JUSIKOM PRIMA (Jurnal Sistem Informasi dan Ilmu Komputer Prima)
Vol. 8 No. 1, August 2024 E-ISSN:2580-2879

Grouping Diseases of Patients at RSU Mitra Medika Bandar Khalippa Medan Using the K-Medoids Clustering Method

Ajeng Kiana Putri[1], Yusuf Ramadhan Nasution[2]
[1,2]North Sumatra State Islamic University, Medan, Indonesia
Email: ajengkyanaputri@gmail.com1 , ramadhannst@gmail.com2

ABSTRACT
The aim of this research is to apply the method*K-Medois*in categorizing the illnesses of patients at the RSUBandar Khalippa Medika PartnersMedan. And to produce a system for grouping patient data based on Rapidminer and Google Colab on patient diseases.Based on the results of research on the application of the K-Medoids algorithm, it was found that the grouping of patient diseases at RSU Mitra Medika Bandar Khalippa used the RapidMiner application with a C0 (High) cluster of 3 diseases, a C1 (Medium) cluster of 6 diseases and a C2 (Low) cluster of 1 disease. Meanwhile, using the Google Colabs application with a C0 (High) cluster of 3 diseases, a C1 (Medium) cluster of 4 diseases and a C2 (Low) cluster of 3 diseases. The results of grouping patient disease data at RSU Mitra Medika Bandar Khalippa using RapidMiner, it was found that the disease with the highest grouping (C0)is a disease*Pulmonary tuberculosis*, Essential Hypertension and Diabetes Mellitus. Whereasgrouping patient disease datawith Google Colabs it was found that the disease with the highest grouping (C0)is a disease*Bronchus Or Lung*, Trachea, Bronchus And Lung and Pleural Effusion.

Keywords: Disease Grouping, RSU, Methods*K-Medois*.

INTRODUCTION

RSU Mitra Medika Bandar Khalippa Medan, located in Medan City, is one of the hospitals under the Ministry of Health whose function is to provide services to the community. Currently, RSU Mitra Medika Bandar Khalippa Medan, in storing patient medical record data, only sorts the diseases most commonly suffered and does not pay attention to addresses, types of participants and gender. Therefore, RSU Mitra Medika Bandar Khalippa Medan needs disease data processing to find out information about diseases that patients often suffer from. With this information, RSU Mitra Medika Bandar Khalippa Medan can take policy action in anticipating treatment and disease prevention by conducting outreach, providing additional doctors and being able to equip RSU facilities according to patient disease clustering.

To produce accurate information, medical record data needs to be processed using data mining. Data mining provides several techniques for producing important information from piles of different data such as clustering techniques. Clustering in data mining is a method of dividing data into groups that have objects with the same characteristics(Ginting et al., 2021). One method that can be used for data mining clustering is K-Medoids. K-Medoids is a partition

algorithm, because it uses the most centralized objects (medoids) in a cluster to become the cluster center of the average value of objects in a cluster. The k-medoids method is more suitable for grouping data(Veronika et al., 2023).

Based on the problems above, this research applies a well-known data mining technique, namely clustering using the k-medoids method to group diseases of RSU patients.Bandar Khalippa Medika PartnersMedan. By knowing the clustering condition of the patient's disease, the RSUBandar Khalippa Medika PartnersMedan is expected to be able to improve services by taking policy action in anticipating treatment, prevention and adding facilities as well as adding doctors according to the patient's illness.

LITERATURE REVIEW

The aim of previous research was to collect reference and comparative data. Additionally, do not draw parallels between this research and other research. Therefore, the researcher used the following findings from previous research in this previous study:The results of research into the grouping of the spread of the new Covid-19 show that people come from various regions in Indonesia. Characteristics of a body temperature above 36.9◦c and accompanied by fever and continuous cough shows one of the characteristic symptoms of Covid-19 (Sindi et al., 2020). Based on the results of this research, it can be concluded that with the K-Medoids method, 34 provinces were obtained from the grouping data, resulting in 21 provinces, namely the low cluster, 12 provinces in the medium cluster, and 1 province in the high cluster from the percentage of allergy immunization in each province (Ningrum et al. ., 2021).

Data Mining

*Data Mining*is an analytical process designed to examine large amounts of data in search of valuable and consistent hidden knowledge(Pulungan et al., 2019). Data Mining is used to search for knowledge contained in large databases so it is often called Knowledge Discovery Database (KDD), namely the stages carried out in extracting knowledge from a collection of data ((Marisa, 2018). The process of extracting hidden information in a large database often uses the terms datamining and knowledge discovery in databases (KDD).

K-Medoids

K-Medoids Clustering known as Partitioning Around Medoids (PAM), is a variant of the k-means method, using the partition clustering method to group a set of n objects into k

clusters(Nahdliyah et al., 2019). This algorithm uses objects in a collection of objects to represent a cluster called medoids.

Rapid Miner

Rapid Miner is open source software. Rapid Miner is a solution for analyzing data mining, text mining and predictive analysis(Luluk Elvitaria, 2017). Rapid Miner uses various descriptive and predictive techniques to provide insights to users so they can make the best decisions.

Google Colab

Google Colab, or Google Colaboratory, is a cloud-based platform that allows users to write and run Python code in the browser. This platform is very popular among data scientists, researchers, and developers because it provides an easy-to-access programming environment without the need for local configuration. Google Colab provides access to free CPU, GPU, and TPU, which makes it very useful for computationally demanding tasks, such as training machine learning and deep learning models(Soen et al., 2022).

METHODS

The research determined the research location at RSU Mitra Medika Bandar Khalippa Medan Jl. Medan - Batang Kuis, Hamlet XI Emplasmen, Bandar Klippa Village, District. Percut Sei Tuan, Deli Serdang Regency, North Sumatra. Research methodology is useful so that the implementation of research can run well and systematically and meet the desired objectives. This research uses the rapidminer application to carry out a data mining process using the K-Medoids method for clustering patient illnesses at RSU Mitra Medika Bandar Khalippa Medan. In the process, the research procedures used were the following stages: Literature Study, Interviews, Data Collection, System Work Design, Application Testing.The following is the research flow as in Figure 1.

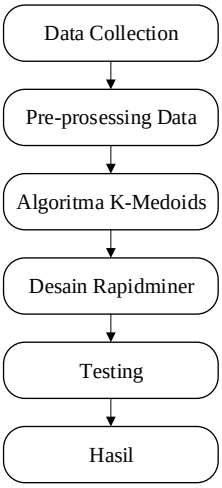

Figure 1. Research Flow Chart

The steps that will be used in this research can be seen in Figure 2.

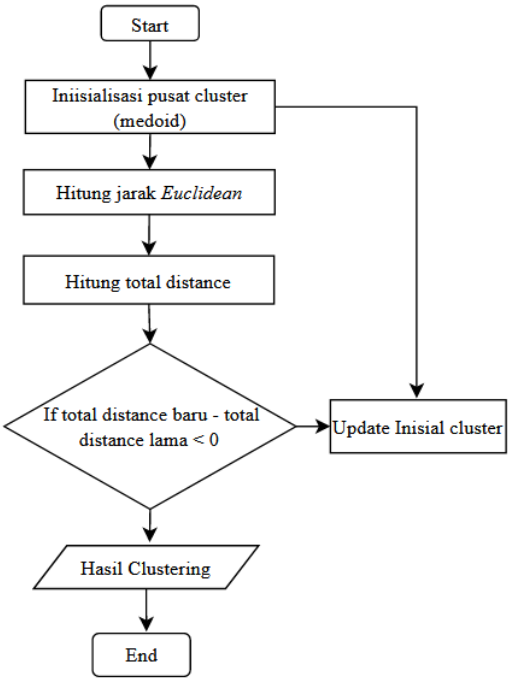

Figure 2. *K-Medoids Flowchart*

RESULTS

Data analysis

The data used in grouping patients' diseases at RSU Mitra Medika Bandar Khalippa using the K-Medoids clustering method is pre-existing data with High, Medium and Low disease types. This aims to make it easier to obtain information that is used to classify patient illnesses. The following is data on diseases and their numbersin January 2023 to December 2023 as in Table 1.

Table 1. Patient Disease DataRSU Mitra Medika Bandar Khalippa

No	Disease Name	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
1	*Neumonia*	64	36	67	37	89	67	37	29	21	32	29	21
2	*Chronic Obstructive Pulmonary*	47	24	44	35	76	80	35	31	12	35	31	12
3	*Asthma*	42	17	56	61	55	44	33	52	14	51	52	14
...	...	...	...	...	...	...	...	...	...	...	...	...	...
10	*Diabetes Mellitus*	42	29	77	73	25	56	73	31	26	57	31	26

Implementation of the K-Medoids Algorithm

In this section, we will explain how to implement the K-Medoids algorithm on data on the diseases most frequently experienced by patients obtained from interviews with parties.RSU Mitra Medika Bandar Khalippa. Before calculating, the results of the K-Medoids method are to determine the grouping of the patient's disease. The types of grouping used in this research are in Table 2.

Table 2. Patient Disease Grouping

No	Clusters	Grouping Type
1	C0	Tall
2	C1	Currently
3	C2	Low

The solution steps using the K-Medoids algorithm include the following:

1. Initialize as many cluster centers as possible from sample data. For selection, each medoid is chosen randomly.*Bronchus Or Lung*,*Chronic Obstructive Pulmonary*And*Essential Hypertension*as the initial medoid as in Table 3.

Table 3. Initial Medoid

Disease Name	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
Bronchus Or Lung	36	31	55	29	49	45	29	61	28	63	61	28
Chronic Obstructive Pulmonary	47	24	44	35	76	80	35	31	12	35	31	12
Essential Hypertension	25	37	67	28	38	10	28	40	34	48	40	34

2. Calculate the value of the closest distance (cost) using the Euclidian Distance equation, namely equation (1). So the calculation to calculate the distance of each object to the initial medoid is as follows:

$D_{\text{Nuemonia, C0}}=\sqrt{}(64\text{-}36)2+(36\text{-}31)2+(67\text{-}55)2+(37\text{-}29)2+(89\text{-}49)2+(67\text{-}45)2+(37\text{-}29)2+(29\text{-}61)2+(21\text{-}28)2+(32\text{-}63)2+(29\text{-}61)2+(21\text{-}28)2=79.19596$

$D_{\text{Nuemonia, C1}}$=√(64-47)2+(36-24)2+(67-44)2+(37-35)2+(89-76)2+(67-80)2 +(37-35)2 +(29-31)2 +(21-12)2 +(32-35)2 +(29-31)2 +(21-12)2 =38.56164

$D_{\text{Nuemonia, C2}}$=√(64-25)2+(36-37)2+(67-67)2+(37-28)2+(89-38)2+(67-10)2 +(37-28)2 +(29-40)2 +(21-34)2 +(32-48)2 +(29-40)2 +(21-34)2 =93.7763296

... .

$D_{\text{Diabetes, C0}}$=√(42-25)2+(29-37)2+(77-67)2+(73-28)2+(25-38)2+(56-10)2 +(73-28)2 +(31-40)2 +(26-34)2 +(57-48)2 +(31-40)2 +(26-34)2 =84.0892383

The overall results can be seen in Table 4:

Table 4. 1st Iteration K-Medoids Algorithm Calculation Results

Disease Name	C0	C1	C2	Proximity
Neumonia	79.19596	38.56164	93.7763296	38.561639
Chronic Obstructive Pulmonary	73.57309	0	96.1405222	0
Asthma	44.81071	61.48984	67.955868	44.810713
Bronchus Or Lung	0	73.57309	54.1479455	0
Trachea, Bronchus And Lung	83.15047	100.2547	129.799846	83.150466
Pleural Effusion	64.93073	66.89544	87.8293801	64.930732
Pulmonary tuberculosis	75.52483	61.7333	86.4291617	61.733297
Congestive Heart Failure	93.31667	106.7286	116.137849	93.316665
Essential Hypertension	53.17894	94.46163	4	4
Diabetes Mellitus	83.28865	89.88882	84.0892383	83.288655
Number of Proximities				**473.79217**

After obtaining the distance results from each object in the 1st iteration, proceed to the 2nd iteration. The new medoid can be seen in Table 5.

Table 5. New Medoid 2nd Iteration

Disease Name	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
Diabetes Mellitus	42	29	77	73	25	56	73	31	26	57	31	26
Pleural Effusion	32	41	54	51	98	56	51	49	38	52	49	38
Congestive Heart Failure	61	61	45	75	56	76	75	57	58	60	57	58

Recalculate the distance from each object in the 2nd Iteration with the new medoid. The same process is carried out in the second iteration. So the results obtained from the entire 2nd iteration can be seen in table 6:

Table 6. 2nd Iteration K-Medoids Algorithm Calculation Results

Disease Name	C0	C1	C2	Proximity
Neumonia	90.13878	59.91661	100.687636	59.916609

Disease Name	C0	C1	C2	Proximity
Chronic Obstructive Pulmonary	89.88882	66.89544	106.371989	66.895441
Asthma	67.66831	65.32993	97.0257698	65.329932
Bronchus Or Lung	83.28865	64.93073	93.4879671	64.930732
Trachea, Bronchus And Lung	133,832	92.21714	107.126094	92.217135
Pleural Effusion	89.61585	0	73.9324016	0
Pulmonary tuberculosis	54.79964	83.04216	88.3062852	54.799635
Congestive Heart Failure	84.77618	75.03333	2	2
Essential Hypertension	84.61087	84.95881	115.775645	84.610874
Diabetes Mellitus	0	89.61585	85.5277733	0
Number of Proximities				**490.70036**

3. Calculate Total Deviation (S)

   After getting the distance value between the 1st iteration and the 2nd iteration, calculate the total deviation (S) by calculating the new total distance value - the old total distance value. If S<0, then swap the objects with the cluster data to form a new set of k objects as medoids.

   S = Total*cost*new – Old total cost =**490.70036**–**473.79217**=**16.90819**

   Because the S value > 0, the iteration process is stopped. So that the members of each cluster are obtained in Table 7.

Table 7. Clustering Results with K-Medoids

No	Disease Name	C0	C1	C2	Proximity	Clutser
1	*Neumonia*	90.13878	59.91661	100.687636	59.916609	C1
2	*Chronic Obstructive Pulmonary*	89.88882	66.89544	106.371989	66.895441	C1
3	*Asthma*	67.66831	65.32993	97.0257698	65.329932	C1
...	...	...	...	...	...	...
10	*Diabetes Mellitus*	0	89.61585	85.5277733	0	C0

Results of Applying the K-Medoids Method

After searching for the closest cluster value, the following are the results of grouping the patient's diseaseRSU Mitra Medika Bandar Khalippabased on 10 disease data using the K-Medoids method as in Table 8.

Table 8. Patient Disease Grouping

No	Disease Name	Clusters	Grouping Type
1	*Neumonia*	C1	Currently
2	*Chronic Obstructive Pulmonary*	C1	Currently
3	*Asthma*	C1	Currently
...	...	...	...
10	*Diabetes Mellitus*	C0	Tall

In Table 8, the grouping of patients' illnesses into high, medium and low categories is explained. The high category has 4 types of disease, the medium category has 3 types of disease

and the low category has 3 types of disease. Based on grouping using the K-Medoids method, it was found that the diseases most often faced by partiesRSU Mitra Medika Bandar Khalippain the period January 2023 to December 2023 is a disease*Pulmonary tuberculosis*, Essential Hypertension and Diabetes Mellitus. Based on the results of the grouping, partiesRSU Mitra Medika Bandar Khalippacan prioritize adding facilities or doctors who treat patients with High or Medium category illnesses.

System Implementation

Based on the analysis in the previous chapter, an implementation and testing process is needed to apply and test the correctness of the results of data processing that has been done manually. For the implementation and testing process you can use one of the Rapidminer application software version 9.8.1 using patient disease data at the RSU Khalippa Bandar Medika Partners. Next is to test the system with rapidminer with the following steps.

1. Open the Rapid Miner Application

   Here is what rapidminer looks like when it is first opened:

Figure 3. Flash Rapidminer display

After the rapidminer flash display then the main menu of rapidminer 9.8.1 appears as follows:

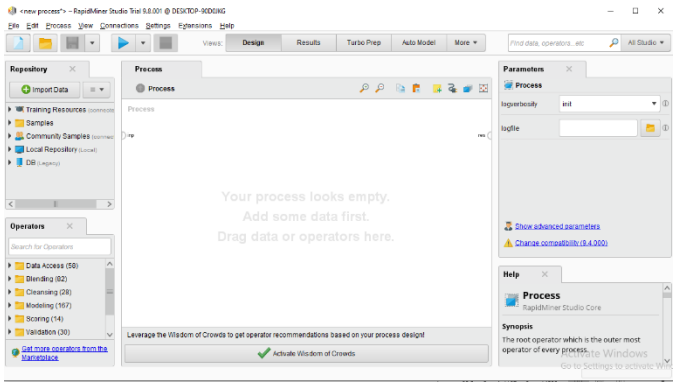

Figure 4. Main Menu Display

2. Data Collection Process

After the menu appears, the next step is to insert the EXEL patient disease data which has been prepared in the following way:

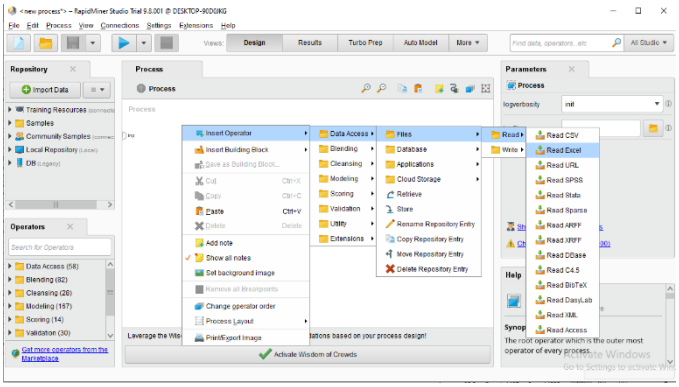

Figure 5. Process of Inserting Patient Disease Data

Based on the image above, inserting patient disease data can be done by right clicking on the rapidminer worksheet then selecting "Insert Operator", "Data Access", Files", "Read" and "Read Excel", so that the following menu appears:

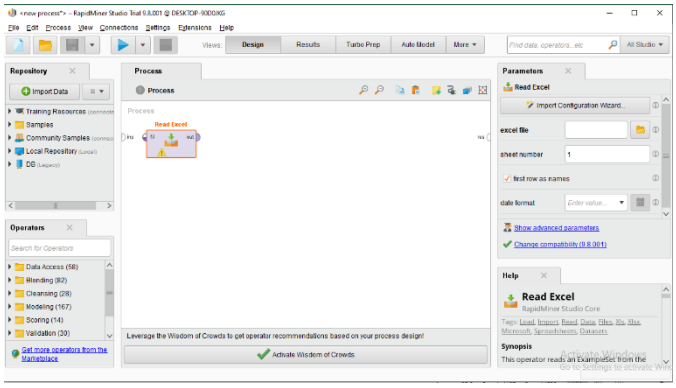

Figure 6. Process State Display After Read Excel

Based on the image above, then select the Excel patient disease data by selecting the "Import Configuration Wizard" button so that the search form displays as follows:

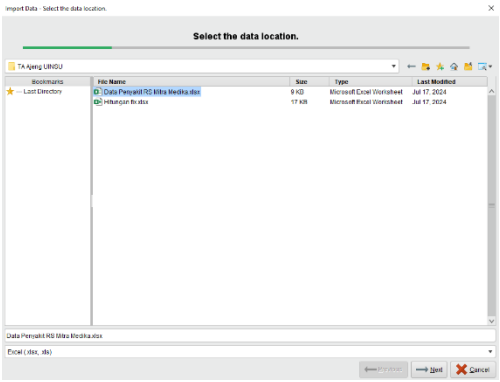

Figure 7. Data Selection Menu

Based on the image above, the next step is to select the Excel patient disease data, then select "Next" so that it appears as follows:

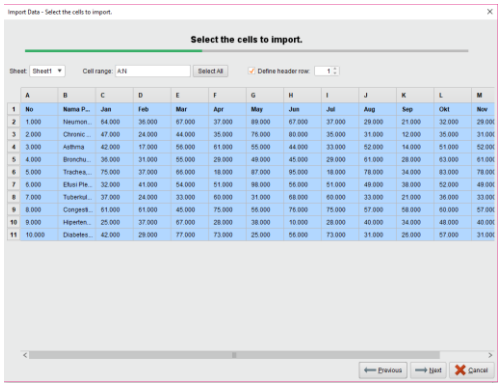

Figure 8. Contents of the Excel File with Data Selection Results

Based on the image above, the application displays the patient's disease data in Excel, then select "Next" so that it appears as follows:

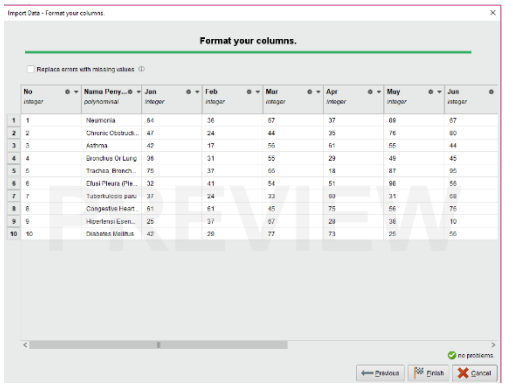

Figure 9. Patient Disease Data Column Format

Based on the image above, then make changes to the columns where the rule number column is used as "ID" and the disease name column becomes "Label", this is because the number and disease name columns are not calculated, so the appearance is as follows:

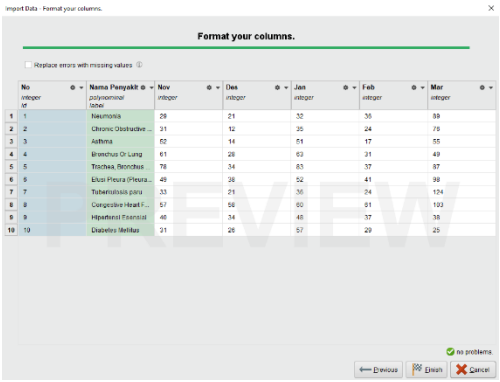

Figure 10. Completed Patient Disease Data Column Format

Based on the image above, the next step is to press the "Finish" button so that the following menu appears:

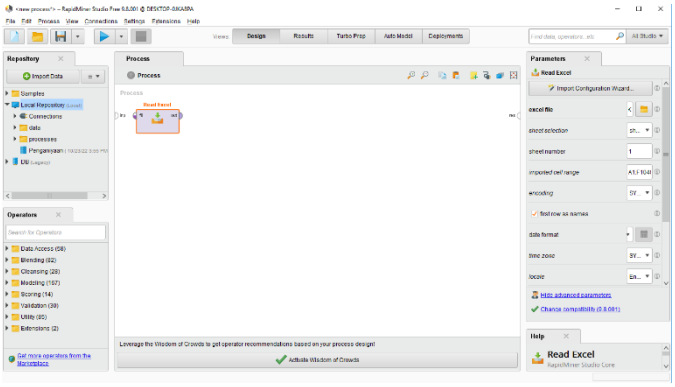

Figure 11. Selection Results Patient Disease Data

3. K-Medoids Method Selection Process

Based on the results of the data selection, the data is then displayed on a worksheet, the next step is to select the method used for the process of grouping patient illnesses at RSU Mitra Medika Bandar Khalippa based on the K-Medoids method as follows:

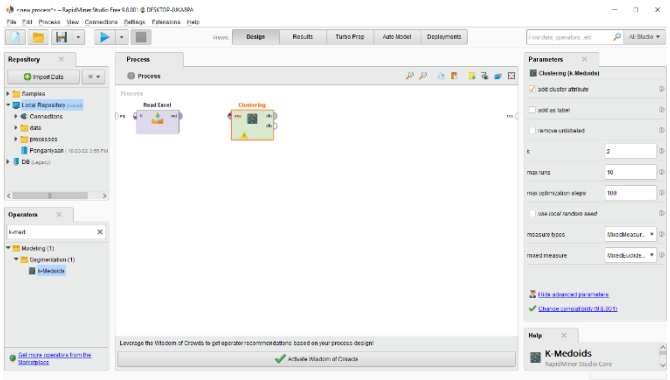

Figure 12. K-Medoids Operator Selection Results

Based on the image above, the process of grouping patient diseases is carried out using the K-Medoids method by selecting the "Operator" modeling tab then Segmentation, k-

medoids. Next, connect the Excel data with the operator data and determine the value K = 3 so that the results are as in Figure 13.

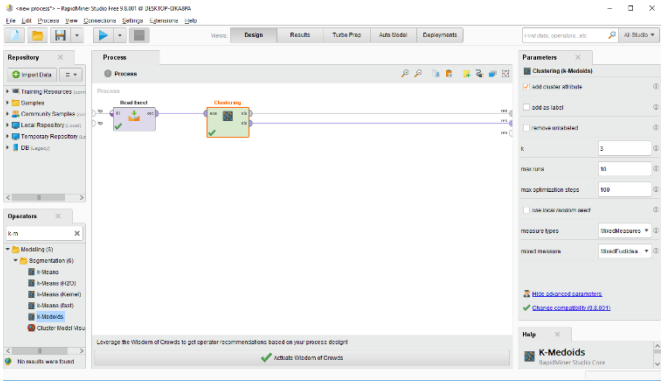

Figure 13. Results of Linking Disease Data and K-Medoids Operators

Based on Figure 13, patient disease data is linked to all k-medoids operators with a value of K=3, where the K value is a type of cluster consisting of High, Medium and Low groupings.

Rapid Miner System Testing

After all the data has been input and connected to the k-medoids operator, the next step is to carry out testing to obtain information on the grouping of patient diseases at RSU Mitra Medika Bandar Khalippa based on data from January 2023 to December 2023 with 3 clusters. Based on Figure 13, then select "Play" to produce clustering as follows:

Open in Turbo Prep Auto Model

...	No	Nama Penyakit	cluster	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Okt	Nov	Des
1	1	Neumonia	cluster_1	64	36	67	37	89	67	37	29	21	32	29	21
2	2	Chronic Obstructive Pulmonary	cluster_1	47	24	44	35	76	80	35	31	12	35	31	12
3	3	Asthma	cluster_1	42	17	56	61	55	44	33	52	14	51	52	14
4	4	Bronchus Or Lung	cluster_1	36	31	55	29	49	45	29	61	28	63	61	28
5	5	Trachea, Bronchus And Lung	cluster_1	75	37	66	18	87	95	18	78	34	83	78	34
6	6	Efusi Pleura (Pleural Effusion)	cluster_1	32	41	54	51	98	56	51	49	38	52	49	38
7	7	Tuberkulosis paru	cluster_0	37	24	33	60	31	68	60	33	21	36	33	21
8	8	Congestive Heart Failure	cluster_2	61	61	45	75	56	76	75	57	58	60	57	58
9	9	Hipertensi Esensial	cluster_0	25	37	67	28	38	10	28	40	34	48	40	34
10	10	Diabetes Mellitus	cluster_0	42	29	77	73	25	56	73	31	26	57	31	26

Figure 14. Results of Clustering Patient Disease Data

Based on the results of patient disease clustering testing using 10 patient disease data at Mitra Medika Bandar Khalippa RSU based on data from January 2023 to December 2023, it was found that cluster 0 grouping contained 3 diseases, cluster 6 grouping contained 3 diseases and cluster 2 grouping contained 3 diseases. as many as 1 disease. Based on the results of the process, a cluster model and graph are obtained as in Figure 15.

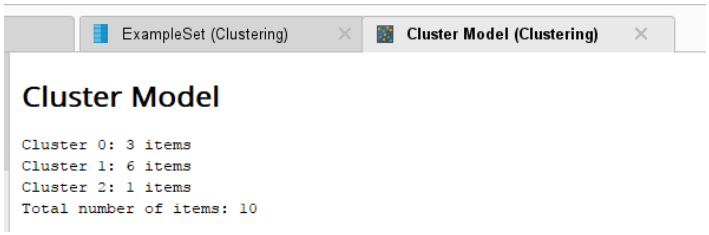

Figure 15. Clustering Model for Patient Disease Data

Based on Figure 15, it is known that in the cluster model, cluster 0 consists of 4 items, cluster 1 consists of 3 items and cluster 2 consists of 3 disease items. The following is a graph of the results of clustering patient disease data as in Figure 16.

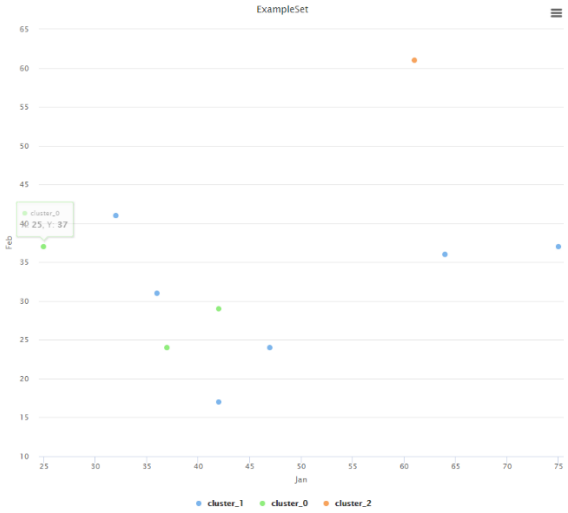

Figure 16. Graph of Clustering Results of Patient Disease Data

Based on Figure 16 graph, the attributes used are the name of the disease and the number of diseases. The graph is a description of the clustering results based on the attributes of the dataset on the number of diseases.

Google Colabs System Testing

After getting the clustering results in Rapidminer, then test the clustering results with Google Colabs using the Python programming language.

1. *Import Libraries*

   To install the necessary libraries such as scikit-learn and pandas for data processing, and pyclustering for the K-Medoids algorithm. The results on Google Colabs are as in Figure 17.

```
import pandas as pd
from pyclustering.cluster.kmedoids import kmedoids
from pyclustering.cluster import cluster_visualizer
from pyclustering.utils import read_sample
from pyclustering.samples.definitions import SIMPLE_SAMPLES
from sklearn.metrics import pairwise_distances
```

Figure 17. Import Python Library

2. Upload and Read Disease Datasets

The next step is to upload and read the disease dataset that has been provided in Excel format as in Figure 18.

```
[14] from google.colab import drive
     drive.mount('/content/drive')
     file_path = '/content/drive/My Drive/data_penyakit.csv'

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

data = pd.read_csv(file_path)
print(data)

   No                  Nama Penyakit  Jan  Feb  Mar  Apr  May  Jun  Jul  \
0   1                       Neumonia   64   36   67   37   89   67   37
1   2  Chronic Obstructive Pulmonary   47   24   44   35   76   80   35
2   3                         Asthma   42   17   56   61   55   44   33
3   4                Bronchus Or Lung   36   31   55   29   49   45   29
4   5      Trachea, Bronchus And Lung   75   37   66   18   87   95   18
5   6  Efusi Pleura (Pleural Effusion)   32   41   54   51   98   56   51
6   7                Tuberkulosis paru   37   24   33   60   31   68   60
7   8        Congestive Heart Failure   61   61   45   75   56   76   75
8   9             Hipertensi Esensial   25   37   67   28   38   10   28
9  10                Diabetes Mellitus   42   29   77   73   25   56   73

   Aug  Sep  Okt  Nov  Des
0   29   21   32   29   21
1   31   12   35   31   12
2   52   14   51   52   14
3   61   28   63   61   28
4   78   34   83   78   34
5   49   38   52   49   38
6   33   21   36   33   21
7   57   58   60   57   58
8   40   34   48   40   34
9   31   26   57   31   26
```

Figure 18. Upload and Read Disease Dataset

3. Clustering Disease Data

The next step is to cluster disease data with K-Medoids based on the command as in Figure 19.

```
X = df.iloc[:, :1].values
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)
kmedoids = KMedoids(n_clusters=3, random_state=42)
kmedoids.fit(X_scaled)
df['Cluster'] = kmedoids.labels_
print(df)
```

Figure 19. K-Medoids Clustering Command

Based on Figure 19, the results of patient disease clustering were found using 10 patient disease data at RSU Mitra Medika Bandar Khalippa based on data from January 2023 to December 2023 as in Figure 20.

	No	Nama Penyakit	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Okt	Nov	Des	Cluster
0	1	Neumonia	64	36	67	37	89	67	37	29	21	32	29	21	2
1	2	Chronic Obstructive Pulmonary	47	24	44	35	76	80	35	31	12	35	31	12	2
2	3	Asthma	42	17	56	61	55	44	33	52	14	51	52	14	2
3	4	Bronchus Or Lung	36	31	55	29	49	45	29	61	28	63	61	28	0
4	5	Trachea, Bronchus And Lung	75	37	66	18	87	95	18	78	34	83	78	34	0
5	6	Efusi Pleura (Pleural Effusion)	32	41	54	51	98	56	51	49	38	52	49	38	0
6	7	Tuberkulosis paru	37	24	33	60	31	68	60	33	21	36	33	21	1
7	8	Congestive Heart Failure	61	61	45	75	56	76	75	57	58	60	57	58	1
8	9	Hipertensi Esensial	25	37	67	28	38	10	28	40	34	48	40	34	1
9	10	Diabetes Mellitus	42	29	77	73	25	56	73	31	26	57	31	26	1

Figure 20. Results of K-Medoids Clustering in Python

Based on Figure 20, it is known that in the cluster model, cluster 0 consists of 3 items, cluster 1 consists of 4 items and cluster 2 consists of 3 disease items.

System Test Results

After testing the patient's disease clustering onRSU Mitra Medika Bandar Khalippa uses Rapidminer and Google Colabs, the following are the comparison results of grouping patient diseases using the K-Medoids method as in Table 9.

Table 9. Test Results

No	Disease Name	*RapidMiner*		*Google Colabs*	
		Clusters	Classification	Clusters	Classification
1	*Neumonia*	*cluster_1*	Currently	*cluster_2*	Low
2	*Chronic Obstructive Pulmonary*	*cluster_1*	Currently	*cluster_2*	Low
3	*Asthma*	*cluster_1*	Currently	*cluster_2*	Low
...	...	...	...	....	....
10	*Diabetes Mellitus*	*cluster_0*	Tall	*cluster_1*	Currently

Based on table 9, grouping of patient diseases inRSU Mitra Medika Bandar Khalippa, expectedpartyRSU Mitra Medika Bandar Khalippacan prioritize adding facilities or doctors who treat patients with High or Medium category illnesses. Identify disease patterns and trends over time. Plan preventative actions and allocate resources more effectively. Determine the need for medical resources such as medical personnel, medicines and medical devices. Optimizing the use of resources and ensuring sufficient availability to deal with the most common diseases.

CONCLUSION

Based on the results of the discussion, the results of this research can be concluded that Based on the application of the K-Medoids algorithm, it was found that the grouping of patient diseases at RSU Mitra Medika Bandar Khalippa used the RapidMiner application with the number of C0 (High) clusters being 3 diseases, C1 (Medium) clusters being 3 diseases. 6 diseases and cluster C2 (Low) with 1 disease. Meanwhile, using the Google Colabs application with a C0 (High) cluster of 3 diseases, a C1 (Medium) cluster of 4 diseases and a C2 (Low) cluster of 3 diseases. The results of grouping patient disease data at RSU Mitra Medika Bandar Khalippa using RapidMiner, it was found that the disease with the highest grouping (C0)is a disease*Pulmonary tuberculosis*, Essential Hypertension and Diabetes Mellitus. Whereasgrouping patient disease datawith Google Colabs it was found that the disease with

the highest grouping (C0)is a disease*Bronchus Or Lung*, Trachea, Bronchus And Lung and Pleural Effusion.

REFERENCES

Andini, AD, & Arifin, T. (2020). Implementation of the k-medoids algorithm for clustering patient disease data in Bandung City Hospital. Responsive Journal: Science and Informatics Research, 2(2), 128-138.

Cahyana, N., & Aribowo, A. (2018). K-Means Data Mining Method for Clustering Public Health Treatment and Service Data. National Seminar on Medical Informatics, 5, 24–31.

Luluk Elvitaria, MH (2017). Predicting the Level of Extracurricular Interest in Abdurrab Health Analysis Vocational School Students Using the C4.5 Algorithm (Case Study: Abdurrab Health Analysis Vocational School). (University Journal of Technology and Information Systems, 2(2), 220–233.

Marisa, F. (2018). Educational Data Mining (Concepts and Applications). Information Technology, 4(2), 90–97. http://ejurnal.stimata.ac.id/index.php/TI/article/view/108/148

Mustofa, Z., & Mood, IS (2018). K-Medoids Clustering Algorithm in E-Government in the Field of Information and Communication. Journal of Technology and Communication, 9, 1–10.

Nahdliyah, MA, Widiharih, T., & Prahutama, A. (2019). k-MEDOIDS CLUSTERING METHOD WITH VALIDATION OF SILHOUETTE INDEX AND C-INDEX (Case Study of the Number of Regency/City Crimes in Central Java in 2018). Gaussian Journal, 8(2), 161–170. https://doi.org/10.14710/j.gauss.v8i2.26640

Pulungan, N., Suhada, S., & Suhendro, D. (2019). Application of the K-Medoids Algorithm to Group Population 15 Years and Over According to Main Job Field. COMIC (National Conference on Information and Computer Technology), 3(1), 329–334. https://doi.org/10.30865/komik.v3i1.1609

Rosa, & Saladin. (2019). Software engineering: structured and object-oriented. Informatics.

Soen, GIE, Marlina, M., & Renny, R. (2022). Implementation of Cloud Computing with Google Collaboratory in the Zoom Participants Data Processing Application. JITU: Journal of Informatic Technology and Communication, 6(1), 24–30. https://doi.org/10.36596/jitu.v6i1.781