568	0,425
P_{14}	0,236	0,396	0,368
P_{15}	0,050	0,504	0,447
P_{16}	0,039	0,513	0,448
P_{17}	0,690	0,156	0,154
P_{18}	0,028	0,462	0,510
P_{19}	0,003	0,449	0,548
P_{20}	0,0450	0,459	0,491

P_{21}	0,606	0,202	0,192
P_{22}	0,915	0,043	0,042

Based on table 4.7 above, the results of grouping based on the value of the membership function are obtained using the fuzzy c-mean clustering method:

1. Member in cluster 1

   Hamparan Perak, Percut Sei Tuan, Sunggal, Tanjung Morawa.

2. Member in cluster 2

   Batang Kuis, Galang, Labuhan Deli, Lubuk Pakam, Pagar Merbau, Pancur Batu, Pantai Labu, Patumbak.

3. Member in cluster 3

   Bangun Purba, Beringin, Sibiru-Biru, Deli Tua, Gunung Meriah, Kutalimbaru, Namorambe, Sibolangit, Sinembah Tanjung Muda Hilir, Sinembah Tanjung Muda Hulu.

4.   CONCLUSION

The problem with the number of recipitiens of assistance provided by the government during the covid-19 period was quite effective and people who received one the four aids provided by the government, namely provincial assistance, district assistance, ministry of social assistance (kemensos), could push the level of assistance slightly. Their welfare during the covid-19 pandemic, although the number of people who received the four aids was less than those who did not, the government provided other aids with more recipitients to people affected by covid-19.

5. REFERENCES

[1] Ahmad Wahbi Fadillah, Andi Tejawati, Novianti Puspitasari. 2018. "Penerapan Fuzzy C-Means pada Curah Hujan di Kalimantan Timur". *Jurnal Teknologi Informasi*. 2(1): 82-89.

[2] Amalyah, F. 2012. "Pengelompokkan Kabupaten/Kota Di Jawa Timur Berdasarkan Indikator Indonesia Sehat 2010. ITS Surabaya: Surabaya.

[3] Aprilia, R., & Sari, R. F. (2017). Implementation of Pagerank Algorithm in Matlab. *Zero: Jurnal Sains, Matematika dan Terapan, 1*(1), 46-53.

[4] Awaliyah, Resky. 2018. "Analisis Clustering Untuk Mengelompokkan Tingkat Kesejahteran Kabupaten/Kota Berdasarkan Sosial Ekonomi Rumah Tangga di Wilayah Provinsi Sulawesi Selatan". Fakultas Sains Dan Teknologi. UIN Alauddin: Makasar.

[5] Fadillah, Ubai. 2011. "Analisis Penyandang Masalah Keseahteraan Sosial Di Indonesia Menggunakan Metode Fuzzy C-Mean Clustering Dan Biplot". Fakultas Sains Dan Teknologi. UIN Syarif Hidayatullah: Jakarta.

[6] Firdausi, Tsaniyah N. 2010. "Proyeksi Tingkat Kemiskinan Di Indonesia (Studi kasus: 30 Provinsi). Semarang.

[7] Irandha, Irma. 2010. Analisa Keluarga Miskin Dengan Menggunakan Metode *Fuzzy C-Means Clustering*. Surabaya.

[8] Pertiwi, F. N., Lubis, R. S., & Widyasari, R. (2021). Population Projection And Factor Analysis Affecting Population Growth In The City Medan Using Non Linear Trends Polynomic Method. *Journal of Mathematics and Scientific Computing With Applications, 2*(1), 40-45.

[9] Rosni. 2017. "Analisis Tingkat Kesejahteraan Masyarakat Nelayan Di Desa Dahari Selebar Kecamatan Talawi Kabupaten Batubara". *Jurnal Geografi*. 9(1): 53-66.

[10] Safa'at Yulianto, Kishera Hilya Hidayatullah. 2014. "Analisis Klaster Untuk Pengelompokkan Kabupaten/Kota Di Provinsi Jawa Tengah Berdasarkan Indikator Kesejahteraan Rakyat". *Jurnal Statistika*. 2(1): 56-63.

[11] Sari, Devani Ariestha. 2016. "Analisis Faktor-faktor Yang Mempengaruhi Kesejahteraan Masyarakat Di Kota Bandar Lampung". Fakultas Ekonomi Dan Bisnis. Universitas Lampung: Bandar Lampung.

[12] Siregar, I. P. S., Widyasari, R., & Prasetya, N. H. (2021). A Weighted Markov Chain Application to Predict Consumer Price Index When Facing Pandemic Covid-19. *Quadratic: Journal of Innovation and Technology in Mathematics and Mathematics Education, 1*(1), 41-46.

[13] Sirojuddin, Achmad. 2016. "Analisis Cluster Pada Kabupaten/Kota Di Provinsi Jawa Timur Berdasarkan Indikator Indeks Pembangunan Manusia". Fakultas Sains Dan Teknologi. UIN Maulana Malik Ibrahim: Malang.

[14] Sri Wahyuni, Yogo Aryo Jatmiko. 2018. "Pengelompokkan Kabupaten/Kota Di Pulau Jawa Berdasarkan Faktor-faktor Kemiskinan Dengan Pendekatan Average Linkage Hierarchical Clustering". *Jurnal Aplikasi Statistika & Komputasi Statistik*. 10(1): 1-8.

[15] Widyasari, R., Cipta, H., & Hasugian, A. H. (2020). Developing Fuzzy-Promethee Method by using AHP Method on Student Achieving Selection. *IJISTECH (International Journal of Information System & Technology)*, *3*(2), 309-314.

[16] Yunita Sari Siregar, Putri Harliana. 2018. "Analisis Perancangan Algoritma Fuzzy C-Means dalam Menetukaan Dosen Pembimbing Tugas Akhir". *Jurnal Teknik* Informatika. 3(1): 181-185.