INTERNATIONAL JOURNAL OF APPLIED ARTIFICIAL INTELLIGENCE AND ROBOTICS
E-ISSN: 3124-1212, DOI: 00.000/XXXX.YYYY.II.NNN
Vol. 2 No. 1 2026 pp. 40-48

A Hybrid K-Means and Agglomerative Hierarchical Clustering Method for Zakat Mal Recipient Classification

Hafidh Asyi[1], Hendrawaty[2], Mulyadi[3*]

[1,2,3] Department of Information Technology and Computers, Lhokseumawe State Polytechnic
Jl. Banda Aceh-Medan Km. 280 Buketrata 24301

Coressponding Author : mulyadi@pnl.ac.id

Article info: Received 18-02-2026, Revised 02-03-2024, Accepted 18-03-2026

This is an open access article under the CC BY-SA license.

Abstract

Grouping and determining zakat mal recipients is an important stage in the accurate and precise distribution of zakat mal. The process of determining zakat mal recipients needs to be taken into account considering that the number of zakat mal recipients in Aceh Besar exceeds the available quota in each zakat mal distribution period in 2022. This research aims to using KnA method in the process of determining zakat mal recipients. This research uses criteria data that has been determined by Baitul Mal Aceh Besar for 17 groups of zakat mal recipients. The KnA method is used to rank zakat recipients mal in the process of determining zakat mal recipients based on criteria for each group. Testing the results of the KnA method was carried out by assessing using original data on five zakat mal recipients for each group. The results of this research show that there are five groups with five zakat mal recipients in the same order with a percentage of 29%, three groups with three zakat mal recipients in the same order with a percentage of 17%, four groups with two zakat mal recipients in the same order with a percentage of 23%, four groups with one zakat mal recipient in the appropriate order with a percentage of 23%, and two groups with zero zakat mal recipients in the appropriate order with a percentage of 11%. Based on the research results, it can be concluded that the use of the KnA method in determining zakat mal recipients has a relatively low level of accuracy with only five of the seventeen groups in the appropriate order for five sample data with an accuracy percentage of 29%.

Keyword: zakat, mal, clustering, recipient, aceh

1. Introduction

Zakat originates from the Arabic word *zaka*, which means good, blessed, growing, pure, and increasing. In Islamic jurisprudence, zakat refers to a specific portion of wealth that must be given, as prescribed by Allah SWT, to those who are entitled to receive it (*mustahiq*) from those obligated to pay it (*muzakki*) [1]. As explained by Sheikh Dr. Yusuf Al-Qardhawi in his book *Fiqh az-Zakah*, *zakat al-māl* encompasses savings in gold and silver, commercial assets, livestock, agricultural produce, processed goods derived from plants and animals, mining and marine products, rental income, professional service earnings, and profits from shares. This is also in accordance with Law No. 23 of 2011 [2]. The distribution of *zakat al-māl* is guided by the Qur'an, specifically in Surah At-Taubah verse 60, which classifies the recipients of zakat into eight categories: the poor (*fuqara'*), the needy (*masakin*), zakat administrators (*amil*), new converts to Islam (*mu'allaf*), those seeking to free captives (*riqab*), debtors (*gharim*), those striving in the cause of Allah (*fi sabilillah*), and travelers in need (*ibnu sabil*).

Pursuant to Aceh Qanun No. 10 of 2018, Baitul Mal Aceh has been established as a non-structural regional institution vested with the authority to manage and develop zakat, waqf, and other religious assets for the welfare of the community, and to serve as guardian-overseer for orphans or custodians of inherited property who lack a guardian under Islamic law. Baitul Mal is organized into four tiers: provincial, regency/city (kabupaten/kota), kemukiman (subdistrict cluster), and gampong (village) [3].

The difficulties encountered by Baitul Mal include the process of aggregating and classifying data and

determining recipients of zakat al-māl based on information provided by gampong officials. This process is considered ineffective and may adversely affect the distribution outcomes for intended beneficiaries. Zakat administrators (amil), who are responsible for managing zakat, frequently experience difficulty in assessing eligible recipients. Therefore, an innovation is required in the form of a computerized system for clustering and identifying zakat al-māl beneficiaries to replace the current system.

The clustering and recipient-determination system is designed to identify prospective zakat al-māl recipients using the KnA method (Combination of K-Means and Agglomerative Hierarchical Clustering), with the expectation that the resulting clusters will be more accurate and appropriate for beneficiary selection. This combination of two clustering techniques leverages the complementary strengths of bottom-up clustering (agglomerative) and top-down clustering (K-Means) [4]. The implemented KnA method also integrates two clustering paradigms simultaneously: hierarchical clustering (Agglomerative Hierarchical Clustering) and partitional clustering (K-Means) [5].

The clustering and recipient-determination system uses the complete dataset of all zakat recipients within a gampong (village) as input. These data are processed using the KnA method (Combination of K-Means and Agglomerative Hierarchical Clustering) to generate a prioritized ranking of zakat recipients. This ranking is intended to serve as a solution for allocating zakat in accordance with the beneficiaries' level of need. The system's results are accessible to all residents of Aceh Besar Regency, thereby making zakat distribution more transparent and reliable. The system is expected to assist the Baitul Mal institution in the processes of clustering and selecting zakat recipients.

2. Research Methodology

2.1. Data Collection

Data were collected to obtain the information required to achieve the objectives of this study. Data collection was carried out through direct interviews with relevant personnel at Baitul Mal Aceh Besar and a research visit to the Badan Baitul Mal Aceh Besar in Kota Jantho. The interviews addressed the current systems for the collection and distribution of zakat al-māl as well as the procedures for determining recipients. The research visit was conducted to gather data on the eligibility criteria for zakat al-māl recipients in each category and the regulations established by the Badan Baitul Mal Aceh Besar in 2022.

2.2. Data Requirements Analysis

The data requirements analysis for this study comprises an examination of the recipient-criteria data for each category of *zakat al-māl* as stipulated by Baitul Mal in 2022, together with the *zakat al-māl* beneficiary data for 2022. There are six recipient categories, which are further divided into several subcategories totaling 17, each with distinct criteria and conditions.

2.3. Functional Requirements Analysis

The system's functional requirements involve four external entities:

1. Admin functional requirements: The admin shall be able to log in and log out of the admin account, manage Baitul Mal staff data, and manage gampong officials' data.
2. Baitul Mal staff functional requirements: Baitul Mal staff shall be able to log in and log out of staff accounts, determine recipients of zakat al-māl and print the determination results by beneficiary category, and manage their biodata/profile.
3. Gampong officials functional requirements: Gampong officials (village officials) shall be able to log in and log out of gampong accounts, manage their gampong's zakat al-māl recipient data, and manage their biodata/profile.
4. Community (public) functional requirements: The community shall be able to view the complete dataset of zakat al-māl recipients.

2.4. Combination Method of K-Means and Agglomerative Hierarchical Clustering (KnA)

The KnA method is a combination of the Agglomerative Hierarchical Clustering and K-Means Clustering methods, designed to leverage the advantages of both bottom-up clustering (agglomerative) and top-down clustering (K-Means). The Agglomerative Hierarchical Clustering method is employed to determine the initial cluster centers for the K-Means algorithm, as the clustering results of K-Means are highly dependent on the selection of the initial centroids [4].

The steps for performing the clustering process using the KnA method are as follows [4]:

1. Begin with N clusters, where each cluster represents a single data object, and construct a symmetric N × N

matrix representing the distances between clusters.

2. Calculate the distance between data points using the Euclidean distance formula:

$$d(x,y) = ||x - y|| = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad (1)$$

Explanation:

x_i = object *x* at the *i* observation

y_i = object *y* at the *i* observation

n = total number of objects

3. Identify the pair of clusters that have the smallest distance value in the distance matrix. Suppose the most similar clusters are *U* and *V*, denoted as dUV.
4. Merge clusters U and V, and label the new cluster as (UV). Then, update the entries in the distance matrix as follows:
   (1). Remove the rows and columns corresponding to clusters U and V.
   (2). Add a new row and column representing the distances between the new cluster (UV) and the remaining clusters.
5. Repeat Steps 2, 3, and 4 a total of N – 1 times. When the algorithm terminates, all objects will belong to a single cluster. Record the identity of the merged clusters at each hierarchical level using the single linkage method, defined as:

$$d(UV)W = min\,(dUW, dVW) \quad (2)$$

Explanation:

dUW = distance between clusters *U* and W

dVW = distance between clusters *V* and W

6. Define the number of clusters (k) to be used for the K-Means method.
7. Select the initial centroids for K-Means clustering based on the results obtained from the Agglomerative Hierarchical Clustering process.
8. Compute the distance between each object and each cluster centroid using the Euclidean distance formula.
9. Assign each object to the cluster whose centroid is nearest. This allocation can be performed using the hard K-Means approach, which explicitly assigns each object to the cluster with the minimum distance to its centroid.
10. Determine the new centroid position using the following formula:

$$C = \frac{1}{n}\sum_{i=1}^{n} x_i + x_{i+1} + x_{i+2} + x_n \quad (3)$$

Explanation:

C = new centroid

x_i = object *x* at the *i* observation

n = number of objects

11. Repeat Steps 8, 9, and 10 if the cluster memberships change after recalculating the distances using the updated centroid positions.

2.5. Flowchart Design

A flowchart is a diagram that represents an algorithm or a sequence of step-by-step instructions. Figure 1 illustrates the design of the KnA method flowchart used in the clustering and determination system for zakat al-māl recipients.

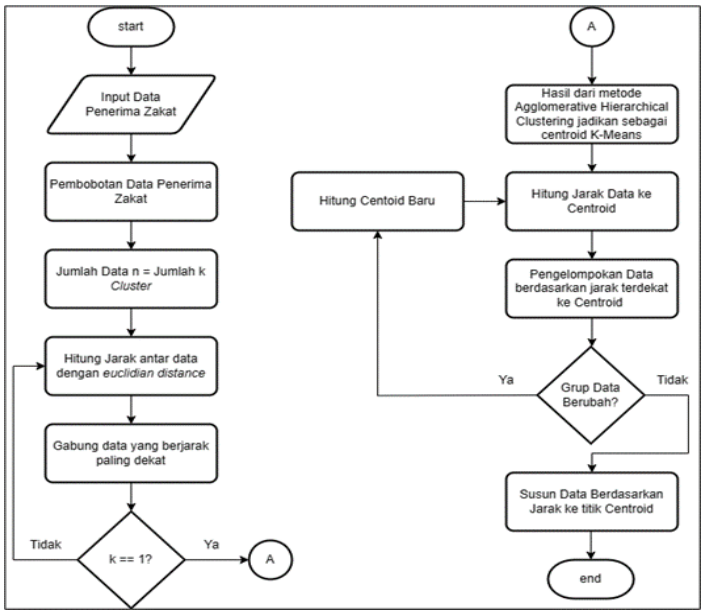

Figure 1. KnA method flowchart

2.6. Entity Relationship Diagram (ERD) Design

A database relationship is established based on a real-world perspective consisting of several objects used to fulfill the data requirements of an organization. Figure 2 illustrates the Entity Relationship Diagram (ERD) design of the clustering and *zakat al-māl* recipient determination system.

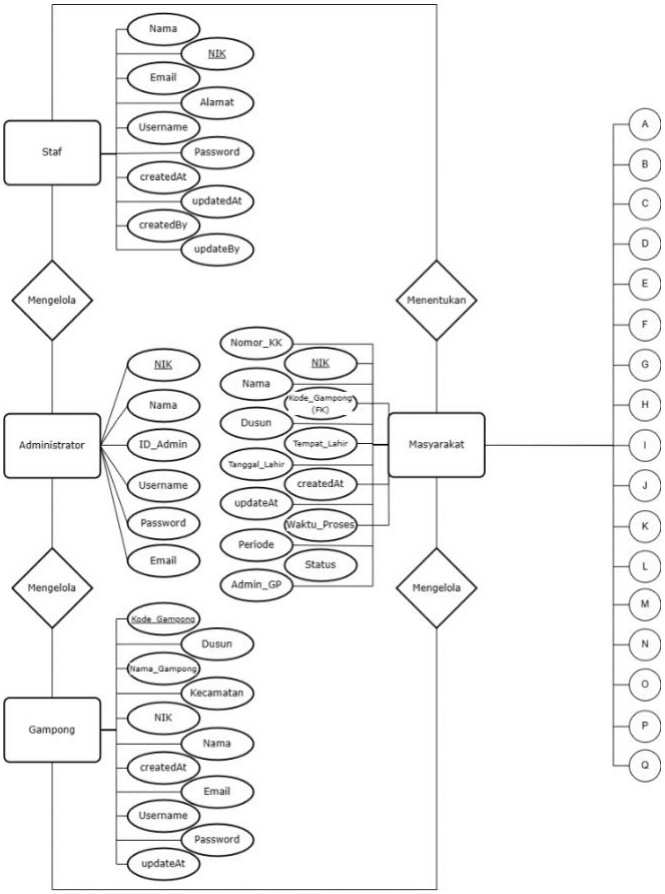

Figure 2. Entity Relationship Diagram System

There are four core tables that are interrelated and seventeen additional tables representing subcategories of zakat al-māl recipients. These supplementary tables store the criteria and detailed information for each category, which are linked to the community table through connecting symbols in the form of circles along with their respective initializations.

2.7. Context Diagram Design

A context diagram represents the highest level of a Data Flow Diagram (DFD) that provides an overview of all inputs and outputs within the system. Figure 3 illustrates the design of the context diagram for the zakat al-māl recipient clustering and determination system.

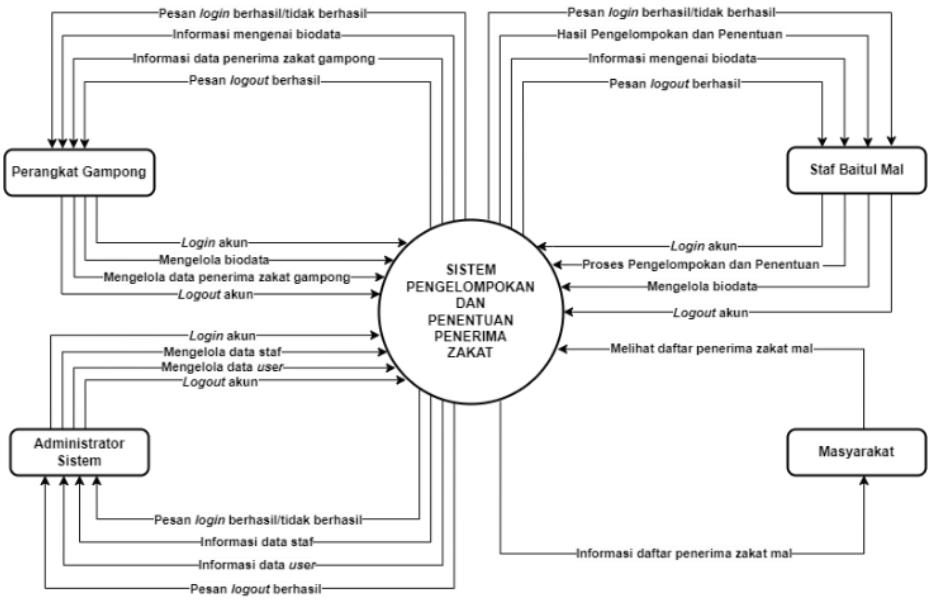

Figure 3. Context Diagram System

3. Result and Discussions

3.1. Implementation of the KnA Method

The implementation of the KnA method in the clustering and determination system for *zakat al-māl* recipients is intended to generate a prioritized list of beneficiaries. This ranking process is carried out by ordering the list of *zakat al-māl* recipients based on their proximity to the centroid point produced by the KnA method, from the closest to the farthest distance.

3.1.1. Manual Calculation of KnA Method

The manual calculation of the KnA method in the clustering and determination system for *zakat al-māl* recipients is conducted as a means of verifying the accuracy of the method implemented in the application. In the KnA approach, the test data are first processed using the Agglomerative Hierarchical Clustering (AHC) algorithm, followed by processing with the K-Means method.

Manual calculations were performed using Microsoft Excel based on the fakir (poor) classification criteria. The following are the steps for implementing the KnA method in the clustering and determination system for zakat mal recipients:

1. The data weighting has been applied in accordance with the predetermined criteria. Table 1 presents the test data used in the manual calculation of the KnA method.

Tabel 1. KnA Method Test Data

No	Nama	K1	K2	K3	K4	K5
1	Agus Salim A.	1	1	3	1	1
2	T. Sulaiman	1	1	3	1	2
3	Hendrik Nasir M.	2	4	2	1	1
4	Munzir Abdullah	1	1	2	1	2
5	Abdullah Umar	2	3	2	1	2
6	Hermawan Hamzah	2	3	1	1	1
7	Tgk. Nyak Umar	2	5	1	1	2
8	FA. Mutia	2	2	3	1	2
9	Nyak Pakeh	2	3	2	1	1
10	Ahmad Marzuki	1	1	2	1	1

Description:
K1: Occupation
Weight: 1 ⇒ Unemployed
Weight: 2 ⇒ Farmer / Fisherman / Laborer
K2: Income
Weight: 1 ⇒ None
Weight: 2 ⇒ Below 100,000
Weight: 3 ⇒ 100,000 – 250,000
Weight: 4 ⇒ 251,000 – 500,000
Weight: 5 ⇒ Above 500,000
K3: Number of Dependents
Weight: 1 ⇒ Person with Disabilities & Widow
Weight: 2 ⇒ Person with Disabilities / Widow
Weight: 3 ⇒ None
K4: Possession of ID Card / Domicile Certificate
Weight: 1 ⇒ Available
Weight: 2 ⇒ Not Available
K5: Possession of Certificate of Poverty
Weight: 1 ⇒ Available
Weight: 2 ⇒ Not Available

2. The initial number of clusters (k) is determined to be equal to the total number of data points (n).
3. The cluster distance is calculated using the Euclidean distance technique as expressed in Equation (1):

Tabel 2. AHC 1st Iteration Cluster Distance

No	1	2	3	4	5	6	7	8	9	10
1	0,00	1,00	3,32	1,41	2,65	3,00	4,69	1,73	2,45	1,00
2	1,00	0,00	3,46	1,00	2,45	3,16	4,58	1,41	2,65	1,41
3	3,32	3,46	0,00	3,32	1,41	1,41	1,73	2,45	1,00	3,16
4	1,41	1,00	3,32	0,00	2,24	2,65	4,24	1,73	2,45	1,00
5	2,65	2,45	1,41	2,24	0,00	1,41	2,24	1,41	1,00	2,45
6	3,00	3,16	1,41	2,65	1,41	0,00	2,24	2,45	1,00	2,45
7	4,69	4,58	1,73	4,24	2,24	2,24	0,00	3,61	2,45	4,36
8	1,73	1,41	2,45	1,73	1,41	2,45	3,61	0,00	1,73	2,00
9	2,45	2,65	1,00	2,45	1,00	1,00	2,45	1,73	0,00	2,24
10	1,00	1,41	3,16	1,00	2,45	2,45	4,36	2,00	2,24	0,00

4. The clusters with the smallest distance are merged into a single cluster, and the weight of the merged cluster is calculated using the single linkage method based on Equation (2).

Tabel 3. Results of the 1st Iteration of the AHC Method

Cluster	K1	K2	K3	K4	K5
1.2	1	1	3	1	1

Based on Table 3, the first iteration merges Cluster 1 and Cluster 2 into a single cluster, as the distance between Cluster 1 and Cluster 2 represents the smallest inter-cluster distance.

5. Repeat Steps 3 and 4 a total of $k - 1$ times. All objects will be contained within a single cluster once the algorithm has completed.

Tabel 4. Results of the 9th Iteration of the AHC Method

Cluster	K1	K2	K3	K4	K5
1.2.4.10.3.9.5.6.8.7	1	1	1	1	1

Based on Table 4, the results of the ninth iteration form a single cluster in the final iteration.

6. The results obtained from the Agglomerative Hierarchical Clustering calculation are used as the initial centroid points for the K-Means clustering method.

Tabel 5. Initial Centroid Point of K-Means Method

Cluster	K1	K2	K3	K4	K5
1	1	1	1	1	1

7. Calculate the distance of each object to the centroid point using the Euclidean distance technique as expressed in Equation (1).

Tabel 6. Results of the 1st Iteration of the K-Means Method

No	Name	Result (Distance)	Sequence
1	Agus Salim Adam	2,00	3
2	T. Sulaiman	2,24	4
3	Hendrik Nasir M.	3,32	9
4	Munzir Abdullah	1,41	2
5	Abdullah Umar	2,65	7
6	Hermawan Hamzah	2,24	4
7	Tgk. Nyak Umar	4,24	10
8	FA. Mutia	2,65	7
9	Nyak Pakeh	2,45	6
10	Ahmad Marzuki	1,00	1

Based on Table 6, the results of the first iteration of the K-Means method can be observed, along with the sequence generated from calculating the distance of each object to the centroid point. All objects are located within the same cluster.

8. Determine the new centroid point using equation (3):

Tabel 7. New Centroid Point of K-Means Method

Iterasi	K1	K2	K3	K4	K5
2	1,6	2,4	2,1	1,0	1,5

9. Recalculate the distance of each object to the centroid point using the Euclidian distance technique using equation (1):

Tabel 8. Results of the 2nd Iteration of the K-Means Method

No	Name	Result (Distance)	Sequence
1	Agus Salim Adam	1,84	8
2	T. Sulaiman	1,84	8
3	Hendrik Nasir M.	1,73	7
4	Munzir Abdullah	1,61	5
5	Abdullah Umar	0,88	1
6	Hermawan Hamzah	1,41	4
7	Tgk. Nyak Umar	2,89	10
8	FA. Mutia	1,17	3
9	Nyak Pakeh	0,88	1
10	Ahmad Marzuki	1,61	5

10. Hasil perhitungan metode k-means kemudian diurutkan berdasarkan jarak terdekat hingga jarak terjauh dengan titik centroid.

Tabel 9. Manual Calculation Results of the KnA Method

Sequence	Name	Result (Distance)
1	Abdullah Umar	0,88
2	Nyak Pakeh	0,88
3	FA. Mutia	1,17
4	Hermawan Hamzah	1,41
5	Munzir Abdullah	1,61
6	Ahmad Marzuki	1,61
7	Hendrik Nasir M.	1,73
8	Agus Salim Adam	1,84
9	T. Sulaiman	1,84
10	Tgk. Nyak Umar	2,89

3.1.2. The results of the KnA method on the system

The test data used for evaluating the KnA method in the system are the same as those employed in the manual calculation of the KnA method conducted using Microsoft Excel, as presented in Table 1.

3.1.3. Testing the results of the KnA method

The evaluation of the KnA method results involves testing the consistency of the ranking generated by the system based on assessments using the original weighted data for each criterion. The data used in the evaluation consist of the top five records from the clustering results of the KnA method in the system for each category. Table 10 presents the results of the KnA method testing derived from the system.

Tabel 10. KnA Method Results Testing of the System

No	Sub-Group	Number of Test Data	Number of Matched Sequences
1	Fakir	5	5
2	Miskin	5	1
3	Miskin Insidentil	5	2
4	Guru Dayah	5	3
5	Santri Dayah	5	1
6	Santri Dayah Diluar Daerah	5	1
7	Anak Yatim	5	0
8	Disabilitas	5	5
9	Pelajar Miskin	5	2
10	Santri Muallaf	5	1
11	Muallaf	5	2
12	Gharimin	5	0
13	Santri Prestasi	5	3
14	Madrasah	5	5
15	Waqaf	5	5
16	Pelajar Diluar Negeri	5	3
17	Ibnu Sabil	5	5

Tabel 11. Percentage of Results of KnA Method from System

No	Number of Groups	Number of Matched Sequences	Percentage
1	5	5	29,41%
2	3	3	17,64%
3	4	2	23,52%
4	4	1	23,52%
5	2	0	11,76%

Based on Table 11, it can be concluded that the application of the KnA method in determining zakat al-māl recipients demonstrates a relatively low level of accuracy, with only five out of seventeen groups correctly classified in order for the five sample data, resulting in an accuracy rate of 29.41%.

4. Result and Discussions

Based on the results and discussion presented, the following conclusions can be drawn:

1. The clustering and determination system for zakat al-māl recipients was successfully designed and developed by implementing the KnA method in accordance with its algorithm. This method serves as the basis for identifying eligible zakat recipients within each category, in line with the regulations established by Baitul Mal Aceh Besar.
2. The results of applying the KnA method in the process of determining *zakat al-māl* recipients are as follows:
   1) The results of the KnA method, in order of accuracy based on evaluation using the original data with five correctly classified test data, correspond to the categories of *fakir* (the very poor), persons with disabilities, *madrasah* operational assistance, *waqf* property certification assistance, and *ibnu sabil* (travelers in need).
   2) The results of the KnA method, based on evaluation using the original data with three correctly classified test data, correspond to the categories of *dayah* teachers, outstanding *santri* (Islamic boarding school students), and students studying abroad.
   3) The results of the KnA method, based on evaluation using the original data with two correctly classified test data, correspond to the categories of incidental poor (*miskin insidentil*), poor students, and *muallaf* (new converts to Islam).
   4) The results of the KnA method, based on evaluation using the original data with one correctly classified test data, correspond to the categories of the poor (*miskin*), *dayah* students, *dayah* students outside Aceh Besar, and *muallaf* students.
   5) The results of the KnA method, based on evaluation using the original data with zero correctly classified test data, correspond to the categories of orphans (*anak yatim*) and *gharimin* (debtors).

REFERENCES

[1] K. U, "Manajemen Zakat Modern," *UIN-Maliki Press,* no. 3, 2010.

[2] Baznas, "Zakat Maal," Baznas, 2019. [Online]. Available: https://baznas.go.id/zakatmaal. [Accessed 06 Januari 2023].

[3] Baitul Mal Aceh, "Profil BMA," 2022. [Online]. Available: https://www.baitulmal.acehprov.go.id/. [Accessed 06 Januari 2023].

[4] I. Alpiana and L. Anifah, "Penerapan Metode KnA (Kombinasi K-Means dan Agglomerative Hierarchical *Cluster*ing) dengan Pendekatan Single Linkage untuk Menentukan Status Gizi pada Balita," *INAJET,* vol. 1, no. 2, 2019.

[5] T. Alfina, B. Santosa and A. R. Barakbah, "Analisa Perbandingan Metode Hierarchical *Cluster*ing, K-Means dan Gabungan Keduanya Dalam *Cluster* Data (Studi Kasus: Problem Kerja Praktek Teknik Industri ITS)," *Jurnal Teknik ITS,* vol. 1, no. 1, 2012.