

P-ISSN: 2774-4574; E-ISSN: 363-4582
TRILOGI, 6(1), Januari-Maret 2025 (53-62)
©2025 Lembaga Penerbitan, Penelitian,
dan Pengabdian kepada Masyarakat (LP3M)
Universitas Nurul Jadid Paiton Probolinggo
DOI: [10.33650/trilogi.v6i1.10666](https://doi.org/10.33650/trilogi.v6i1.10666)



Penggunaan Algoritma K-Means Berbasis Website dalam Menentukan Kelompok Pembinaan Al-Qur'an di Pondok Pesantren Islamiyah Syafiiyah Sumberanyar

Zainur Rahman

Universitas Nurul Jadid, Indonesia
rahmanzainur491@gmail.com

Moh. Ainol Yaqin

Universitas Nurul Jadid, Indonesia
ainolyaqin09@unuja.ac.id

M. Syafiih

Universitas Nurul Jadid, Indonesia
m.syafii@unuja.ac.id

Abstract

This study aims to develop a web-based application system using the K-Means algorithm to cluster students in Al-Quran training at Pondok Pesantren Islamiyah Syafi'iyah Sumberanyar, Paiton, Probolinggo. The system is designed to improve efficiency and accuracy in determining training groups based on students' Quranic reading skills, including Tajwid, fluency, and memorization. The research employs a quantitative approach with a data mining method using the K-Means algorithm. Data from 824 students were collected and processed through the stages of Knowledge Discovery in Databases (KDD), which include data collection, data cleaning, data transformation, and clustering using the K-Means algorithm. The clustering results were evaluated using the Davies-Bouldin Index to measure clustering quality. The findings indicate that the K-Means algorithm significantly enhances efficiency and accuracy in grouping students compared to manual methods. The students were successfully classified into three skill categories: "Ulya" (advanced), "Wustha" (intermediate), and "Ula" (beginner). This clustering process provides an objective guide for assigning students to appropriate training groups and supports the personalization of the learning process. Additionally, the implementation of web-based technology facilitates easier data management and greater accessibility for administrators. These findings are not only relevant to this pesantren but also have the potential to be applied to other Islamic educational institutions to optimize training management.

Keywords: K-Means Algorithm; Clustering; Al-Quran Training; Web-Based Technology; Data Mining.

Abstrak

Penelitian ini bertujuan untuk mengembangkan sistem aplikasi berbasis web menggunakan algoritma K-Means untuk mengelompokkan santri dalam pembinaan Al-Quran di Pondok Pesantren Islamiyah Syafi'iyah Sumberanyar, Paiton, Probolinggo. Sistem ini dirancang

untuk meningkatkan efisiensi dan akurasi dalam menentukan kelompok pembinaan berdasarkan kemampuan membaca Al-Quran santri, meliputi Tajwid, kelancaran, dan hafalan. Metode penelitian yang digunakan dalam penelitian ini adalah pendekatan kuantitatif dengan metode data mining menggunakan algoritma K-Means. Data sebanyak 824 santri dikumpulkan dan diproses melalui tahapan Knowledge Discovery in Databases (KDD), yang mencakup pengumpulan data, pembersihan data, transformasi data, serta proses clustering menggunakan algoritma K-Means. Evaluasi hasil klasterisasi dilakukan menggunakan Indeks Davies-Bouldin untuk mengukur kualitas pengelompokan. Hasil penelitian menunjukkan bahwa algoritma K-Means secara signifikan meningkatkan efisiensi dan akurasi dalam pengelompokan dibandingkan dengan metode manual. Santri berhasil dikelompokkan ke dalam tiga kategori keterampilan: "Ulya" (tinggi), "Wustha" (menengah), dan "Ula" (pemula). Proses klasterisasi ini memberikan panduan objektif dalam menetapkan kelompok pembinaan yang sesuai bagi setiap santri, serta mendukung personalisasi proses pembelajaran. Selain itu, implementasi teknologi berbasis web memungkinkan pengelolaan data yang lebih mudah dan aksesibilitas yang lebih baik oleh administrator. Temuan ini tidak hanya relevan bagi pesantren ini tetapi juga berpotensi diterapkan pada lembaga pendidikan Islam lainnya untuk mengoptimalkan pengelolaan proses pembinaan.

Katakunci: Algoritma K-Means; Klasterisasi; Pembinaan Al-Quran; Teknologi Berbasis Web; Data Mining.

1 Pendahuluan

Pondok Pesantren Islamiyah Syafi'iyah Sumberanyar, yang terletak di Paiton Probolinggo, memainkan peran penting dalam penyampaian pendidikan Islam dengan fokus pada persaudaraan Islam (Syafiih & Hatima Indah Arifin, 2023). Salah satu kegiatan pendidikan utama di Islamiyah Syafi'iyah Sumberanyar adalah pelatihan Al-Quran bagi para santri. Salah satu komponen penting dari pelatihan ini adalah pengelompokan santri berdasarkan keterampilan dan kebutuhan individu mereka. Saat ini, pengelompokan ini dilakukan secara manual, memerlukan pengetahuan mendalam tentang kemampuan masing-masing santri dan tujuan pembelajaran yang berbeda dari kelompok-kelompok yang berbeda. Namun, metode manual rentan terhadap kesalahan manusia dalam penilaian santri dan penugasan kelompok, selain itu juga memakan waktu dan tenaga (Hutagalung, 2022).

Untuk meningkatkan efisiensi dan akurasi penugasan kelompok dalam pelatihan Al-Quran, penelitian ini mengusulkan penggunaan teknologi dan metode analisis data, khususnya data mining untuk mengidentifikasi pola yang relevan dan pengelompokan menggunakan algoritma K-Means (Santi & Februariyanti, 2023). Penelitian ini menekankan perlunya meningkatkan efisiensi dan akurasi penugasan kelompok dalam pelatihan Al-Quran (Br Sembiring et al., 2022). Evaluasi objektif kemampuan santri menimbulkan tantangan karena karakteristik unik masing-masing santri, yang memerlukan pertimbangan seperti kemampuan membaca Al-Quran, pemahaman

Tajwid, dan kelancaran membaca. Selain itu, terdapat berbagai hasil pembelajaran yang diharapkan di antara kelompok-kelompok pelatihan; beberapa santri mungkin memerlukan bantuan tambahan dengan konsep-konsep dasar Al-Quran sedangkan yang lain siap untuk pelajaran yang lebih lanjut (Adillah et al., 2022).

Dalam skenario ini, teknik data mining dan pengelompokan menawarkan metode yang lebih objektif dan efisien untuk penugasan kelompok (Soliani & Juanita, 2022). Analisis data dapat mengungkap pola dalam hasil ujian pelatihan Al-Quran, yang dapat digunakan untuk menilai kemampuan santri (Syafiih et al., 2021). Pengelompokan, terutama dengan algoritma K-Means, memungkinkan pengelompokan homogen berdasarkan karakteristik data yang diidentifikasi (Mirantika, 2021). Implementasi teknologi dan metode analisis data dalam penelitian ini bertujuan untuk mengatasi tantangan yang dihadapi Pesantren Islamiyah Syafi'iyah Sumberanyar. Dengan memanfaatkan teknologi, proses pengelompokan dapat menjadi lebih cepat, lebih akurat, dan lebih objektif, sehingga meningkatkan motivasi santri melalui pengalaman pelatihan yang dipersonalisasi sesuai dengan kebutuhan individu (Esmi Nur Fitri et al., 2023). Rencana implementasi mencakup penggunaan algoritma K-Means untuk analisis data, dan pengembangan antarmuka web dengan PHP untuk kemudahan akses dan penggunaan oleh administrator dan peserta pelatihan (Subecz, 2021). Database SQL akan digunakan untuk memastikan penyimpanan data yang terorganisir dan mudah diakses (Montero, 2022).

Penelitian ini bertujuan untuk mengatasi masalah-masalah khusus di Pesantren Islamiyah Syafi'iyah Sumberanyar dan berkontribusi pada pengembangan sistem pelatihan yang lebih responsif dan efisien di lembaga pendidikan serupa (Likmi & Barat, 2024). Selain itu, wawasan dari studi terkait memperkuat dasar penelitian ini. Sebagai contoh, Ai Rohmah, Falentino Sembiring, dan Adhitha Erfina (2021) menyoroti penggunaan algoritma K-Means untuk mengklasifikasikan tingkat hambatan belajar dalam konteks pembelajaran jarak jauh selama pandemi COVID-19 (Rohmah et al., 2021). Demikian pula, Dian Permata Sari (2021) menerapkan algoritma K-Means untuk menentukan tingkat penyebaran pandemi COVID-19 di Sumatera Barat, mengelompokkan daerah berdasarkan kasus positif dan penyebaran virus (Sari, 2021). Selain itu, Achmad Dimiyati (2023) fokus pada penggunaan K-Means untuk mengevaluasi nilai akademik santri di TPQ Darussalamah berbasis Framework Laravel (Dimiyati, 2023). Studi-studi ini secara kolektif mendukung adopsi teknik data mining seperti K-Means untuk meningkatkan praktik pendidikan (Anggara¹ et al., 2024). Integrasi pengelompokan K-Means ke dalam pengaturan pendidikan telah menunjukkan potensi dalam berbagai studi, termasuk aplikasinya dalam memprediksi kinerja akademis santri, menganalisis data akademik santri, dan mengklasifikasikan prestasi belajar santri (Hendrastuty, 2024). Aplikasi-aplikasi seperti itu mengilustrasikan fleksibilitas K-Means dalam analisis data pendidikan dan potensinya untuk mengatasi tantangan kompleks dalam penilaian santri dan manajemen pendidikan (Palevi & Indra, 2024). Dengan memanfaatkan metodologi ini, lembaga pendidikan dapat menyederhanakan proses, mempersonalisasi pengalaman belajar, dan mendorong keberhasilan akademis (Novita et al., 2023). Penelitian ini membangun fondasi tersebut untuk mengusulkan pendekatan berbasis data guna meningkatkan pelatihan AI-Quran di Islamiyah Syafi'iyah Sumberanyar dan lembaga serupa.

2 Metode

Sistem aplikasi Penentuan Kelompok Pelatihan AI-Quran dikembangkan menggunakan bahasa pemrograman PHP dengan kerangka kerja Laravel dan basis data SQL. Metodologi Knowledge Discovery in Databases (KDD) digunakan untuk menganalisis data secara terstruktur guna mengekstraksi informasi baru (Ramadanti & Muslih, 2022). Proses data mining melibatkan

penerapan metode clustering untuk mengungkap informasi tersembunyi. Tujuan utama adalah menganalisis hasil tes pelatihan AI-Quran yang dilakukan oleh Biro Urusan Agama Pesantren Islamiyah Syafi'iyah Sumberanyar, Paiton, Probolinggo.

Penelitian ini mencakup beberapa tahapan yang menggambarkan desain penelitian: (a) Pengumpulan Data, (b) Memahami Literatur Klasterisasi K-Means pada Data, (c) Pengolahan Data, (d) Pengembangan Situs Web, (e) Langkah-langkah Klasterisasi Data, dan (f) Evaluasi Hasil Klasterisasi. Gambar 1 di bawah ini menggambarkan bagan alir untuk menentukan kluster menggunakan algoritma K-Means.



Gambar 1. Langkah-langkah Penelitian

2.1. Pengumpulan Data

Data yang dikumpulkan terdiri dari catatan hasil tes pelatihan AI-Quran yang dilakukan oleh Bagian Pembinaan Al-Qur'an di Pesantren Islamiyah Syafi'iyah Sumberanyar, Paiton, Probolinggo. Sebanyak 824 catatan hasil tes dikumpulkan, masing-masing berisi skor individu untuk setiap surah yang diuji. Data yang dikumpulkan mencakup skor Tajwid, kelancaran, dan hafalan untuk setiap santri yang terlibat dalam kegiatan pelatihan AI-Quran.

2.2. Memahami Literatur K-Means Pada Data

Setelah data dikumpulkan dengan benar, langkah selanjutnya adalah memahami konsep dan teori yang mendasari pengelompokan data menggunakan metode Klasterisasi K-Means (Sinaga & Yang, 2020). Proses memahami literatur terkait penerapan Klasterisasi K-Means pada data yang dikumpulkan melibatkan beberapa langkah:

- a) **Tinjauan Literatur:** Melakukan pencarian dan tinjauan literatur serta penelitian terkait Klasterisasi K-Means dengan memeriksa artikel, buku, dan publikasi ilmiah yang membahas konsep dasar, algoritma, dan

- aplikasi Klasterisasi K-Means di berbagai bidang (Hidayat, 2015).
- Konsep Matematika:** Meneksplorasi dasar-dasar matematika Klasterisasi K-Means, termasuk perhitungan jarak antara titik data dan centroid, serta prinsip-prinsip optimasi yang mendasari algoritma tersebut. Ini memberikan pemahaman mendalam tentang cara kerja algoritma dan efisiensinya dalam pengelompokan data (Juli, 2023).
 - Analisis Studi Kasus:** Menganalisis studi kasus dan implementasi praktis Klasterisasi K-Means dalam berbagai konteks dan sektor dengan meninjau bagaimana metode ini telah berhasil digunakan dalam analisis data dunia nyata, mencatat kekuatan dan keterbatasannya (Esmi Nur Fitri et al., 2023).
 - Evaluasi Metode:** Menilai relevansi dan potensi penerapan metode K-Means untuk menganalisis data yang dikumpulkan dengan menentukan apakah metode ini dapat secara efektif mengungkap pola atau wawasan signifikan dalam data (Nabila et al., 2021).

2.3. Pengolahan Data

Tahap pengolahan data dilakukan untuk membersihkan, mengubah, dan mempersiapkan data sehingga siap untuk proses clustering. Langkah-langkah yang akan dilakukan meliputi:

- Menghapus data yang tidak lengkap atau tidak relevan.
- Normalisasi data untuk memastikan semua variabel memiliki skala yang seragam.
- Penanganan data outlier, jika diperlukan.
- Konversi data ke format yang sesuai untuk pemrosesan oleh algoritma K-Means.

2.4. Pembuatan Basis Website

Selanjutnya, Pembangunan interface website dilakukan untuk membuat kebutuhan sistem. Sistem yang dibuat berbasis website menggunakan bahasa pemrograman PHP dengan Framework Laravel dan database SQL agar dapat dengan mudah diakses untuk perhitungan penentuan kelompok dari pembinaan Al-Qur'an. Implementasi Algoritma K-Means Clustering pada Framework Laravel merupakan tahap perhitungan dan pengaplikasian terhadap data yang telah diolah.

2.5. Tahapan Klustering Data

Selanjutnya, dilakukan tahapan clustering dengan algoritma k-means pada data tersebut. Tahapan ini dilakukan melalui uji coba clustering data ulya secara manual menggunakan excel dengan format .xlsx dan menggunakan python pada aplikasi googlecolab guna memastikan kesamaan dari akurasi data yang diperoleh. Adapun langkah-langkah yang dilakukan sebagai berikut:

- Menentukan data yang akan dicluster.
- Menentukan banyak cluster (k) yang akan dibentuk.

- Menentukan pusat cluster (centroid) awal secara acak.
- Menghitung jarak setiap data centroid menggunakan rumus Euclidean Distance.

$$D_{(i,j)} = \sqrt{(P_{1i} - Q_{1j})^2 + (P_{2i} - Q_{2j})^2 + \dots + (P_{ki} - Q_{kj})^2}$$

Dimana nilai $D_{(i,j)}$ merupakan jarak data ke i ke pusat cluster j , nilai X_{ki} adalah Data ke i pada atribut data ke k , dan X_{kj} ini merupakan titik pusat ke j pada atribut ke k .

- Mengelompokkan data ke dalam cluster dengan jarak yang paling pendek ke pusat cluster.
- Menghitung nilai pada pusat cluster yang didapatkan. Rumus untuk menghitung nilai pusat cluster disajikan pada sebuah persamaan.

$$C_i = \frac{P_1 + P_2 + P_3 + \dots + P_n}{\sum P}$$

Dimana C_i merupakan pusat cluster data, P_n adalah data ke n pada atribut data ke n , dan $\sum P$ merupakan jumlah data.

- Melakukan Iterasi berulang hingga posisi data tidak mengalami perubahan.

2.6. Evaluasi Hasil Klasterisasi

Setelah proses klasterisasi selesai, hasilnya akan dievaluasi menggunakan rumus perhitungan Indeks Davies-Bouldin. Perhitungan Indeks Davies-Bouldin didasarkan pada prinsip memaksimalkan jarak antar klaster sambil meminimalkan jarak intra-klaster (Muttaqin & Defriani, 2020). Oleh karena itu, dapat dinyatakan bahwa semakin kecil nilai Indeks Davies-Bouldin, semakin optimal skema klasterisasi. Evaluasi ini akan membantu menentukan jumlah klaster yang optimal dan mengevaluasi kualitas klasterisasi yang dihasilkan oleh algoritma K-Means (Gustientiedina et al., 2019). Rumus untuk menghitung Indeks Davies-Bouldin dapat disajikan sebagai berikut:

$$DBI = \frac{1}{k} \sum_{i=1}^k \{R_i\}$$

dimana $R_i = \max_{j=1, \dots, k, i \neq j} R_{ij}$, $R_{ij} = \frac{S_i + S_j}{d_{ij}}$

dan $S_1 = \left[\frac{1}{n} \sum_{x \in n_i} d^2(x, v_i) \right]^{\frac{1}{2}}$

Dalam persamaan ini, simbol k mewakili jumlah klaster, dan $R_{i,j}$ adalah ukuran kesamaan antara nilai n_i dan n_j . Simbol S_i mewakili ukuran dispersi dari klaster ke- i , dimana $i = 1, 2, 3, \dots, k$. Simbol d_{ij} menunjukkan jarak antara centroid dari klaster i dan centroid dari klaster j ($d_{ij} = d_{ji}$). Simbol n_i menunjukkan jumlah anggota dalam klaster i , dimana $i = 1, 2, 3, \dots, k$. Terakhir, v_i mewakili nilai centroid dari klaster n_i .

Ditulis dalam bentuk paragraf. terdiri dari uraian tentang jenis penelitian, pengumpulan

data, sumber data, jenis data, dan analisis data. Ditulis dalam bentuk paragraf. terdiri dari uraian tentang jenis penelitian, pengumpulan data, sumber data, jenis data, dan analisis data. Ditulis dalam bentuk paragraf.

3 Hasil dan Diskusi

Hasil analisis prediksi pengelompokan bimbingan Al-Qur'an menggunakan algoritma K-Means dilakukan dengan mempertimbangkan skor tajwid, skor fashohah, dan skor hafalan. Data dari 824 santri yang mengikuti tes bimbingan Al-Qur'an dikumpulkan dan digunakan dalam proses pengelompokan.

Hasil analisis menunjukkan bahwa bimbingan Al-Qur'an dapat dikategorikan ke dalam tiga kelompok berbeda berdasarkan tingkat kemahiran mereka dalam membaca Al-Qur'an. Kelompok-kelompok ini diberi label "ulya", "wustha", dan "ula" sesuai dengan karakteristik masing-masing kelompok. Berikut ini adalah hasil implementasi situs web berbasis algoritma K-Means dalam menentukan kelompok bimbingan Al-Qur'an.

3.1 Pengumpulan Data

Pada tahap ini, 824 data poin yang terkumpul digunakan. Data ini mencakup penilaian dari setiap surah yang diuji dalam bimbingan Al-Qur'an. Selanjutnya, skor dari setiap tes diakumulasi untuk mendapatkan total skor dari data tersebut

3.2 Pengolahan Data

Pada tahap ini, pengolahan data dilakukan untuk mempersiapkan data. Pengolahan data melibatkan langkah-langkah seperti pembersihan data, normalisasi, dan pelabelan atau inisialisasi data.

a) Pembersihan Data

Pada tahap pembersihan data, baris data yang tidak relevan dihapus untuk perhitungan K-Means. Baris data ini termasuk kolom yang tidak digunakan dalam analisis, seperti Wilayah, Daerah, Lembaga, dan nilai dari beberapa surah yang diuji, termasuk Surah At-Takasur, Al-'asr, Al-Humazah, Al-Fil, Quraisy, Al-Ma'un, Al-Kautsar, Al-Kafirun, An-Nasr, Al-Lahab, Al-Ikhlash, Al-Falaq, dan An-Nas. Ini dilakukan karena perhitungan K-Means hanya menggunakan total skor dari semua surah yang diuji.

Tabel 1. Hasil Pembersihan Data yang Tidak Relevan.

NO	NIUP	NAMA	NILAI		
			T	F	H
1	12020311 438	Achmad	39	39	39
		Sulton	0	0	0
		Amiruddin K			
2	12020711 426	Afif Dwi Ainul	39	39	38
		Yaqin	0	0	8
3	12020111 253	Ahmad Adi	38	38	38
		Saputra	8	8	8
4	12020511 458	Ahmad	39	39	38
		Daniel	0	0	8
		Mateen Wafa			
5	12020511 218	Ahmad	39	39	38
		Misbahus Sururi	6	6	8
....
82 2	11720503 113	Moh. Fahri	30	31	39
		Raihan Gomila	5	0	0
82 3	11720302 242	Zainul	33	36	36
		Muhaimin Fakhri	5	0	5
82 4	11820701 873	Zacky	33	33	36
		Kamaliyah Zafran	5	5	5

b) Normalisasi Data

Selanjutnya, normalisasi data tidak dilakukan sebagai langkah persiapan sebelum proses klusterisasi menggunakan algoritma K-Means. Dalam hal ini, karena data yang akan diproses sudah memiliki rentang skala yang seragam untuk setiap fitur, dan tidak ada fitur yang mendominasi proses klusterisasi.

3.3 Pembuatan Basis Website

Langkah berikutnya dalam penelitian ini adalah pengembangan dasar situs web menggunakan bahasa pemrograman PHP, kerangka kerja Laravel, dan basis data SQL. Tujuan dari situs web ini adalah untuk menyajikan data yang telah diproses dalam format yang mudah dipahami dan digunakan. Situs web ini juga diharapkan dapat memfasilitasi administrator dalam mengakses data yang dihasilkan dari perhitungan algoritma K-Means. Berikut adalah tampilan dan fitur situs web algoritma K-Means.

3.4 Perhitungan Klusterisasi Data

Dalam penelitian ini, proses klusterisasi dilakukan berdasarkan data yang terkait dengan Skor Tajwid, Skor Fashohah, dan Skor Hafalan. Setelah berhasil memproses data, ditentukan bahwa objek akan dikelompokkan menjadi 3

klaster. Pemilihan k=3 atau c=3 untuk klasterisasi didasarkan pada analisis yang menunjukkan bahwa data dapat dikategorikan secara efektif ke dalam tiga kelompok yang berbeda. Keputusan ini didasarkan pada distribusi dan pola yang diamati dalam skor, yang menunjukkan tiga klaster yang dapat diidentifikasi dengan tingkat kefasihan Tajwid, kejelasan Fashohah, dan keterampilan Hafalan yang berbeda. Dengan memilih k=3, penelitian ini bertujuan untuk memberikan segmentasi data yang bermakna, memungkinkan wawasan yang lebih jelas tentang hubungan antara variabel dan memfasilitasi interpretasi hasil yang lebih mendalam. Dengan demikian, k=3 sangat penting untuk mencapai pemahaman yang terstruktur tentang bagaimana skor ini berkorelasi dan berbeda di seluruh kelompok yang dikelompokkan. Berikut adalah proses klasterisasi data untuk setiap model perhitungan.

a) Menentukan Pusat Klaster Awal

Pada klaster awal, pusat klaster ditentukan secara acak. Dalam percobaan pertama (Iterasi 1), 3 titik data dipilih secara acak sebagai titik pusat awal untuk menghitung jarak dari semua kelompok klaster yang akan dibentuk.
 Jumlah Klaster : 3 (Ulya, Wustha, Ula)
 Jumlah Data : 824 Data Santri
 Jumlah Variabel: 3 (Tajwid, Fashohah, Hafalan)

Tabel 2. Data Awal Pusat Iterasi Pertama

NIUP	NAMA	T	F	H	CLUSTER
12020311438	Achmad Sulton Amirudin K	39	39	39	C1
12020711426	Afif Dwi Ainul Yaqin	39	39	38	C2
12020111253	Ahmad Adi Saputra	38	38	38	C3

Dari Tabel 2, terlihat bahwa setiap data memiliki nilai pusat klaster.

C1 = (390 390 390)
 C2 = (390 390 388)
 C3 = (388 388 388)

b) Perhitungan Jarak Centroid

Pada tahapan ini, jarak antara setiap titik data dan pusat klaster ditentukan menggunakan Jarak Euclidean. Perhitungan jarak dari setiap titik data ke pusat klaster pertama dijelaskan di bawah ini:

Untuk Klaster 1 (1):

$$\sqrt{((390 - 390)^2 + (390 - 390)^2 + (390 - 390)^2)}$$

Untuk Klaster 2 (2):

$$\sqrt{((390 - 390)^2 + (390 - 390)^2 + (388 - 390)^2)}$$

Untuk Klaster 3 (3):

$$\sqrt{((388 - 390)^2 + (388 - 390)^2 + (388 - 390)^2)}$$

Proses ini dilanjutkan untuk semua titik data dan semua klaster. Setelah mendapatkan jarak ini dalam iterasi pertama, mereka dicatat dalam Tabel 3.

Tabel 3. Hasil Perhitungan Iterasi Pertama.

N O	NAMA	C1	C2	C3
1	Achmad Sulton Amirudin K	0	2	3,46822162
2	Afif Dwi Ainul Yaqin	2	0	2,82842713
3	Ahmad Adi Saputra	3,46822162	2,82842713	0
...
82	Moh. Fahri Raihan Gomila	116,743308	113,916636	113,916636
82	Zainul Muhaimin Fakhri	66,7382949	64,2028037	64,2028037
82	Zacky Kamaliyah Zafran	81,111035	78,4028061	78,4028061

Selanjutnya, iterasi kedua melibatkan perhitungan posisi centroid baru. Ini dicapai dengan mencari rata-rata data dalam setiap klaster. Rata-rata untuk setiap atribut (Tajwid, Fashohah, dan Hafalan) dihitung dan disajikan sebagai centroid baru dalam Tabel 4.

Tabel 4. Centroid Baru dalam Iterasi Kedua.

CENTRO ID BARU	T	F	H
C1	431,968254	432,269882	390,055556
C2	420,731343	417,940299	372,701493

C3	333,8173 52	331,1461 19	371,1735 16
----	----------------	----------------	----------------

Setelah menghitung centroid baru, jarak dari setiap titik data ke centroid ini dievaluasi ulang menggunakan metode Jarak Euclidean. Langkah ini mengulangi metodologi yang dijelaskan dalam iterasi awal. Iterasi kedua dimulai dengan menghitung Jarak Euclidean antara setiap titik data dan centroid yang diperbarui dari setiap kluster. Prosedur ini mirip dengan pendekatan yang dilakukan pada Iterasi pertama.

3.5 Evaluasi Hasil Pengelompokan

Proses pengelompokan untuk kelompok pembelajaran Al-Qur'an menggunakan algoritma k-means menghasilkan jumlah data yang sesuai dalam setiap kluster setelah iterasi kesepuluh. Dari perhitungan data 824 santri, terbentuk kelompok-kelompok berikut:

- a) Kluster pertama (C1), mewakili kelompok dengan nilai ulya, terdiri dari 292 santri.
- b) Kluster kedua (C2), mewakili kelompok dengan nilai wustha, terdiri dari 342 santri.
- c) Kluster ketiga (C3), mewakili kelompok dengan nilai ula, terdiri dari 190 santri.

Tabel 5. Kelompok Kluster Pertama (Ulya).

N O	NAMA	NILAI			CLUSTER
		T	F	H	
1	Brilliant Saputera Pratama M.	42 6	42 2	38 8	C1
2	Muhammad Syarif Hasan Ubaidillah	43 2	42 4	39 0	C1
3	Rafa El Asyraf Tegar Maharaja	43 5	43 5	39 0	C1
...
14 5	Mochammad Rizal Rofiqi	42 5	41 5	38 5	C1
14 6	Ahmad Imong Budiono	43 0	40 8	39 0	C1

Tabel 6. Kelompok Kluster Kedua (Wustha).

N O	NAMA	NILAI			CLUSTER
		T	F	H	
1	Achmad Sulton Amiruddin K	39 0	39 0	38 8	C2
2	Afif Dwi Ainul Yaqin	39 0	39 0	39 0	C2
3	Ahmad Adi Saputra	38 8	38 8	38 8	C2
4	Ahmad Daniel Mateen Wafa	39 0	39 0	38 8	C2
5	Ahmad Misbahus Sururi	39 6	39 6	38 8	C2
...
17 0	Zainul Muhaimin Fakhri	33 5	36 0	36 5	C2
17 1	Zacky Kamaliyah Zafran	33 5	33 5	36 5	C2

Tabel 7. Kelompok Kluster Ketiga (Ula).

NAMA	NILAI			CLUSTER
	T	F	H	
Raditya Farrel Bayu Prasasta	315	319	284	C3
Vicky Fiera Alfiansyah	298	320	370	C3
Rif'at Syawqi	310	315	390	C3
Riski Septia Ramadani	300	320	380	C3
Robithurrahman	320	265	385	C3
.....
Vikri Aska Tufail	350	270	370	C3
Moh. Fahri Raihan Gomila	305	310	390	C3

Studi ini berhasil mengembangkan sistem aplikasi berbasis web menggunakan algoritma K-Means untuk mengelompokkan santri dalam kegiatan belajar Al-Qur'an di Pondok Pesantren Islamiyah Syafi'iyah Sumberanyar, Paiton Probolinggo. Hasil penelitian menunjukkan bahwa penggunaan algoritma K-Means efektif dalam mengelompokkan santri, sehingga meningkatkan efisiensi dan akurasi dalam proses pembelajaran. Pada bagian ini, para penulis akan membahas temuan dari studi ini lebih lanjut, membandingkannya dengan penelitian serupa, serta memberikan interpretasi dan implikasi dari hasil yang diperoleh.

Salah satu temuan utama dari studi ini adalah efektivitas algoritma K-Means dalam mengelompokkan santri berdasarkan kemampuan membaca Al-Qur'an mereka, termasuk kelancaran dalam tajwid dan skor hafalan. Algoritma ini berhasil mengelompokkan santri ke dalam tiga kategori keterampilan: "Ulya," "Wustha," dan "Ula." Pengelompokan ini memberikan panduan objektif dan akurat dalam menentukan kelompok belajar yang tepat untuk setiap santri. Hasil ini konsisten dengan penelitian yang dilakukan oleh Achmad Dimiyati (2023), yang juga menemukan bahwa algoritma K-Means efektif dalam mengevaluasi skor akademik santri di TPQ Darussalamah. Tabel di bawah ini membandingkan hasil studi ini dengan beberapa penelitian serupa:

Tabel 8. Perbandingan Hasil Penelitian

STUDI	KONTEKS	METODE	TEMUAN UTAMA
Penelitian ini	Mengelompokkan santri berdasarkan kemampuan Qur'an di Pondok Pesantren Islamiyah Syafi'iyah Sumberanyar	K-Means	Efektif dalam mengelompokkan santri ke dalam tiga kategori keterampilan: "Ulya," "Wustha," dan "Ula".
Achmad Dimiyati (2023)	Menilai skor akademik santri di TPQ Darussalamah	K-Means	Algoritma K-Means efektif dalam mengelompokkan santri berdasarkan skor akademik. Temuan ini mendukung kesimpulan studi kami bahwa K-Means dapat

digunakan untuk klasifikasi pendidikan.

Ai Rohmah, Falentino Sembiring, and Adhitha Erfina (2021)	Mengklasifikasi tingkat hambatan belajar dalam pembelajaran jarak jauh selama pandemi COVID-19	K-Means	Algoritma K-Means membantu mengidentifikasi tingkat hambatan belajar yang dihadapi santri. Studi ini menyoroti fleksibilitas K-Means dalam berbagai konteks pendidikan, mirip dengan penelitian kami.
Dian Permatasari (2021)	Menentukan tingkat penyebaran pandemi COVID-19 di Sumatera Barat	K-Means	Algoritma K-Means efektif dalam mengelompokkan wilayah berdasarkan kasus positif dan penyebaran virus. Penelitian ini menunjukkan bahwa K-Means dapat digunakan dalam analisis data epidemiologi, meskipun berbeda dari konteks pendidikan tetapi menunjukkan fleksibilitas algoritma tersebut.

Implementasi teknologi dan metode analisis data dalam studi ini bertujuan untuk mengatasi tantangan yang dihadapi Pondok Pesantren Islamiyah Syafi'iyah Sumberanyar. Dengan

memanfaatkan teknologi, proses pengelompokan menjadi lebih cepat, lebih akurat, dan lebih objektif, sehingga meningkatkan motivasi santri melalui pengalaman belajar yang dipersonalisasi dan disesuaikan. Studi ini menunjukkan bahwa dengan pengelompokan yang tepat, santri dapat menerima pelatihan sesuai dengan kemampuan mereka, yang pada gilirannya meningkatkan hasil belajar.

Pengembangan sistem berbasis web ini juga memudahkan akses dan penggunaan oleh administrator dan peserta. Sistem ini tidak hanya menguntungkan Pondok Pesantren Islamiyah Syafi'iyah Sumberanyar tetapi juga berpotensi untuk diterapkan di lembaga pendidikan Islam lainnya. Dengan sistem ini, proses pembelajaran Al-Qur'an dapat dilakukan dengan lebih efisien, memungkinkan guru untuk lebih fokus pada pengajaran daripada tugas administratif.

4 Kesimpulan

Penelitian ini menunjukkan bahwa penggunaan algoritma K-Means sangat efektif dalam mengelompokkan santri berdasarkan kemampuan membaca Al-Qur'an mereka. Penelitian ini menganalisis data 824 santri, dievaluasi dalam hal tajwid (pengucapan), fashahah (kelancaran), dan hafalan. Data ini kemudian diproses dan dikelompokkan menggunakan algoritma K-Means. Hasil penelitian menunjukkan bahwa santri dapat dikategorikan ke dalam tiga tingkat kemahiran: "Ulya", "Wustha", dan "Ula". Kategori-kategori ini memberikan panduan yang jelas untuk menentukan kelompok pelatihan yang sesuai untuk setiap santri, yang pada akhirnya meningkatkan efisiensi dan akurasi proses pelatihan sesuai dengan kurikulum yang telah ditetapkan. Metode pengelompokan ini memungkinkan pelatihan yang lebih tepat dengan mengelompokkan santri berdasarkan kesamaan dalam kemampuan membaca Al-Qur'an mereka. Penggunaan teknologi berbasis web juga memudahkan akses dan penggunaan sistem oleh administrator, membuat proses tersebut lebih cepat, lebih akurat, dan lebih objektif.

Secara keseluruhan, penelitian ini tidak hanya mengatasi tantangan yang dihadapi Pondok Pesantren Islamiyah Syafi'iyah Sumberanyar dalam mengorganisir kelompok pelatihan Al-Qur'an, tetapi juga berkontribusi pada pengembangan sistem pelatihan yang lebih responsif dan efisien untuk lembaga serupa di masa depan. Untuk penelitian mendatang, disarankan untuk menguji sistem ini di berbagai

lembaga pendidikan Islam dengan karakteristik yang berbeda untuk memvalidasi keandalannya. Selain itu, integrasi sistem penilaian otomatis yang lebih canggih dapat lebih meningkatkan akurasi pengelompokan.

5 Referensi

- Adillah, D., Manurung, N., & ... (2022). Implementation of K-means Clustering Analysis to Determine Barriers to Online Learning Case Study: Swasta Yapendak Tinjowan Junior High School. *Jurnal Teknik Informatika ...*, 3(3), 519–525.
- Anggara, R., Defit, S., Hendrik³, B., & Komputer, F. I. (2024). *Implementasi K-Means Clustering Dalam Analisa Soal Ujian CBT Universitas Baiturrahmah*. 5(2), 577–586.
- Br Sembiring, S. N., Winata, H., & Kusnasari, S. (2022). Pengelompokan Prestasi Siswa Menggunakan Algoritma K-Means. *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, 1(1), 31. <https://doi.org/10.53513/jursi.v1i1.4784>
- Dimiyati, A. (2023). *E-ISSN : XX-XX-XX ejurnal . itsnupekalongan . ac . id / sandtree E-ISSN : XX-XX-XX*. 1(1), 39–54.
- Esmi Nur Fitri, M. Hafidz Ariansyah, Sri Winarno, Fikri Budiman, Asih Rohmani, Zeniarja, J., & Edi Sugiarto. (2023). Performance of K-Means Clustering and Knn Classifier in Fish Feed Seller Determination Models. *Jurnal Teknik Informatika (Jutif)*, 4(3), 485–491. <https://doi.org/10.52436/1.jutif.2023.4.3.725>
- Gustientiedina, G., Adiya, M. H., & Desnelita, Y. (2019). Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 5(1), 17–24. <https://doi.org/10.25077/teknosi.v5i1.2019.17-24>
- Hendrastuty, N. (2024). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Dalam Evaluasi Hasil Pembelajaran Siswa. *Jurnal Ilmiah Informatika Dan Ilmu Komputer (Jima-Ilkom)*, 3(1), 46–56.
- Hidayat, M. M. (2015). Data Mining Data mining. *Mining of Massive Datasets*, 2(January 2013), 5–20.
- Hutagalung, J. (2022). Pemetaan Siswa Kelas Unggulan Menggunakan Algoritma K-Means Clustering. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(1), 606–620. <https://doi.org/10.35957/jatisi.v9i1.1516>
- Juli, B. (2023). *Jurnal Kajian dan Terapan Matematika*. 8, 114–128.

- Likmi, S., & Barat, J. (2024). *APPLICATION OF DATA MINING TO MEASURE THE EFFECTIVENESS OF THE ISLAMIC BOARDING SCHOOL ' S INDEPENDENT CURRICULUM BASED ON LEARNING ACHIEVEMENT USING THE CLUSTERING METHOD*. 3(5), 514–519. <https://doi.org/10.58344/jws.v3i5.595>
- Mirantika, N. (2021). Penerapan Algoritma K-Means Clustering Untuk Pengelompokan Penyebaran Covid-19 di Provinsi Jawa Barat. *Nuansa Informatika*, 15(2), 92–98. <https://doi.org/10.25134/nuansa.v15i2.4321>
- Montero, Z. D. (2022). Customer Grouping for Customer Relationship Management Optimization with the K-Means Algorithm. *Journal of Computer Scine and Information Technology*, 98–105. <https://doi.org/10.35134/jcsitech.v8i4.46>
- Muttaqin, M. R., & Defriani, M. (2020). Algoritma K-Means untuk Pengelompokan Topik Skripsi Mahasiswa. *ILKOM Jurnal Ilmiah*, 12(2), 121–129. <https://doi.org/10.33096/ilkom.v12i2.542.121-129>
- Nabila, Z., Isnain, A. R., Permata, & Abidin, Z. (2021). Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means. *Jurnal Teknologi Dan Sistem Informasi (JTSI)*, 2(2), 100.
- Novita, R., Khomarudin, A. N., Aulia, R., Jamaluddin, J., Yuditihwa, A., & Ayuri, A. (2023). Penerapan Algoritma K-Means dan Analisisnya untuk Menentukan Kebijakan Strategis Penyelesaian Studi Mahasiswa. *Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika Dan Komputer)*, 22(2), 401. <https://doi.org/10.53513/jis.v22i2.8461>
- Palevi, M. R., & Indra, Z. (2024). Implementasi Algoritma K-Means Clustering Dengan Pendekatan Active Learning Pada Siswa SMA Untuk Menentukan Jurusan Ke Perguruan Tinggi. *Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika Dan Komputer)*, 23(1), 26. <https://doi.org/10.53513/jis.v23i1.9553>
- Ramadanti, E., & Muslih, M. (2022). Penerapan Data Mining Algoritma K-Means Clustering Pada Populasi Ayam Petelur Di Indonesia. *Rabit: Jurnal Teknologi Dan Sistem Informasi Univrab*, 7(1), 1–7. <https://doi.org/10.36341/rabit.v7i1.2155>
- Rohmah, A., Sembiring, F., & Erfina, A. (2021). Implementasi Algoritma K-Means Clustering Analysis untuk Menentukan Hambatan Pembelajaran Daring (Studi Kasus: SMK Yaspim Gegerbitung). *Seminar Nasional Sistem Informasi Dan Manajemen Informatika*, 290–298.
- Santi, S., & Februariyanti, H. (2023). Implementation of Clustering on Tweet Uploading Side Effects of Covid-19 Post Vaccination Using K-Means Algorithm. *Jurnal Teknik Informatika (Jutif)*, 4(4), 779–786. <https://doi.org/10.52436/1.jutif.2023.4.4.704>
- Sari, D. P. (2021). Implementasi Algoritma K-Means Dalam Menentukan Tingkat Penyebaran Pandemi Covid-19 Di Sumatera Barat. *Computer Based Information System Journal*, 9(1), 50–56. <https://doi.org/10.33884/cbis.v9i1.3646>
- Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. *IEEE Access*, 8, 80716–80727. <https://doi.org/10.1109/ACCESS.2020.2988796>
- Soliani, I., & Juanita, S. (2022). Grouping the Prevalence of Disease Cases By Age in Bandung City Hospitals Using K-Means. *Jurnal Teknik Informatika (Jutif)*, 3(6), 1647–1654. <https://doi.org/10.20884/1.jutif.2022.3.6.430>
- Subecz, Z. (2021). Web-development with Laravel framework. *Gradus*, 8(1), 211–218. <https://doi.org/10.47833/2021.1.csc.006>
- Syafiih, M., Abdullah, M. D., Purwanto, M. R., Shodiq, M. J., Hijrah, M., Putra Mawardi, M. A., & Nasirudin, M. Y. (2021). Pembuatan Website Biro Pendidikan Nurul Jadid dengan Sistem Framework sebagai Media Informasi Pendidikan di Pesantren. *GUYUB: Journal of Community Engagement*, 2(1), 157–169. <https://doi.org/10.33650/quyub.v2i1.2120>
- Syafiih, M., & Hatima Indah Arifin, N. (2023). Sistem Informasi Monitoring Target Capaian Pembinaan Al-Qur'an Di Wilayah Pondok Pesantren Nurul Jadid Berbasis Web. *Jar's*, 1(2), 23.