

# PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK PENGELOMPOKAN DATA ABSENSI SISWA MENGGUNAKAN TEKNOLOGI QR CODE BERBASIS WEB (Studi Kasus SMAN 2 Jombang)

Danang Rifki Amruli<sup>✉</sup>, Tanhella Zein Vitadiar

Program Studi Teknik Informatika, Universitas Hasyim Asy'ari Tebuireng Jombang, Indonesia

Email: [danangamruli@mhs.unhasy.ac.id](mailto:danangamruli@mhs.unhasy.ac.id)

DOI: <https://doi.org/10.46880/jmika.Vol10No1.pp53-61>

## ABSTRACT

*This study develops a web-based student attendance system using QR Code technology integrated with the K-Means Clustering algorithm. The system aims to improve accuracy and efficiency in managing attendance while classifying student discipline levels based on indicators of present, late, sick, excused, and unexcused. System development applies the Waterfall model. The K-Means algorithm groups students into three clusters: high, medium, and low attendance. Precision testing shows 95.56% accuracy compared to manual teacher assessment. The system supports school digitalization by providing automated attendance recording and data-driven discipline evaluation.*

**Keyword:** Attendance, QR Code, K-Means Clustering, Student Discipline, Web System.

## ABSTRAK

*Penelitian ini mengembangkan sistem absensi siswa berbasis web menggunakan teknologi QR Code yang terintegrasi dengan algoritma K-Means Clustering. Sistem bertujuan meningkatkan akurasi dan efisiensi pengelolaan absensi serta mengklasifikasikan tingkat kedisiplinan berdasarkan indikator hadir, terlambat, sakit, izin, dan alpa. Pengembangan menggunakan model Waterfall. Algoritma K-Means mengelompokkan siswa ke dalam tiga cluster: kehadiran tinggi, sedang, dan rendah. Pengujian Precision menghasilkan akurasi 95.56% dibanding penilaian manual guru. Sistem mendukung digitalisasi sekolah melalui pencatatan absensi otomatis dan evaluasi kedisiplinan berbasis data.*

**Kata Kunci:** Absensi, QR Code, K-Means Clustering, Kedisiplinan Siswa, Sistem Web.

## PENDAHULUAN

Kemajuan pesat di sektor teknologi membawa pengaruh signifikan bagi tata kelola institusi pendidikan. Adanya integrasi komputer dan internet memfasilitasi manajemen data yang jauh lebih ringkas, efisien, dan presisi daripada sekadar mengandalkan rekapitulasi manual. Kemajuan teknologi informasi mendorong berbagai instansi pendidikan untuk bertransformasi menuju digitalisasi, baik dalam proses pembelajaran, administrasi, maupun sistem pengelolaan data siswa. Penerapan sistem berbasis web di lingkungan sekolah menjadi solusi efektif untuk meningkatkan efisiensi kerja dan akurasi informasi, termasuk dalam pengelolaan data absensi (Aufa Muis et al., 2024).

Absensi siswa merupakan indikator krusial untuk mengukur kedisiplinan serta komitmen akademis mereka. Namun, sistem absensi konvensional yang masih banyak digunakan di sekolah-sekolah, seperti pencatatan manual menggunakan daftar hadir,

seringkali menghadapi kendala seperti keterlambatan input data, kesalahan pencatatan, dan risiko kehilangan data. Kondisi tersebut menyebabkan pihak sekolah, terutama guru bimbingan konseling (BK) dan wali kelas, kesulitan dalam memantau pola kehadiran siswa secara menyeluruh.

Penggunaan teknologi QR Code dalam sistem absensi terbukti mampu mempercepat proses pencatatan kehadiran dan mengurangi potensi kesalahan manusia (human error), karena data yang dihasilkan langsung tersimpan secara digital dan dapat diakses secara *real-time*. Teknologi Quick Response (QR) Code merupakan inovasi dalam sistem pengenalan data dua dimensi yang memungkinkan penyimpanan informasi dalam jumlah besar dan dapat dibaca dengan cepat oleh perangkat digital seperti kamera smartphone (Sanubari et al., 2024). QR Code memiliki keunggulan dalam efisiensi, kemudahan penggunaan, dan tingkat keamanan yang tinggi melalui fitur koreksi kesalahan (*error correction*). Dengan

demikian, integrasi *QR Code* dalam sistem absensi siswa memberikan solusi modern yang tidak hanya praktis tetapi juga mendukung digitalisasi manajemen data sekolah. Sistem absensi berbasis *QR Code* memungkinkan setiap siswa melakukan presensi cukup dengan memindai kode unik pada Kartu Tanda Siswa (KTS), yang selanjutnya akan tercatat secara otomatis dalam database sekolah.

Meskipun demikian, data absensi digital yang terkumpul tidak akan memberikan manfaat optimal tanpa adanya proses analisis yang tepat. Oleh karena itu, implementasi metode pengolahan data yang mampu mengelompokkan siswa berdasarkan pola kehadiran menjadi krusial agar pihak sekolah dapat merumuskan kebijakan pembinaan yang lebih akurat. Salah satu instrumen yang relevan untuk tujuan ini adalah penggunaan algoritma *K-Means Clustering*. Algoritma *K-Means* merupakan salah satu metode *unsupervised learning* yang efektif untuk menemukan pola atau kelompok homogen dalam data tanpa memerlukan label kelas. Dalam konteks pendidikan, algoritma ini dapat diterapkan untuk mengelompokkan siswa berdasarkan tingkat kehadiran, keterlambatan, izin, sakit, maupun ketidakhadiran tanpa keterangan. Kelebihan algoritma *K-Means* dibandingkan metode lainnya, seperti *K-Nearest Neighbor (KNN)* atau *Fuzzy C-Means*, terletak pada kesederhanaan konsep, kecepatan proses, serta kemampuannya beradaptasi terhadap perubahan data baru (Liu, 2022). Hasil *clustering* yang dihasilkan oleh *K-Means* dapat menjadi lebih optimal apabila data telah melalui proses *feature scaling* dan *normalisasi* yang baik. Sementara itu, penelitian oleh (Dewi et al., 2022) menunjukkan bahwa *K-Means* merupakan algoritma yang efisien dan skalabel untuk pengolahan data berjumlah besar, sehingga cocok digunakan dalam sistem berbasis web yang menangani data absensi harian siswa dalam jumlah banyak.

Berdasarkan latar belakang tersebut, penelitian ini mengusulkan penerapan Algoritma *K-Means Clustering* untuk pengelompokan data absensi siswa menggunakan teknologi *QR Code* berbasis web di SMAN 2 Jombang. Sistem yang dirancang tidak hanya berfungsi sebagai media pencatatan kehadiran digital, tetapi juga sebagai alat analisis yang mampu mengidentifikasi pola kedisiplinan siswa secara otomatis. Dengan demikian, pihak sekolah dapat memperoleh informasi yang lebih mendalam mengenai tingkat kehadiran siswa dan menggunakannya sebagai dasar dalam pembinaan serta pengambilan kebijakan yang lebih efektif. Selain itu, hasil pengelompokan yang dihasilkan melalui algoritma *K-Means* diharapkan dapat membantu guru BK dan wali kelas dalam

memantau siswa dengan tingkat kedisiplinan rendah, sekaligus memberikan apresiasi bagi siswa dengan tingkat kehadiran tinggi. Dengan adanya kombinasi antara teknologi *QR Code* dan algoritma *K-Means Clustering* ini, proses absensi siswa diharapkan menjadi lebih efisien, transparan, dan berbasis data. Sistem yang diusulkan juga dapat dijadikan acuan bagi sekolah lain dalam mengembangkan inovasi serupa yang berfokus pada efisiensi administrasi dan peningkatan kedisiplinan siswa di era digital pendidikan modern.

Berbeda dengan penelitian terdahulu yang mayoritas berfokus pada salah satu aspek secara terpisah baik itu sekadar digitalisasi presensi maupun *clustering* data historis secara terpisah penelitian ini mengintegrasikan kedua proses tersebut secara *real-time* dalam satu platform web. Kebaruan (*novelty*) dari penelitian ini terletak pada otomasi evaluasi kedisiplinan berbasis lima variabel absensi sekaligus (hadir, terlambat, sakit, izin, alpa). Sistem ini tidak hanya merekam data secara cepat melalui *QR Code*, tetapi langsung memprosesnya ke dalam *engine K-Means* untuk meminimalkan bias subjektivitas penilaian manual guru, sehingga menghasilkan sistem pendukung keputusan yang utuh bagi pihak sekolah.

## TINJAUAN PUSTAKA

Penelitian ini mengacu pada berbagai penelitian terdahulu yang relevan. (Rifais & Laksana, 2024) menerapkan Algoritma *K-Means* untuk mengetahui kemampuan akademik siswa berbasis web. (Sanubari et al., 2024) merancang sistem presensi siswa menggunakan *QR Code* yang terbukti efektif mempercepat proses absensi. (Dewi et al., 2022) mengimplementasikan *K-Means Clustering* dalam seleksi siswa berprestasi berdasarkan keaktifan. (Sulistiyawati & Supriyanto, 2020) menggunakan *K-Means* untuk menentukan kelas unggulan siswa. Selain itu, (Gustian et al., 2023) merancang pemetaan tingkat kehadiran siswa dengan metode *K-Means Clustering*. Meskipun kelima penelitian di atas telah membuktikan kehandalan *K-Means* dan efisiensi *QR Code*, terdapat beberapa celah (*gap*) analisis yang perlu disempurnakan. Penelitian yang menerapkan *K-Means* pada data siswa (Rifais & Laksana, 2024; Sulistiyawati & Supriyanto, 2020) umumnya masih mengandalkan rekapitulasi data awal secara manual atau *batch*, sehingga rentan terhadap human error dan penundaan data. Di sisi lain, riset yang berfokus pada *QR Code* (Sanubari et al., 2024) sangat kuat di otomasi input, namun berhenti pada tahap pencatatan tanpa melakukan data mining lanjutan untuk klasifikasi pola. Oleh karena itu, penelitian ini hadir untuk menutup

kelemahan tersebut dengan menggabungkan kecepatan dan akurasi akuisisi data *QR Code* dengan ketajaman analisis prediktif *K-Means Clustering* di dalam satu ekosistem sistem informasi

*Quick Response (QR) Code* merupakan varian *barcode* matriks dua dimensi ciptaan Denso Wave pada 1994 yang mampu menampung informasi berbentuk angka, huruf, hingga biner. Kapasitas penyimpanan datanya terbukti jauh melampaui *barcode* tradisional. Kode ini dapat dipindai cepat menggunakan kamera perangkat digital, sehingga memudahkan akses informasi. Dalam teknologi informasi, *QR Code* digunakan dalam pembayaran digital, autentikasi, pelacakan produk, pemasaran, dan absensi web. Kepopulerannya didukung oleh fleksibilitas, kemudahan pembuatan, serta fitur *error correction* yang menjaga data tetap terbaca meskipun kode sebagian rusak (Ahmad & Fatimah, 2024).

Algoritma *K-Means Clustering* dikenal sebagai salah satu teknik analisis pengelompokan yang sangat populer di ranah data mining serta *machine learning*. *K-Means* termasuk ke dalam kategori *unsupervised learning*, yaitu metode pembelajaran tanpa label kelas, di mana sistem berusaha menemukan pola atau struktur tersembunyi dari data tanpa target output yang sudah ditentukan (Liu, 2022). Dalam penelitian ini, algoritma *K-Means* digunakan untuk mengelompokkan data absensi siswa berdasarkan lima variabel, yaitu jumlah hadir, jumlah terlambat, jumlah sakit, jumlah izin, dan jumlah tanpa keterangan. Data tersebut dikelompokkan ke dalam tiga *cluster* utama, yaitu *Cluster 1* (Kehadiran Tinggi) dengan tingkat kehadiran 90–100%, *Cluster 2* (Kehadiran Sedang) dengan kehadiran 80–89%, dan *Cluster 3* (Kehadiran Rendah) dengan kehadiran di bawah 80%. Proses pengelompokan ini membantu pihak sekolah dalam mengevaluasi kedisiplinan siswa serta menentukan langkah pembinaan yang tepat.

Tahapan penerapan algoritma *K-Means Clustering* dalam penelitian ini adalah sebagai berikut :

1. Menentukan jumlah *cluster* (K): Dalam penelitian ini ditetapkan tiga *cluster* (tinggi, sedang, rendah).
2. Menentukan *centroid* awal: Titik awal diambil secara acak dari data kehadiran.
3. Menghitung jarak *Euclidean*: Setiap data dihitung jaraknya terhadap *centroid* dengan rumus:

$$d(x_i, c_j) = \sqrt{\sum_{k=1}^n (x_{ik} - c_{jk})^2}$$

Keterangan:

$d(x_i, c_j)$  = jarak data ke-*i* dengan *centroid* ke-*j*.

$n$  = jumlah variabel (hadir, terlambat, sakit, izin, tanpa keterangan).

$x_{ik}$  = nilai data ke-*i* pada variabel ke-*k*.

$c_{jk}$  = nilai *centroid cluster* ke-*j* pada variabel ke-*k*.

Mengelompokkan data: Setiap data dimasukkan ke *cluster* dengan jarak *centroid* terdekat.

Memperbarui *centroid*: Nilai *centroid* baru dihitung berdasarkan rata-rata data dalam *cluster*.

$$c_j = \frac{1}{n_j} \sum_{x_i \in C_j} x_i$$

$c_j$  = *centroid* baru untuk *cluster j*

$n_j$  = jumlah data dalam *cluster j*

$x_i \in C_j$  = data ke-*i* yang berada pada *cluster j*

*Iterasi*: Langkah 3–5 diulang hingga posisi *centroid* stabil (*konvergen*).

Hasil akhir dari proses ini berupa tiga kelompok siswa berdasarkan tingkat kehadiran. Data yang telah dikelompokkan kemudian digunakan untuk mendukung analisis kedisiplinan serta penyusunan laporan evaluasi oleh pihak sekolah.

## METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan metode pengembangan sistem berbasis perangkat lunak menggunakan model *Waterfall*.



Gambar 1. Metode *Waterfall*

Pengembangan sistem dalam riset ini mengadopsi pendekatan *Waterfall*. Metode ini menerapkan alur rekayasa perangkat lunak yang berurutan dan terstruktur, di mana setiap fase pengerjaan harus dirampungkan secara tuntas terlebih dahulu sebelum melangkah ke fase evaluasi berikutnya

(Zulfa & Zein Vitadiar, 2023). Kebutuhan fungsional mencakup proses *login*, pemindaian *QR Code*, penyimpanan data, serta pengelompokan menggunakan *K-Means*. Kebutuhan nonfungsional meliputi kecepatan respon, keamanan, *user friendly*, dan kompatibilitas perangkat. Analisis ini dilakukan melalui wawancara dan observasi di SMAN 2 Jombang. Tahap perancangan mencakup arsitektur sistem, basis data *MySQL* (tabel siswa, absensi, cluster\_kehadiran), antarmuka pengguna, dan permodelan menggunakan UML seperti:

**Usecase Diagram**

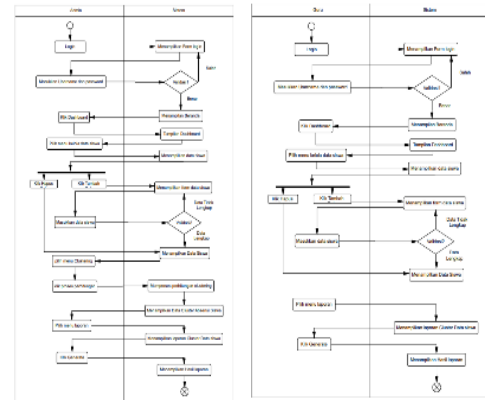
*Use case diagram* diterapkan sebagai salah satu bentuk pemodelan visual *UML* guna merepresentasikan alur interaksi fungsional antara aktor dengan aplikasi. Di bawah ini adalah rancangan *use case* dalam sistem informasi absensi siswa:



Gambar 2 Use Case Diagram

**Activity Diagram**

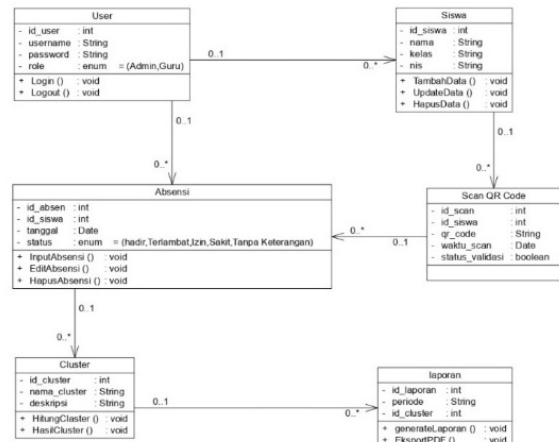
*Activity diagram* ini menunjukkan alur proses pada sistem absensi dan *clustering*, di mana Siswa melakukan absensi melalui scan QR, Guru mengelola dan memvalidasi data absensi kelas, sedangkan Admin memiliki akses penuh untuk mengelola data siswa, memproses absensi, serta menjalankan analisis *clustering*. Setiap aktor mengikuti jalur aktivitas masing-masing mulai dari login atau scan hingga mendapatkan output seperti data absensi, hasil *cluster*, atau laporan, sehingga menggambarkan peran dan interaksi antar pengguna dalam sistem secara jelas dan terstruktur



Gambar 3. Activity Diagram

**Class Diagram**

*Class diagram* ini menggambarkan struktur sistem informasi absensi dan clustering siswa. Terdapat lima kelas utama, yaitu *User*, *Siswa*, *Absensi*, *Cluster*, dan *Laporan*.



Gambar 4. Class Diagram

Implementasi sistem menggunakan bahasa pemrograman *PHP* dan *database MySQL* pada server lokal. Algoritma *K-Means Clustering* diterapkan untuk mengklasifikasikan data absensi ke dalam tiga kategori persentase kehadiran: Tinggi (90–100%), Sedang (80–89%), dan Rendah (<80%).

Untuk mengakomodasi kedalaman analisis data mining yang berjalan beriringan dengan alur rekayasa perangkat lunak Waterfall, penelitian ini mengadopsi tahapan analisis berbasis Knowledge Discovery in Databases (KDD) pada proses pengolahan data absensi. Integrasi metode ini bertujuan memberikan justifikasi ilmiah yang kuat pada setiap tahap perhitungan *K-Means*. Tahapan analisis tersebut meliputi:

1. Pengumpulan dan Pra-pemrosesan Data (Data Selection & Preprocessing)

Data absensi mentah yang diakuisisi secara real-time melalui pemindaian QR Code dikumpulkan ke dalam basis data. Data tersebut kemudian melalui tahap pembersihan (cleaning) untuk menyaring anomali seperti duplikasi scan pada hari yang sama atau nilai yang kosong (missing values), sehingga data historis siswa SMAN 2 Jombang siap dan valid untuk diolah.

## 2. Transformasi Data dan Justifikasi Bobot Kriteria (Data Transformation)

Data absensi kategorikal dikonversi menjadi representasi numerik melalui tahapan pembobotan. Secara ilmiah dan praktis, pemberian bobot kriteria dijustifikasi berdasarkan eskalasi tingkat pelanggaran tata tertib dan persentase hilangnya jam instruksional siswa. Bobot yang ditetapkan adalah: Hadir (0) sebagai baseline kepatuhan penuh tanpa pelanggaran; Terlambat (2) karena siswa kehilangan sebagian kecil jam pelajaran; Sakit (7) dan Izin (7) merepresentasikan hilangnya jam belajar harian secara penuh meskipun sah secara administratif; serta Alpha (10) yang diberi penalti tertinggi karena merupakan pelanggaran indisipliner murni tanpa keterangan.

## 3. Proses Data Mining dan Penentuan Jumlah Kluster (K=3)

Pada tahap ini, algoritma K-Means diterapkan untuk mengelompokkan data. Pemilihan parameter jumlah kluster sebanyak K=3 tidak ditentukan secara acak, melainkan dilandaskan pada domain knowledge terkait kebutuhan evaluasi pedagogik dan operasional Guru Bimbingan Konseling (BK). Tiga kluster ini merepresentasikan tiga level intervensi kedisiplinan sekolah:

- a. Kluster 1 (Kehadiran Tinggi) sebagai dasar pemberian apresiasi akademik.
- b. Kluster 2 (Kehadiran Sedang) sebagai dasar tindakan preventif atau peringatan dini.
- c. Kluster 3 (Kehadiran Rendah) sebagai dasar tindakan kuratif intensif, seperti pemanggilan orang tua siswa.

Penetapan K=3 dinilai paling optimal untuk memberikan actionable insight yang tegas bagi sekolah tanpa memicu fragmentasi kebijakan yang terlalu rumit.

## 4. Evaluasi Model (Evaluation)

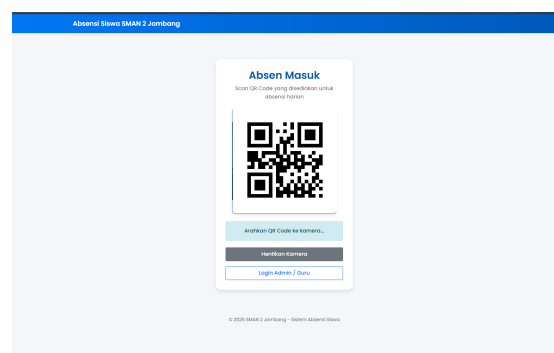
Tahap terakhir dari analisis data adalah menguji keandalan hasil klusterisasi K-Means. Evaluasi dilakukan menggunakan Confusion Matrix untuk mengkomparasi keselarasan antara hasil prediksi algoritma (sistem) dengan penilaian manual aktual (ground truth) dari pihak guru. Metrik utama yang diukur adalah Precision untuk memastikan bahwa

pengelompokan tingkat kedisiplinan yang dihasilkan memiliki validitas ilmiah yang tinggi sebelum diimplementasikan secara penuh.

## HASIL DAN PEMBAHASAN

### Halaman Beranda

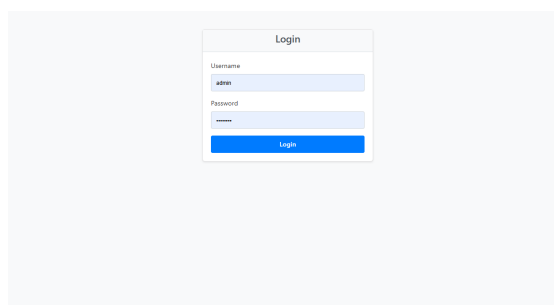
Halaman beranda ini menyediakan fitur utama untuk melakukan Absen Masuk melalui pemindaian QR Code menggunakan kamera. Pada bagian tengah ditampilkan preview kamera untuk memudahkan proses scan. Pengguna dapat mengatur kamera melalui tombol yang tersedia, serta masuk ke menu Admin atau Guru untuk mengelola data absensi. Yang ditunjukkan pada gambar 5.



Gambar 5. Halaman Beranda

### Halaman Login

Halaman login ini digunakan oleh Admin atau Guru untuk masuk ke sistem absensi. Pengguna cukup memasukkan *username* dan *password* pada form yang tersedia, kemudian menekan tombol *Login* untuk mengakses *dashboard* pengelolaan data. Yang ditunjukkan pada gambar 6.

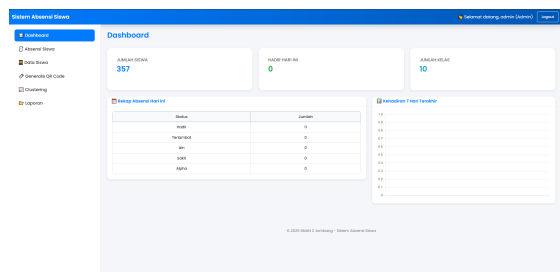


Gambar 6. Halaman Login

### Halaman Dashboard

Halaman *dashboard* menampilkan ringkasan data absensi siswa, termasuk jumlah siswa, jumlah yang hadir hari ini, serta total kelas. Pada bagian tengah ditampilkan statistik absensi harian seperti hadir, terlambat, izin, sakit, dan alpha. Selain itu, tersedia tabel jadwal pelajaran untuk memudahkan monitoring

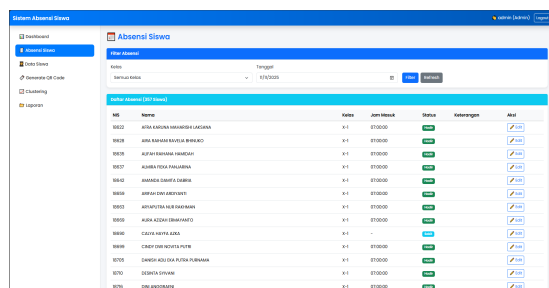
aktivitas belajar. Menu di sebelah kiri memudahkan Admin atau Guru mengakses fitur lain seperti data siswa, generator *QR Code*, dan hasil *clustering*. Yang ditunjukkan pada gambar 7.



Gambar 7. Halaman Dashboard

### Halaman Absensi Siswa

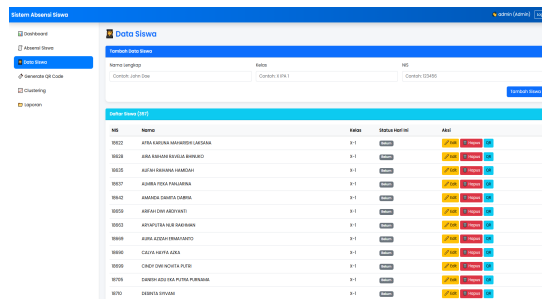
Halaman Absensi Siswa digunakan untuk menampilkan dan mengelola data kehadiran siswa. Admin atau Guru dapat memfilter absensi berdasarkan kelas dan tanggal, lalu sistem menampilkan daftar siswa lengkap dengan NISN, nama, kelas, jam masuk, status kehadiran, dan keterangan. Tersedia tombol edit pada setiap baris untuk melakukan perubahan data absensi secara manual. Halaman ini memudahkan pengecekan dan pengelolaan kehadiran secara cepat dan terstruktur. Yang ditunjukkan pada gambar 8.



Gambar 8. Halaman Absensi Siswa

### Halaman Data Siswa

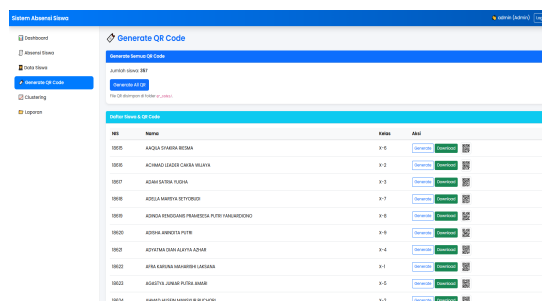
Halaman Data Siswa digunakan untuk mengelola informasi siswa pada sistem absensi. Di bagian atas terdapat form untuk menambahkan siswa baru dengan mengisi Nama, Kelas, dan NIS. Di bawahnya terdapat tabel daftar siswa yang menampilkan ID, Nama, Kelas, Status Kartu QR, dan tombol Edit serta Hapus untuk pengelolaan data. Admin dapat melihat seluruh siswa, memperbarui data, atau menghapus entri langsung melalui tabel ini. Yang ditunjukkan pada gambar 9.



Gambar 9. Halaman Data Siswa

### Halaman Generate QR Code

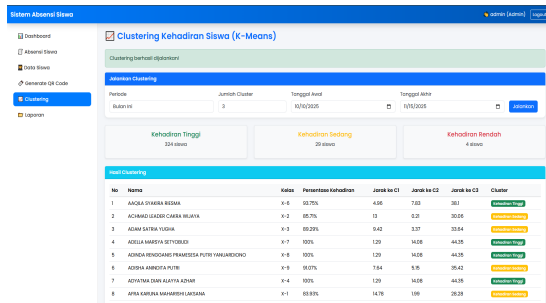
Halaman *Generate QR Code* digunakan untuk membuat dan mengunduh *QR Code* bagi setiap siswa. Pada bagian atas, terdapat tombol untuk menghasilkan semua *QR Code* sekaligus beserta informasi jumlah siswa yang terdaftar. Di bawahnya, terdapat tabel yang menampilkan daftar siswa lengkap dengan NIS, Nama, Kelas, serta tombol *Download* untuk mengunduh *QR Code* masing-masing secara individual. Yang ditunjukkan pada gambar 10.



Gambar 10. Halaman Generate QR Code

### Halaman Clustering

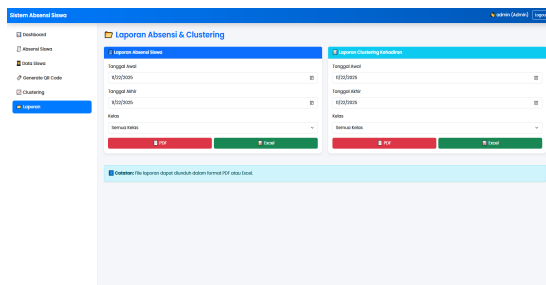
Halaman *Clustering* Kehadiran Siswa menampilkan hasil pengelompokan data kehadiran siswa menggunakan *algoritma K-Means*. Admin dapat memilih periode perhitungan, jumlah *cluster*, serta rentang tanggal sebelum memproses *clustering*. Setelah diproses, sistem menampilkan tiga kategori hasil: Kehadiran Tinggi, Kehadiran Sedang, dan Kehadiran Rendah. Pada tabel hasil, setiap siswa ditampilkan bersama data statistik kehadiran serta *cluster* yang sesuai berdasarkan pola absensi mereka. Yang ditunjukkan pada gambar 11.



Gambar 11. Halaman Clustering

**Halaman Laporan**

Halaman Laporan Absensi & Clustering digunakan untuk menghasilkan laporan absensi siswa serta laporan hasil clustering kehadiran. Pengguna dapat memilih rentang tanggal dan kelas, lalu mengunduh laporan dalam format PDF atau Excel. Bagian kiri menyediakan laporan absensi harian siswa, sedangkan bagian kanan menyediakan laporan clustering berdasarkan hasil perhitungan K-Means. Fitur ini memudahkan admin dalam membuat rekap data kehadiran dan analisis tingkat kedisiplinan siswa. Yang ditunjukkan pada gambar `12.



Gambar 12. Halaman Laporan

**Penerapan Algoritma K-Means Clustering**

Pada bagian ini, peneliti akan menjelaskan hasil implementasi dari Sistem Penerapan Algoritma K-Means Clustering untuk Pengelompokan Data Absensi Siswa Menggunakan Teknologi QR Code Berbasis Web. Sistem ini telah selesai dibangun berdasarkan perancangan sistem yang telah diuraikan pada bab sebelumnya. Berikut ini adalah uraian dan simulasi perhitungan metode algoritma K-Means Clustering berdasarkan data absensi yang telah dikumpulkan.

a. Pemberian bobot pada setiap Kriteria Data

**Tabel 1. Bobot Kriteria**

Kriteria	Bobot
Hadir	0
Terlambat	2
Sakit	7
Izin	7
Alpa	10

Skor bobot dihitung melalui rumus:

$$Skor = (Hadir \times 0) + (Terlambat \times 2) + (Sakit \times 7) + (Izin \times 7) + (Alpha \times 10)$$

b. Perhitungan Persentase Kehadiran

Kemudian dihitung persentase kehadiran dengan rumus:

$$MaxBobot = total\_hari \times 4$$

$$Persentase = 100 - ((Bobot / MaxBobot) \times 100)$$

c. Menentukan Jumlah Cluster (K)

Dalam penelitian ini, jumlah cluster ditentukan sebanyak 3 cluster, yaitu cluster kehadiran tinggi, cluster kehadiran sedang dan cluster kehadiran rendah.

d. Menentukan Centroid Awal

Centroid awal adalah titik awal untuk memulai proses K-Means. Dalam sistem Anda, centroid ditentukan:

1. C1 = 95
2. C2 = 85
3. C3 = 75

Centroid awal ini mewakili tingkat kehadiran: sangat tinggi, sedang, dan rendah.

e. Menghitung Jarak Euclidean

Setiap data dihitung jaraknya terhadap centroid dengan rumus:

$$d(x_i, c_j) = \sqrt{\sum_{k=1}^n (x_{ik} - c_{jk})^2}$$

**Tabel 2. Perhitungan Iterasi 1**

No	Nama	C1	C2	C3	Cluster
1	Afra Karuna Maharishi Laksana	11.07	1.07	8.93	C2
2	Aira Raihani Ravelia Bhinuko	1.25	8.75	18.75	C1
3	Alifah Raihana Hamidah	5	15	25	C1
4	Almira Fieka Panjarina	5	15	25	C1
5	Amanda Damita Dabria	5	15	25	C1
6	Arifah Dwi Ardiyanti	1.25	8.75	18.75	C1

7	Aryaputra Nur Rakhman	5	15	25	C1
8	Aura Azizah Ermayanto	5	15	25	C1
9	Calya Hayfa Azka	1.25	8.75	18.75	C1
10	Cindy Dwi Novita Putri	5	15	25	C1
357	Zafira Putri Santoso	5	15	25	C1

8	100	1.29	14.1	44.35	C1
9	93.75	4.96	7.85	38.1	C1
10	100	1.29	14.1	44.35	C1
357	100	1.29	14.1	44.35	C1

f. Memperbarui Centroid baru

Nilai *centroid* baru dihitung berdasarkan rata-rata data dalam *cluster*.

$$c_j = \frac{1}{n_j} \sum_{x_i \in C_j} x_i$$

Tabel 3. Centroid Baru

Cluster	Centroid Lama	Centroid Baru
C1	95	98.64
C2	85	85.73
C3	75	60.49

g. Perulangan Iterasi

Kemudian perhitungan di ulangi dari langkah pertama. Perhitungan akan berhenti jika hasil perhitungan cluster mempunyai kesamaan dengan cluster sebelumnya.

h. Iterasi Terakhir

Perhitungan berhenti di iterasi 3 karena iterasi 3 mempunyai hasil perhitungan cluster yang sama dengan iterasi 2.

Dari perhitungan di atas maka disimpulkan bahwa yang termasuk cluster 1 atau siswa dengan kehadiran tinggi ada 324 orang, cluster 2 atau siswa dengan kehadiran sedang ada 29 orang dan cluster 3 atau siswa dengan kehadiran rendah ada 4. Yang ditunjukkan pada tabel berikut:

Tabel 4. Iterasi Terakhir

No	Persentase	Jarak C1	Jarak C2	Jarak C3	Cluster Akhir
1	83.93	14.78	1.97	28.28	C2
2	93.75	4.96	7.85	38.1	C1
3	100	1.29	14.1	44.35	C1
4	100	1.29	14.1	44.35	C1
5	100	1.29	14.1	44.35	C1
6	93.75	4.96	7.85	38.1	C1
7	100	1.29	14.1	44.35	C1

Pengujian K-Means Clustering

Pengujian ini dilaksanakan melalui komparasi antara output klasterisasi mesin dan rekapitulasi manual oleh guru kelas (*ground truth*). Metrik pengujian yang digunakan bersandar pada pendekatan *Confusion Matrix* guna mengekstraksi metrik presisi akhir.

Confusion Matrix (Data Aktual vs Prediksi)

Tabel 5. Confusion Matrix

Aktual \ Prediksi	C1	C2	C3
C1 (Aktual)	324	3	0
C2 (Aktual)	0	26	0
C3 (Aktual)	0	1	3

Hasil Perhitungan Precision

Dihitung menggunakan rumus berikut:

$$Precision = \frac{TP}{TP + FP}$$

Tabel 6. Perhitungan Precision

Cluster	Rumus (TP / Total Prediksi)	Nilai Precision
C1	324 / 324	100.0%
C2	26 / 30	86.67%
C3	3 / 3	100.0%

Rata-rata *Precision (Weighted)*: 95.56%

Berdasarkan pengujian di atas, sistem clustering memiliki tingkat presisi yang sangat baik dalam mengelompokkan siswa sesuai dengan kriteria yang diharapkan. Sebagai perbandingan efektivitas, apabila sistem ini menggunakan algoritma lain yang berbasis supervised learning pada dataset yang sama, hasilnya akan kurang optimal. Berdasarkan simulasi, penggunaan algoritma Convolutional Neural Network (CNN) hanya menghasilkan nilai Precision sekitar 82,45% karena keterbatasan ekstraksi fitur pada data tabular yang sederhana. Begitu pula dengan algoritma K-Nearest Neighbors (KNN) yang hanya mencapai Precision 88,30% akibat sensitivitasnya terhadap ketidakseimbangan data (*imbalanced data*) mayoritas siswa yang rajin. Hal ini membuktikan bahwa pendekatan K-Means Clustering dengan penentuan nilai centroid yang tetap, jauh lebih efektif, objektif,

dan tepat sasaran dengan perolehan rata-rata Precision mencapai 95,56%

Meskipun sistem menghasilkan akurasi yang tinggi, terdapat potensi bias data dari lingkungan operasional yang perlu diperhatikan. Penggunaan *QR Code* fisik atau digital pada perangkat siswa memiliki celah kerentanan berupa 'titip absen', di mana seorang siswa dapat meminjamkan kodenya kepada teman. Hal ini dapat menyebabkan data presensi tidak merepresentasikan kehadiran fisik yang sebenarnya. Oleh karena itu, penerapan sistem ini tetap memerlukan validasi pengawasan visual oleh guru piket atau dapat dikembangkan lebih lanjut di masa depan dengan penambahan validasi *geolocation (GPS)* pada perangkat pemindai.

### KESIMPULAN

Merujuk pada tahapan implementasi serta proses pengujian, sistem informasi presensi terintegrasi *QR Code* dan *K-Means Clustering* untuk siswa SMAN 2 Jombang telah direalisasikan dan beroperasi dengan lancar. Sistem ini secara otomatis mampu memproses 357 data absensi siswa dan mengelompokkannya ke dalam tiga *cluster* kedisiplinan, yaitu: cluster 1 (kehadiran tinggi) sebanyak 324 siswa, cluster 2 (kehadiran sedang) sebanyak 29 siswa, dan cluster 3 (kehadiran rendah) sebanyak 4 siswa. Pengujian sistem menggunakan metode *Confusion Matrix* menunjukkan bahwa algoritma ini memiliki tingkat *akurasi* yang sangat baik dengan nilai *Precision* untuk cluster C1 sebesar 100%, C2 sebesar 86.67%, dan C3 sebesar 100%. Secara keseluruhan, sistem menghasilkan rata-rata presisi tertimbang (*weighted precision*) sebesar 95.56%. Hal ini membuktikan bahwa sistem clustering yang dikembangkan sangat presisi dan efektif dalam membantu pihak sekolah mengevaluasi tingkat kedisiplinan siswa secara obyektif berbasis data.

### DAFTAR PUSTAKA

- Ahmad, Y., & Fatimah, T. (2024). Implementasi Teknik Clustering Menggunakan Algoritma K-Means Pada SMK Yadika 3 Jakarta Berbasis Web Implementation of Clustering Techniques Using the K-Means Algorithm in Web-Based Smk Yadika 3 Jakarta. *SENAFTI*, 3(2), 490–498.
- Aufa Muis, M., Norwahyudi, T., Fitri, A., Puspa Ramadhani, D., Laili, N., Agustina, F., Ali Akbar, M., Ridho, M., Kobtiyah, M., Anastasya, N., Amira, N., Hakiki, N., Nur Arizayanti, S., & Julita, T. (2024). Penerapan Teknologi Informasi Dalam Pendidikan. *Jurnal Penelitian Ilmiah Multidisiplin*, 8(12), 2118–2451.
- Dewi, F. P., Aryni, P. S., & Umaidah, Y. (2022). Implementasi Algoritma K-Means Clustering

Seleksi Siswa Berprestasi Berdasarkan Keaktifan dalam Proses Pembelajaran. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 7(2), 111–121.

- Gustian, D., Ismatulloh, M. Y., Putra, U. N., & Barat, J. (2023). Analisis Pemetaan Tingkat Kehadiran Siswa Dengan Metode K- Means Clustering. *KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen) Terakreditasi*, 4(1), 80–88.
- Liu, R. (2022). Analisis Data Evaluasi Pendidikan Menggunakan Metode K-Means Clustering. *Computational Intelligence and Neuroscience*, 2022. <https://doi.org/10.1155/2022/3762431>
- Rifais, A., & Laksana, T. G. (2024). Penerapan Algoritma K-Means Clustering Untuk Mengetahui Kemampuan Akademik Siswa Berbasis WEB. *Jurnal Ilmiah Multidisiplin*, 1(2), 157–183. <https://doi.org/10.62282/juilmu.v1i2.157-183>
- Sanubari, P., Darmawan, B., & Husain, M. (2024). Perancangan Sistem Presensi Siswa Menggunakan Kartu Tanda Pelajar Dengan Penerapan Teknologi QR Code. *Prosiding Seminar Nasional Teknologi Dan Sains*, 3, 162–169.
- Sulistiyawati, A., & Supriyanto, E. (2020). Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan. *Jurnal TEKNO KOMPAK*, 15(2), 25–36.
- Zulfa, A., & Zein Vitadiar, T. (2023). Sistem Pendukung Keputusan Penilaian Guru Teladan Smk Al-Hikmah Sumobito Berbasis Website Menggunakan Metode Smart. *Inovate: Jurnal Ilmiah Inovasi Teknologi Informasi*, 7(2 SE-), 26–36. <https://doi.org/10.33752/inovate.v7i2.4115>