JURNAL RISET INFORMATIKA **P-ISSN: 2656-1743 |E-ISSN: 2656-1735**
Vol. 5, No. 1. December 2022 DOI: https://doi.org/10.34288/jri.v5i1.180
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

Implementation of the K-Means Clustering for Teacher Performance Assessment Grouping (PKG) at MI Bani Hasyim Cerme

Bagus Firmansyah[1*], Umi Chotijah[2]

Informatics Engineering, Engineering
University of Muhammadiyah Gresik
Gresik, Indonesia
bagusfirman860@gmail.com[1*], umi.chotijah@umg.ac.id[2]
(*) Correponding Author

Abstract

Evaluation of teacher performance at MI Bani Hasyim Cerme still uses the manual method. Using office applications such as excel and word results in a significant accumulation of data that makes it difficult for school principals to calculate scores and evaluate the results of clustering or teacher performance scores, so it is wasteful of energy, time, and cost. The k-Means clustering method is expected to facilitate the clustering process of teacher performance values as a source of information and make it easy for school principals to classify teacher performance results. This research aims to obtain clustering values on teacher performance assessment data and to replace the teacher performance assessment system at MI Bani Hasyim, which was previously carried out conventionally into a web-based system. The results of this study are the clustering values of teacher performance assessment and a web-based teacher performance appraisal system. It is expected to facilitate the process of evaluating teacher performance in the Bani Hasyim primary school in the future.

Keywords: teacher, teacher performance assessment, K-Means clustering, MI Bani Hasyim, web system

Abstrak

Penilaian kinerja guru di MI Bani Hasyim Cerme masih menggunakan cara manual. Menggunakan aplikasi perkantoran seperti excel dan word yang berakibat penumpukan data yang sangat banyak sehingga menyulitkan kepala sekolah dalam menentukan skor penilaian dan mengevaluasi pengelompokan atau nilai kinerja guru dengan cara yang boros tenaga, waktu dan uang. Metode K-Means clustering diharapkan dapat mempermudah proses clustering nilai kinerja guru sebagai sumber informasi dan memudahkan kepala sekolah dalam mengklasifikasikan hasil kinerja guru. Penelitian ini bertujuan untuk mendapatkan nilai klusterisasi pada data penilaian kinerja guru serta menggantikan sistem penilaian kinerja guru pada MI Bani Hasyim yang sebelumnya dilakukan secara konvensional menjadi sistem berbasis web. Hasil dari penelitian ini adalah nilai clastering penilaian kinerja guru dan sistem penilaian kinerja guru berbasis web sehingga diharapkan dapat mempermudah proses penilaian kinerja guru pada MI Bani Hasyim kedepannya.

Kata kunci: guru, penilaian kinerja guru, K-Means clustering, MI Bani Hasyim, Sistem web

INTRODUCTION

Measuring an educational institution's performance is critical. Performance measurement is carried out to evaluate and plan future education appropriately, especially on teachers' performance as executors and even as spearheads of education. Various types of information are required to ensure that education and learning services are delivered effectively, efficiently, and accountable. Improving educational quality must always measure its performance through various information, task control, funding reports, and the, most important, teacher performance reports because teachers play a very strategic role in determining educational quality, which necessitates legal personality and professional ability requirements and can be held accountable (Muhiddinur, 2019).

Data mining is a method of data processing used to discover hidden patterns in data. This data mining method's data processing results can be used to make future decisions. Data mining entails, in essence, data collection and selection, data pre-processing, data analysis (including visualization of results), interpretation of findings, and knowledge application. Data mining is the process of extracting patterns from data using specific algorithms. The process uses detailed analysis, automatically looking for simple patterns in large amounts of data (Ndehedehe et al., 2013; Ong, 2013; Schuh et al., 2019).

Clustering is the process of grouping objects with similar properties into object classes. K-Means is one of the clustering methods that can be used in this problem. As a method of non-hierarchical data clustering that groups data into one or more clusters. Data with the same characteristics are grouped in one cluster, while data with different characteristics are grouped in another. This method is used to categorize teachers and employees based on data from student, teacher, and employee questionnaires. This method is used because it is an interactive method that is simple to interpret, apply, and dynamic on scattered data (Han et al., 2012; Hughes, 2012; Ong, 2013).

The Manhattan distance is commonly used for measurement because it is simple to calculate and understand and more appropriate for some problems, such as calculating the absolute difference between the coordinates of two objects (Pribadi et al., 2022; Yaniar, 2011).

David L. Davies and Donald W. Bouldin invented the Davies Bouldin Index (DBI) in 1979. The Davies-Bouldin Index maximizes inter-cluster distance while attempting to minimize the distance between points within a cluster. If the maximum inter-cluster distance exists, it indicates that the similarities between each cluster have increased slightly, making the differences between clusters more visible. If the minimal intra-cluster distance indicates that each object in the cluster has level similarity, then the characteristics of the high level (Bates & Kalita, 2016; Sartika & Jumadi, 2019).

Teacher performance is still evaluated manually at Bani Hasyim Primary School, using office applications such as Excel and Word. The results of the performance appraisal instrument generate a large number of documents for each teacher. Thus, even during the storage process, teachers and school principals will struggle to determine the results of calculating scores and evaluating the results of clustering or teacher performance scores, wasting time and money (Faisal et al., 2020; Lopis, 2016).

In previous research by (Panjaitan et al., 2015) and (Sukrianto, 2016), a study was conducted on teacher performance clustering using the K-Means Clustering method, which resulted in the classification of teacher performance into five clusters: bad cluster, poor cluster, moderate cluster, poor cluster well, and perfect cluster.

In previous research by (Imantika et al., 2019), The K-Means clustering method has been described as being used to divide teachers and employees into groups based on the value of the questionnaire results. The Analytical Hierarchy Process (AHP) method is then used to prioritize teachers' and employees' choices from various alternatives.

Related research was also carried out by (Nurzahputra et al., 2017). This paper, titled Application of the K-Means Algorithm for Lecturer Assessment Clustering Based on the Student Satisfaction Index, used the results of 146 student satisfaction questionnaires for all lecturers in the study program totaling 12 lecturers. The K-Means clustering method was used in this study, with good and poor clusters. The total centroid score for the excellent cluster is 17,099 (5 good lecturers), and the total centroid score for the poor cluster is 15,874. (7 bad lecturers).

Previous research differs from this research in that the authors used the K-Means Clustering method to classify teacher performance scores at MI Bani Hasyim over the last five years, then added a graph to monitor the development of teacher performance scores over the last five years so that it can be seen whether the teacher is improving or deteriorating. K-Means clustering is expected to facilitate the clustering process of teacher performance values as a source of information and make it easy for school principals to classify teacher performance results (Faisal et al., 2020).

RESEARCH METHODS

Types of research

This research uses a quantitative method that is systematic and uses mathematical models.

Time and Place of Research

This research was conducted at Bani Hasyim Primary School, and the time of research was from August 2022 to November 2022.

Research Target / Subject

The target of this research is the performance value of teachers at Bani Hasyim Primary School.

Procedure

1. Problem Identification

Problem identification is the first step in applying the K-Means Clustering method. Problem identification aims to determine the appropriate data be analyzed using the K-Means Clustering method to classify teachers based on performance scores.

P-ISSN: 2656-1743 |E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v5i1.180
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

2. Data, Instruments, and Data Collection Techniques

The data used in this study is the Bani Hasyim primary school teachers, The techniques used for data collection include the following:

a) Field Research

In field research, researchers directly visit research sites and collect data needed for research. Field research was conducted directly by interviewing the Bani Hasyim primary school principal to obtain the required teacher information.

b) Literature Research

Literature research is carried out by collecting references from journals or academic books related to the issues discussed.

3. Data Processing

This stage is carried out to create raw data that will be processed into quality data. This is done in order to obtain more accurate results with the use of the K-Means clustering method.

4. Data Analysis

This stage is carried out based on the results of observations and data collection carried out. System requirements analysis is carried out to determine the features to be used in the system.

RESULTS AND DISCUSSION

The K-means algorithm is one of the partitional algorithms since it is based on defining the initial centroid value, allowing the initial number of groups to be determined (Madhulatha, 2012). The K-means algorithm uses an iterative procedure to create database clusters. After receiving the desired number of initial clusters as input, it generates the final centroid point as output. The centroid's starting point will be chosen randomly by the K-means method's pattern k. The initial cluster centroid candidates can influence the total number of iterations needed to find the cluster centroid. In order to design the algorithm in a way that will produce higher performance, we must identify the centroid cluster, which can be seen from the high initial data density (Eltibi & Ashour, 2011; Hung et al., 2005; Saranya & Punithavalli, 2011).

The K-Means algorithm's final output will be a centroid point, which is what it is intended to do. Each dataset object joins a cluster once the K-Means iteration is complete. The cluster value is calculated by looking through all the items to locate the cluster closest to the object. Based on the shortest distance, the K-means algorithm will cluster data points in a dataset (Bangoria et al., 2013). The distance to all of the data from the original centroid value, which was randomly selected as the starting point, was determined using the Euclidean Distance calculation. Data that are close to the centroid will group. This process is repeated until no change exists in any group(Chaturvedi et al., 2013). According to this study, the authors grouped by using four variables shown in table 1.

Table 1. Criteria Data

Code	Criteria
K1	Pedagogic
K2	Personality
K3	Social
K4	Professional

This calculation uses the performance values of 8 teachers, which are initialized with the letters A to H. Then, for the year of performance evaluation, it is initialized with one as 2018 and 5 as 2022. For example, data A1 was Mrs. Istianah's performance value in 2018, data A2 was Mrs. Istianah's performance value in 2019, and so on, until data H5. Table 2 shows the initialization of the teacher code, and Table 3 shows the data for teacher performance scores calculated using the K-Means Clustering method.

Table 2. Teacher Data

Code	Teacher Name
A	Istianah, S.Pd.I
B	Dwi Yuniartiningtyas W, S.Pd
C	Muslimah, S.Pd.I
D	Mar'atus Sholihah, S.Ag
E	Siti Qoniah, S.Pd.I
F	Ni'matul Karimah, S.Pd.I
G	Winanto, S.Pd
H	Muhammad Irwan, S.Pd.I

Table 3. Teacher Performance Assessment Data

Code	K1	K2	K3	K4
A1	3,3	4	3,5	2,5
A2	3,1	3,3	4	3,5
A3	3,3	3	4	3,5
A4	3,4	3	4	3,5
A5	3,6	3,7	3,5	3
...	...	...	...	...
...	...	...	...	...
H1	3	4	3,5	1,5
H2	3,1	3	3	3
H3	3,4	3,3	4	3,5
H4	3,4	3,7	3,5	3,5
H5	3,7	3,7	3	3,5

Teacher performance appraisal data is processed using the K-Means Clustering method, which will then be grouped into 4 clusters, namely

"Very Good," "Good," "Enough," and "Poor," which is shown in table 4.

Table 4. Score

Score
Very Good
Good
Enough
Poor

1. Determine the number of K clusters

According to this study, 4 clusters were selected randomly with pedagogic, personality, social, and professional variables.

2. Determine the initial value of the midpoint (centroid) randomly

Based on this, the authors determine that the initial centroid is done randomly, as seen in table 5.

Table 5. Initial Centroids

Initial Centroid				
Cluster	K1	K2	K3	K4
C1	3,7	3	3	3
C2	3,7	3,3	4	3
C3	3,9	4	4	3,5
C4	2,7	3,7	4	3

3. Calculate each data's distance to the centroid with the Manhattan distance formula shown in formula 1.

$d_{Manhattan}(x,y) = \sum_{i=1}^{n} |x_i - y_i|$ ....(1)

The following example is calculated from A1 data to 4 centroids with Manhattan distance. Where data A1 was obtained previously through initialization in table 2 in the calculation becomes x1 and four centroids consisting of c1, c2, c3, and c4.

a) calculation of data A1 against centroid 1

$$d(x1,c1) = \sum_{i=1}^{r} |x1i - c1i|$$
$$= |3{,}3 - 3{,}7| + |4 - 3| + |3{,}5 - 3| + |2{,}5 - 3| = 2{,}4$$

b) calculation of data A1 against centroid 2

$$d(x1,c2) = \sum_{i=1}^{r} |x1i - c2i|$$
$$= |3{,}3 - 3{,}7| + |4 - 3{,}3| + |3{,}5 - 4| + |2{,}5 - 3| = 2{,}1$$

c) calculation of data A1 against centroid 3

$$d(x1,c3) = \sum_{i=1}^{r} |x1i - c3i|$$
$$= |3{,}3 - 3{,}9| + |4 - 4| + |3{,}5 - 4| + |2{,}5 - 3{,}5| = 2{,}1$$

d) calculation of data A1 against centroid 4

$$d(x1,c4) = \sum_{i=1}^{r} |x1i - c4i|$$
$$= |3{,}3 - 2{,}7| + |4 - 3{,}7| + |3{,}5 - 4| + |2{,}5 - 3| = 1{,}9$$

The calculation of A1 data for the centroid above obtained the lowest value in the calculation of A1 data for the fourth centroid, which is equal to 1.9, so that A1 data will be entered into the fourth cluster, and so on for A2 data to H5 data.

4. Assigns each data to the nearest cluster

The following is the result of calculating the iteration distance; the shortest distance for each data to the centroid is shown in the table in yellow, and the closest centroid is the cluster that the data follows, which can be seen in table 6.

Table 6. Iteration 1 Distance Calculation Results

Code	Iteration 1			
	C1	C2	C3	C4
A1	2,4	2,1	2,1	1,9
A2	2,43333	1,1	1,46667	1,2
A3	1,9	1,2	1,6	1,8
A4	1,8	1,1	1,5	1,9
A5	1,26667	0,9	1,63333	1,4
...	...	...	...	...
...	...	...	...	...
H1	3,7	3,4	3,4	2,6
H2	0,6	1,9	3,3	2,1
H3	2,13333	0,8	1,16667	1,5
H4	1,96667	1,6	1,33333	1,7
H5	1,16667	1,8	1,53333	2,5

5. Defining a new Centroid

The average value of each variable in each cluster can be used to calculate the new centroid shown in table 7.

Table 7. New Centroid

New Centroid				
Cluster	K1	K2	K3	K4
C1	3,48	3,26	3,06	3,06
C2	3,51	3,47	3,79	3,17
C3	3,49	3,85	3,82	3,73
C4	3,11	3,88	3,44	2,69

The objective function change value is still over the threshold in the first iteration. Thus the

P-ISSN: 2656-1743 |E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v5i1.180
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

calculation will continue until the objective function change value is below the threshold in the following phase, which involves lowering the initial objective function value. The results of the objective function computation and variations in the objective function's value for each completed iteration are displayed in table 8.

Table 8. Objective Function Change

Iteration	Objective Function	Objective Function Change
1	14,277	985,723
2	10,638	3,639
3	10,638	0

The calculation halts at the third iteration in line with the results' goal function change value in table 8. The third-iteration change in the goal function, which has a value of 0, is significant enough to surpass the threshold. The outcome of the third iteration calculation is shown in table 9.

Table 9. Iteration 3 Distance Calculation Results

Code	Iteration 3			
	C1	C2	C3	C4
A1	2,00	1,77	1,98	0,16
A2	1,70	0,84	1,29	2,45
A3	1,61	0,97	1,42	2,59
A4	1,60	0,87	1,32	2,55
A5	1,36	1,00	1,32	1,09
...	...	...	...	...
...	...	...	...	...
H1	3,30	3,07	3,28	1,46
H2	0,64	1,77	2,79	2,17
H3	1,49	0,54	0,99	2,22
H4	1,32	0,94	0,82	1,39
H5	1,12	1,70	1,42	2,07

The cluster center or centroid obtained is the centroid in the last iteration, namely the centroid in the 3rd iteration. The final centroid is shown in table 10.

Table 10. Last Centroid

Cluster	Last Centroid			
	K1	K2	K3	K4
C1	3,36	3,22	3	3,17
C2	3,42	3,4	3,85	3,2
C3	3,5	3,81	3,83	3,75
C4	3,37	3,96	3,44	2,5

In this study, researchers have determined four criteria for evaluating teacher performance, as shown in Table 4. The four criteria are initialized into 4 clusters by sorting the average of each cluster on the last centroid shown in table 11, followed by the clustering results in table 12.

Table 11. Score Initialization

Score	Initialization
Very Good	C3
Good	C2
Enough	C4
Poor	C1

Table 12. Results of Teacher Performance Assessment Clustering

Code	Cluster	Score
A1	C4	Enough
A2	C2	Good
A3	C2	Good
A4	C2	Good
A5	C2	Good
...	...	...
...	...	...
H1	C4	Enough
H2	C1	Poor
H3	C2	Good
H4	C3	Very Good
H5	C1	Poor

Table 12 shows that teachers have very good, good, enough, and poor scores. Furthermore, teachers with low scores will be included in the training for improving teacher performance assessments at Bani Hasyim Primary School.

Davies-Bouldin Index Validity

The Davies-Bouldin index seeks to minimize distances between cluster points while maximizing distances between clusters (dense). The davies-Bouldin index's lowest value will indicate the ideal number of clusters, which falls within the range of (0, 1).

The distance of each data point from the centroid and the mean value is calculated to provide calculations for the SSW in the first stage. The results of estimating the SSW value using the K-Means computations are shown in table 13.

Table 13. SSW Calculation Results

Cluster	SSW
C1	0,52
C2	0,52
C3	0,46
C4	0,42

The next step is calculating the SSB (Sum of Square Between Cluster) values to gauge how far clusters are from one another apart. To do this, measure the distance between a cluster's centroids. The results of estimating the SSB value are shown in table 14.

Table 14. SSB Calculation Results

SSB	Cluster			
	1	2	3	4
1	0,00	0,87	1,18	1,09
2	0,87	0,00	0,69	0,99
3	1,18	0,69	0,00	1,32
4	1,09	0,99	1,32	0,00

Evaluation of the ratio (Rij), which seeks to determine the DBI value for each cluster, comes next. Each cluster's ratio value (DBI) is used to evaluate the DBI of the entire cluster. A good cluster has the smallest density value and the highest possible separation value. The results of estimating the DBI value using the K-Means computations are shown in table 15.

Table 15. DBI Calculation Result

R	Cluster				R max	DBI
	1	2	3	4		
1	0,00	1,19	0,84	0,86	1,19	1,24
2	1,19	0,00	1,42	0,94	1,42	
3	0,84	1,42	0,00	0,66	1,42	
4	0,86	0,94	0,66	0,00	0,94	

The ratio with the most significant value is chosen to find the average, resulting in a DBI value of 1.24235.

System Implementation

1. System login page

When the user enters the system, he or she will see the display shown in Figure 1. The user is asked to log in using the email and password previously created. If the user has not registered, he will not be able to enter the system.

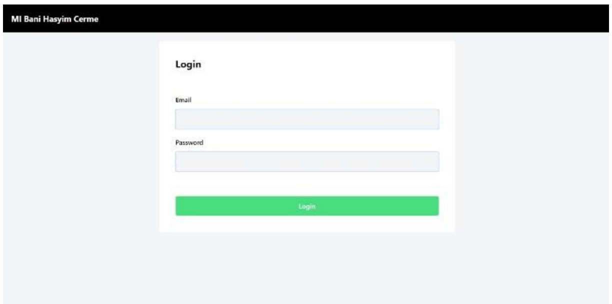

Figure 1. login page

2. Teachers data page

After the user logs into the system, the teacher data page will appear. Users can add, edit and delete teacher data through the teacher data page shown in figure 2.

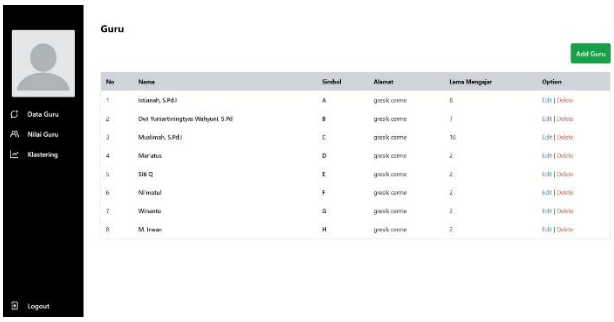

Figure 2. Teachers data page

3. Teachers score page

On this teacher's score page shown in figure 3, there are teacher performance scores from year to year for the last five years which include pedagogic, personality, social, and professional.

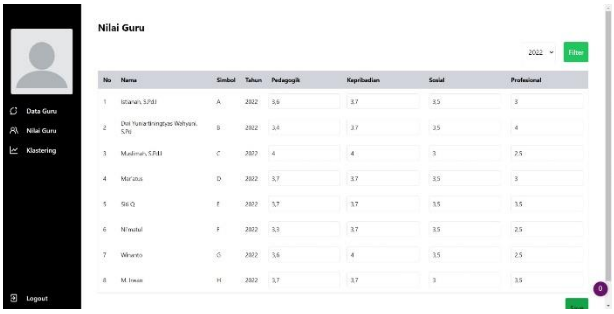

Figure 3. Teachers score page

4. Clustering page

Furthermore, on the clustering page shown in figure 4, there are several features, such as the range of years that will be calculated with K-Means clustering, then the user can choose which centroid will be used to perform the calculation.

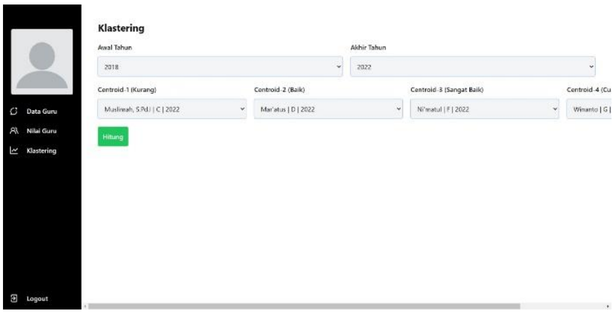

Figure 4. Clustering page

5. Calculation process page

After selecting the year range and centroid, the user will be directed to the calculation process

P-ISSN: 2656-1743 |E-ISSN: 2656-1735
DOI: https://doi.org/10.34288/jri.v5i1.180
Accredited rank 4 (SINTA 4), excerpts from the decision of the DITJEN DIKTIRISTEK No. 230/E/KPT/2023

page shown in figure 5. Here, the user can see K-Means calculations starting from the first iteration to the last iteration.

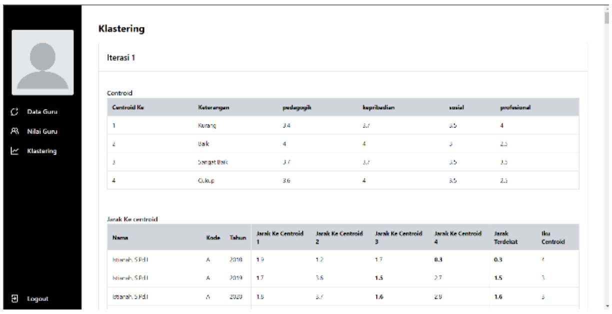

Figure 5. Calculation process page

6. Clustering results page

The clustering results page shown in figure 6 contains the clustering results from each teacher over five years. On this page, the user can get conclusions about which teachers get good grades and which teachers get poor grades so that training and workshops can be conducted for teachers who get poor grades to improve teacher performance appraisal.

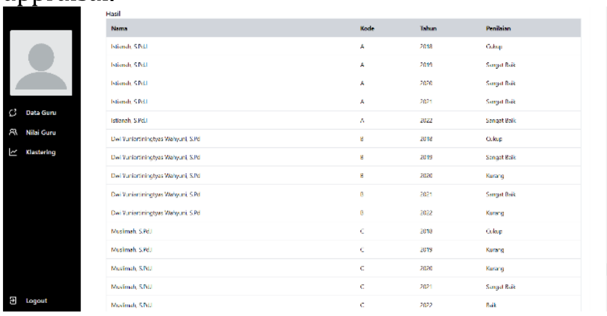

Figure 6. Clustering results page

7. Monitoring page

The monitoring page shown in figure 7 contains a graph of each teacher's performance calculation score in the last five years. Here, users can monitor the progress of each teacher's performance.

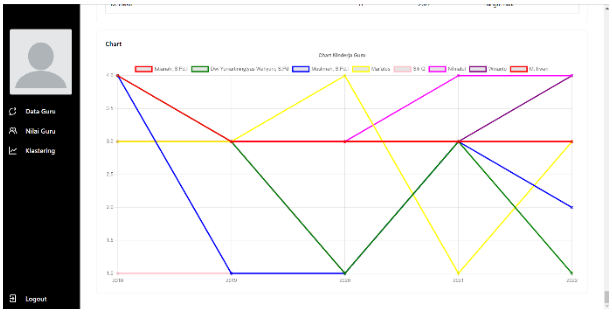

Figure 7. Monitoring page

CONCLUSION AND SUGGESTION

Conclusion

The authors' conclusions from the research include classifying teacher performance evaluations at MI Bani Hasyim based on four assessment categories, pedagogic, personality, social, and professional. Teachers' performance assessments are grouped into very good, good, enough, and poor. The iteration process carried out in this study obtained three iterations and the results of the tests that were carried out, then formed teacher group data with excellent ratings consisting of 12 (twelve) teacher data, teacher group data with good ratings consisting of 10 (ten) teacher data, teacher group data with enough assessment consisting of 9 (nine) teacher data, and teacher group data with poor assessment consisting of 9 (nine) teacher data.

Suggestion

The K-Means method should also be compared with other approaches to make more accurate clustering.

References

Bangoria, B., Mankad, N., & Pambhar, V. (2013). A survey on Efficient Enhanced K-Means Clustering Algorithm. *IJSRD -International Journal for Scientific Research & Development|*, *1*(9online), 2321–2613.

Bates, A., & Kalita, J. (2016). Counting clusters in twitter posts. *ACM International Conference Proceeding Series*, *04-05-Marc*. https://doi.org/10.1145/2905055.2905295

Chaturvedi, N., & Rajavat, A. (2013). An Improvement in K-mean Clustering Algorithm Using Better Time and Accuracy. *International Journal of Programming Languages and Applications*, *3*(4), 13–19. https://doi.org/10.5121/ijpla.2013.3402

Eltibi, M. F., & Ashour, W. M. (2011). Initializing K-Means Clustering Algorithm using Statistical Information. *International Journal of Computer Applications*, *29*(7), 51–55.

Faisal, A., Basri, B., & Sari, C. R. (2020). Sistem Informasi Pengklasifikasian Hasil Belajar Peserta Didik Dengan Teknik Data Mining Metode K-Nearest Neighbor (K-Nn). *Journal Peqguruang: Conference Series*, *2*(1), 79. https://doi.org/10.35329/jp.v2i1.1386

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques. 3rd ed*. Morgan Kaufmann Publishers.

Hughes, A. (2012). Strategic Database Marketing. 4th ed. *McGraw Hill Professional*.

Hung, M. C., Wu, J., Chang, J. H., & Yang, D. L. (2005). An efficient k-means clustering algorithm using simple partitioning. *Journal of Information Science and Engineering*, *21*(6), 1157–1177.

Imantika, D., Bachtiar, F. A., & Rokhmawati, R. I. (2019). Penerapan metode k-means clustering dan analytical hierarchy process (ahp) untuk pengelompokan kinerja guru dan karyawan pada sma brawijaya smart school. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer J-PTIIK*, *3*(8), 7382–7390. http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/5958

Lopis, M. Y. (2016). *Rancang Bangun Sistem Penilaian Kinerja Guru Studi Kasus : SMK N 1 Bancak Peneliti : Program Studi Pendidikan Teknik Informatika dan Komputer Fakultas Teknologi Informasi Universitas Kristen Satya Wacana*. *702011057*.

Madhulatha, S. (2012). An overview of clustering methods. In *IOSR Journal of Engineering* (Vol. 2, Issue 4, pp. 719–725). https://doi.org/https://doi.org/10.48550/arXiv.1205.1117

Muhiddinur, K. (2019). *Guru : Suatu Kajian Teoritis dan Praktis* (viii). AURA. http://repo.iainbukittinggi.ac.id/id/eprint/131

Ndehedehe, C., Simeon, O., & Ekpa, A. (2013). Spatial Image Data Mining Using K-Means Analysis: A Case Study of Uyo Capital City, Nigeria. *International Journal of Advanced Research*, *1*(8), 1–6.

Nurzahputra, A., Muslim, M. A., & Khusniati, M. (2017). Penerapan Algoritma K-Means Untuk Clustering Penilaian Dosen Berdasarkan Indeks Kepuasan Mahasiswa. *Techno.Com*, *16*(1), 17–24. https://doi.org/10.33633/tc.v16i1.1284

Ong, J. O. (2013). Implementasi Algotritma K-means clustering untuk menentukan strategi marketing president university. *Jurnal Ilmiah Teknik Industri*, *vol.12, no*(juni), 10–20. https://doi.org/https://doi.org/10.23917/jiti.v12i1.651

Panjaitan, M., & Sitompul, D. (2015). Implementasi Algoritma K-Means Dan Analytic Hierarchy Process (AHP) Untuk Klasterisasi Guru Dan Memilih Guru Terbaik (Studi Kasus : SMA Santo Yoseph Medan). *Fasilkom-Ti Usu*, *1–13*.

Pribadi, W. W., Yunus, A., & Sartika Wiguna, A. (2022). Perbandingan Metode K-Means Euclidean Distance Dan Manhattan Distance Pada Penentuan Zonasi Covid-19 Di Kabupaten Malang. *Jurnal Mahasiswa Teknik Informatika)*, *6*(2), 493–500.

Saranya, & Punithavalli. (2011). An Efficient Centroid Selection Algorithm for K-Means Clustering. *International Journal of Management, IT and Engineering*, 130–140.

Sartika, D., & Jumadi, J. (2019). *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS) Clustering Penilaian Kinerja Dosen Menggunakan Algoritma K-Means (Studi Kasus: Universitas Dehasen Bengkulu)*. 703–709. https://seminar-id.com/semnas-sainteks2019.html

Schuh, G., Reinhart, G., Prote, J. P., Sauermann, F., Horsthofer, J., Oppolzer, F., & Knoll, D. (2019). Data mining definitions and applications for the management of production complexity. *Procedia CIRP*, *81*, 874–879. https://doi.org/10.1016/j.procir.2019.03.217

Sukrianto, D. (2016). Penerapan Data Mining Untuk Kinerja Dosen Menggunakan Metode K–Means Clustering (Studi Kasus Di Amik Mahaputra Riau). *Jurnal PI-Cache*, *5, No 1*(Dm), 54–63.

Yaniar, N. S. (2011). Perbandingan Ukuran Jarak pada Proses Pengenalan Wajah Berbasis Principal Component Analysis (PCA). *Proceeding Seminar Tugas Akhir Jurusan Teknik Elektro FTI-ITS*, 1–6.