Strategy for Improving and Empowering MSMEs Through Grouping Using the AHC Method

Lisna Zahrotun[1]*, Yosyadi Rizkika Amanatullah[2], Utaminingsih Linarti[3], Anna Hendry Soleliza Jones[4]
Informatics, Faculty of Industrial Technology[1], [2], [4]
Industrial Engineering, Faculty of Industrial Technology [3]
Yogyakarta, Indonesia
lisna.zahrotun@tif.uad.ac.id[1]*, yosyadi111@gmail.com[2], utaminingsih.linarti@ie.uad.ac.id [3], annahendri@tif.uad.ac.id[4]

***Abstract*— The high number of immigrants in the city of Yogyakarta has resulted in increased opportunities for Micro, Small and Medium Enterprises (MSMEs) in the Culinary and Crafts sector. The large number of MSMEs creates increasingly higher competitiveness. Apart from that, the large amount of data collected by the Department of Cooperatives and MSMEs, which reaches thousands, causes the Department to have difficulties in efforts to improve and empower these MSMEs. Grouping is one method that can be used as a strategy in mapping MSMEs, especially in efforts to improve and empower MSMEs through training conducted by the Department. The aim of this research is to group MSMEs using the Agglomerative Hierarchical Clustering (AHC) method in an effort to achieve strategies for improving and empowering MSMEs. The focus of this research is MSMEs in the craft sector and MSMEs in the culinary sector. The results of this research provide 2 group recommendations from a total of 1336 Culinary MSME data and 3 group recommendations from a total of 145 Craft MSME data. The silhouette score test results in the Culinary Sector are in the strong structure category with a value of 0.79 and in the Crafts Sector are in the Medium Structure category with a value of 0.615. From the number of groups in the two MSMEs, strategies were obtained to improve and empower MSMEs, especially those with a turnover of less than 10 million, marketing objectives within the Yogyakarta area, and not having capital assistance from the government.**

Keywords— Crafts, Culinary, Gglomerative Hierarchical Clustering; strategy

I. REDEMPTION

Micro, Small and Medium Enterprises (MSMEs) play an important role in Indonesia's growth. MSMEs are one of the buffers of the country's economy in facing the crisis [1]. In 1998 many large businesses had to close due to the ongoing economic crisis in Indonesia. However, the MSME sector was able to survive in facing the crisis. To survive and the number continues to increase, most MSMEs produce consumer goods and services by measuring the sensitivity of consumers to low income. In addition, most MSMEs do not get capital from banks. In addition, MSMEs also have limited capital and a competitive market [2]. Culinary is a processed product in the form of dishes such as side dishes, food, and drinks. Each region has its own taste, so it's no wonder each region has different culinary traditions [3]. Likewise, the field of handicrafts in addition to playing an important role in introducing and maintaining the sustainability of an area, crafts also have important potential in developing community attitudes and entrepreneurship [4]. The grouping of MSMEs has been carried out in an effort to carry out strategies to increase sales [5], MSME promotion strategy [6][7],and improvement of advertising services [8].

Yogyakarta is one of the tourist sites in Indonesia with many visitors of 3,898,951 people in July 2022 based on data from the Yogyakarta City Tourism Office. The large number of immigrants to Yogyakarta opens business opportunities for the people of Yogyakarta, especially in the field of Culinary and Handicrafts. The Yogyakarta Government seeks to conduct training and empowerment of existing MSMEs as stated in PERDA No.9 of 2017. Based on data from the Yogyakarta Cooperative and MSME Office, there are 2,769 MSME data in 2021-2022, including MSMEs in the Fashion sector totaling 195 MSMEs, 399 MSMEs in the service sector, 145 MSMEs in the craft sector, 1336 MSMEs in the culinary sector, and 694 MSMEs in other fields. From the MSME data that has been collected at this time, the Yogyakarta Cooperative and MSME Office has not managed MSME data in the culinary and handicraft sector easily and quickly. The large amount of data collected by the Cooperatives Service, which reaches thousands, causes the department to have difficulty in improving and empowering MSMEs, especially in the culinary and craft sectors. It is a challenge for the Department of Cooperatives and MSMEs to be able to determine strategies for improvement and empowerment efforts using the data that has been obtained. So the aim of this research is to carry out strategies in efforts to improve and empower MSMEs, especially crafts and culinary arts, using clustering techniques. The results of this grouping are important because they are used as recommendations in determining which MSME priorities need to be further improved. Apart from that, the grouping results are important because they can be used as a basis for providing recommendations for implementation, design and evaluation in developing and empowering MSME data in Yogyakarta, one example of which is training for MSMEs.

Clustering is one solution that can be used in grouping data. One of the methods used in clustering is *Agglomerative Hierarchical Clustering (AHC). AHC* be Group analysis methods that seek to establish a group hierarchy [9]. Dendogram itself is an image of the distribution of grouping

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v13i1.2021, Copyright ©2024
Submitted : November 21, 2023, Revised : December 18, 2023, Accepted : December 27, 2023, 2024, Published : February 15, 2024

results in hierarchical form. This AHC has the advantage of being able to describe the proximity between data with a dendrogram. In addition, it can also produce *Cluster* quality with high accuracy [10].

Some previous studies using the AHC method include: in grouping poverty rates in East Kalimantan. The results of this study are able to provide an overview of the distribution model of poverty data [11]. The application of AHC in segmenting barbershop customers, where in its research uses 2 distance parameters, namely *Single Linkage* and *average linkage* [12]. AHC is also carried out in detecting communities on Facebook social media [13], modeling flooded areas in East Java with an accuracy of 0.92 [14], and grouped pre-college academic history data and student graduation data [15][16]. By looking at the ability of the AHC method in grouping data in previous studies, in this study the AHC method was used in grouping MSME data.

The purpose of this research is to carry out a strategy to improve and empower MSMEs, especially in the crafts and culinary fields, through grouping using AHC. Where the results of the grouping will show the priorities of MSMEs that will receive empowerment from the Cooperative Department, one of which is through organizing training.

II. RESEARCH METHODOLOGY

A. Agglomerative Hierarchical Clustering Single Linkage

Hierarchical grouping (*Hierarchical clustering*) is a method of group analysis that seeks to establish a group hierarchy [17]. *Agglomerative Hierarchical Clustering* is a method of grouping hierarchies with a bottom-up approach (*bottom up*). Here are the steps in a hierarchical grouping [18]:

a. Calculate the proximity matrix based on the type of distance used,
b. Repeat steps 3 through 4, until only 1 group remains,
c. Combine two nearby groups based on the specified proximity parameters.
d. Update the proximity matrix to reflect the closeness between the new group and the original merged group.

To calculate the distance between *clusters* , *Euclidean Distance* is used in Equation 1.

$$||U - V||_2 = \sqrt{\sum_1^n (Ui - Vi)^2} \quad (1)$$

Euclidean distance is a metric or mathematical concept used to measure the distance or length between two points in Euclidean space.

Information:
Ui = *value* U data *training*.
Vi = *V value* in testing data.

To determine the group used *Single Linkage* with Equation 2.

$$d_{UV} = \{d_{UV}\}, d_{UV} \in D \quad (2)$$

Information:
dUV = distance between the nearest neighbors of the group (you and V).
D = proximity matrix.

Single linkage is the process of forming a hierarchy in single linkage clustering begins with each point being considered as a single cluster. Then, the points or clusters that have the shortest distance between them are combined into one new cluster. This process continues, and the closest clusters will continue to be combined until all points or data are combined into one large group.

B. Silhouette Coefficient

Silhouette coefficient is a validation that combines two elements, namely: *Cohesion* and separation. Value of *Silhouette Coefficient* It has a range from -1 to 1 [19]. Value 1 is categorized in *Cluster* either or *Cluster* What is formed is the result of *Cluster* which is good, while a value of -1 indicates that the result *Cluster* The less well formed, the closer to 0 indicates the fewer documents the correct grouping [20].

.Stages of Research

The stages in conducting this research are shown in the flowchart in Figure 1.

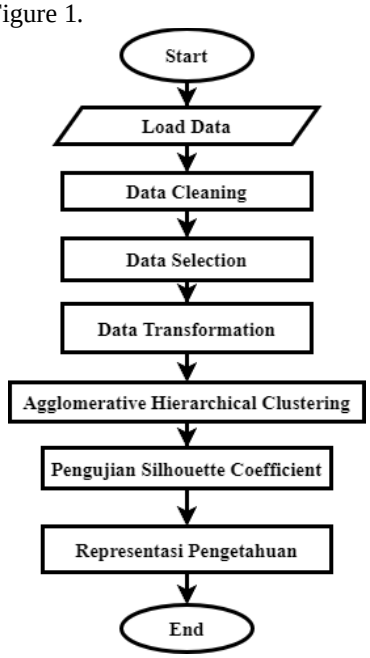

Fig 1. Stages of Research

1. Load Data
   At this stage the *dataset* is in the form of *an excel file* that will be used and loaded into the program for processing.
2. Data Cleaning
   At this stage, data cleaning is carried out from *missing value data* or *noise*. One way is to delete empty rows. For example, deleting data that does not yet have a Turnover-Annually.
3. Data Selection
   At this stage, data selection is carried out from many data based on the needs for subsequent processing. For example, do not select variables that are not related to the data to be used. Like not using variable Name and NIK.
4. Data Transformation
   At this stage, *string data* is converted to 0 and 1 forms with *One-Hot encoding* so that the data can be used or

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v13i1.2021, Copyright ©2024
Submitted : November 21, 2023, Revised : December 18, 2023, Accepted : December 27, 2023, 2024, Published : February 15, 2024

clustered. The value 0 does not represent there and 1 represents there. For example, the Turnover-Year column has a value in the form of a string, which is less than 10 million and more than 10 million. Then a new column will be formed, namely columns less than 10 and columns more than 10 million. If there is a value of less than 10 million, it will be given a value of 1 in the column less than 10 million, and the column more than 10 million will be 0.

5. Agglomerative Hierarchical Clustering Single Linkage *Method*
   This stage is useful for grouping existing objects into *clusters* based on the closest distance or similarity between objects. *The* clusters formed will be merged again based on the closest distance or similarity between the *cluster or* cluster with objects. Then the process will be repeated until one *cluster* remains.
6. Silhouette Coefficient Testing
   To find out whether the *resulting cluster* is good or then testing will be done using *silhouette coefficient. Clusters with* a silhouette coefficient *value of 1 are* categorized *as clusters* that are formed are good cluster results *while a value of -1 indicates that the results of* clusters formed are not good. Getting closer to 0 indicates the less data the correct grouping is.
7. Knowledge Representation
   This stage is carried out analysis of cluster *results* and *silhouette coefficient* testing by analyzing the results of each *cluster* is expected to provide knowledge to be utilized.

III. RESULTS AND DISCUSSION

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

A. Data Collection Results

II. The data obtained in the Culinary sector amounted to 1339 MSMEs and the Craft sector 145 MSMEs for the 2021-2022 period. *The dataset* is in the form of excel with the names "Culinary Field.xlsx" and "Field of Cultivation.xlsx" with 40 variables, namely No, Ref. OSS, NIK, Legkap Name, Date of Birth, Age, Gender, Education, Phone No, e-Mail, Province, Regency, District, Village, Street Name, Business Name, NIB, NIB Issue Date, Business Establishment Date, Province, Regency, District, Village, Street Name, Coordinates, Field, Sector, Activity, Product, Destination, Land/Building Ownership Status, Electronic Media Facilities, Government Assistance Capital, People's Business Credit Loans, Year-Round Turnover, Health Insurance Ownership, Men, Women.

B. Research Steps

This study has several stages of data mining including the following:

a. Load Data

This stage is carried out to contain excel data, namely data on Culinary MSMEs and Handicraft MSMEs. The list of Culinary MSME data is shown in Table 1. Table 1 displays data that includes 40 variables with a total of 1339 MSMEs in the culinary field.

Table 1. MSME Data

No.	Age	Gender	Recent Education	Phone Number	...	Average Age of Workers	Form Status
1 2							
3	41	P	SMA	089620770xxx	...	-	Verified
4	53	P	SMA	085718338xxx	...	35-50 years	Verified
5	45	P	0	081226224xxx	...	35-50 years	Verified
6	40	L	SMA	089630999xxx	...	25-35 years	Verified
7	-	-	JUNIOR	085743611xxx	...	-	Verified
8	27	L	-	081779533xxx	...	35-50 years	Verified
9	20	P	SMK	0856292 7xxx	...	-	Verified
.....	...	..	.....		...	.....	.....
1339	64	P	SD	0899513 8xxx	...	-	Verified

b. Data *Cleaning*

This stage is done to clean the data in the *Dataset* from data that is noise or cannot be used in the process. The results of the data cleaning process are shown in Table 2. From the results of the cleaning process on Culinary data with 1339 data, there was a reduction in data to 1336 data, and in Craft data with 148 data, there was a reduction to 145. This happens because there is a missing value or empty data in the initial data obtained.

Table 2. The results of the MSME data cleaning process

No.	Age	Gender	Recent Education	Phone Number	...	Average Age of Workers	Form Status
3	41	P	SMA	089620770xxx	...	-	Verified
4	53	P	SMA	085718338xxx	...	35-50 years	Verified
5	45	P	0	081226224xxx	...	35-50 years	Verified
6	40	L	SMA	089630999xxx	...	25-35 years	Verified
7	-	-	JUNIOR	085743611xxx	...	-	Verified
8	27	L	-	081779533xxx	...	35-50 years	Verified
9	20	P	SMK	085629	...	-	Verifie

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v13i1.2021, Copyright ©2024
Submitted : November 21, 2023, Revised : December 18, 2023, Accepted : December 27, 2023, 2024, Published : February 15, 2024

				27xxx	. . . . .	d
.....	...	..	.....		.....	.....
1339	64	P	SD	089951 38xxx	-	Verified

c. Data *Selection*

This stage is done to select the variable to be used from all variables in the
Datasets. From the selection process of the 40 variables contained in Point 3.1, 12 variables were selected. The results of data selection are shown in Table 3. The results of the variable selection are Last Education, Business Establishment Date, Business Activities, Marketing Objectives, Land/Building Ownership Status, Electronic Media Facilities, Government Assistance Capital, People's Business Credit Loans, Yearly Turnover, Health Insurance Ownership, Male Labor, Female Labor.

Table 3. Results of the *MSME data* selection process

education	Date of business establishment	Business activities	....	Male Workers	Female workforce
SMA	July 24, 2012	Sales, Production	....	0	0
SMA	February 16, 2016	Sales	....	0	2
0	January 07, 2022	Sales, Production	....	1	1
SMA	June 22, 2020	Sales	....	1	0
JUNIOR	February 04, 2005	Sales	....	0	0
-	February 20, 2019	Sales, Production	....	3	3
SMK	June 28, 2003	Sales, Production	....	0	0

a. Data Transformation

Some of the variables that must be transformed are seen in Table 4. In this table consists of three columns, namely column no, variable and variable result. Variables describe the type of variable to be transformed, while the variable results are the results of the elaboration of the variability to carry out the one-hot encoding process.

Table 4. Transformed variables

No	Variable	Variable Results and One-Hot Encoding
1	Recent Education	-, 0, D1, D2, D3, D4, S1, S2, S3, ELEMENTARY, HIGH SCHOOL, VOCATIONAL SCHOOL, JUNIOR HIGH SCHOOL
2	Business Activities	Sales, Production
3	Objective of Marketing	Within DIY area, Within Yogyakarta City area, Within Java island area, Within Java island area (scattered), Overseas (export), Overseas (usual type of shipment)
4	Land/Building Ownership	Other, Magersari (customary), Owned, Rent
	Status	
5	Electronic Media Facilities	-, Facebook, Gojek, Grab, Instagram, Shopee, Tokopedia, Twitter, WhatsApp, Others
6	Government Assistance Capital	-, DIY Local Government, Central Government, Yogyakarta City Government
7	People's Business Credit Loan	-, Bank, Cooperative, Other, Government
8	Turnover per Year	Less than 10 million, 10 million to 25 million, 40 million to 55 million, 55 million to 70 million, 70 million to 85 million, 85 million to 100 million, 100 million to 120 million, 120 million to 150 million, More than 150 million
9	Health Insurance Ownership	-, Private Insurance, BPJS

The results of MSME data transformation are shown in Table 5. Table 5 displays MSME data after data transformation with references contained in Table 4. For example, the last Education variable will change to variables D1, D2, D3, D4, S1, S2, S3, SD, SMA, SMK, SMP, as well as for other vaiabel

Table 5. Data transformation data results

No.	**D1**	**D2**	**D3**	**D4**	.....	**Labour_P**	**Bussines_age**
3	0	0	0	0	.....	0	10
4	1	0	0	0	.....	2	6
5	0	0	0	0	.....	1	0
6	0	1	0	0	.....	0	2
7	0	0	0	0	.....	0	17
8	0	0	0	1	.....	3	3
9	0	0	0	0	.....	0	19
.....	0	0	0	0	.....	0	1
1339	0	0	0	0	.....	0	32

d. Grouping using AHC

From *the transformation dataset in table 5, a grouping process will be carried out using AHC with the following stages:*

1) Calculate the distance between data using *Euclidean distance*,
In this process will produce the proximity values between the data represented in Table 6. In Table 6, it can be seen that the distance between MSME 1 and MSME 1 is 0, meaning this is the same data, while for MSME 1 and MSME 2 it has a distance of 5.38 and for ajrak between other MSMEs it can be seen in Table 6.

Table 6. Results of calculating the distance between data using *Euclidean distance*

	MSME s 1	**MSME s 2**	**MSME s 3**	**MSME s 4**	**MSMEs 5**	**MSME s 6**	**MSME s 7**
MSME s 1	0	5,385	10,677	8,774	7,810	8,660	9,695
MSME s 2	5,385	0	7,141	5,477	11,747	5,477	13,527
MSME s 3	10,677	7,141	0	3,605	17,406	5,744	19,442

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v13i1.2021, Copyright ©2024
Submitted : November 21, 2023, Revised : December 18, 2023, Accepted : December 27, 2023, 2024, Published : February 15, 2024

MSMEs 4	8,774	5,477	3,605	0	15,362	5,656	17,521
MSMEs 5	7,810	11,747	17,406	15,362	0	15,165	4,582
MSMEs 6	8,660	5,477	5,744	5,656	15,165	0	16,881
MSMEs 7	9,695	13,527	19,442	17,521	4,582	16,881	0

2) Grouping with AHC single *linkage algorithm* (closest distance).
Based on the data in Table 6, a grouping process was carried out using the single linked AHC method. From table 6, it can be seen that the closest distance is UMK 3 and MSME 4, therefore MSME 3 and UMK 4 join in 1 cluster. The first iteration of the cluster is shown in Table 7.

Table 7. Produce group matrices (MSMEs 3 and MSMEs 4)

	MSMEs 3, MSMEs 4	**MSMEs 1**	**MSMEs 2**	**MSMEs 5**	**MSMEs 6**	**MSMEs 7**
MSMEs 3, MSMEs 4	0	8,774	3,605	15,362	5,656	17,521
MSMEs 1	8,774	0	5,385	7,810	8,660	9,695
MSMEs 2	3,605	5,385	0	11,747	5,477	13,527
MSMEs 5	15,362	7,810	11,747	0	15,165	4,582
MSMEs 6	5,656	8,660	5,477	15,165	0	16,881
MSMEs 7	17,521	9,695	13,527	4,582	16,881	0

The results of Table 7 will be carried out the process of finding the closest distance between data and then combining them into one group. For the final results of grouping by specifying the number of clusters 2 shown in Table 8. Table 8 shows that after the process of combining data that has the closest distance, the last 2 groups consisting of MSMEs 1, 2, 6, 3 and 4 are members of one group and MSMEs 5 and 7 are members of 1 group.

Table 8. Results of grouping 2 clusters

	MSMEs 1, MSMEs 2, MSMEs 6, MSMEs 3, MSMEs 4	**MSMEs 5, MSMEs 7**
MSMEs 1, MSMEs 2, MSMEs 6, MSMEs 3, MSMEs 4	0	7,81024968
MSMEs 5,MSMEs 7	7,810249676	0

C. Implementation

This implementation of *aggromerative hierarchical clustering single linkage* produces a website-based application in grouping MSMEs data in Yogyakarta City.

1) Preprocessing Page
This page is used to display the results of datasets that have been cleaned of missing value and noise data. The cleaning view can be shown in Figure 2. This figure shows MSME data that has been carried out in the cleaning process.

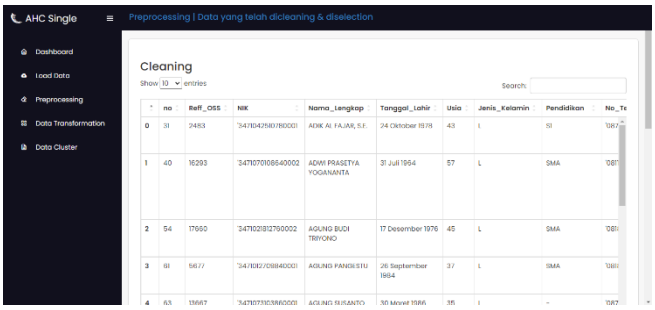

Fig 2. Page preprocessing cleaning section

The display of the results of the selection process is shown in Figure 3. This picture displays the data from the selection process, from 40 variables of the initial data carried out by the selection process into 12 variables that will be used in this grouping process.

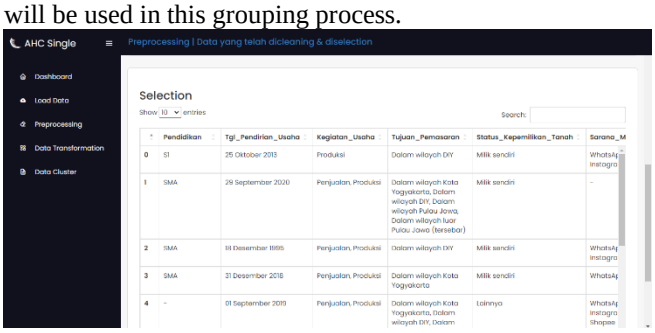

Fig 3. Preprocessing page selection section

2) Transformation Page
This transformation is used to display the transformed data according to Table 4. The transformed data page is shown in Figure 4. This display shows variables that have changed according to the one Hot Encoding process.

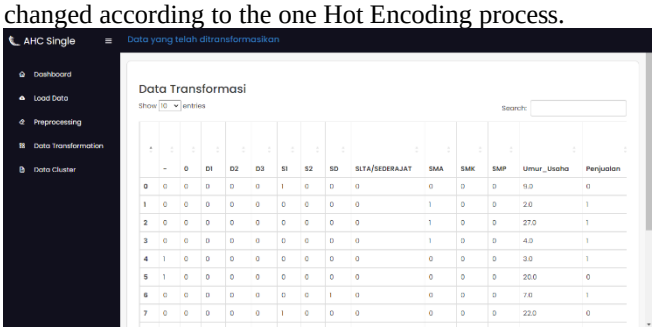

Fig 4. Transformationon page

3) Grouping Results page using AHC method
This page displays a dendogram of the results of grouping using the AHC method using Equation 1 and Equation 2. In this grouping process using euclidean distance and single linked dendrogram display 2 clusters shown in Figure 5. In this picture, you can see 2 colors, namely blue and red which show the distribution of cluster 1 and cluster 2.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v13i1.2021, Copyright ©2024
Submitted : November 21, 2023, Revised : December 18, 2023, Accepted : December 27, 2023, 2024, Published : February 15, 2024

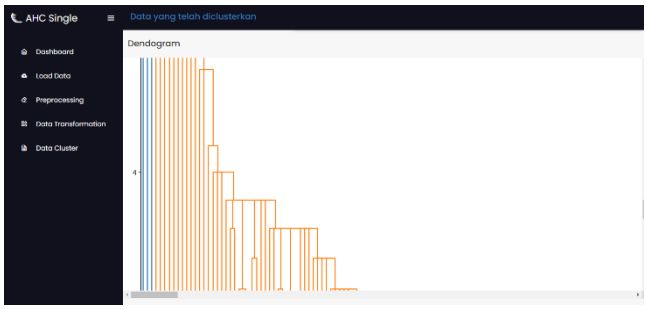

Fig 5. Dendrogram section grouping page

Meanwhile, the data for each cluster produced is shown in Figure 6. In Figure 6. This can be seen that there is an addition to the column on the far right, namely the cluster, the content of this column is the identification of the data entered into how many clusters.

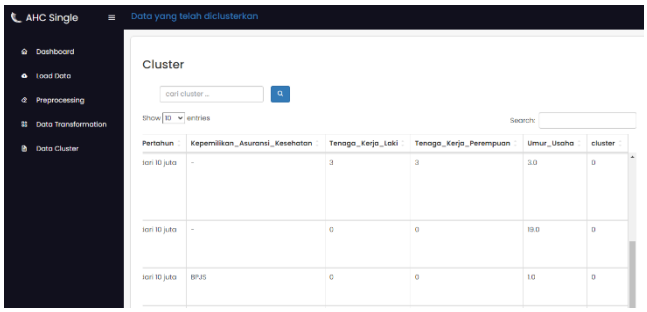

Fig 6. Grouping page clustering section

4) Testing of clustering results

Data that has been processed using the AHC single linkage method, obtained system test results as in Table 9. Table 9 shows the value of accuracy and execution time of MSME data in the culinary and handicraft fields using the AHC single linkage method

Table 9. System Test Results

Number of Clusters	Culinary Data		Craft Field Data	
	Accuracy with silhouette Coefficient	Time (seconds)	Accuracy with silhouette Coefficient	Time (seconds)
2	0.793	57,33	0.610	13,57
3	0.710	73,10	0.615	16,49
4	0.622	104,02	0.350	16,30
5	0.443	58,41	0.310	15,04
6	0.383	55,89	0.314	16,50
7	0.342	64,89	-0.004	15,24
8	0.337	55,29	0.237	16,10
9	0.336	57,61	0.238	14,52
10	0.297	56,92	0.119	16,74

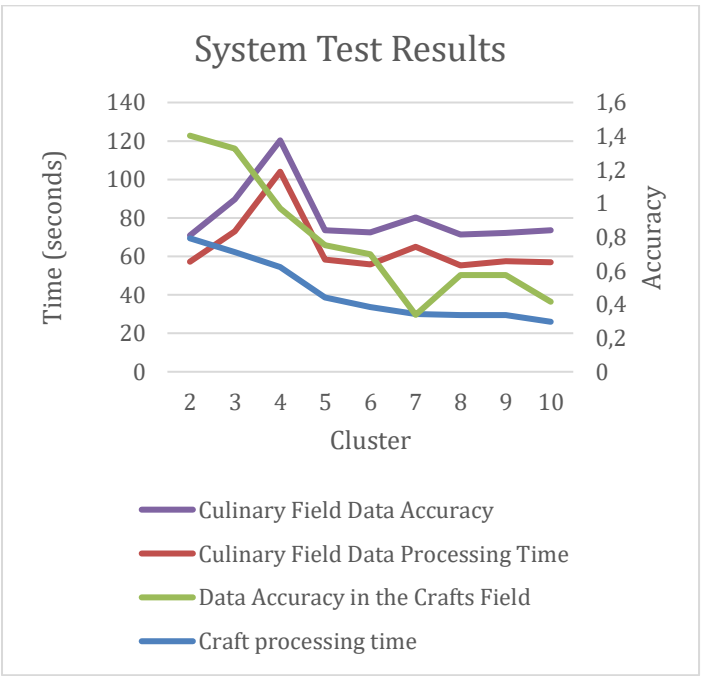

Fig 7. System Test Results

From Figure 7, it can be seen that the test results from experiments 2 to 10 clusters can be seen that the data of MSMEs in the culinary sector is the more the number of clusters, the accuracy value produced is smaller, as well as MSMEs in the field of seedling crafts many clusters, the accuracy is also getting smaller, but in cluster 8 the accuracy increases from cluster 7. As for the execution time during the data clustering process of culinary MSMEs and craft MSMEs have something in common, namely a horizontal pattern is seen, but the execution time of MSME data in the craft sector is longer than that of culinary MSMEs

C. Performance of the AHC Method of this study

In this study, the AHC method produced a good performation or entered in *strong structure*, this can be proven through testing using *Silhouette Score* obtained a value of 0.79 with the amount of data used 1336 culinary MSME data, but for craft MSME data with a total of 145 data got an accuracy value of 0.615 which is included in the category *medium structure* [20].

D. Performance of the AHC Method of this study

From the number of suitable clusters, namely using the method used, knowledge representation can be carried out, namely:

1. Culinary MSMEs produce 2 clusters, namely cluster 0 and cluster 1. Where cluster 0 shows a more dominant number of members with a business age of 6.8 years, educators of high school and vocational business owners, domestic and foreign marketing objectives have 1-4 employees and also have BPJS insurance. As for Cluster 1, the number is smaller with marketing purposes only covering the Yogyakarta area, not yet having a work force and BPJS insurance. Even so, these two clusters still have a turnover below 10 million.
2. Craft MSMEs produce 3 clusters, namely cluster 0, cluster

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v13i1.2021, Copyright ©2024
Submitted : November 21, 2023, Revised : December 18, 2023, Accepted : December 27, 2023, 2024, Published : February 15, 2024

1 and cluster 2. Where clusters 0 and 1 show the education of high school and S1 business owners, domestic and foreign marketing goals, have 2-24 employees and already have BPJS. However, the two clusters have differences in turnover, namely for cluster 0 the turnover is still below 10 million while for cluster 1 the turnover has reached 100-120 million. For Cluster 2 education, business owners are still in elementary school, marketing objectives include the Yogyakarta area, do not yet have employees and BPJS insurance.

IV. CONCLUSION

Based on the results of the research on the Application of the Single Linkage AHC Method in Grouping MSME Data in the Culinary and Handicraft Sector in 2021 – 2022 In Yogyakarta City, it can be concluded that the research resulted in an application of grouping MSME data in the Culinary and Handicraft sector using the AHC Single Linkage method. There are 1336 Culinary MSME data and 145 Craft MSME data with silhouette coefficient testing obtained the best accuracy in each MSME, namely 0.793 and 0.615. The number of clusters in culinary MSMEs is 2 clusters, while MSMEs in the craft sector are 3 clusters. From this number of clusters, it can be used as a strategy for selecting MSMEs that are worthy of getting priority for empowerment and improvement through programs from the department through training organized by the Department of Cooperatives. These MSMEs are especially MSMEs with a turnover of less than 10 million, marketing purposes within the Yogyakarta area, and do not have capital assistance from the government.

ACKNOWLEDGMENTS

The author would like to thank the Institute for Research and Community Service of Universitas Ahmad Dahlan for providing grants for this research through the 2023/2024 Applied Research scheme.

REFERENCES

[1] D. Solihin, A. Ahyani, K. Karolina, L. Pricilla, and ..., "Pelatihan Pemasaran Online Berbasis Digital Untuk Meningkatkan Penjualan Bisnis Online Pada Umkm Di Desa Cicalengka …," *Dedik. …*, vol. 2, no. 3, pp. 307–311, 2021.

[2] R. Wilda, A. Arifin, and A. Hasan, "Penyusunan Laporan Keuangan Berdasarkan Standar Akuntansi Keuangan Entitas Mikro Kecil dan Menengah (Studi Kasus pada UMKM Gabba Kitchen di Kabupaten Barru)," *J. Mirai Manag.*, vol. 7, no. 2, pp. 181–184, 2022, [Online]. Available: https://doi.org/10.37531/mirai.v7i2.2075

[3] Y. Tresnawati and K. Prasetyo, "Pemanfaatan Digital Marketing Bagi Usaha Mikro Kecil dan Menengah Bisnis Kuliner," *J. New Media Commun.*, vol. 1, no. 1, pp. 43–57, 2022, doi: 10.55985/jnmc.v1i1.5.

[4] N. L. W. S. Telagawathi, N. M. Suci, and K. K. Heryanda, "Implikasi Kewirausahaan Terhadap Digitalisasi Ekonomi dan Ekonomi Kemanusiaan UMKM Kerajinan Tenun di Provinsi Bali," *J. Ilm. Manaj.*, vol. 11, no. 2, pp. 228–240, 2021.

[5] D. Rawat, R. K. Mittal, and V. S. Aggarwal, "Cluster Development Approach in India: an Antidote for Micro, Small, and Medium Enterprises," *Indian J. Commer. Manag. Stud.*, vol. VIII, no. 2, pp. 19–29, 2017, doi: 10.18843/ijcms/v8i2/03.

[6] D. Astuti, A. R. Iskandar, and A. Febrianti, "Penentuan Strategi Promosi Usaha Mikro Kecil Dan Menengah (UMKM) Menggunakan Metode CRISP-DM dengan Algoritma K-Means Clustering," *J. Informatics, Inf. Syst. Softw. Eng. Appl.*, vol. 1, no. 2, pp. 60–72, 2019, doi: 10.20895/inista.v1i2.71.

[7] L. Zahrotun, S. H. Nugroho, U. Linarti, and A. H. S. Jones, "Analisis Persebaran UMKM Bidang Jasa Menggunakan Metode AHC Complete Linkage," *Kesatria J. Penerapan Sist. Inf. (Komputer dan Manajemen)*, vol. 4, no. 2, pp. 255–265, 2023.

[8] M. T. Satya, N. K. Rahayu, and A. Fidayan, "Pengklasifikasian Karakteristik Iklan UMKM Menggunakan Analisis Cluster," *Jesya (Jurnal Ekon. Ekon. Syariah)*, vol. 3, no. 2, pp. 218–231, 2020, doi: 10.36778/jesya.v3i2.169.

[9] I. D. Id, *MACHINE LEARNING: Teori, Studi Kasus dan Implementasi Menggunakan Python*, 1st ed. Riau: UR Press, 2021.

[10] P.-N. Tan, M. Steinbach, A. Karpatne, and V. Kumar, *Introduction to Data Mining*, Global Edi. United Kingdom: Person Education Limited, 2019.

[11] E. Suherman, "Agglomerative Hierarchical Clustering Dengan Berbagai Pengukuran Jarak Dalam Mengklaster Daerah Berdasarkan Tingkat Kemiskinan," vol. 5, no. 1, pp. 978–979, 2019.

[12] R. P. Justitia, N. Hidayat, and E. Santoso, "Implementasi Metode Agglomerative Hierarchical Clustering Pada Segmentasi Pelanggan Barbershop (Studi Kasus : RichDjoe Barbershop Malang)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 3, pp. 1048–1054, 2021.

[13] I. W. Rahayu *et al.*, "Hierarchical Clustering Untuk Deteksi Komunitas Pada Media Sosial Facebook Analysis and Implementation of Agglomerative Hierarchical Clustering Algorithm for Community Detection in Social Media Facebook," *e-Proceeding Eng.*, vol. 5, no. 1, pp. 1460–1468, 2018.

[14] R. O. Pratikto and N. Damastuti, "Klasterisasi Menggunakan Agglomerative Hierarchical Clustering Untuk Memodelkan Wilayah Banjir," *JOINTECS (Journal Inf. Technol. Comput. Sci.*, vol. 6, no. 1, p. 13, 2021, doi: 10.31328/jointecs.v6i1.1473.

[15] B. Harli Trimulya Suandi As and L. Zahrotun, "Penerapan Data Mining Dalam Mengelompokkan Data Riwayat Akademik Sebelum Kuliah dan data Kelulusan Mahasiswa Menggunakan Metode Agglomerative Hierarchical Clustering(Implementation Of Data Mining In Grouping Academic History Data Before Students And Stud," *J. Teknol. Informasi, Komput. dan Apl.*, vol. 3, no. 1, pp. 62–71, 2021, [Online]. Available: http://jtika.if.unram.ac.id/index.php/JTIKA/

[16] L. Zahrotun, U. Linarti, B. H. T. Suandi As, H. Kurnia, and L. Y. Sabila, "Comparison of K-Medoids Method and Analytical Hierarchy Clustering on Students' Data Grouping," *Int. J. Informatics Vis.*, vol. 7, no. 2, pp. 446–454, 2023, doi: 10.30630/joiv.7.2.1204.

[17] K. E. T. Luthfi, *Algoritma Data Mining*, 1st ed. Yogyakarta: C.V Andi Offset, 2019.

[18] E. Prasetyo, *DATA MINING-Konsep dan Aplikasi Menggunakan MATLAB*, 1st ed. Yogyakarta: Andi Offset, 2012.

[19] M. Liang, *Data Mining: Concepts, Models, Methods, and Algorithms*, vol. 36, no. 5. 2004. doi: 10.1080/07408170490426107.

[20] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data An Introduction to Cluster Analysis*. Canada: John Wiley & Sons, Inc, Hoboken, New Jersey.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v13i1.2021, Copyright ©2024
Submitted : November 21, 2023, Revised : December 18, 2023, Accepted : December 27, 2023, 2024, Published : February 15, 2024