Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics
Multidisciplinary : Rapid Review : Open Access Journal **Vol. 6, No. 2, May 2024, pp.65-77 e-ISSN: 2656-8624**

RESEARCH ARTICLE **OPEN ACCESS**

Manuscript received December 18, 2023; revised January 20, 2024; accepted January 12, 2024; date of publication Mei 23, 2024
Digital Object Identifier (**DOI**): https://doi.org/10.35882/ijeeemi.v6i2.346

How to cite: Hermanto , Ade Hendi, and Aminatuz Zuhriyah, "Implementation of the Web-Based K-Means Clustering Algorithm on Hypertension Levels in the Elderly at the Bungah District Health Center", Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics, vol. 6, no. 1, pp. 65-77, Mei. 2024.

Implementation of the Web-Based K-Means Clustering Algorithm on Hypertension Levels in the Elderly at the Bungah District Health Center

Hermanto , Ade Hendi, and Aminatuz Zuhriyah

Department of Informatics Engineering, Qomaruddin University, Jl. Raya Bungah No. 1, Gresik, 61152, Indonesia

Corresponding author: Hermanto (hermanto25@uqgresik.ac.id)

ABSTRACT This research was carried out at the community health center in Bungah sub - district with the aim that so that patients receive appropriate treatment, a web-based system is needed to group hypertension data into several classes (clusters). The method used in data grouping is the K-Means Clustering method . In this study, 100 hypertension data were grouped into 4 clusters . The hypertension data used consists of two attributes, namely systole and diastole, the working mechanism is to normalize the data first, then the system groups the data into groups that have the same characteristics. The system will display the results of the clustering process which consists of four (4) clusters , namely cluster 1 Isolated Systolic Hypertension, cluster 2 Grade 1 (mild hypertension), cluster 3 Grade 2 (moderate hypertension), and cluster 4 Grade 3 (high hypertension). The data used for clustering is 100 data from blood pressure checks of patients at the Community Health Center using a blood pressure monitor. The results of the last iteration of 100 hypertension data in the system were used as parameters for calculating the level of effectiveness, by comparing the data from the clustering test results to the data from diagnosis a which produced an effectiveness value of 80%.

INDEX TERMS ***Hypertension, K-Means Clustering, Web Based.***

I. INTRODUCTION

Physical changes in the elderly [1]will be a health problem that makes them susceptible to disease, such as cardiovascular changes, decreased elasticity of blood vessels, changes in respiration, namely decreased strength of the respiratory muscles, as well as changes in hearing and changes in vision [2]. Hypertension or high blood pressure is a condition where the systolic pressure reaches above 140 mmHg and the diastolic pressure above 90 mmHg [3]. This chronic disease often appears without symptoms, resulting in increased mortality rates, and is often referred to as the silent killer [4]. The Bungah District Community Health Center (Puskesmas) is one of the health institutions that handles elderly patients, who have a high potential for having a history of hypertension. Based on data from examinations of elderly patients from 11 October to 20 October 2022 , data was obtained that 100 patients had hypertension, with the percentage ratio between male and female patients being 39% to 61%. In order for patients with hypertension to receive appropriate treatment, it is necessary to group data on hypertension in the elderly so as to minimize the emergence of degenerative diseases caused by hypertension. The process of grouping this data requires a clustering method.

Clustering is a grouping method based on a measure of closeness (similarity) [5] . Clustering is different from group, group means the same group. Clusters do not always have to be the same, but the grouping is based on the closeness of an existing sample characteristic, one of which is by

using the Euclidean distance formula [6] . K-Means clustering is a non-hierarchical data clustering method that attempts to partition existing data into one or more clusters/groups. This method partitions data into clusters/groups so that data that has the same characteristics is grouped into the same cluster [7] .

The K-Means method is a fast and efficient method that can be used in data clustering. The K-Means clustering method has several advantages, namely that the method is easy to implement and run, the time required to carry out the learning method is relatively fast, it is easy to adapt, and the method is also commonly used [8] .

This research draws on several previous research studies . One of them is entitled " Implementation of the *K-Means Clustering Algorithm* in Determining Ideal Body Weight " [9] This research discusses the application of the *K-Menas Clustering method* to determine ideal body weight criteria. These criteria are based on data on the patient's physical condition, including calculating *the body mass index* (BMI), the size of the patient's frame and the patient's daily calorie needs based on the patient's age and daily activities. The results of this research are that the *K-Means Clustering algorithm* can be used to determine ideal body weight which consists of 3 *clusters* based on BMI and frame, namely normal BMI and large frame, moderate obesity BMI and medium frame, and severe obesity BMI and small frame.

Meanwhile, the aim of this research is to create a web - based application for grouping high blood pressure (Hypertension) in the elderly at *the* Bungah District Health Center into 4 c *increase the level* of hypertension and measure the effectiveness of using the K - Means Clustering method .

II. MATERIALS AND METHODS

A. Clustering

Data clustering can be divided into two objectives [10] , namely clustering for understanding and clustering for use. If the goal is understanding, the clusters formed must capture the natural structure of the data. Usually the clustering process for this purpose is only an initial process to then continue with this work such as summarization (average, standard deviation), class labeling in each group to then be used as classification training data, and so on. Meanwhile, if the goal is for use, usually the main goal is to find *a cluster prototype* that is most representative of the data and provides an abstraction for each data object in *the cluster* where the data is located. inside it. *Clustering can be differentiated according to cluster* structure , data membership in *the cluster* and data compactness in *the cluster*.

B. K-Means Clustering

There are two types of data *clustering* that are often used in the data grouping process, namely *hierarchical* (hierarchical) and *non-hierarchical* (non-hierarchical) data *clustering* . *K-Means* is a non-hierarchical data *clustering method* that partitions data into one or more clusters/groups, so that data that has the same characteristics is grouped in the same *cluster* and data that has different characteristics is grouped into other groups [11] .

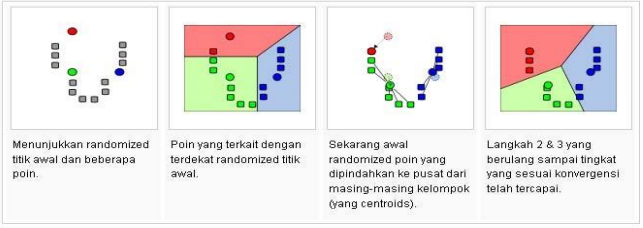

Figure 1. Illustrations of 2 Dimensional K-Means Clustering Actions

It is stated that the *K-Means method* is a fast and efficient method that can be used in *data clustering* . The K-Means method was first introduced by MacQueen JB in 1976. This method is one of the *non-hierarchical methods* that is commonly used. This method is included in the partition technique *which* divides or separates objects into k separate areas.

K-Means, each object must be included in a certain group, but in a certain stage of the process, objects that have been included in one group, at the next stage will move to another group [12]

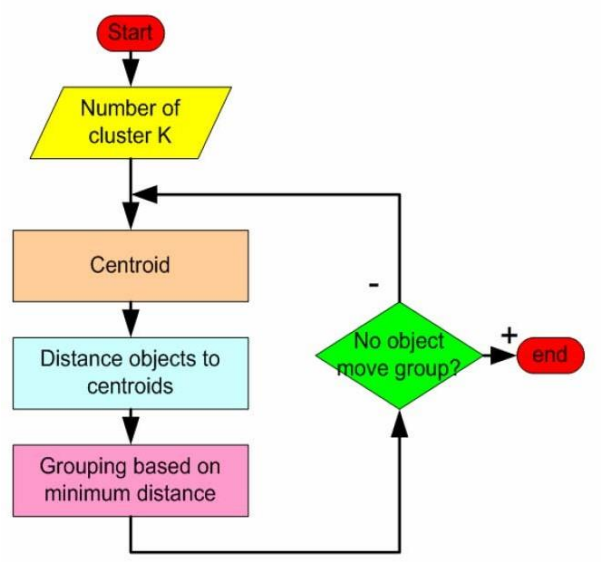

Figure 2. Flowchart of the K-Means Clustering Algorithm

C. Hypertension Classification

Classification of hypertension according to The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation and the Treatment of High Blood Pressure [13].

TABLE 1
Classification of Hypertension according to JNC

Category	Systolic (mmHg)	Diastolic (mmHg)
Optimal	115 or less	75 or less
Normal	Less than 120	Less than 80
Prehypertension	120 – 139	80 – 89
Stage I Hypertension	140 – 159	90 – 99
Stage II Hypertension	More than 160	More than 100

WHO (World Health Organization) and ISH (International Society of Hypertension) [14]groups hypertension as follows:

TABLE 2
Classifications of Hypertension According to WHO - ISH

Category	Blood pressure	
	Systole (mmHg)	Diastole (mmHg)
Optimal	<120	<80
Normal	<130	<85
Normal – high	130 – 139	85 – 89
Grade 1 (Mild Hypertension)	140 – 159	90 – 99
Sub - group : border	140 – 149	90 – 94
Grade 2 (moderate hypertension)	160 – 179	100 – 109
Grade 3 (Severe Hypertension)	>180	>110
Isolated systolic hypertension	≥ 140	<90
Sub - group : border	140 – 149	<90

Association in January 2007 launched guidelines for treating hypertension in Indonesia, taken from state guidelines developed and neighboring countries by referring to the results of JNC and WHO [15].

TABLE 3
Classifications of Hypertension Based on the Consensus of the Indonesian Hypertension Association

Category	Blood pressure	
	Systole (mmHg)	Diastole (mmHg)
Normal	<120	<80
Prehypertension	120 – 139	80 – 90
Stage 1 hypertension	140 – 159	90 – 99
Stage 2 hypertension	>160	>100
Isolated Systolic Hypertension	≥ 140	<90

D. Data Modeling

Clustering process uses data on high blood pressure (Hypertension) in the Elderly, which consists of two criteria, namely systole and diastole . The data will be normalized first and then grouped into the number of clusters determined according to the results of the *clustering process* . The process of finding the initial *centroid value* is as follows:

Calculate the mean (average value) of all high blood pressure data for each cluster (Hypertension) using equation 3. [16]

$$\overline{X} = \frac{\Sigma x}{n}$$

Where $\overline{X}$ is mean (average), $\sum_x$ is the respective results of adding up the systole and diastole data, and n is the total number of systole and diastole data respectively

The process of grouping high blood pressure (Hypertension) data based on WHO (World Health Organization) and ISH (International Society of Hypertension) can be seen in Table 4.

TABLE 4.
Grouping of Hypertension

Clusters	Category	Blood pressure	
		Systole (mmHg)	Diastole (mmHg)
C 2	Grade 1 (Mild Hypertension)	140 – 159	90 – 99
	Sub - group : border	140 – 149	90 – 94
C 3	Grade 2 (moderate hypertension)	160 – 179	100 – 109
C 4	Grade 3 (Severe Hypertension)	>180	>110
C 1	Isolated systolic hypertension	>140	<90
	Sub - group : border	140 – 149	<90

The systole and diastole blood pressure data obtained cannot be processed directly because there are quite large numbers between the Systole and Diastole variables. This large difference in distance or number size can make the grouping process difficult. One solution used to reduce the magnitude of the numbers between variables is to normalize the numbers in the systole and diastole variables using equation 4. [17]

$$\text{normalization value} = \frac{(\text{initial value} - \text{minimum valuel})}{(\text{maximum value} - \text{minimum valuel})}$$

The values of the systole and diastole variables will be normalized into the range 0 – 1. Normalization of the numbers for each variable is very necessary before the process of calculating the centroid value by the K-Means algorithm so that no parameter dominates in calculating the distance between data The stages carried out for the normalization process are:

Finding the maximum and minimum values for the systole variable (X)

Find the maximum and minimum values for the diastole variable (Y)

Calculate the normalization value using Eq. (1)

Example of systole data normalization process:

$$X_{11} = \frac{(X_{sistole1} - X_{min})}{(X_{maks} - X_{min})} = \frac{(157 - 140)}{(220 - 140)} = 0.21$$

$$X_{12} = \frac{(X_{sistole2} - X_{min})}{(X_{maks} - X_{min})} = \frac{(167 - 140)}{(220 - 140)} = 0.34$$

$$X_{13} = \frac{(X_{sistole3} - X_{min})}{(X_{maks} - X_{min})} = \frac{(152 - 140)}{(220 - 140)} = 0.15$$

$$X_{14} = \frac{(X_{sistole4} - X_{min})}{(X_{maks} - X_{min})} = \frac{(167 - 140)}{(220 - 140)} = 0.34$$

$$X_{15} = \frac{(X_{sistole5} - X_{min})}{(X_{maks} - X_{min})} = \frac{(186 - 140)}{(220 - 140)} = 0.58$$

Diastole data normalization process:

$$Y_{21} = \frac{(Y_{diastole1} - Y_{min})}{(Y_{maks} - Y_{min})} = \frac{(92 - 60)}{(116 - 60)} = 0.57$$

$$Y_{22} = \frac{(Y_{diastole2} - Y_{min})}{(Y_{maks} - Y_{min})} = \frac{(107 - 60)}{(116 - 60)} = 0.84$$

$$Y_{23} = \frac{(Y_{diastole3} - Y_{min})}{(Y_{maks} - Y_{min})} = \frac{(85 - 60)}{(116 - 60)} = 0.45$$

$$Y_{24} = \frac{(Y_{diastole4} - Y_{min})}{(Y_{maks} - Y_{min})} = \frac{(89 - 60)}{(116 - 60)} = 0.52$$

$$Y_{25} = \frac{(Y_{diastole5} - Y_{min})}{(Y_{maks} - Y_{min})} = \frac{(109 - 60)}{(116 - 60)} = 0.88$$

The same calculation is carried out until the high blood pressure (Hypertension) data reaches – 100. Clustering Process This stage will apply the K-Means method to group data. The results of this grouping will then carry out a test data validation stage against the diagnostic data where the results of the culstering in the system will be calculated for its effectiveness.

The K-Means Culstering algorithm for the level of hypertension in the elderly is as follows:
The number of clusters formed as a k value is four (k=4) Allocate systole and diastole data into groups randomly Generate 4 random values for the initial cluster center (centroid) from the data in each group. The centroid of criterion 1 is systole and the centroid of criterion 2 is diastole. The centroid location of each group is taken from the average (mean) of all data values for each feature. Calculation of the average value (mean) uses equation 3.1. The following are the results of calculating random values for the initial cluster center:

TABLE 5
Initial Cluster

Clusters	Initial Centroid	
	C. Systole	C. Diastole
Cluster 1 (HST)	0.05	0.33
Cluster 2 (Grade 1)	0.07	0.53
Cluster 3 (Grade 2)	0.35	0.58
Cluster 4 (Grade 3)	0.63	0.74

Calculate the distance of each high blood pressure (Hypertension) data that has been input to each centroid using the Eucledian Distance formula until the closest distance of each data is found to the centroid using equation 2. :

$$d(x_i, y_j) = \sqrt{(x_{ia} - y_{ja})^2 + (x_{ib} - y_{jb})^2}$$

Where :
xi : data on high blood pressure (Hypertension)
yj : centroid of the jth cluster
xia : normalized systole data
xib : normalized diastole data
yja : the value of criterion 1 from the jth cluster centroid is the random value of systole data
yjb : the 2nd criterion value from the jth cluster centroid is the random value of diastole data

An example of calculating the distance to the 1st data in each cluster is:

$$d(x_1, C_1) = \sqrt{(\text{systole}_1 - \text{systole}_{c1})^2 + (\text{diastole}_1 - \text{diastole}_{c1})^2}$$
$$= \sqrt{(0.21 - 0.05)^2 + (0.57 - 0.33)^2}$$
$$= 0.29$$
$$d(x_1, C_2) = \sqrt{(\text{systole}_1 - \text{systole}_{c2})^2 + (\text{diastole}_1 - \text{diastole}_{c2})^2}$$
$$= \sqrt{(0.21 - 0.07)^2 + (0.57 - 0.53)^2}$$
$$= 0.15$$
$$d(x_1, C_3) = \sqrt{(\text{systole}_1 - \text{systole}_{c3})^2 + (\text{diastole}_1 - \text{diastole}_{c3})^2}$$
$$= \sqrt{(0.21 - 0.35)^2 + (0.57 - 0.58)^2}$$
$$= 0.13$$
$$d(x_1, C_4) = \sqrt{(\text{systole}_1 - \text{systole}_{c4})^2 + (\text{diastole}_1 - \text{diastole}_{c4})^2}$$
$$= \sqrt{(0.21 - 0.63)^2 + (0.57 - 0.74)^2}$$
$$= 0.45$$

The same equations and calculations were applied to 100 high blood pressure data to get the distance for each data in each cluster.
Group each high blood pressure (Hypertension) data based on its proximity to the centroid (smallest distance).
For example, the calculation results from data distance to 1 have a distance of 0.29 to cluster 1. Cluster 2 has a distance of 0.15. Cluster 3 has a distance of 0.13. And in cluster 4 it has a distance of 0.45. Of the 4 clusters, blood pressure data for hypertension 1 has the shortest distance to cluster 3. Therefore blood pressure data for hypertension 1 is included in cluster 3. The same steps are applied to the next 100 data to carry out grouping in iteration 1. Updates the centroid value. The new centroid value is obtained from the average of the cluster in question using equation 2.2 .

$$y_j(t+1) = \frac{1}{N_{sj}} \sum_{j \in Sj} x_j$$

Where :
$y_j(t+1)$: new centroid at iteration (t+1)
N_{sj} : lots of hypertension blood pressure data in the Sj cluster

Repeat from steps 2 to 5, if there is still hypertension blood pressure data that moves clusters. Until the members of each cluster nothing changes.

If step 7 has been fulfilled, then the cluster center value (yj) in the last iteration will be used as a parameter to determine the classification of high blood pressure (Hypertension) levels in the Elderly.

Effectiveness Process of Clustering Results

After the clustering process, the next stage is the effectiveness calculation process. In the process of calculating this effectiveness, it will be determined how much the level of effectiveness of the data from the clustering test results in the last iteration is with the data from the diagnosis by one of the parties at the elderly clinic at the Bungah District Health Center. The process of calculating the level of effectiveness uses equation 3.3 [18]

$$\text{Effectiveness } (\%) = \frac{\text{amount of classified data}}{\text{amount of tested data}} \text{ x } 100\%$$

E. Unified Modeling Language(UML)

UML (Unified Modeling Language) is a language for determining, visualizing, constructing and documenting *artifacts* (parts of information that are used or produced in a software creation process, *artifacts* can be models, descriptions or other software) [19] . In this research, the system design uses several types of UML, including:

F. Use Case Diagrams

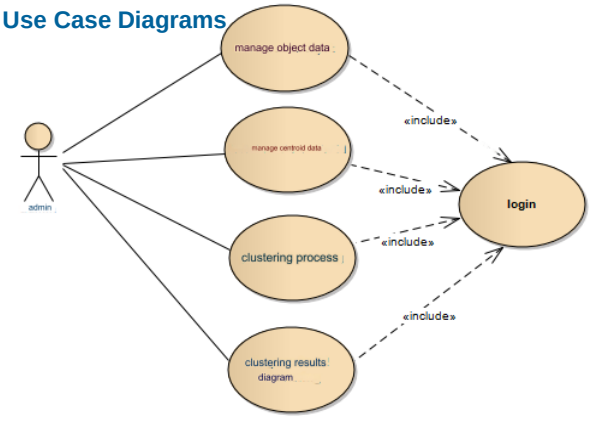

Figure 3. Use Case Diagrams

explain about use case diagrams from the system[20] , there is one *user* , namely admin, there are 4 entities including managing object data, managing *centroid data* , processing the K Means *Clustering method* , and managing clustering results diagrams . Before processing data, admins are required to carry out a login process first so that admins can carry out data processing activities .

G. Activity Diagram login

Activity diagrams admin login [21], admin enters the *login menu* then the system will display *the username* and column password , if an error occurs when filling in the username and *password* , it will return to the login menu. If filling in is successful , the system will display the main page and the admin will enter the main page.

H. Activity Diagram Manage Hypertension Data

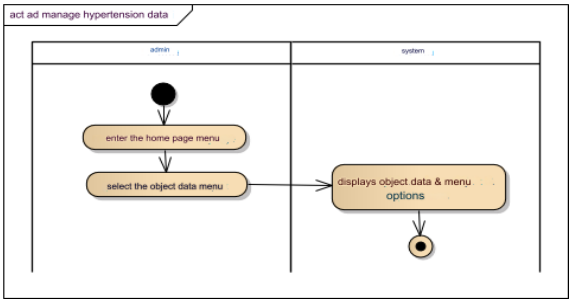

Figure 4. Activity Diagrams Manage Hypertension Data

Activity diagrams manage hypertension data [22], the first step is for the admin to enter the *home page menu* then *select* the object data menu and the system will display the object data & menu options.

I. Activity Diagram Centroid Data Input

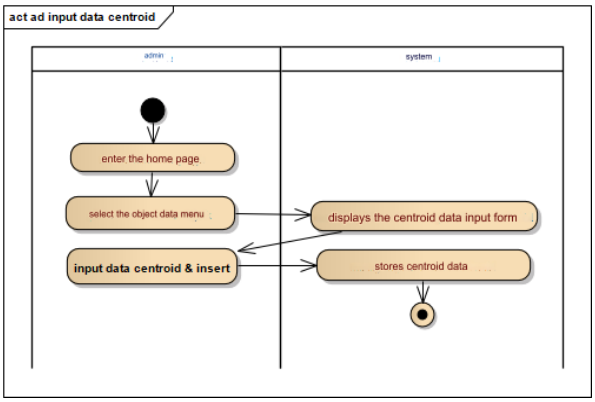

Figure 5. Activity Diagram Centroid Data Input

Activity diagram input centroid data [23], admin enters the *home page* menu then selects the object data menu , then the system displays

the *centroid* data and other menu options , the system displays the centroid data input form, then the admin inputs the centroid *data* & select *insert* dat a and s i stem to save the data *centroid* .

J. Clustering Process Activity Diagram

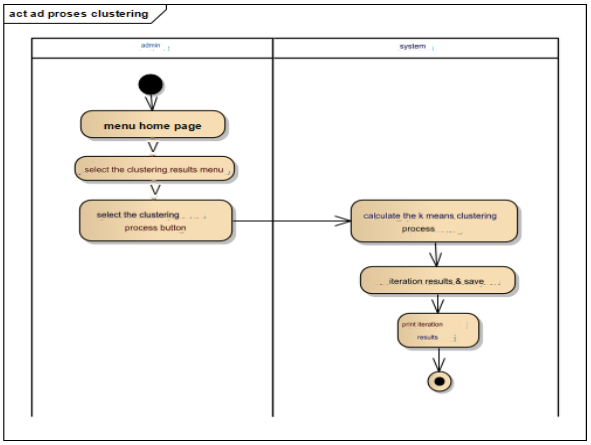

Figure 6. Activity Diagram Clustering Process

Activity diagrams clustering process [24]**,** the admin's first step is to enter *the home page* menu then select the *clustering* results menu , click the *clustering* process *button* , *then the system* will process the calculation of hypertension data using the K Means *Clustering method* and display the results of the calculation iteration .

K. Admin Login Sequence Diagram

Sequence Admin Login Diagram [25], there are several processes, namely the admin enters the login menu then the system will display the username and password columns, if an error occurs when filling in the username and password then it will return to the login menu, if the filling is successful then the system will display the main page and the admin enters the page main.

L. Sequence Diagrams Manage Hypertension Data

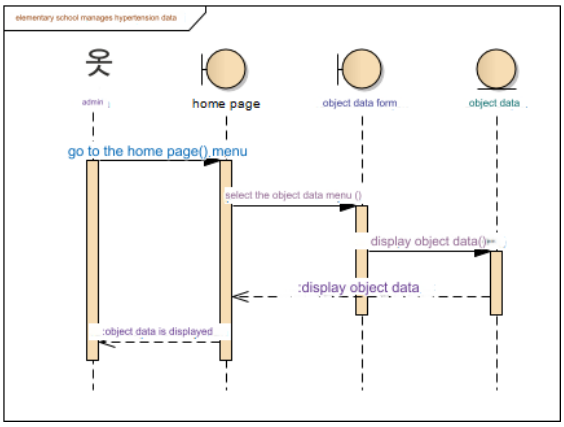

Figure 7. Sequence Diagram for Managing Hypertension Data

Sequence diagram for managing hypertension data [26]**,** the first step is that the admin enters the homepage menu then selects the object data menu and the system will display the object data & menu options.

M. Sequence Diagram Centroid Data Input

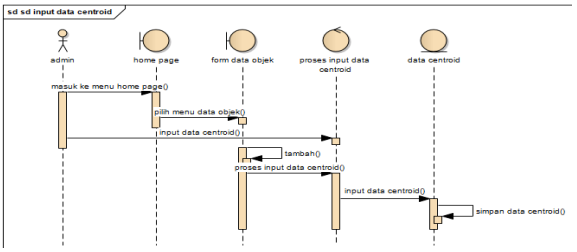

Figure 8. Sequence Diagram of Centroid Data Input

Sequence diagram input centroid data [27]**,** admin enters the *home page menu* then selects the object data menu, then the system displays centroid data and other menu options , the system displays *the centroid* data *input form* then *admin* inputs *centroid* data & select *insert* dat a and s i stem to save the data *centroid* .

N. Clustering Process Sequence Diagram

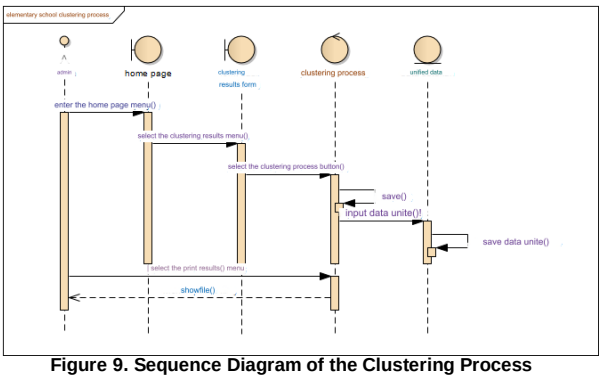

Figure 9. Sequence Diagram of the Clustering Process

Sequence *diagram clustering* process [28], the admin's first step is to enter the *home page* menu then select the *clustering* results menu , click the *clustering* process *button* , then the system will process hypertension data calculations using the K Means *Clustering method* and displays the results of calculation iterations , and there is a print results button to print *clustering results* .

O. Class Diagrams

Class diagram K Means Clustering Application of Hypertension Levels in the Elderly at the Bungah District Health Center [29]. There are six class diagrams . Namely the admin class , centroid class, diagram class, diagram_centroid class, object class, and unite class . In the admin class there is a username and password which is the identity of the admin. Admin must have *a username* and *password* to enter, view and delete data. The admin *class* consists of two attributes, namely *username* and *password* .

III. RESULTS

The results of the implementation are the application of the program to a system that has been researched previously. The results of the implementation of this system are applied in accordance with the design that was previously made. This implementation stage is the stage of placing the system or program that has been created. A web-based K-Means *Clustering* application for hypertension levels in the elderly so that it can be applied properly. The system is implemented in a web-based display, namely by using the PHP programming language and MySQL *database* , as well as XAMPP as the implementation server.

Implementation of the K-Means Clustering Application on Hypertension Levels in the Elderly. As follows:

1. **Login menu,** showing the Login page contains the admin username and password columns. Admin fills in the username and password fields that have been created in phpMyAdmin, then click login to enter the main page.
2. **Home Page Menu,**
   The home page menu shows that after the admin logs in, he goes to the main home page. On this page there is an explanation about data mining.
3. **The Hypertension Level menu,** shows the Hypertension Level page, on the Hypertension Level page there is an explanation of the level of hypertension according to WHO and ISH, and the patient's age limit.

Figure 10. Hypertension levels

4. **All Data menu ,** shows the All Data page, this menu displays a form that the admin can use to input hypertension data and the initial centroid data of the data to be processed. In this menu there are several action buttons that admins can use to process data, namely: Add data: to add data
   Browse data : alternative add button data via xls *file* .

   Erase data : to delete the data you want to delete

Figure 11. All Data

5. **Clustering Results Menu** , shows the *Clustering Results page* In the menu there is a *button* , when the admin clicks the button the system will process the data that was previously input using the K-Means *clustering formula* .

Figure 12. Clustering Process Button

The following is Figure 14. showing the interface display page after the admin clicks the button. In the clustering results display there is a button to print the clustering results.

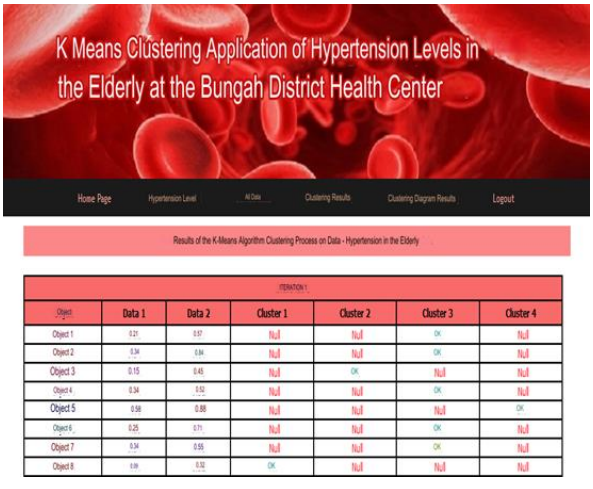

Figure 13. Clustering Results

6. **Clustering Diagram Results menu** shows the Diagram Results page, in the menu there is a *button* , when the admin clicks the button the system will display the results of the *clustering diagram* .

Figure 14. Clustering Diagram Results

The following are the results of the *clustering diagram* :

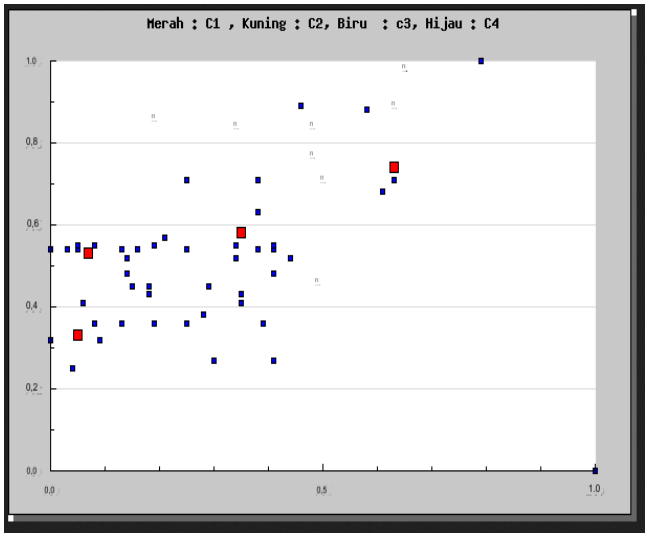

Figure 15. Clustering Diagram Results

7. **Hypertension Dta Calculation Results using the K-Meanas Clustering Method**

The following are the results of clustering calculations from system implementation of 100 hypertension data which were normalized first and then grouped using the K-Means Clustering method based on their proximity to the centroid center.

The data used to calculate effectiveness is data from diagnostic results from examinations of elderly high blood pressure (Hypertension) patients, namely patients aged 45 years to 90

years and over at the Bungah Gresik District Health Center who have been classified by one of the parties at the Bungah District Health Center which is located at the Elderly Clinic. and data from the last iteration of the clustering process, the data used in the clustering process uses hypertension data without using class labels.

The amount of diagnostic data used was 100 high blood pressure (Hypertension) data consisting of 5 isolated Systolic Hypertension high blood pressure data, 48 high blood pressure data with Grade 1 (Mild Hypertension), 36 data with Grade 2 (Moderate Hypertension).), and 11 hypertension data with Grade 3 (High Hypertension), as well as 100 high blood pressure test data which were grouped using the K-Means Clustering method which was divided into four clusters.

The effectiveness calculation stage is the stage where the data from the last iteration of the clustering test using the K-Means Clustering method on the diagnosis result data which has been classified by the Elderly Polytechnic of the Bungah District Health Center is calculated for its effectiveness using the equation. The calculation results are as follows:

Effectiveness = 80/100 x 100%=80%

With the results of these effectiveness calculations, there are 80 data from the last iteration of the clustering process using the K-Means Clustering method which is the same as the diagnostic data that has been classified by the Elderly Policing at the Bungah District Health Center. The results of the effectiveness calculation may change as the hypertension data changes in the clustering process. Can be seen in the table below.

Table 6. Data on the Effectiveness of the Last Iteration Cluster Results

N NO	C 1 (HST)	C 2 (G1)	C 3 (G2)	C 4 (G3)	MATCH
1		Ok			
2		Ok			1
3	ok				
4		Ok			1
5				ok	1
6		Ok			1
7		Ok			1
8	ok				1
9	ok				1
10	ok				1
11		Ok			1
12		Ok			1
13				ok	1
14	ok				
15		ok			
16	ok				
17		ok			1
18			ok		1
19	ok				
20		ok			1
21			ok		1
22		ok			1
23			ok		1
24		ok			1
25				ok	
26		ok			1
27		ok			1
28			ok		
29			ok		1
30	ok				
31	ok				
32	ok				1
33	ok				
34			ok		1
35	ok				
36			ok		1
37		ok			1
38				ok	
39		ok			1
40		ok			1
41	ok				
42		ok			1
43		ok			1
44		ok			1
45		ok			1
46		ok			1
47			ok		1
48			ok		1
49	ok				

50	ok			1
51	ok			1
52			ok	1
53			ok	1
54	ok			
55			ok	1
56			ok	1
57	ok			
58	ok			1
59		ok		1
60	ok			1
61	ok			1
62	ok			
63	ok			1
64			ok	1
65	ok			1
66		ok		1
67	ok			1
68	ok			1
69	ok			1
70		ok		1
71		ok		1
72	ok			1
73			ok	
74			ok	1
75		ok		1
76		ok		1
77	ok			1
78	ok			1
79	ok			1
80		ok		1
81		ok		1
82			ok	1
83	ok			1
84	ok			1
85	ok			1
86			ok	1
87	ok			1
88		ok		1
89	ok			
90	ok			1
91	ok			1
92	ok			1
93	ok			1
94		ok		1
95	ok			1
96	ok			1
97		ok		1
98			ok	1
99	ok			1
100		ok		1

This test obtained 5 iteration results using calculations effectiveness of 100 test data using the K-Means *Clustering method* in the last iteration Based on the diagnostic results data, an effectiveness value of 80% was obtained. From these effectiveness values, 80 test result data were obtained which were in the same group as the result data .

IV. DISCUSSIONS

Based on the results of this research, an application for grouping hypertension data was obtained using the K-Means *Clustering algorithm.* As for Grouping data on the level of hypertension in the elderly consists of two criteria, namely *systole* and *diastole* which are divided into four (4) *clusters* , namely *cluster* 1 Isolated Systolic Hypertension, *cluster* 2 *Grade* 1 (mild hypertension), *cluster* 3 *Grade* 2 (moderate hypertension), and *cluster* 4 *Grade* 3 (high hypertension). From this test, 5 iteration results were obtained using calculations effectiveness of 100 test data using the K-Means *Clustering method* in the last iteration Based on the diagnostic results data, an effectiveness value of 80% was obtained. Based on the results of the effectiveness value, the K-Means *Clustering method is used* This is effective enough to be used as a support in determining the level of hypertension in the elderly.

The results of this research were compared with another research entitled "Implementation of the K-Means Clustering Algorithm in Determining Ideal Body Weight". [22]The data used to calculate ideal body weight is based on the patient's height, body weight and frame size, while to calculate daily calorie needs is based on the patient's age, weight, height, body weight and activity. Then the patient data is classified according to

body mass index (BMI), Basal Metabolic Rate (BMR), age and skeletal size based on the physical data of the patient that has been taken previously. used the K-Means Clustering method with 3 clusters based on BMI and frame, namely normal BMI and large frame, moderate obesity BMI and medium frame, and severe obesity BMI and small frame. There is a difference, namely using 4 clusters and 3 clusters.

This application has a drawback, namely that there is no special form for the data normalization process, and a special form for checking *centroid data* so that the data normalization process and the *centroid data checking process* cannot be done automatically.

So this clustering research makes it easier for health workers to group hypertension data in the elderly at the Bungah District Health Center.

V. CONCLUSION

The aim of this research is to create a web-based application for grouping high blood pressure (hypertension) data in the elderly at the Bungah sub-district health center into several hypertension level clusters and measuring the effectiveness of using the K-Means Clustering method.

The conclusion that can be drawn based on the results of the discussion is that the application of hypertension data grouping using the K-Means *Clustering algorithm has* an effectiveness value of 80%. From these effectiveness values, 80 test result data were obtained which were in the same group as the diagnosis result data. Based on the results of the effectiveness value, the K-Means *Clustering method is used* This is effective enough to be used as a support in determining the level of hypertension in the elderly.

For further work development of this system, further research is needed which is useful for completing existing deficiencies. These include: 1) It is hoped that the application can be developed further by adding other features that might be even more effective, such as adding a data normalization form feature, a data check feature to make it easier to determine the *centroid point* . 2) It is best that in the *clustering process* , the criteria used are added, because there are still many other criteria based on hypertension factors that can be used as research, for example food intake factors, daily patient activities, hereditary factors (race), medication, stress levels and others. - other.

VI. REFERENCES

[1] S. Rahman, "Faktor-faktor yang mendasari stres pada lansia," *J. Penelit. Pendidik.*, vol. 16, no. 1, 2016.
[2] D. Kurnianto, "Menjaga kesehatan di usia lanjut," *Jorpres J. Olahraga Prestasi*, vol. 11, no. 2, 2015.
[3] A. Yonata and A. S. P. Pratama, "Hipertensi sebagai faktor pencetus terjadinya stroke," *J. Major.*, vol. 5, no. 3, pp. 17–21, 2016.
[4] B. Nuraini, "Risk factors of hypertension," *J. Major.*, vol. 4, no. 5, 2015.
[5] J. Irani, N. Pise, and M. Phatak, "Clustering techniques and the similarity measures used in clustering: A survey," *Int. J. Comput. Appl.*, vol. 134, no. 7, pp. 9–14, 2016.
[6] S. T. Wierzchoń and M. A. Kłopotek, *Modern algorithms of cluster analysis*, vol. 34. Springer, 2018.
[7] D. Ardiada, P. A. Ariawan, and M. Sudarma, "Evaluation of Supporting Work Quality Using K-Means Algorithm," *IJEET Int J Eng Emerg Technol*, vol. 3, no. 1, pp. 52–55, 2018.
[8] S. J. Russell and P. Norvig, *Artificial intelligence a modern approach*. London, 2010.
[9] D. Maulani, "Implementasi Algoritma K-Means Clustering Dalam Menentukan Berat Badan Ideal," *Implementasi Algoritma K-Means Clust. Dalam Menentukan Berat Badan Ideal*, 2013.
[10] D. Dutta, P. Dutta, and J. Sil, "Data clustering with mixed features by multi objective genetic algorithm," presented at the 2012 12th International Conference on Hybrid Intelligent Systems (HIS), IEEE, 2012, pp. 336–341.
[11] P. Govender and V. Sivakumar, "Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019)," *Atmospheric Pollut. Res.*, vol. 11, no. 1, pp. 40–56, 2020.
[12] M. Hariyanto and R. T. Shita, "Clustering pada Data Mining untuk Mengetahui Potensi Penyebaran Penyakit DBD Menggunakan Metode Algoritma K-Means dan Metode Perhitungan Jarak Euclidean Distance," *SKANIKA Sist. Komput. Dan Tek. Inform.*, vol. 1, no. 1, pp. 117–122, 2018.
[13] Joint National Committee on Prevention, Treatment of High Blood Pressure, and National High Blood Pressure Education Program, *Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure*, vol. 6. Public Health Service, National Institutes of Health, National Heart, Lung …, 1997.
[14] G. S. Stergiou *et al.*, "A universal standard for the validation of blood pressure measuring devices: Association for the Advancement of Medical Instrumentation/European Society of Hypertension/International Organization for Standardization (AAMI/ESH/ISO) Collaboration Statement," *Hypertension*, vol. 71, no. 3, pp. 368–374, 2018.

[15] L. D. Y. Pramana, "Faktor-Faktor Yang Berhubungan Dengan Tingkat Hipertensi Di Wilayah Kerja Puskesmas Demak II," 2016.
[16] N. F. Hastuti, "Pemanfaatan metode k-means clustering dalam penentuan penerima beasiswa," 2013.
[17] B. Ulum, "Penerapan Metode Clustering Dengan Algoritma K-Means Pada Pengelompokan Jurusan Data Calon Siswa Baru Di Smk Al-Ishlah Cikarang Utara," *J. Transform. Mandalika Jtm e-issn 2745-5882 p-issn 2962-2956*, vol. 3, no. 2, pp. 5–17, 2022.
[18] Y. Permatasari, "Klasifikasi Risiko Bahaya Kehamilan dengan Metode Fuzzy C-Means," 2012.
[19] A. D. Wahyudi, "Sistem Informasi Inventory Gudang Dalam Ketersediaan Stok Barang Menggunakan Metode Buffer Stok," *J. Inform. Dan Rekayasa Perangkat Lunak*, vol. 1, no. 2, pp. 174–182, 2020.
[20] D. Beimel and E. Kedmi-Shahar, "Improving the identification of functional system requirements when novice analysts create use case diagrams: the benefits of applying conceptual mental models," *Requir. Eng.*, vol. 24, pp. 483–502, 2019.
[21] H. Hartatik, F. Alfiana, and S. Wulandari, "Designing A Web-Based Activity Management System For Enhanced Data Analytics Performance," *Int. J. Multidiscip. Res. Lit.*, vol. 2, no. 3, pp. 306–315, 2023.
[22] R. Band *et al.*, "Intervention planning for a digital intervention for self-management of hypertension: a theory-, evidence-and person-based approach," *Implement. Sci.*, vol. 12, pp. 1–13, 2017.
[23] X. YANG, L. David, X. XIA, and J. SUN, "Condensing class diagrams with minimal manual labeling cost.(2016). COMPSAC 2016: Proceedings of the 40th IEEE Annual International Computers," presented at the Software and Applications Conference, Atlanta, Georgia, 2016, pp. 10–14.
[24] P. Sapna and H. Mohanty, "Clustering test cases to achieve effective test selection," in *Proceedings of the 1st Amrita ACM-W Celebration on Women in Computing in India*, 2010, pp. 1–8.
[25] A. I. Nur *et al.*, "User scenario 2 seq: Automated system to generate sequence diagram from user stories scenario," presented at the AIP Conference Proceedings, AIP Publishing, 2023.
[26] C. Krittanawong, A. S. Bomback, U. Baber, S. Bangalore, F. H. Messerli, and W. Wilson Tang, "Future direction for using artificial intelligence to predict and manage hypertension," *Curr. Hypertens. Rep.*, vol. 20, pp. 1–16, 2018.
[27] J. Zhang, W. Chen, M. Gao, and G. Shen, "K-means-clustering-based fiber nonlinearity equalization techniques for 64-QAM coherent optical communication system," *Opt. Express*, vol. 25, no. 22, pp. 27570–27580, 2017.
[28] N. F. FAHRUDIN, "Sequence clustering in process mining for business process analysis using k-means," *MIND Multimed. Artif. Intell. Netw. Database J.*, vol. 6, no. 1, pp. 16–30, 2021.
[29] N. A. A. Nasir, "Proceeding Of International Conference On Post-Covid Healthcare, Medical Research And Education 2022 (Ps1-S102) & Proceeding Of 35th Annual Scientific Meeting Of The Malaysian Society Of Pharmacology And Physiology (MSPP 2022)(pS103-S199)," *J. Clin. Health Sci.*, vol. 7, no. 2 (S), pp. 1–199, 2022.
[30] D. Maulani, "Implementasi Algoritma K-Means Clustering Dalam Menentukan Berat Badan Ideal," *Implementasi Algoritma K-Means Clust. Dalam Menentukan Berat Badan Ideal*, 2013.