SMATIKA: STIKI Informatics Journal
Accredited Sinta 4 by the Ministry of Higher Education, Science and Technology Republic of Indonesia Decree No.79/E/KPT/2023
E-ISSN: 2580-6939 || P-ISSN: 2087-0256

smatika
STIKI Informatika Jurnal

Implementation of K-Means Clustering with Attribute Adjustment and Cluster Validation in a Web-Based TPQ Student Assessment Information System

Aura Firdausiyah[1], Sri Widoyoningrum[2]

[1,2]Hasyim Asy'ari University, Faculty of Information Technology, Information Systems Study Program, Irian Jaya Street No. 55, Cukir, Diwek District, Jombang Regency, East Java 61471, Indonesia

Keywords

Grouping; Information System; K-Means; Student Assessment; TPQ

Corresponding Author:
aurafirdausiyah2@mhs.unhasy.ac.id

Abstract

This study aims to implement K-Means Clustering with attribute-level adjustment in a web-based TPQ student assessment information system and evaluate the quality of the resulting clustering results. The study was conducted at TPQ Ashfiya' Kauman Ngoro Jombang using a dataset of 40 students at the Al-Qur'an level obtained from learning assessment data. The variables used include fashohah, tajwid, ghorib and musykilat, as well as voice and melody. The clustering process was carried out using the K-Means algorithm to group students based on similarity in scores, while attribute adjustment was applied in the system design to prevent unused assessment components from being interpreted as low ability scores. Evaluation of cluster quality was carried out using the Sum of Squared Error (SSE), Silhouette Score, and Davies-Bouldin Index (DBI). Based on the evaluation results, K=3 was selected because it formed an interpretable grouping structure and was supported by the Elbow Method. The results show that the application of the K-Means method in the information system can help group students' abilities in a more structured and objective manner and support assessment data management by TPQ teachers. However, the results of this study are a case study in the TPQ environment studied and have not been intended for generalization to a wider population.

1. Introduction

The development of information technology has had a significant impact on various fields, including education. The use of information systems in educational institutions can increase the effectiveness of data management, speed up administrative processes, and support more accurate decision-making [1]. Web-based information systems also enable real-time, integrated, and more efficient data management compared to manual processes [2].

Al-Qur'an Education Park (TPQ), or Al-Qur'an Learning Center, is a non-formal educational institution that plays a role in developing the ability to read the Qur'an from an early age. One of the learning methods used in TPQ is the Tilawati method, which emphasizes assessment aspects such as fluency, tajwid, ghorib and musykilat, as well as voice and melody [3]. Student assessment is an important component in determining students' abilities and learning development.

Cite: Firdausiyah, A., et al., (2026). Implementation of K-Means Clustering with Attribute Adjustment and Cluster Validation in a Web-Based TPQ Student Assessment Information System. SMATIKA : STIKI Informatika Jurnal,16(2). doi: https://doi.org/10.32664/smatika.v16i02.2386

However, at the Ashfiya' Kauman Ngoro Jombang TPQ, the assessment process is still conducted manually using assessment books. This slows down data processing and potentially leads to recording errors. Furthermore, study group assignments are still subjective based on teacher observations, so the resulting groupings are not entirely objective and consistent.

These problems indicate the need for a data-driven approach to group students more objectively. One technique that can be used is data mining with a clustering approach. Clustering is a technique for grouping data based on the level of similarity of certain characteristics [4].

Some commonly used clustering algorithms include K-Means, Hierarchical Clustering, DBSCAN, and Fuzzy C-Means. K-Means has advantages in terms of simplicity, computational efficiency, and suitability for medium to large sized numerical datasets [5].

Previous research shows that K-Means has been widely used in education. Prior studies applied K-Means to group students' tahfidz and tahsin abilities [6]. Other work used K-Means to segment students' abilities based on academic grades [7]. Related research also applied web-based K-Means to cluster Qur'an learning groups [8] . However, these studies still have limitations, namely not considering differences in attribute characteristics at each learning level and not conducting comprehensive cluster evaluations.

In the Tilawati method, each learning level has a different assessment component. Certain attributes are not used at certain levels, resulting in a value of 0 in the dataset. If all attributes are processed without adjustment, a value of 0 can be misinterpreted by the clustering algorithm as low ability, which can bias the clustering results.

Furthermore, most previous studies have focused solely on the clustering process without comprehensively validating the clustering quality. Therefore, evaluations using the Sum of Squared Error (SSE), Silhouette Score, Davies-Bouldin Index (DBI), and Elbow Method are necessary to ensure optimal and objective clustering quality.

1.1 Research Gaps and Contributions

Based on a review of previous studies, several gaps have been identified. First, previous research has not considered the differences in assessment attribute characteristics at each Tilawati learning level, so all data is treated homogeneously in the grouping process. This can lead to bias, as zero scores may stem from unused attributes, rather than low student ability.

Second, some previous studies have focused on the application of clustering methods to generate student groups, but not many have integrated this grouping process into a web-based information system that can be used directly in managing assessment data and making learning decisions in TPQ environments. Therefore, this study implements the K-Means method in a web-based information system to support the process of grouping student abilities in a more structured and objective manner.

Third, most studies only focus on the grouping process without strong integration into a web-based information system that can be used directly by TPQ administrators and teachers to manage assessment data in real time.

Fourth, the evaluation of clustering results in previous studies is still limited, so it does not provide comprehensive validation of the quality of the resulting clusters.

Based on these gaps, this study contributes in several ways. First, this study applies the K-Means Clustering method with attribute adjustments based on Tilawati's learning level to reduce potential data bias. Second, this study integrates the clustering method into a web-based TPQ student assessment information system so that it can be used directly in the administration and learning evaluation process. Third, this study evaluates cluster quality using SSE, Silhouette Score, Davies-Bouldin Index (DBI), and the Elbow Method to obtain more objective and valid clustering results.

Therefore, this research is expected to contribute to increasing the objectivity of student grouping, assisting TPQ teachers in determining learning strategies that suit students' abilities, and strengthening the application of data mining in web-based educational information systems.

2. Research methods

This research was conducted at TPQ Ashfiya' Kauman Ngoro Jombang, a Qur'anic educational institution based on the Tilawati method. This study aims to implement the K-Means method in a web-based student assessment information system to support more objective and systematic grouping of student abilities.

2.1 Data collection

Data were collected using several techniques, including observation, interviews, and documentation. Observations were conducted directly during the learning and assessment process at the TPQ. Interviews were conducted with teachers to obtain information about the assessment process and student study group assignments. Documentation was conducted by collecting student grade data used in the grouping process.

2.2 Research Dataset

The research dataset used for clustering evaluation consisted of 40 students at the Al-Qur'an level at TPQ Ashfiya' Kauman Ngoro Jombang. The dataset was selected because students at this level have complete assessment components in the Tilawati method, allowing the clustering process to use the same set of attributes across all alternatives.

Although the dataset is limited to one TPQ and one learning level, the assessment scores still show variation in students' Qur'anic reading ability. Therefore, the dataset is sufficient for demonstrating the implementation of K-Means Clustering and cluster validation in the developed information system.

The assessment variables used include fashohah, tajwid, ghorib and musykilat, as well as voice and melody. In the system design, attribute-level adjustment is provided because each Tilawati learning level may use different assessment components. This adjustment ensures that unused attributes at certain levels are not interpreted as low ability scores by the clustering algorithm.

Table 1 presents the assessment variables based on student learning levels:

Table 1. Assessment variables

Learning Level	Assessment Variables
Volumes 1–2	Fashohah, Tajwid
Volumes 3–5	Fashohah, Tajwid, Voice and Melodys
Volume 6 and the Qur'an	Fashohah, Tajwid, Ghorib and Musykilat, Voice and Melodys

Attribute adjustments were made because some assessment components are not implemented at certain Tilawati learning levels. In the system, a score of 0 for unused attributes is treated as a non-applicable value rather than as a low ability score. This mechanism is intended to reduce bias when the system is used for other learning levels, while the empirical clustering evaluation in this study focuses on complete Al-Qur'an level data.

2.3 Data Preprocessing

Data preprocessing is performed before the clustering process to ensure the data meets the analysis requirements. This stage includes data completeness checks, attribute selection, and data format adjustments to ensure it can be processed using the K-Means algorithm.

The dataset used in this study consisted of assessment data from 40 students at the Al-Qur'an level at TPQ Ashfiya' Kauman Ngoro Jombang. The variables used included fashohah, tajwid, ghorib and musykilat, as well as voice and melody. Because all empirical data came from the same learning level, all four attributes were processed consistently in the K-Means calculation.

The pre-processing stage is carried out to ensure that the data used has a consistent structure so that the clustering process can produce groupings that are more representative of the students' abilities.

2.4 Research Stages

The research stages were conducted systematically using the System Development Life Cycle (SDLC) approach, starting from problem identification to system testing. In this study, each SDLC stage was implemented as follows :

1. Problem identification was carried out to analyze the manual assessment and grouping process at TPQ Ashfiya' Kauman Ngoro Jombang. At this stage, the researchers identified the main problems, including slow data processing, potential recording errors, and subjective study group determination. The results of this stage became the basis for defining the need for a web-based assessment and clustering system.

2. Data collection was carried out by obtaining student assessment data used as the research dataset. The data were collected through documentation of TPQ assessment records and confirmed through interviews with teachers. The collected data were then checked to ensure that the attributes were relevant to the Tilawati assessment components used in the clustering process.

3. System requirements analysis was carried out to determine the functional requirements and data requirements of the information system. The functional requirements included login, student data management, score input, clustering processing, cluster result display, and reporting. The data requirements included student identity, learning level, assessment scores, cluster labels, and grouping results.

4. System design was conducted by preparing the system flow, database structure, user interface layout, and clustering process flow before implementation. The design stage ensured that assessment data could be entered, processed, and displayed systematically. The system design also accommodated attribute-level adjustment so that non-applicable assessment components would not bias the clustering process.

5. The implementation of the K-Means method was carried out by integrating the clustering algorithm into the web-based system. Student assessment scores were used as input data, and the system calculated distances using Euclidean Distance, updated centroids iteratively, and assigned students to the nearest cluster. The final cluster labels were then mapped into learning groups that could be interpreted by TPQ teachers.

6. Clustering evaluation was conducted using four methods: SSE, Silhouette Score, Davies-Bouldin Index (DBI), and the Elbow Method. SSE was used to measure within-cluster compactness, Silhouette Score to evaluate cluster separation, and DBI to assess cluster quality based on intra-cluster and inter-cluster distances. The Elbow Method was applied by comparing SSE values across K=2, K=3, and K=4 to determine the most appropriate number of clusters for system implementation.

7. System testing was carried out using Black Box Testing and User Acceptance Testing (UAT). Black Box Testing was used to verify whether each feature produced the expected output under valid and invalid input conditions. UAT was used to obtain an initial indication of user acceptance regarding ease of use, interface clarity, processing speed, functional suitability, and user satisfaction.

2.5 K-Means Method

K-Means is a clustering algorithm used to group data based on the level of similarity of certain characteristics. This algorithm works by dividing data into several clusters based on the shortest distance to the cluster centroid [9].

The stages of the K-Means method are as follows :

1. Determine the number of clusters (K).
2. Determine the initial center point (centroid) randomly.
3. Calculate the distance of each data point to the center point (centroid) using Euclidean Distance.
4. Assign data to the cluster with the shortest distance.
5. Calculate the new centroid based on the average of the cluster members.
6. Repeat the calculation process until the center point (centroid) no longer changes.

Calculation of distance using Euclidean distance as follows:

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

Information:
d = distance between data point and center point (centroid)
xi = data value
yi = centroid value
n = number of attributes

2.6 Grouping Evaluation

Clustering evaluation was carried out using Sum of Squared Error (SSE), Silhouette Score, Davies-Bouldin Index (DBI), and Elbow Method to determine cluster quality [10],[11].

2.6.1 Sum of Squared Errors (SSE)

The SSE is used to measure the proximity of data points to the cluster centroid. The smaller the SSE value, the better the cluster quality.

$$SSE = \sum (x_i - c_i)^2 \quad (2)$$

2.6.2 Silhouette Score

Silhouette Score is used to measure the level of similarity of data in the same cluster compared to other clusters. Silhouette Score ranges from -1 to 1. Values closer to 1 indicate better cluster quality [12].

2.6.3 Davies-Bouldin Index (DBI)

The Davies-Bouldin index is used to measure the degree of separation between clusters. A lower DBI value indicates better cluster quality.

2.6.4 Elbow Method

The Elbow method is used to determine the optimal number of clusters by comparing the SSE values across several clusters. The point at which the SSE decline begins to slow down indicates the optimal number of clusters [13].

2.7 System Implementation

The student assessment information system was developed as a web-based system using PHP and MySQL. The system includes features for managing student data, entering grades, processing groupings, and generating student grouping reports.

1) Login Page

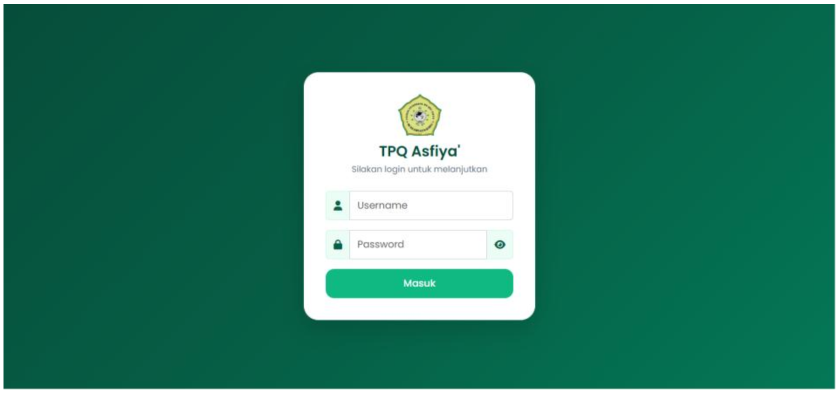

Figure 1. Login Page TPQ Student

The login page is the initial access to the TPQ student grouping information system. Users must enter their registered username and password for authentication.
If the data is correct, the user will be redirected to the main page (dashboard). If incorrect, the system will deny access and display an error message. This page serves as a security measure for system access.

2) Dashboard Page

Figure 2. Dashboard Page TPQ Student

The dashboard is the main page displayed after a user successfully logs in. This page serves to present a summary of the main information of the system as a whole.
The information displayed includes the number of students, grades, clustering results, and active volumes. The system also displays data visualizations in graphical form to provide an overview of student data distribution and clustering results.
With the dashboard, users can obtain important information quickly and concisely.

3) Student Data Page

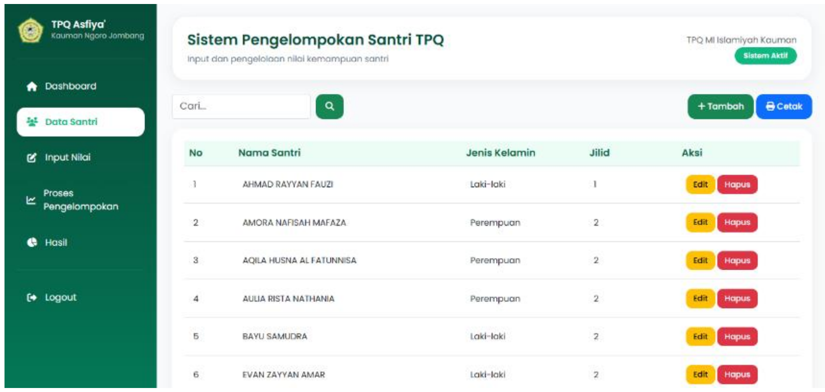

Figure 3. Student Data Page TPQ Student

The student data page is used to manage student information, which forms the basis for the assessment and grouping process using the K-Means method. On this page, administrators can manage data, such as adding, changing, deleting, and searching student data.
Student data is displayed in a structured manner in tables containing student identity information. A search feature is also provided to facilitate fast and efficient data access.
This page supports the data management process to be more organized and effective in supporting the clustering process.

4) Value Input Page

Figure 4. Value Input Page TPQ Student

The score input page is used to manage student assessment data, which forms the basis for the grouping process using the K-Means method. The data managed includes several assessment aspects, such as fluency, tajwid, ghorib and musykilat, as well as Voice and Melody.
The entered grades are used as input for the automatic student clustering process. Furthermore, the system ensures that the entered data complies with the applicable assessment requirements at the TPQ.
This page supports more structured and efficient management of grade data in the process of analyzing and grouping students.

5) Clustering Process Page

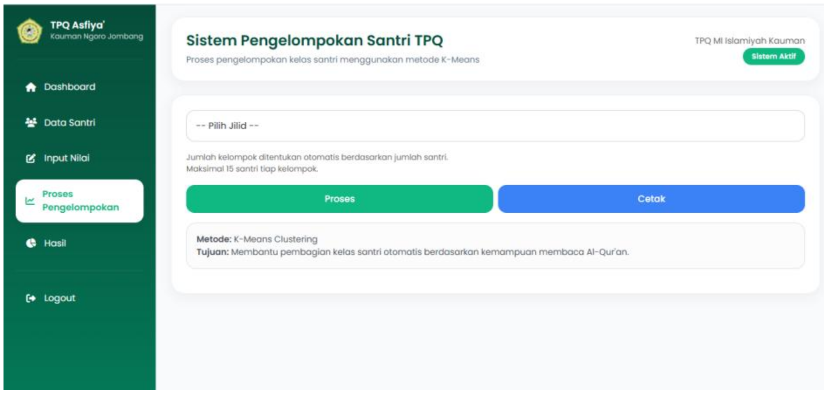

Figure 5. Clustering Process Page TPQ Student

The grouping process page is used to cluster student data using the K-Means method based on the inputted grades. The grouping process is carried out automatically by the system to generate student groups based on their Quranic reading ability.
The results of the process are displayed in the form of a table showing the grouping results and a summary of the number of members in each cluster.
This page helps in producing a more structured division of student groups according to their respective ability levels.

6) Results Page

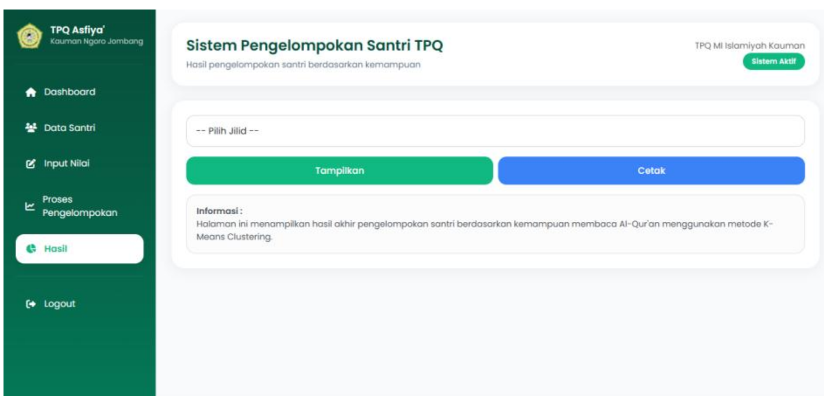

Figure 6. Result Page TPQ Student

The clustering results page displays the final results of the student data clustering process using the K-Means method. The results are a division of student groups based on their Quran reading ability.
The grouping results are presented in the form of a table containing the names of students, clusters, and study groups, as well as visualization in the form of graphs to provide an overview of data distribution.
This page makes it easier for users to understand clustering results in a clearer, more structured, and more informative way.

2.8 System Testing

System testing is conducted using Black Box Testing and User Acceptance Testing (UAT). Black Box Testing is used to ensure that system functions operate according to user requirements. UAT is used to measure user acceptance of the system based on ease of use, system interface, clustering speed, and functional suitability.

3. Results and Discussion

3.1 Research Data

The research dataset consists of data from students at the Tilawati (Quranic) learning level. This dataset was used to analyze the assessment characteristics of students based on aspects of fluency, tajwid, ghorib (sounding), and musykilat (sounding), as well as Voice and Melody, according to the Tilawati (Quranic) learning level.

Table 2. Data on Students at the Al-Qur'an Level

No	Name	Fashohah	Tajwid	Ghorib & Musykilat	Voice and Melody
1	ACHMAD AFIFURROHMAN ANTOVANO	75	73	99	86
2	ADELYA NUR FADHZILLAH	86	89	99	86
3	AHMAD DAFFA HAMIDUN MAJID SUHARTOYO	79	56	99	86
4	AHMAD RAMADHONI	89	86	99	86
5	AHMAD RIJAL DWI RAMADHAN	96	91	99	86
6	AHMAD RISQIANSYAH HABIBI	86	83	90	86
7	AHMAD ZAHID	93	94	99	86
8	AHMADAHU FAWWAS AL MALIKI	96	98	99	86
9	AINUN NAJIYYA	99	87	90	86
10	AISYAH HUMAIRA	86	91	99	86
11	AL-AZIZAH PUTRI	77	97	90	86
12	ALIMUL HAKIM FIL HISYAM	75	83	80	99
13	AQIFA NAYLA R	93	98	99	86
14	AULIA MAWELLA	75	84	99	86
15	AULIYAH PUTRI ARDHA	89	96	90	86
16	AYRA RIFDA	96	93	99	86
17	AZIZ NUR HAFIDHOH	82	98	90	86
18	BILQIS NUR AZIZAH	75	88	90	86
19	BINTANG AHMAD	72	76	90	86
20	DAFFA ALFARIZI	82	93	99	86
21	FATHUR RAHMAN	68	84	90	86
22	FATIMAH ZAHRA	93	97	99	86
23	FAUZAN ADHIMA	96	96	90	60
24	KHANZA QAMILA	89	76	60	86
25	MAYLA ROSITA	73	91	99	86
26	MUHAMMAD ALFATH NUR HAFIDDZ	75	73	99	86
27	MUHAMMAD RAYYAN ALFAREZI	89	90	99	99
28	MUHAMMAD THORIK SULTAN HAKIM	54	98	99	71
29	MUTROVINA SALSABILA PUTRI	82	67	99	86
30	NABILA PUTRI RAMADHANI	77	81	99	86
31	NAURA FITRI RAMADHANI	93	96	99	86
32	NUR AISYAH PUTRI	72	78	80	86
33	RAFIF AKBAR PRATAMA	96	96	90	86
34	SALSABILA AZZAHRA	79	85	99	86
35	SASKI NADIA	89	96	99	86
36	SEKAR ALIYA MAHESTARI	89	81	99	86
37	SUSENINA	72	79	80	86
38	TALITHA YUMNA RAUDHATUL JANNAH	79	80	99	86
39	ZAHRA NUR AULIAWATI	93	93	99	86
40	ZAHRA SYAKHILA SHOLEKHANI	89	96	90	86

3.2 Clustering Process Using K-Means

3.2.1 Cluster Determination

The number of clusters is determined automatically based on the amount of data using the approach:

$$k = \left\lceil \frac{n}{15} \right\rceil \quad (3)$$

With a total of 40 students as data, the value of K = 3 was obtained so that the data was grouped into three clusters.

3.2.2 Determination of Initial Centroid

The initial centroids in this study were determined based on initial data representing the characteristics of the dataset. The centroid selection served as the starting point for the initialization of the K-Means algorithm to begin the clustering iteration process.

Next, the centroid will be updated iteratively based on the average of the members of each cluster until a convergent condition is obtained.

Although this initialization method is simple and does not use optimization techniques such as K-Means++ or repeated random initialization, the final clustering results are still obtained through an iterative process that updates the centroids at each step, thereby reducing the initial influence on the final results.

Table 3 shows the initial centroids used in the clustering process.

Table 3. Initial Centroid

Cluster	Fashohah	Tajwid	Ghorib & Musykilat	Voice and Melody
Cluster 1	75	73	99	86
Cluster 2	86	89	99	86
Cluster 3	79	56	99	86

3.2.3 Distance Calculation Using Euclidean Distance

The distance calculation is done using the Euclidean distance with the following formula:

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad (4)$$

Information:

- $d(x,y)$= distance between data points
- x_i= data value
- y_i= centroid value

3.2.4 Results of Iteration 1

The results of the distance calculation in the first iteration are presented in Table 4.

Table 4. Results of Iteration 1

No	Name	Cluster Distance 1	Cluster Distance 2	Cluster Distance 3	Cluster Results
1	ACHMAD AFIFURROHMAN ANTOVANO	0.00	19.42	17.46	Cluster 1
2	ADELYA NUR FADHZILLAH	19.42	0.00	34.01	Cluster 2
3	AHMAD DAFFA HAMIDUN MAJID SUHARTOYO	17.46	34.01	0.00	Cluster 3
4	AHMAD RAMADHONI	19.10	4.24	31.06	Cluster 2

No	Name	Cluster Distance 1	Cluster Distance 2	Cluster Distance 3	Cluster Results
5	AHMAD RIJAL DWI RAMADHAN	27.66	10.20	39.12	Cluster 2
6	AHMAD RISQIANSYAH HABIBI	15.26	11.66	29.21	Cluster 2
7	AHMAD ZAHID	27.59	8.94	40.31	Cluster 2
8	AHMADAHU FAWWAS AL MALIKI	32.65	13.45	45.61	Cluster 2
9	AINUN NAJIYYA	29.41	16.64	37.85	Cluster 2
10	AISYAH HUMAIRA	21.26	2.00	36.14	Cluster 2
40	ZAHRA SYAKHILA SHOLEKHANI	27.87	7.62	42.94	Cluster 2

3.2.5 Centroid Update

After the first iteration is complete, a new centroid is calculated based on the average value of the members in each cluster.

The latest centroid results are shown in Table 5.

Table 5. New Centroid Results

Cluster	Fashohah	Tajwid	Ghorib & Musykilat	Voice and Melody
Cluster 1	74.87	84.20	89.60	85.87
Cluster 2	91.14	93.10	96.43	85.38
Cluster 3	77.75	67.25	99.00	86.00

3.2.6 Final Iteration

In the final iteration, the system recalculates the distance of each data point to the new centroid obtained from the previous iteration. The calculation uses Euclidean distance to determine the closest distance between the data point and the cluster centroid. The calculation stops at the 6th iteration because the cluster results are the same as the previous iteration.

The final iteration results are shown in Table 6.

Table 6. Final Iteration Results

No	Name	Cluster Distance 1	Cluster Distance 2	Cluster Distance 3	Cluster Results
1	ACHMAD AFIFURROHMAN	14.62	25.91	6.37	Cluster 3
2	ADELYA	15.34	7.09	23.26	Cluster 2
3	AHMAD DAFFA	30.01	39.12	11.32	Cluster 3
4	AHMAD RAMADHONI	17.07	7.87	21.87	Cluster 2
5	AHMAD RIJAL	24.11	5.91	29.95	Cluster 2
6	AHMAD RISQIANSYAH	11.21	13.04	19.93	Cluster 1
7	AHMAD ZAHID	22.65	3.36	30.79	Cluster 2
8	AHMADAHU FAWWAS	26.93	7.39	35.76	Cluster 2
9	AINUN NAJIYYA	24.30	11.86	30.37	Cluster 2
10	AISYAH HUMAIRA	16.08	6.15	25.14	Cluster 2
40	ZAHRA SYAKHILA	18.42	7.40	32.16	Cluster 2

3.2.7 Centroid Update in Final Iteration

After the final iteration is complete, the system recalculates new centroids based on the average value of the members in each cluster.

The centroid results from the second iteration are shown in Table 7.

Table 7. Final Iteration Centroid Results

Cluster	Group	Fashohah	Tajwid	Ghorib & Musykilat	Voice and Melody
Cluster 1	Al-Qur'an B	74.87	84.20	89.60	85.87
Cluster 2	Al-Qur'an A	91.14	93.10	96.43	85.38
Cluster 3	Al-Qur'an C	77.75	67.25	99.00	86.00

3.2.8 Interpretation of Grouping Results

Based on the final centroid, each cluster has distinct characteristics. Cluster 2 (Qur'an A) shows the highest average score, thus representing students with better Quran reading skills. Cluster 1 (Qur'an B) is at an intermediate level, while Cluster 3 (Qur'an C) shows lower value variation across several attributes.

Cluster naming is done based on the final centroid value ranking automatically by the system.

3.2.9 Final Grouping Results

Based on the K-Means clustering process, the results of student clustering are as follows:

Table 8. Final Grouping Results

Cluster	Group	Number of Students
Cluster 1	Al-Qur'an B	15
Cluster 2	Al-Qur'an A	21
Cluster 3	Al-Qur'an C	4

3.2.10 Visualization of Grouping Results

To clarify the clustering results, visualization was performed based on the final centroid values of each cluster. The visualization was used to show the differences in characteristics between clusters based on the assessment attributes used in the study, namely fluency, tajwid, and ghorib and musykilat.

Based on the visualization results, Cluster 2 (Qur'an A) shows the highest centroid values for most attributes, thus representing a group of students with relatively better Qur'an reading abilities. Cluster 1 (Qur'an B) shows intermediate abilities, while Cluster 3 (Qur'an C) shows different value variations for several assessment attributes.

This visualization is used to help interpret clustering results and show the differences in characteristics between clusters more clearly than tabular presentation alone.

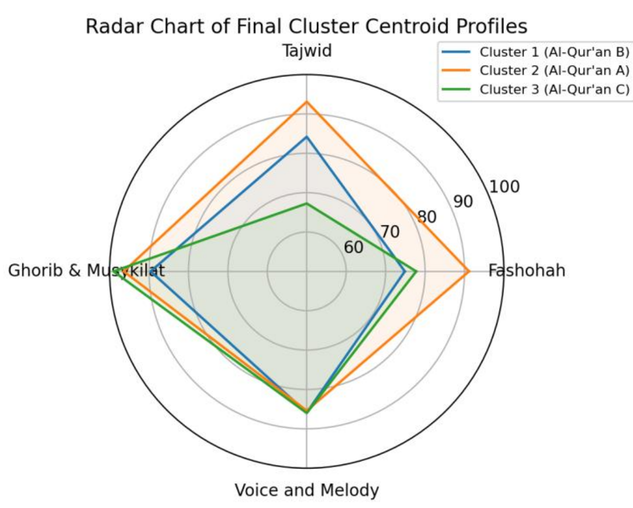

Figure 7. Radar chart of final cluster centroid profiles

3.3 Grouping Evaluation

The evaluation of the number of clusters was conducted using all 40 student data at the Al-Qur'an level. The evaluation was conducted to determine the most appropriate number of clusters based on the quality of the resulting groupings. The evaluation methods used included the Sum of Squared Error (SSE), Silhouette Score, and Davies-Bouldin Index (DBI).

3.3.1 Sum of Squared Errors (SSE)

The Sum of Squared Errors (SSE) is used to measure the closeness of the data to the centroid in each cluster. A smaller SSE value indicates that the cluster members are more homogeneous.

$$SSE = \sum_{i=1}^{n} (x_i - c_i)^2 \qquad (5)$$

Based on the test results using the K-Means method on a dataset of 40 students, the results obtained are as in Table 9.

Table 9. SSE Calculation Results

K	SSE
2	2783.52
3	1861.74
4	1348.91

Based on these results, the SSE value continues to decrease as the number of clusters increases. The largest decrease occurs from K=2 to K=3 and begins to slow down thereafter, indicating an elbow point at K=3.

3.3.2 Silhouette Score

The Silhouette Score is used to measure the proximity of data within a cluster and the separation between clusters. The Silhouette Score ranges from -1 to 1. Values closer to 1 indicate better cluster quality.

The Silhouette Score results are shown in Table 10.

Table 10. Silhouette Score Calculation Results

K	Silhouette
2	0.421
3	0.398
4	0.361

Based on the evaluation results, the highest Silhouette Score value was obtained at K=2. This value indicates that cluster separation at K=2 is relatively better compared to other cluster numbers.

3.3.3 Davies-Bouldin Index (DBI)

The Davies-Bouldin index is used to measure the degree of similarity between clusters. The lower the DBI value, the better the cluster quality because the distance between clusters is more pronounced.

The DBI test results are shown in Table 11.

Table 11. Davies-Bouldin Index Calculation Results

K	DBI
2	0.812
3	0.903
4	0.776

Based on the evaluation results, the smallest DBI value was obtained at K=4, which indicates a better level of separation between clusters.

3.3.4 Evaluation of the Elbow Method

The Elbow method is used to determine the optimal number of clusters based on the pattern of SSE value decline. The number of clusters is determined by observing changes in the rate of SSE decline at each K value.

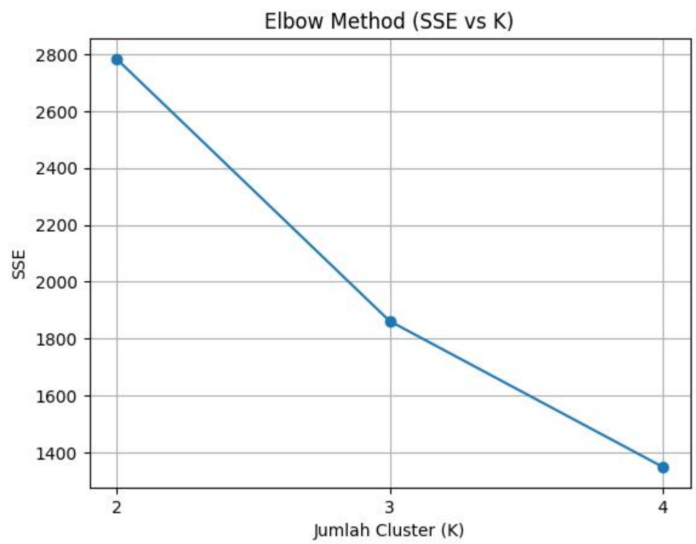

Figure 8. Elbow Method Graph (SSE vs. Number of Clusters)

The graph shows that the SSE value decreased from 2,783.52 at K=2 to 1,861.74 at K=3 and again to 1,348.91 at K=4. The largest decrease occurred from K=2 to K=3, while after K=3 the rate of decrease began to slow down.

To clarify the interpretation of the results, an Elbow graph is used which shows the relationship between the number of clusters and the SSE value.

The graph shows a decrease in SSE values at K=2, K=3, and K=4. The most significant decrease occurs from K=2 to K=3, then begins to level off at K=3 to K=4, so that K=3 is interpreted as an elbow point.

The overall cluster evaluation results are shown in Table 12.

Table 12. Comparison of Grouping Evaluation Results

K	SSE	Silhouette Score	DBI	Interpretation
2	2,783.52	0.421	0.812	Better cluster separation based on Silhouette Score
3	1,861.74	0.398	0.903	Shows elbow point based on SSE decline pattern
4	1,348.91	0.361	0.776	Provides the best DBI value

The evaluation results show that each method produces different recommendations for the number of clusters. The Elbow method, using SSE, shows a change point at K=3, the Silhouette Score shows the best separation quality at K=2, and the Davies-Bouldin Index shows the best value at K=4.

In this study, the number of clusters used was set at K=3 because it was chosen based on the results of the Elbow method and considering the ease of interpretation of the clustering results in the system being developed.

3.4 Cluster Interpretation and Educational Implications

The clustering results show that the K-Means method can group students based on their Quranic reading ability more objectively using assessment data. Each group has different ability characteristics in aspects of fluency, tajwid, ghorib and musykilat, as well as voice and melody.

Based on the final centroid results, Cluster 2 had a higher average score than Cluster 1 in almost all assessment attributes. This indicates that students in Cluster 2 have better Quran reading skills, both in reading fluency, Tajwid accuracy, and Tilawati melody quality. Meanwhile, Cluster 1 showed a lower average score and therefore requires further guidance in several basic aspects of Quran reading.

The largest differences between groups were seen in the fashohah and tajwid attributes. Students in Group 1 tended to have lower fashohah scores than students in Group 2. This indicates that some students still experience difficulties in reading fluency, pronouncing the hijaiyah letters, and correct pronunciation of the makharijul letters. Furthermore, lower tajwid scores indicate that the application of reading rules such as mad, ghunnah, ikhfa, and idgham still needs improvement.

In the attributes of ghorib and musykilat, as well as voice and melody, the differences between groups were not as significant as the differences in fluency and tajwid. This indicates that basic Quranic recitation skills are the primary factor that most strongly influences the grouping process. Therefore, fluency and tajwid can be considered the attributes that most clearly differentiate students' abilities in this study.

The clustering results can be used as a basis for forming student study groups at TPQ (Teaching and Religious Education Institutions) so that the learning process can be tailored to the needs of each group. This grouping helps teachers determine the focus of mentoring based on the characteristics of Quranic reading abilities indicated by the clustering results.

Students in Cluster 2 can be provided with learning reinforcement through activities to improve reading quality, practice the Tilawati song more consistently, deepen their understanding of the ghorib and musykilat material, and regularly evaluate their reading to maintain learning outcomes. Learning in this group can be directed toward improving the accuracy and quality of their reading.

Meanwhile, students in Cluster 1 can receive more intensive support through gradual reading exercises, reading repetition (drilling), pronunciation exercises for the letters, reinforcement of basic tajwid (recitation), and repeated reading habits in small groups. Teachers can also conduct periodic progress evaluations based on assessment results obtained during each learning period.

The application of grouping results helps TPQ teachers develop learning strategies more objectively than manual grouping based solely on observation. Integrating a web-based information system with the K-Means method also supports faster, more structured, and more consistent student ability evaluation.

These findings are consistent with prior work showing that the K-Means method is effective for grouping students according to the similarity of their assessment scores [6]. The results also align with studies reporting that K-Means clustering can segment students more objectively than manual, observation-based assessment [7]. This study extends that body of work by pairing the clustering process with a four-indicator evaluation protocol, SSE, Silhouette Score, Davies-Bouldin Index, and the Elbow Method, so that cluster quality is examined more systematically than in prior studies that report clustering outcomes without a comparable validation procedure [17].

The disagreement among the four evaluation indicators observed in this study is not unique to the TPQ dataset. Comparative studies of cluster validity measures report that the Elbow Method, Silhouette Score, and Davies-Bouldin Index frequently recommend different numbers of clusters for the same dataset, because each index emphasizes a different property of cluster structure: within-cluster compactness for SSE and the Elbow Method, separation between clusters for the Silhouette Score, and the ratio of intra-cluster to inter-cluster distance for DBI [16]. This pattern reinforces the argument that a single validity index is insufficient for determining an optimal K value, and that combining several indices, as implemented in this study, offers a more defensible basis for selecting K than relying on any one metric in isolation.

The attribute-adjustment mechanism applied in this study addresses a preprocessing issue that is often overlooked in educational clustering research: a non-applicable score can be misread by a distance-based algorithm as evidence of low ability rather than as structurally missing information. This concern parallels findings in the wider clustering literature, where unresolved missing or non-applicable values are shown to distort centroid estimation and bias cluster assignment unless they are explicitly treated during preprocessing [15]. By encoding unused Tilawati attributes as non-applicable rather than as zero-value scores, the system developed in this study reduces this source of bias at the data-management level, complementing existing missing-value handling techniques used in general-purpose K-Means applications.

Beyond the TPQ context, these results add to a growing body of educational data mining research showing that K-Means clustering can meaningfully differentiate learners according to behavioral and performance-based attributes, from disposition analysis in online learning environments [18] to fusion-based clustering of university students' engagement patterns [19] and machine-learning-based grouping of learner performance levels [20]. What the present study contributes to this literature is a demonstration that the same clustering logic can be applied in a non-formal, community-based religious education setting, where assessment components differ across instructional levels and where a purely quantitative clustering procedure must be adapted through domain-specific preprocessing rules. This suggests that the value of K-Means for student grouping is not confined to formal schooling contexts, but extends to non-formal and religiously oriented educational institutions once attribute heterogeneity across learning levels is properly accounted for.

From a theoretical standpoint, these results extend data-driven decision-making frameworks in education, which describe how systematically collected and analyzed data can inform instructional and administrative decisions [23], to a setting that has traditionally relied on subjective, observation-based judgment. Embedding K-Means clustering and its evaluation metrics directly into the assessment information system operationalizes data-driven decision-making at the classroom level: TPQ teachers no longer need to rely solely on informal impressions of student ability, but can draw on a validated, quantitative grouping that is recalculated automatically as new assessment data are entered. This positions the system as a form of lightweight learning analytics for non-formal education, an area in which data-driven decision-making frameworks have so far been discussed primarily in the context of formal schooling.

For practice, the clustering results give teachers a concrete basis for differentiated instruction rather than a generic recommendation to personalize learning. Reviews of differentiated instruction report that ability-based, macro-adaptive grouping is one of the more common organizational strategies teachers use to accommodate diverse learning needs, provided that group assignment is grounded in valid and interpretable data [22]. The fashohah and tajwid subscores identified in this study as the strongest differentiators between clusters give TPQ teachers a specific instructional target: Cluster 1 students benefit from focused drilling on makharijul huruf and basic tajwid rules, while Cluster 2 students are better served by refinement-oriented practice in Tilawati melody and reading fluency. Curriculum coordinators can use the same cluster outputs to allocate teaching time and assign mentors according to each group's specific weaknesses, rather than treating the TPQ cohort as a single homogeneous class.

The usability results reported later in this study should be read alongside these clustering findings, since a clustering-based grouping system only benefits students to the extent that teachers actually adopt it in daily use. Prior research on technology acceptance in educational information systems shows that perceived ease of use and system usability are strong predictors of teachers' and students' intention to continue using a system beyond initial adoption [21]. The relatively high UAT scores obtained in this study, discussed in the following section, suggest that the attribute-adjusted K-Means system is not only analytically sound but also practically acceptable to its intended users, a precondition for the clustering method to have any sustained effect on TPQ learning practice.

This study nonetheless has several limitations. The dataset is limited to 40 students at a single Al-Qur'an learning level in one TPQ, so the clustering results describe the characteristics of this particular cohort rather than TPQ students in general. Because all empirical attributes were fully populated at this learning level, the study could not directly test the attribute-adjustment mechanism against clustering behavior at levels with structurally absent components; this remains an assumption grounded in system design and preprocessing logic rather than an empirically validated evaluation. The disagreement among SSE, Silhouette Score, and DBI in selecting the optimal K also means that the choice of K=3 in this study reflects the Elbow Method combined with interpretability considerations, rather than unanimous agreement across all validity indices. In addition, the initial centroids were selected without an optimization technique such as K-Means++, which may affect the stability of results across repeated runs even though the iterative update process reduces this influence. Finally, the User Acceptance Testing involved only five respondents, following established usability guidance for early-stage evaluation, but this sample size limits the generalizability of the acceptance findings.

Future research could address these limitations in several ways. Replicating the clustering and evaluation procedure across multiple TPQ institutions and across Tilawati learning levels with structurally missing attributes would provide a stronger empirical test of the attribute-adjustment mechanism. Applying centroid initialization techniques such as K-Means++ or repeated random initialization with stability analysis would clarify how sensitive the clustering results are to the starting configuration. Tracking the same cohort of students across multiple assessment periods would allow researchers to examine whether cluster membership is stable over time or shifts as students' Qur'anic reading ability develops, indicating whether the system's groupings support long-term instructional planning rather than a single-point-in-time snapshot. Finally, incorporating a larger and more diverse UAT sample together with a validated technology-acceptance survey instrument would allow the usability findings reported in this study to be tested with the same rigor as the clustering results themselves.

3.5 System Testing

System testing was conducted to ensure that all features within the student assessment information system operated as expected. This study used black box testing and user acceptance testing (UAT).

3.5.1 Black Box Testing

Black Box Testing is performed to verify system functionality based on input and output without directly examining the program code. Testing is performed on key system features such as login, student data management, grade input, clustering, and clustering results reporting.

The results of the Black Box testing are shown in Table 13.

Table 13. Black Box Test Results

No	Testing Scenario	Input	Expected results	Actual Results	Status
1	System login	Valid username and password	The system displays the dashboard	Dashboard successfully displayed	Succeed
2	Input student data	Student identity data	Data saved successfully	Data is stored in a database	Succeed
3	Input student grades	The value of fashohah, tajwid, ghorib and musykilat, as well as sounds and songs	Value saved successfully	Values are stored in the database	Succeed
4	Clustering process	Student grade data	The system displays the cluster results.	Cluster results are successfully displayed	Succeed
5	Print clustering results report	Clustering result data	The system displays the report	Report successfully displayed	Succeed
6	Login with invalid data	Incorrect username or password	The system denies access and displays an error message.	The system displays a login failed message.	Succeed

7	Input student data with blank data	One of the identity columns is not filled in	The system displays a warning and the data is not saved.	The system refuses to save data	Succeed
8	Input value outside the range	The value exceeds the maximum limit or is negative	The system displays validation and rejects data.	The system displays an error message	Succeed
9	Clustering process with insufficient data	The dataset does not meet the clustering requirements.	The system displays a notification that the process cannot be executed.	The system displays a warning to the user	Succeed

Based on the results of the Black Box testing, all of the system's main functions can operate as required under both normal and invalid input conditions. The test results indicate that the system is capable of handling data validation and providing appropriate responses to input errors, thus supporting the reliability of the data management and grouping processes.

3.5.2 User Acceptance Testing (UAT)

User Acceptance Testing (UAT) was conducted to obtain an initial overview of user acceptance of the developed system. Due to the preliminary nature of this study, UAT was conducted with five respondents consisting of TPQ teachers and system users. This number follows established usability guidance indicating that five users can reveal most usability problems in an early-stage evaluation; therefore, the results should be interpreted as indicative rather than statistically generalizable [14].

The assessment was carried out using a Likert scale from 1 to 5 with descriptions as in Table 15.

UAT assessment indicators include:

1. Ease of use of the system
2. System interface
3. Speed of grouping process
4. Functional suitability of the system
5. User satisfaction

The calculation of the UAT percentage is done using the following formula:

$$UAT\ Percentage\ = \frac{Obtained\ Score}{Maximum\ Score} \times 100\% \qquad (6)$$

The UAT results are shown in Table 14.

Table 14. User Acceptance Testing (UAT) Results

NO	Assessment Aspects	Percentage
1	Ease of use of the system	88%
2	System interface	85%
3	Speed of grouping process	90%
4	Functional suitability of the system	92%
5	User satisfaction	89%

Based on the test results, an average percentage of 88.8% was obtained. This result indicates initial acceptance of the system by the respondents involved, particularly in terms of ease of use, functional suitability, and speed of the grouping process. However, these results should be interpreted with caution because the limited number of respondents is not sufficient to draw general conclusions regarding the system's usability.

For further research, it is recommended to involve a larger and more diverse number of respondents so that the results of the system usability evaluation are more representative and have a higher level of reliability.

4. Conclusion

This study successfully implemented the K-Means method in a web-based TPQ student assessment information system to support the process of grouping students' abilities based on the attributes of fluency, tajwid, ghorib and musykilat, as well as Voice and Melody. The integration of the clustering method into the system allows for a more structured assessment data management process and assists decision-making in the formation of study groups.

Based on the evaluation results using the Sum of Squared Error (SSE), Silhouette Score, Davies-Bouldin Index (DBI), and Elbow method, it was found that each method provides a different recommendation for the number of clusters. In this study, K = 3 was used as the number of clusters appropriate to the needs of system implementation and interpretation of the clustering results. The analysis results showed that the attributes of fluency and tajwid were the factors that most differentiated the abilities between groups of students.

The contribution of this research lies in the integrated application of K-Means Clustering with attribute-level adjustment in a web-based TPQ student assessment information system. The study also contributes by validating the clustering results using SSE, Silhouette Score, DBI, and the Elbow Method, so that the grouping decision is supported not only by system implementation but also by cluster quality evaluation.

This research still has limitations because it was conducted on a limited dataset and within the context of a case study within the TPQ environment studied. Therefore, the results are not intended for broader generalization. Future research could focus on testing the stability of clustering results across different assessment periods, developing learning recommendation features based on cluster results, and evaluating the system's implementation in long-term use.

5. References

[1] M. Azuddin and others, “Design and Implementation of Relational Database for Academic Information System,” *Int. Trans. Artif. Intell.*, vol. 3, no. 2, pp. 171–180, 2025, doi: 10.33050/italic.v3i2.801.

[2] L. D. Fitrani and A. C. Puspitaningrum, “Utilization of Unified Modeling Language (UML) in Academic Information System Design Based on the OOAD Method,” *J. Acad. Syst. Inf.*, vol. 12, pp. 614–625, 2023, doi: 10.32520/stmsi.v12i2.2871.

[3] Tilawati Center, “Tilawati Metode Baca Al-Qur’an Terpercaya di Indonesia,” 2026. [Online]. Available: https://tilawati.org/

[4] P. S. Hasugian and J. R. Sagala, “Application of Data Mining for Student Grouping Based on Academic Grades with the K-Means Algorithm,” *Informatics Comput. Sci. Rev.*, vol. 3, no. 3, pp. 262–268, 2022.

[5] I. D. Setiawan and A. Triayudi, “Application of Data Mining Using the K-Means Clustering Algorithm for Major Allocation in Senior High Schools,” vol. 5, no. 2, pp. 380–392, 2024, doi: 10.47065/josyc.v5i2.4970.

[6] S. Dewi and S. Fatimah, “Implementation of the K-Means Algorithm Clustering Method to Determine Tahfidz and Tahsin Groups at the Al-Ma’soem Islamic Boarding School,” vol. 6, no. 1, pp. 10–18, 2023, doi: 10.32627/aims.v6i1.701.

[7] A. Firdaus and J. Heikal, “Student Segmentation Using K-Means Clustering: Case Study of Citamulia Quran Memorization Elementary School (SDTQ),” vol. 5, no. 11, pp. 1420–1430, 2024, doi: 10.59141/japendi.v5i11.5922.

[8] Z. Rahman, “Use of Website-Based K-Means Algorithm in Determining Al-Qur’an Guidance Groups at

the Islamic Boarding School of Syafiiyah Sumberanyar," vol. 6, no. 1, pp. 53–62, 2025, doi: 10.33650/trilogi.v6i1.10666.

[9] S. Natalia, B. Sembiring, H. Winata, and S. Kusnasari, "Grouping Student Achievement Using the K-Means Algorithm," vol. 1, pp. 31–40, 2022, doi: 10.53513/jursi.v1i1.4784.

[10] N. D. Rahayu, A. H. Anshor, and I. Afriantoro, "Application of Data Mining for Mapping High-Achieving Students Using the K-Means Clustering Method," vol. 6, pp. 71–83, 2024, doi: 10.53842/juki.v6i1.474.

[11] M. S. Hasibuan, A. H. Lubis, and M. N. Sari, "Comparison of DBSCAN and K-Means Clustering Algorithms in Grouping the Best Students," vol. 5, pp. 301–309, 2024, doi: 10.37373/infotech.v5i2.1457.

[12] B. Janghel and A. Ambhaikar, "Comparative Study and Analysis of Student Outcomes Using Clustering Technique," vol. 9, no. 2, 2021.

[13] P. Zhou, L. Du, and X. Li, "Adaptive Consensus Clustering for Multiple," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 10, pp. 10251–10264, 2023, doi: 10.1109/TKDE.2023.3264970.

[14] J. Nielsen, *Usability Engineering*. San Francisco, CA: Morgan Kaufmann, 1993. doi: 10.1016/B978-0-08-052029-2.50007-3.

[15] N. Suguna and K. G. Thanushkodi, "Predicting Missing Attribute Values Using k-Means Clustering," *J. Comput. Sci.*, vol. 7, no. 2, pp. 216–224, 2011, doi: 10.3844/jcssp.2011.216.224.

[16] M. Diantika, A. Rusgiyono, and B. Saputra, "Comparison of Silhouette, Elbow, and Gap Statistics Optimization Methods in Determining the Number of Best Clusters in K-Means Clustering Analysis," *Int. J. Math. Comput. Res.*, vol. 13, no. 12, pp. 5953–5958, 2025, doi: 10.47191/ijmcr/v13i12.03.

[17] R. Liu, "Data Analysis of Educational Evaluation Using K-Means Clustering Method," *Comput. Intell. Neurosci.*, vol. 2022, Art. no. 3762431, 2022, doi: 10.1155/2022/3762431.

[18] S. Bharara, S. Sabitha, and A. Bansal, "Application of learning analytics using clustering data mining for students' disposition analysis," *Educ. Inf. Technol.*, vol. 23, no. 2, pp. 957–984, 2018, doi: 10.1007/s10639-017-9645-7.

[19] W. Chang, X. Ji, Y. Liu, Y. Xiao, B. Chen, H. Liu, and S. Zhou, "Analysis of university students' behavior based on a fusion k-means clustering algorithm," *Appl. Sci.*, vol. 10, no. 18, p. 6566, 2020, doi: 10.3390/app10186566.

[20] R. Vankayalapati, K. B. Ghutugade, R. Vannapuram, and B. P. S. Prasanna, "K-Means Algorithm for Clustering of Learners Performance Levels Using Machine Learning Techniques," *Rev. Intell. Artif.*, vol. 35, no. 1, pp. 99–104, 2021, doi: 10.18280/ria.350112.

[21] A. Revythi and N. Tselios, "Extension of technology acceptance model by using system usability scale to assess behavioral intention to use e-learning," *Educ. Inf. Technol.*, vol. 24, no. 4, pp. 2341–2355, 2019, doi: 10.1007/s10639-019-09869-4.

[22] A. E. Smale-Jacobse, A. Meijer, M. Helms-Lorenz, and R. Maulana, "Differentiated Instruction in Secondary Education: A Systematic Review of Research Evidence," *Front. Psychol.*, vol. 10, Art. no. 2366, 2019, doi: 10.3389/fpsyg.2019.02366.

[23] E. B. Mandinach, "A Perfect Time for Data Use: Using Data-Driven Decision Making to Inform Practice," *Educ. Psychol.*, vol. 47, no. 2, pp. 71–85, 2012, doi: 10.1080/00461520.2012.667064.