Journal of Information Systems and Informatics
Vol. 6, No. 1, March 2024 e-ISSN: 2656-4882 p-ISSN: 2656-5935
DOI: 10.51519/journalisi.v6i1.653 Published By DRPM-UBD

Assessing Employee Performance in the Information Technology Department Using K-Means Clustering: A Case Study Approach

Afelia Jozalin Elisa Sakalessy[1], Hindriyanto Dwi Purnomo[2]

[1,2]Informatics, Faculty of Information Technology, Satya Wacana Christian University, Diponegoro Street No. 52 – 60, Salatiga, Central Java, Indonesia 50711
Email: [1]672018103@student.uksw.edu, [2]hindriyanto.purnomo@uksw.edu

Abstract

This research explores the utilization of the K-Means method for evaluating the performance of employees within the Information Technology Department at PT Nusa Halmahera Minerals. The research leverages data from Manage Engine to analyze diverse parameters of employee performance. By employing the K-Means method, employees are categorized based on specific characteristics, facilitating a deeper understanding of each individual's contribution and potential in attaining company objectives. The implementation of the K-Means method aims to offer a more objective perspective on employee performance, empowering companies to make informed decisions in the enhancement and development of their human resources.

Keywords: K-Means Method, Employee Performance, Data Analysis, Human Resource Management, Performance Evaluation, Company.

1. INTRODUCTION

The rapid evolution of technology from year to year necessitates that all organizations embrace technology to streamline their operations. Information technology plays a pivotal role in enhancing company performance and efficiency by fostering innovative thinking among employees, leading to the development and advancement of new technologies. Beyond sophisticated technology, companies also require a skilled workforce to execute tasks effectively. Employee performance stands as a primary focus within any company, with each company aiming for exemplary performance from its employees to meet organizational targets. Employee performance entails the proficient execution of duties and responsibilities. Thus, VoOleh advocates for the regular evaluation and assessment of employee performance to gauge their professionalism and contributions accurately. Such evaluations serve as a basis for decisions related to salary increments, bonuses, career advancements, and employment terminations [1].

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

PT Nusa Halmahera Minerals (PT NHM) operates in the mining sector under a Contract of Work with the Indonesian Government signed on April 28, 1997. It manages the Gosowong Gold Mine situated in North Halmahera Regency, North Maluku Province, Eastern Indonesia. With several divisions catering to various aspects of its operations, PT NHM ensures the smooth functioning of its activities. Among these divisions is the Information Technology Department, where the author conducted research.

The research conducted by the author utilizes Manage Engine data to assess the performance of employees within PT NHM's Information Technology Department. Manage Engine data, derived from a system recording departmental activities, includes metrics such as completed tickets, time spent on ticket resolution, and ticket types. Prior to analysis using the K-Means algorithm, the Manage Engine data undergoes preprocessing in Microsoft Excel, employing the Knowledge Discovery in Database (KDD) process. The K-Means algorithm facilitates data grouping based on similarity, aiding in the evaluation of employee performance due to its ease of implementation, quick processing time, and testable accuracy values [2].

The study titled "Application of the K-Means Algorithm for Clustering Medicine Data at Pekanbaru District Hospital" explores drug usage patterns using variables such as drug name, unit, and quantity per month from January to December. Employing the Elbow Criterion method, the research determines the optimal number of clusters ($k = 3$). This research serves as a reference for the author, particularly due to its calculation of the k value using the Euclidean Distance formula and adoption of the Knowledge Discovery In Database (KDD) process for database processing. Moreover, the research employs cluster value testing, including the Davies Bouldin Index (DBI) [3].

Another study, "Employee Performance Analysis Using the K-Means Algorithm at the Central Bengkulu Regency Education and Culture Service," investigates employee performance based on various criteria. Similarly, it utilizes the K-Means method for clustering, albeit with different tools and methodologies. Employing the Waterfall research method and utilizing Data Flow Diagrams (DFD) and Entity Relationship Diagrams (ERD), the research aims to classify employee performance quality. Notably, the research incorporates cluster value testing using the Davies Bouldin Index (DBI) to ensure robustness [4].

In the study "Clustering Using the K-Means Method to Determine the Nutritional Status of Toddlers," toddler nutritional status is assessed using height and weight indicators via the K-Means method. This research, akin to the author's work, employs two variables for analysis. It employs the Euclidean Distance equation for distance calculation and includes a normalization process. Unlike the author's

research, this study compares results with Growth Chart data, yielding similar outcomes [5].

Lastly, the research "Implementation of Rapidminer Using the K-Means Method (Case Study: Measles Immunization for Toddlers Based on Province)" adopts k = 3 as a reference value and utilizes the Euclidean Distance equation for distance calculation. Focused on measles immunization for toddlers, the research aims to highlight provinces requiring heightened attention through cluster analysis. Considering the research's objectives, it would be beneficial to assess the accuracy of cluster values, such as using the Davies Bouldin Index (DBI) [6].

2. METHODS

Knowledge Discovery in Databases (KDD) is a methodology utilized for extracting insights from analyzed database outcomes. The results obtained can be used as basic knowledge (Knowledge Base) which can be used as a basis for making decisions [7]. The KDD process can be described as follows:

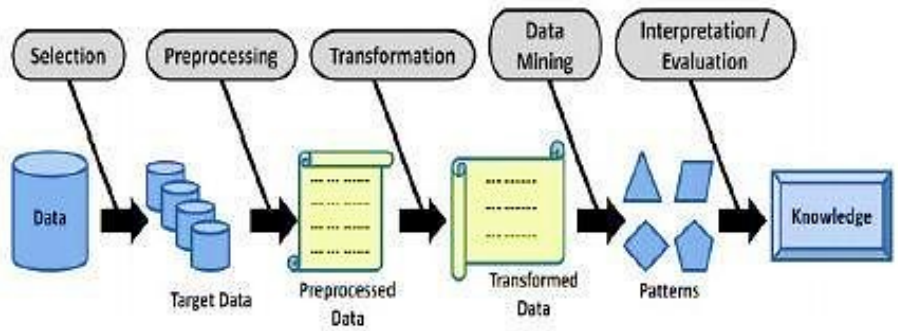

Figure 1. Knowledge Discovery Process in Database

Selectionis the initial step or stage in the KDD process. In this process, variables are selected to be used so that there are no similarities or repetitions in the database management process. In the pre-processing stage there are 2 stages, namely:

1) Data Cleaning, eliminate unused data and clean data from missing values, noise data and irrelevant data.
2) Data Integration, in this process a unique entity identification process occurs.
3) Transformation, changing the appropriate data format in data mining management because several methods in data mining, such as the K-Means method used, require a special data format so that it can be processed.

4) Data Mining, the process of selecting the algorithm or method to be used. In this research, the K-Means Clustering method was used.
5) Evaluation / Interpretation, identifying the patterns used so that you can find out or assess the level of accuracy of existing data studies.
6) Knowledge, the results of data processing that have been in the form of patterns or have been visualized can be presented to the user. In this process, users gain new knowledge which will be used as a basis for making decisions.

K-Means is one of the stages in the Knowledge Discovery in Databases (KDD) process, specifically within the realm of Data Mining. In K-Means, cluster values (k) are initially determined randomly, with the resulting centroids temporarily serving as the center of each cluster [8]. This algorithm is widely adopted for clustering due to its simplicity, scalability for large datasets, efficiency in computation, ease of implementation, and its capability to minimize the distance between data points and their respective clusters [9]. K-Means can only be used for data whose attributes are in numeric format. The steps for calculating using the K-Means algorithm are as follows.

1) Determine the cluster value (k) randomly.
2) Determine the cluster (k) to the initial center point (centroid) randomly. To determine the ith cluster as shown in Equation 1.

$$M_k = \left(\frac{1}{x_k}\right)\sum_{i=1}^{nk} X_{ik} \quad (1)$$

Where *Mk* is centroid point of the K-th cluster, *Xk, nk* is the amount of data in the K-th cluster. *Xik* is IK data in the Kth cluster.

3) Calculate the distance from each object of each cluster to each centroid using the Euclidean Distance formula as shown in Equation 2.

$$d(p,q) = \sqrt{((p1-q1)^2 + ((p2-q2)^2 + \ldots + ((p1-q1)^2)} \quad (2)$$

Where *d(p,q)* is Distance between facilities p and q. *p1, q1* is Coordinates for facility 1. *p2, q2* is Coordinates for facility 2.

4) Classifying data based on the closest distance to the centroid.
5) Calculate again the distance between the cluster center point and the cluster members. The cluster center point is the average value of the object data in a cluster.
6) Repeat step 3 if the new position of the center point (Centorid) is not the same.

The process of this research as shown in Figure 1 as follow.

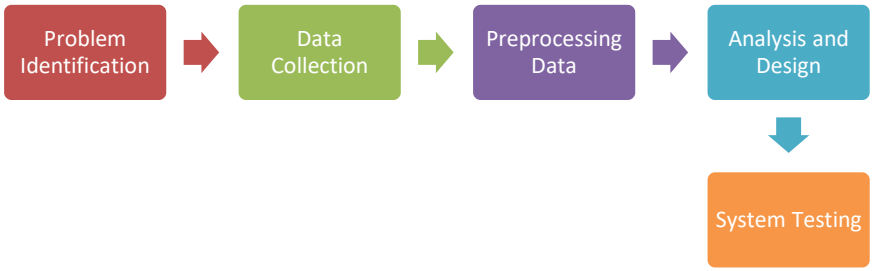

Figure 2. Research Stages

1) The first stage in this research is identifying existing problems and establishing the goals to be achieved which are then implemented in the author's research entitled "Implementation of the K-Means Method to Know the Performance of Information Technology Department Employees Based on Manage Engine Data at PT Nusa Halmahera Minerals".
2) At the data collection stage, the author collected Manage Engine data for January 2022 – August 2022 at the Information Technology Department of PT Nusa Halmahera Minerals for 3 (three) months and saved the data in the form of an Excel file.
3) After the data was collected, the author chose to pre-process the data. The data pre-processing process includes:
   a) Data cleaning
      At this stage the author determines the relevant attributes to be analyzed and then removes or deletes the attributes that are not used.
   b) Data normalization
      At this stage the author determines the appropriate data type for each attribute and determines the attribute that will become the role.
4) In the analysis stage, the author carried out calculations using the K-Means method.
5) In this stage, the author carried out tests using Rapid Miner Studio to find out whether the calculations carried out by the author were appropriate or not.
6) If the calculation results between the author and Rapid Miner Studio are appropriate, the author will use visualization in Rapid Miner Studio to make it easier for users to receive new knowledge to use as a basis for decision making.

3. RESULTS AND DISCUSSION

The Manage Engine data used by the author is data from January 1 2022 to August 2 2022, totaling 3,563 raw data. Before entering the Clustering stage, variables from the Manage Engine data are selected so that the right variables are obtained to be used as research references. The variables selected are Technician Name, Closed

Ticket Value, and Working Time Value in hours. These two variables are considered appropriate to be used as research references because the work process in the Information Technology Department of PT Nusa Halmahera Minerals uses a Ticketing system. After the variable selection stage is carried out, the data is processed to obtain the Total Closed Ticket value and the Total Working Time value in hours per employee.

Table 1. Dataset for measuring employee performance

Technician Name	Total Closed Tickets	Total Working Time (Hours)
Strong Nugroho	445	2685.30
Ceppy Christiantosa	129	2933.13
Abdullah Ibn Albaar	498	2710.65
Alfian	14	213.12
Faldi Risman	141	2830.87
Ardianto Toloa	310	702.58
Irwan	305	1126.87
Gustria Nareswara	21	562.78
Novelion	36	1112.90
PT Traksi	71	4832.43
Alowisius Amelga	6	385.80
Arfiansyah	28	1391.37
Nabillah Anwar	281	3141.43
Haedar Awal	69	928.37
Hendy Cahyadi	15	285.13
Iwan Sanwani	17	1216.12
Tahir	9	128.97
Pierre Gerungan	57	2360.97
Muhammad Rizki	153	1423.38
Melianus Mairuhu	61	2892.05
Greggy Dirk Maitimoe	449	5151.25
Rifyan Marzuki	46	1053.12
Yusup Maolana	34	780.93
Not known	15	24.77
Suprapto	3	4.70
Donny The	6	116.23
Muhammad Ali Imran Samsuri	200	2852.23

This data pre-processing stage includes cleaning data from missing values, duplicate data, and data normalization. At this stage, data types are adjusted to the attributes that will be used in the RapidMiner tool.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Table 2. Data Transformation

Attribute	Data Type	Information
Technician	Polynominal	Technician Name
Total Closed Tickets	Integer	Total Closed Tickets
Total Working Time (Hours)	Real	Total Working Time in hours

This stage is a calculation stage using the K-Means method based on the selected data. Data calculations at this stage will be grouped into 3 clusters, namely, high employee performance (C1), medium employee performance (C2), and low employee performance (3). In the K-Means algorithm there is a midpoint or centroid value to determine the cluster center point. The point values are as follows:

Table 3. Initial Data Centroid (Iteration 1)

Centroid	Total Closed Tickets	Total Working Time (Hours)
1 (Very Good)	498	2710.65
2 (Good)	61	3492.05
3 (Fair)	3	4.70

After the midpoint or centroid value is obtained, the selected data can be grouped into 3 clusters and the shortest distance search process can be carried out. The results of searching for the shortest distance and grouping data into 3 clusters are as follows:

Table 4. First Iteration Results

Technician Name	Total Closed Tickets	Total Working Time (Hours)	C1	C2	C3	Min
Strong Nugroho	445	2685.30	58.75051	893.4772	2716,797	58.75051
Ceppy Christiantosa	129	2933.13	430.8826	563,038	2931,144	430.8826
Abdullah Ibn Albaar	498	2710.65	0	895,296	2750,854	0
Alfian	14	213.12	2543,999	3279.27	208.7079	208.7079
Faldi Risman	141	2830.87	376.6976	666.0056	2829,535	376.6976
Ardianto Toloa	310	702.58	2016,848	2800,558	762.4249	762.4249
Irwan	305	1126.87	1595,499	2377,736	1162,095	1162,095
Gustria Nareswara	21	562.78	2200,195	2929.54	558.3746	558.3746
Novelion	36	1112.90	1663,204	2379,281	1108,692	1108,692
PT Traksi	71	4832.43	2164,323	1340,421	4828,213	1340,421
Alowisius Amelga	6	385.80	2376.34	3106,737	381.1129	381.1129
Arfiansyah	28	1391.37	1400,503	2100,943	1386,893	1386,893

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Technician Name	Total Closed Tickets	Total Working Time (Hours)	C1	C2	C3	Min
Nabillah Anwar	281	3141.43	482.3518	413.9228	3149.03	413.9228
Haedar Awal	69	928.37	1833,187	2563,696	926.0228	926.0228
Hendy Cahyadi	15	285.13	2473.14	3207,247	280.6911	280.6911
Iwan Sanwani	17	1216.12	1570,029	2276,359	1211,499	1211,499
Tahir	9	128.97	2627,586	3363,485	124.4125	124.4125
Pierre Gerungan	57	2360.97	562.8139	1131.09	2356,886	562.8139
Muhammad Rizki	153	1423.38	1332,697	2070,711	1426,592	1332,697
Melianus Mairuhu	61	2892.05	473.1543	600	2887,934	473.1543
Greggy Dirk Maitimoe	449	5151.25	2441,092	1703,963	5165.84	1703,963
Rifyan Marzuki	46	1053.12	1718,057	2438,979	1049,299	1049,299
Yusup Maolana	34	780.93	1984,717	2711,251	776.8532	776.8532
Not known	15	24.77	2728,967	3467,588	23.38195	23.38195
Suprapto	3	4.70	2750,853	3487,832	0.001111	0.001111
Donny The	6	116.23	2640,656	3376,265	111.5748	111.5748
Muhammad Ali Imran Samsuri	200	2852.23	329,924	654.7415	2854,341	329,924

Table 5. First Iteration Grouping

Technician Name	C1	C2	C3
Strong Nugroho	x		
Ceppy Christiantosa	x		
Abdullah Ibn Albaar	x		
Alfian			x
Faldi Risman	x		
Ardianto Toloa			x
Irwan			x
Gustria Nareswara			x
Novelion			x
PT Traksi		x	
Alowisius Amelga			x
Arfiansyah			x
Nabillah Anwar		x	
Haedar Awal			x
Hendy Cahyadi			x
Iwan Sanwani			x
Tahir			x
Pierre Gerungan	x		
Muhammad Rizki	x		
Melianus Mairuhu	x		

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Technician Name	C1	C2	C3
Greggy Dirk Maitimoe		x	
Rifyan Marzuki			x
Yusup Maolana			x
Not known			x
Suprapto			x
Donny The			x
Muhammad Ali Imran Samsuri	x		

Table 6. Centroid of 2nd Iteration Data

Centroid	Total Closed Tickets	Total Working Time (Hours)
1 (Very Good)	211	2586.07
2 (Good)	267	4375.04
3 (Fair)	58	627.11

Table 7. 2nd Iteration Results

Technician Name	Total Closed Tickets	Total Working Time (Hours)	C1	C2	C3	Min
Strong Nugroho	445	2685.30	254.6297	1699,088	2094,189	254.6297
Ceppy Christiantosa	129	2933.13	356.5013	1448,494	2307,105	356.5013
Abdullah Ibn Albaar	498	2710.65	313.33	1680,343	2129,416	313.33
Alfian	14	213.12	2381,078	4169,605	416.3641	416.3641
Faldi Risman	141	2830.87	254.4685	1549,304	2205,306	254.4685
Ardianto Toloa	310	702.58	1886,116	3672,707	262.7003	262.7003
Irwan	305	1126.87	1462,263	3248,394	557.2982	557.2982
Gustria Nareswara	21	562.78	2032,144	3820,184	74.39577	74.39577
Novelion	36	1112.90	1483,472	3270,307	486.3056	486.3056
PT Traksi	71	4832.43	2250,688	497.62	4205,343	497.62
Alowisius Amelga	6	385.80	2209,756	3997,768	246.9278	246.9278
Arfiansyah	28	1391.37	1208,565	2993,229	764.8607	764.8607
Nabillah Anwar	281	3141.43	559.8173	1233,685	2524,161	559.8173
Haedar Awal	69	928.37	1663,734	3452,355	301.4446	301.4446
Hendy Cahyadi	15	285.13	2309.23	4097,662	344.7158	344.7158
Iwan Sanwani	17	1216.12	1383,554	3168,799	590.4587	590.4587
Tahir	9	128.97	2465,355	4253,903	500.5837	500.5837
Pierre Gerungan	57	2360.97	272.4611	2024,991	1733,858	272.4611
Muhammad Rizki	153	1423.38	1164,111	2953,856	801.8766	801.8766
Melianus Mairuhu	61	2892.05	340,547	1497,228	2264,942	340,547

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Technician Name	Total Closed Tickets	Total Working Time (Hours)	C1	C2	C3	Min
Greggy Dirk Maitimoe	449	5151.25	2576,241	797.2626	4540,973	797.2626
Rifyan Marzuki	46	1053.12	1541,757	3329,265	426,187	426,187
Yusup Maolana	34	780.93	1813,748	3601.65	155.7432	155.7432
Not known	15	24.77	2568,756	4357,565	603.9024	603.9024
Suprapto	3	4.70	2589,699	4378,305	624.8678	624.8678
Donny The	6	116.23	2478,291	4266,796	513.5538	513.5538
Muhammad Ali Imran Samsuri	200	2852.23	266.3674	1524,279	2229,626	266.3674

Table 8. 2nd Iteration Grouping

Technician Name	C1	C2	C3
Strong Nugroho	x		
Ceppy Christiantosa	x		
Abdullah Ibn Albaar	x		
Alfian			x
Faldi Risman	x		
Ardianto Toloa			x
Irwan			x
Gustria Nareswara			x
Novelion			x
PT Traksi		x	
Alowisius Amelga			x
Arfiansyah			x
Nabillah Anwar	x		
Haedar Awal			x
Hendy Cahyadi			x
Iwan Sanwani			x
Tahir			x
Pierre Gerungan	x		
Muhammad Rizki			x
Melianus Mairuhu	x		
Greggy Dirk Maitimoe		x	
Rifyan Marzuki			x
Yusup Maolana			x
Not known			x
Suprapto			x
Donny The			x
Muhammad Ali Imran Samsuri	x		

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Table 9. Centroid of 3rd Iteration Data

Centroid	Total Closed Tickets	Total Working Time (Hours)
1 (Very Good)	239	2796.54
2 (Good)	260	4991.84
3 (Fair)	64	673.95

Table 10. 3rd Iteration Results

Technician Name	Total Closed Tickets	Total Working Time (Hours)	C1	C2	C3	Min
Strong Nugroho	445	2685.30	254.6297	1699,088	2094,189	254.6297
Ceppy Christiantosa	129	2933.13	356.5013	1448,494	2307,105	356.5013
Abdullah Ibn Albaar	498	2710.65	313.33	1680,343	2129,416	313.33
Alfian	14	213.12	2381,078	4169,605	416.3641	416.3641
Faldi Risman	141	2830.87	254.4685	1549,304	2205,306	254.4685
Ardianto Toloa	310	702.58	1886,116	3672,707	262.7003	262.7003
Irwan	305	1126.87	1462,263	3248,394	557.2982	557.2982
Gustria Nareswara	21	562.78	2032,144	3820,184	74.39577	74.39577
Novelion	36	1112.90	1483,472	3270,307	486.3056	486.3056
PT Traksi	71	4832.43	2250,688	497.62	4205,343	497.62
Alowisius Amelga	6	385.80	2209,756	3997,768	246.9278	246.9278
Arfiansyah	28	1391.37	1208,565	2993,229	764.8607	764.8607
Nabillah Anwar	281	3141.43	559.8173	1233,685	2524,161	559.8173
Haedar Awal	69	928.37	1663,734	3452,355	301.4446	301.4446
Hendy Cahyadi	15	285.13	2309.23	4097,662	344.7158	344.7158
Iwan Sanwani	17	1216.12	1383,554	3168,799	590.4587	590.4587
Tahir	9	128.97	2465,355	4253,903	500.5837	500.5837
Pierre Gerungan	57	2360.97	272.4611	2024,991	1733,858	272.4611
Muhammad Rizki	153	1423.38	1164,111	2953,856	801.8766	801.8766
Melianus Mairuhu	61	2892.05	340,547	1497,228	2264,942	340,547
Greggy Dirk Maitimoe	449	5151.25	2576,241	797.2626	4540,973	797.2626
Rifyan Marzuki	46	1053.12	1541,757	3329,265	426,187	426,187
Yusup Maolana	34	780.93	1813,748	3601.65	155.7432	155.7432
Not known	15	24.77	2568,756	4357,565	603.9024	603.9024
Suprapto	3	4.70	2589,699	4378,305	624.8678	624.8678
Donny The	6	116.23	2478,291	4266,796	513.5538	513.5538
Muhammad Ali Imran Samsuri	200	2852.23	266.3674	1524,279	2229,626	266.3674

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Table 11. 3rd Iteration Grouping

Technician Name	C1	C2	C3
Strong Nugroho	x		
Ceppy Christiantosa	x		
Abdullah Ibn Albaar	x		
Alfian			x
Faldi Risman	x		
Ardianto Toloa			x
Irwan			x
Gustria Nareswara			x
Novelion			x
PT Traksi		x	
Alowisius Amelga			x
Arfiansyah			x
Nabillah Anwar	x		
Haedar Awal			x
Hendy Cahyadi			x
Iwan Sanwani			x
Tahir			x
Pierre Gerungan	x		
Muhammad Rizki			x
Melianus Mairuhu	x		
Greggy Dirk Maitimoe		x	
Rifyan Marzuki			x
Yusup Maolana			x
Not known			x
Suprapto			x
Donny The			x
Muhammad Ali Imran Samsuri	x		

After analyzing and calculating the data, implementation was carried out using the Rapid Miner tool to test the grouping that had been carried out. In this process, Performance Vector testing is also carried out to determine the best cluster value using the Davies Bouldin Index value.

Journal of Information Systems and Informatics
Vol. 6, No. 1, March 2024
p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

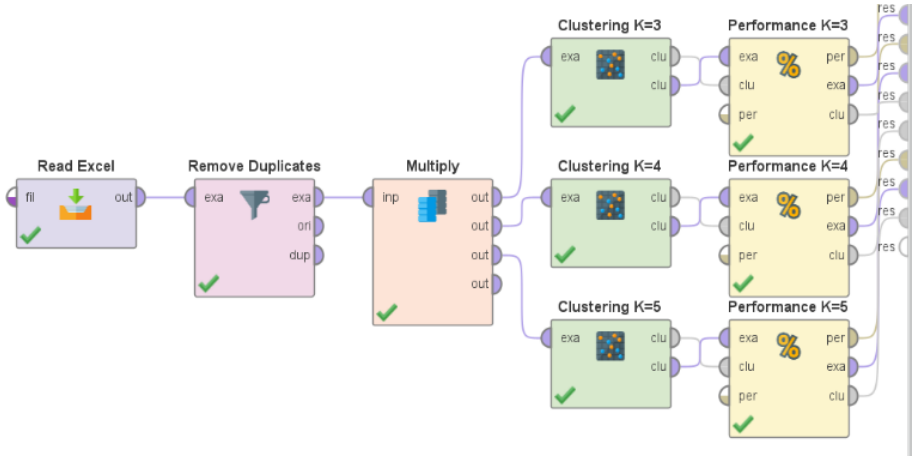

Figure 3. Testing in RapidMiner

1	1	Kukoh Nugroho	cluster_0	445	2685.300
2	2	Ceppy Christi...	cluster_0	129	2933.133
3	3	Abdullah Ibn...	cluster_0	498	2710.650
4	4	Alfian	cluster_2	14	213.117
5	5	Faldi Risman	cluster_0	141	2830.867
6	6	Ardianto Toloa	cluster_2	310	702.583
7	7	Irwan	cluster_2	305	1126.867
8	8	Gustria Nare...	cluster_2	21	562.783
9	9	Novelion	cluster_2	36	1112.900
10	10	PT Traksi	cluster_1	71	4832.433
...					
...					
21	21	Greggy Dirk ...	cluster_1	449	5151.250
22	22	Rifyan Marzuki	cluster_2	46	1053.117
23	23	Yusup Maolana	cluster_2	34	780.933
24	24	Tidak Diketahui	cluster_2	15	24.767
25	25	Suprapto	cluster_2	3	4.700
26	26	Donny The	cluster_2	6	116.233
27	27	Muhammad A...	cluster_0	200	2852.233

Figure 4. Grouping Results in RapidMiner

From Figure 4 the grouping done manually is the same as the grouping results using RapidMiner. The next process is testing Vectpr Performance using the Davies Bouldin Index value. In this process, the author chose to test data grouping with the values k=3, k=4, and k=5.

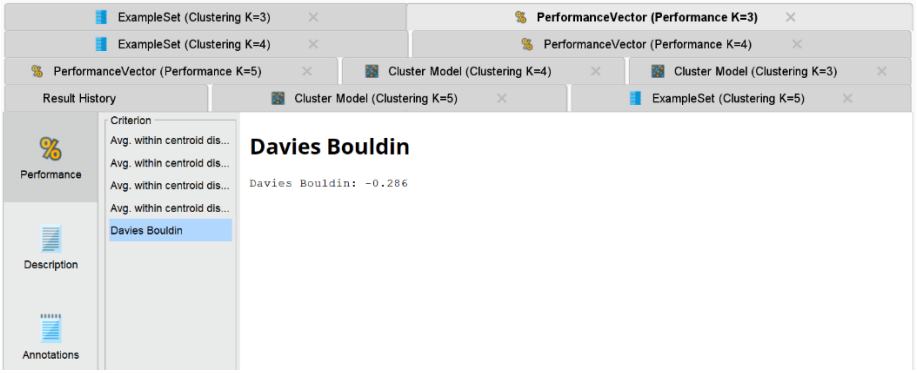

Figure 5. Davies Bouldin Index value k=3

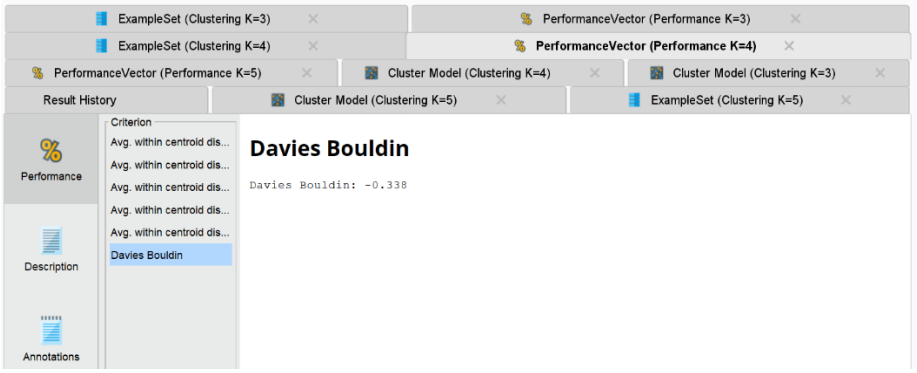

Figure 6. Davies Bouldin Index value k=4

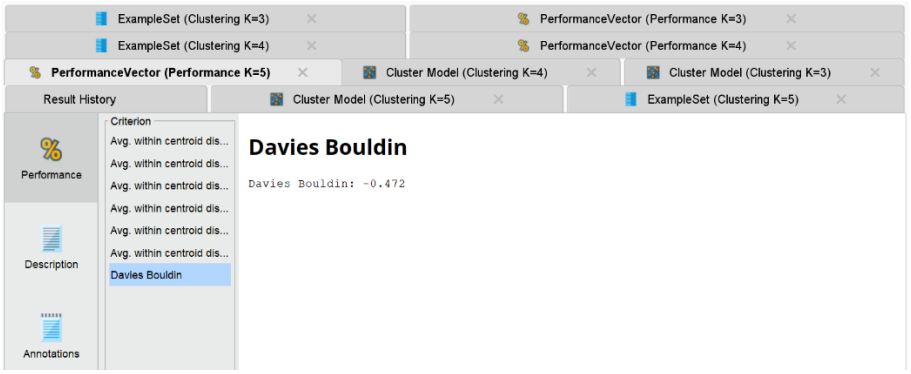

Figure 7. Davies Bouldin Index value k=5

From figures 5, 6 and 7, you can see the values and graphs of performance testing using the Davies Bouldin Index as follows:

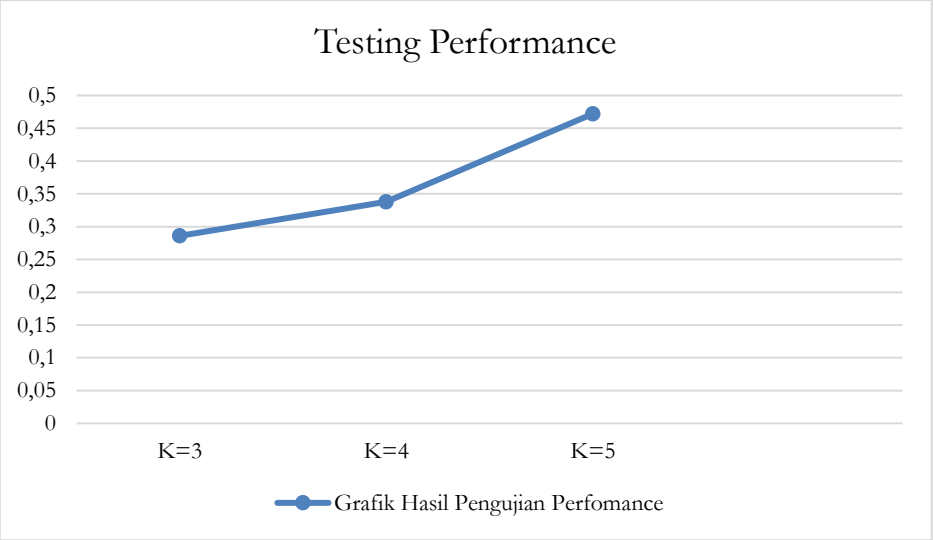

Figure 8. Graph of Performance Testing Results Using DBI

Where K=3 : -0.286 (Figure 5), K=4 : -0.338 (Figure 6), and K=5 : -0.472 (Figure 7). From these results, it can be concluded that the best Performance Vector is found in k=3 or clustering of 3 groups with a Performance Vector value of 0.286 because the DBI value is considered good if it is close to 0. Based on the results obtained, it can be concluded that the Information Technology Department of PT Nusa Halmahera Minerals has 8 employees with very good performance, 2 employees with good performance, and 17 employees with fairly good performance.

4. CONCLUSION

Based on the calculations and tests conducted, the following conclusions can be drawn (1) The K-Means algorithm proves effective in measuring employee performance within the Information Technology Department of PT Nusa Halmahera Minerals through clustering. The results reveal 8 employees categorized as "very good" (Cluster I), 2 as "good" (Cluster II), and 17 as "fairly good" (Cluster III). (2) Evaluation using the Davies Bouldin Index indicates that the optimal Performance Vector consists of 3 clusters, achieving a high accuracy level of 0.286. Consequently, this study adopts a value of k=3, delineating performance levels as "very good," "good," and "fairly good."

REFERENCES

[1] P. P. L. N. Dewi, I. N. Purnama, and N. W. Utami, "Application of Data Mining for Clustering Lecturer Performance Assessments Using the K-Means Algorithm (Case Study: STMIK Primakara)," *Asian Information Technology Scientific Journal*, vol. 16, no. 2, 2022.

[2] D. Sartika and J. Jumadi, "Clustering Lecturer Performance Assessment Using the K-Means Algorithm (Case Study: Dehasen University Bengkulu)," *in Proc. National Seminar on Computer & Science Technology (SAINTEKS)*, 2019, pp. 703–709.

[3] Z. Rialmi, I. Noviyanti, and R. Akbar, Effect of Job Satisfaction on Employee Performance (Study of PT Antarmitra Sembada, West Jakarta Logistics Division). *Journal of Research in Business, Economics, and Education*, vol. *2*, no. 5, pp.1076-1084, 2020.

[4] F. N. R. F. Aziz, B. D. Setiawan, and I. Arwani, "Implementation of the K-Means Algorithm for Clustering Student Academic Performance," *J. Pengemb. Technol. Inf. and Comput. Sci.*, vol. 2, no. 6, pp. 2243–2251, 2018.

[5] R. Sandra, S. Entin, and A. Nani, "Clustering Quality of Employee Performance in Chemical Companies Using the K-Means Algorithm," *Budidarma Media Informatics Journal*, vol. 5, no. 2, pp. 573-582, 2021.

[6] T. A. Ibnu, S. Anindita, and P. Novianti, "Grouping Employee Performance Quality Using the K-Means Method," *Komputika: Journal of Computer Systems*, vol. 11, no. 2, pp. 131-141, 2022.

[7] S. Setiawan, "Cluster Analysis Using the K-Means Algorithm to Find Out the Capabilities of Employees in the IT Field at CV Roxed LTD," *Pelita Informatics Journal*, vol. 7, no. 3, 2019.

[8] Gustientiedina, M. H. Adiya, and Y. Desnelita, "Application of the K-Means Algorithm for Clustering Medicine Data at Pekanbaru District Hospital," *National Journal of Information Technology and Systems*, vol. 5, no. 1, pp. 17-24, 2019.

[9] W. S. Riyani, A. Wanto, and A. P. Windarto, "Implementation of Rapidminer Using the K-Means Method (Case Study: Measles Immunization in Toddlers Based on Province)," *KOMIK (National Conference on Information and Computer Technology)*, vol. 2, no. 1, 2018.

[10] D. Sartika and J. Jumadi, "Clustering Lecturer Performance Assessment Using the K-Means Algorithm (Case Study: Dehasen University Bengkulu)," *in Proc. Sainteks*, pp. 703–709, 2019.

[11] S. Setiawan, "Cluster Analysis Using the K-Means Algorithm to Determine the Ability of Employees in the IT Field on CV. Roxed Ltd," *J. Pelita Inform.*, vol. 18, pp. 80–86, 2019.

[12] D. Imantika et al., "Application of the K-Means Clustering Method and Analytical Hierarchy Process (AHP) for Grouping Teacher and Employee

Performance at Sma Brawijaya Smart School," *J. Pememb. Technol. Inf. and Comput. Sci. J-Ptiik*, vol. 3, no. 8, pp. 7382–7390, 2019.

[13] Wikarno, R. Malani, and B. Suprapty, "Comparison of K-Means and Fuzzy C-Means Methods for Grouping Employees Based on Performance Values and Employee Discipline Levels," *in Proc. Semin. Computer Science. And Technol. Inf.*, vol. 3, no. 1, pp. 45–52, 2018.

[14] O. Panggi, L. Yulianti, and J. Fredricka, "Employee Performance Analysis Using the K-Means Algorithm at the Central Bengkulu Regency Education and Culture Service," *Infotama Media Journal*, vol. 17, no. 2, 2021.

[15] E. Saputra and Y. Nataliani, "Analisis Pengelompokan Data Nilai Siswa untuk Menentukan Siswa Berprestasi Menggunakan Metode Clustering K-Means", *J. Inf. Syst. Informatics,* vol. 3, no. 3, pp. 424-439, 2021.