

## BAB II

### TINJAUAN PUSTAKA

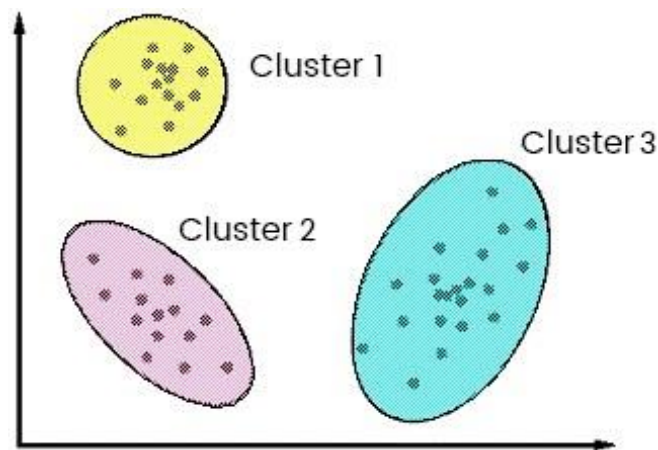
#### 1.1 Dasar Teori

##### 1.1.1 Data Mining

*Data mining* atau *knowledge discovery in databases* (KDD) adalah proses menemukan pola, hubungan, atau informasi bermakna dari kumpulan data besar menggunakan teknik statistik, pembelajaran mesin, dan basis data. Proses *data mining* meliputi pemilihan data, pra-pemrosesan (*cleaning*, transformasi, reduksi dimensi), pemodelan (penerapan teknik mining seperti klasifikasi, *clustering*, asosiasi), evaluasi model, dan presentasi pengetahuan. Dalam aplikasi praktis, *data mining* dipakai untuk segmentasi pelanggan, deteksi anomali, prediksi, dan rekomendasi [3].

##### 1.1.2 Clustering

*Clustering* adalah teknik pembelajaran tanpa pengawasan (*unsupervised learning*) yang membagi sekumpulan objek menjadi kelompok-kelompok sedemikian rupa sehingga objek di dalam kelompok (*cluster*) memiliki kemiripan tinggi, sedangkan objek antar *cluster* berbeda. *Clustering* tidak memerlukan label kelas sebelumnya sehingga cocok untuk eksplorasi data dan segmentasi. Metode *clustering* meliputi algoritma partisi (*K-Means*, *K-Medoids*), hirarkis (*agglomerative*, *divisive*), *density-based* (DBSCAN), dan model-based (*Gaussian Mixture Models*). Pemilihan metode bergantung pada bentuk data, ukuran dataset, dan tujuan analisis [4].



**Gambar 2.1 Diagram Clustering**  
 Sumber : <https://assets.janbasktraining.com/>

### 1.1.3 Algoritma *K-Means*

*K-Means* adalah merupakan salah satu metode dalam *Data Mining* yang dapat mengelompokkan data atau *Clustering* sebuah data kedalam bentuk satu cluster atau lebih *cluster* sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu *cluster* yang sama dan data dengan karakteristik yang berbeda dikelompokkan ke dalam kelompok berbeda yang lainnya. Sarwono mengemukakan secara detail, algoritma *K-Means* adalah sebagai berikut [3]:

1. Menentukan nilai  $k$  sebagai jumlah cluster yang ingin di bentuk.
2. Menentukan nilai acak atau random untuk pusat *cluster* awal *centroid* sebanyak  $k$ , untuk menghitung jarak setiap data input terhadap masing-masing *centroid* dengan menggunakan rumus jarak *Euclidean Distance* yaitu :

$$d(x_i, \mu_j) = \sqrt{\sum (x_i - \mu_j)^2} = \sqrt{\sum (x_i - u_j)^2} \quad (1)$$

Dimana:

$x_i$  = data kriteria

$\mu_j$  = *centroid* pada *cluster* ke- $j$ s

3. Mengelompokkan setiap data berdasarkan kedekatannya dengan *centroid* atau mencari jarak terkecil.
4. Memperbaharui nilai centroid baru, nilai *centroid* baru di peroleh dari rata-rata cluster yang bersangkutan dengan menggunakan rumus yaitu :

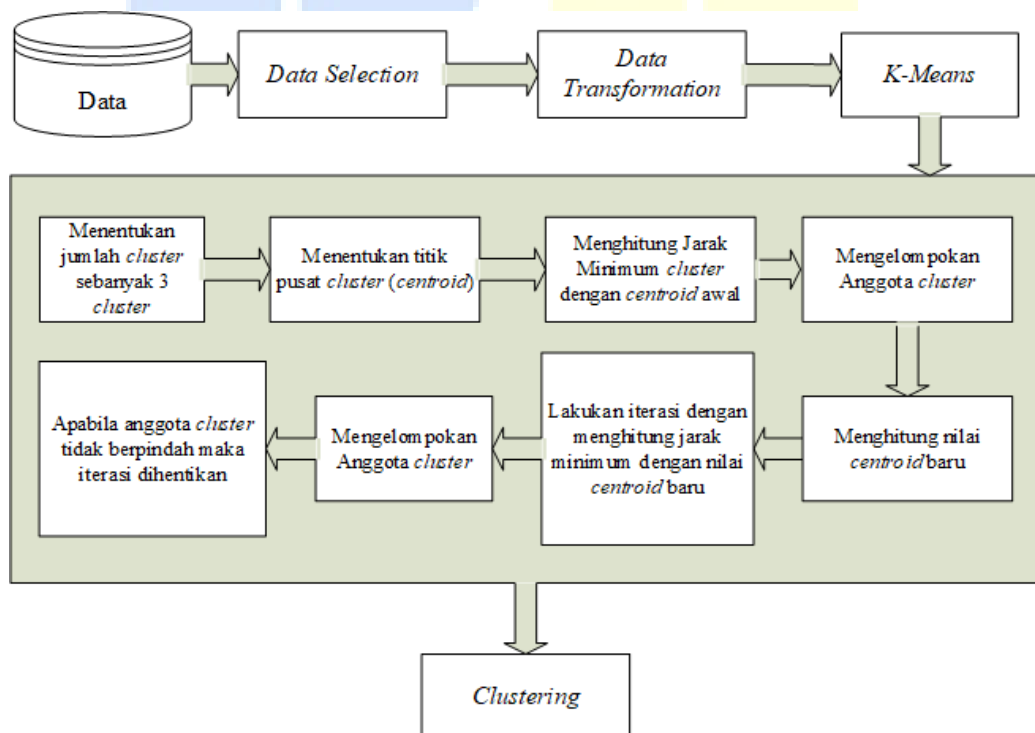
$$\mu_j(t+1) = \frac{1}{N_{sj}} \sum_{j \in S_j} x_j \quad (2)$$

Keterangan :

$\mu_j(t+1)$  = *centroid* baru pada iterasi (t+1)

$N_{sj}$  = Data pada *cluster*  $S_j$

5. Apabila data setiap *cluster* belum berhenti, lakukan perulangan dari langkah 2 hingga 5, sampai anggota tiap *cluster* tidak ada yang berubah



Gambar 2.2 Blok Diagram *K-Means Clustering*

Perbaikan pada *K-Means* modern mencakup inisialisasi  $k$  yang lebih baik (mis. *K-Means++*), normalisasi fitur sebelum *clustering*, dan pemilihan  $k$  dengan metode *Elbow* atau *Silhouette*. Dalam aplikasi perpustakaan sekolah, fitur yang sering dipakai untuk *K-Means* meliputi frekuensi peminjaman per kategori, jumlah buku dipinjam per siswa, durasi pinjam, dan atribut demografis (kelas/jurusan). Evaluasi kluster harus disertai interpretasi yang dapat diterjemahkan ke tindakan praktis oleh pustakawan. Studi-studi terkini (2021–2025) menunjukkan *K-Means* cocok sebagai metode awal segmentasi minat baca karena sederhana dan mudah diintegrasikan ke *backend* aplikasi *web* [5].

#### **1.1.4 Konsep Manajemen Data Akademik Dan Minat Belajar**

Manajemen data akademik adalah proses mengelola informasi siswa, seperti nilai, absensi, dan perkembangan belajar, agar data tersimpan dengan baik, mudah diakses, dan dapat digunakan dalam pengambilan keputusan pendidikan. Pada lembaga seperti MDTA, pengelolaan data yang baik penting untuk memantau perkembangan santri dan mengenali pola belajar mereka [6]. Tujuan utama manajemen data akademik ialah menyediakan informasi yang akurat dan terstruktur untuk mendukung evaluasi pembelajaran. Dengan data yang dikelola secara sistematis, madrasah dapat menilai kedisiplinan, pemahaman materi, dan minat belajar santri secara lebih objektif. Sebaliknya, data yang tidak teratur akan menyulitkan madrasah dalam melakukan pemantauan dan intervensi pembelajaran [7]. Jika data akademik dikelola dengan baik, madrasah dapat mengelompokkan santri berdasarkan tingkat minat belajar serta memahami faktor-faktor yang

memengaruhinya. Hal ini selaras dengan konsep pengendalian data akademik untuk mendukung perencanaan pembelajaran yang lebih tepat sasaran

### **1.1.5 Sistem Informasi Pendidikan**

Apabila jumlah siswa dan data akademik semakin bertambah, pengelolaan informasi secara manual melalui catatan kertas atau *spreadsheet* menjadi tidak efisien, rawan kesalahan, dan sulit diperbarui secara cepat. Karena itu, banyak lembaga pendidikan mulai menggunakan Sistem Informasi Pendidikan (*Educational Information System / Academic Information System*) [8]. Sistem Informasi Pendidikan adalah aplikasi berbasis komputer yang digunakan untuk mencatat data siswa, nilai, absensi, serta perkembangan belajar, dan menghasilkan laporan akademik secara otomatis. Sistem ini memungkinkan guru dan pengelola madrasah memantau nilai dan kehadiran siswa secara *real-time* sehingga proses evaluasi dapat dilakukan dengan lebih akurat [9]. Manfaat sistem ini antara lain mempercepat input data, mengurangi kesalahan pencatatan, memudahkan pelacakan perkembangan siswa, serta membantu pihak sekolah atau madrasah dalam merencanakan program pembelajaran berdasarkan data yang lebih terstruktur [10].


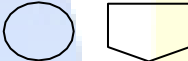



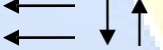


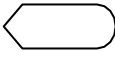

### **1.1.6 Alat Bantu Perancangan Sistem**

#### **1.1.6.1 Aliran Sistem Informasi**

Aliran sistem informasi menggambarkan bagaimana data dikumpulkan, diproses, dan disajikan menjadi informasi yang berguna bagi organisasi. Data

pertama-tama dikumpulkan dari berbagai sumber (*input*), kemudian diolah melalui sistem dan prosedur tertentu (*proses*), dan hasilnya disampaikan kepada pengguna dalam bentuk laporan atau tampilan informasi (*output*). Umpan balik (*Feedback*) dari informasi yang dihasilkan digunakan untuk menyempurnakan proses berikutnya, sehingga sistem tetap efektif dalam mendukung pengambilan keputusan, koordinasi, dan pengelolaan operasional [11]. Berikut Simbol dari aliran sistem informasi.

**Tabel 2.1 Simbol Aliran Sistem Informasi**

No	Nama	Simbol	Keterangan
1	Proses		Untuk proses pengolahan data
2	Penghubung		Untuk menghubungkan sambungan aliran
3	Dokumen		Digunakan untuk operasi <i>Input</i>
4	Arsip		Merupakan arsip data yang dihasilkan
5	Proses manual		Untuk proses pengolahan data secara manual
6	Aliran Sistem		Untuk arah pengaliran data proses
7	Basis Data		Untuk media penyimpanan secara terkomputerisasi
8	Pita Kertas		Untuk menunjukkan <i>Input/Output</i> menggunakan pita kertas
9	Display		Untuk menampilkan <i>Output</i> kelayar monitor
10	Manual <i>Input</i> keyboard		Untuk manual <i>Input</i> menggunakan keyboard

Sumber: [14]

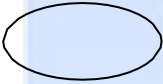

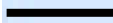

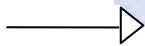
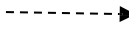
### 1.1.6.2 Unified Modeling Language (UML)

*Unified Modeling Language (UML)* adalah sekumpulan standar diagram yang digunakan untuk memvisualisasikan, merancang, dan mendokumentasikan setiap tahap pengembangan sistem, mulai dari analisis hingga implementasi, dengan cara yang mudah dipahami melalui. Berikut adalah jenis-jenis UML [12]:

#### 1. Use Case Diagram

*Sequence Diagram* adalah diagram yang menunjukkan bagaimana objek-objek dalam suatu *class* saling berinteraksi dan bekerja sama dalam menjalankan suatu proses atau fungsi. Berikut adalah simbol dari *Sequence Diagram* [12].

**Tabel 2.2 Simbol Use Case Diagram**

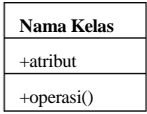





No.	Simbol	Nama	Keterangan
1		<i>Use Case</i>	Deskripsi dari urutan aksi-aksi yang ditampilkan sistem untuk menghasilkan sesuatu bagi aktor.
2		Aktor ( <i>actor</i> )	Mewakili peran orang, proses, atau sistem lain yang berinteraksi dengan sistem informasi yang akan dibuat.
3		Asosiasi ( <i>Association</i> )	Menghubungkan antara aktor dan <i>use case</i> yang ikut berperan pada <i>use case</i> .
4		Ekstensi ( <i>Extend</i> )	Menunjukkan relasi <i>use case</i> ke sebuah <i>use case</i> dimana <i>use case</i> dapat berdiri sendiri
5		Generalisasi ( <i>Generalization</i> )	Menunjukkan hubungan antara dua buah <i>use case</i> dimana objek anak berbagi struktur data dan perilaku dari objek yang ada di atasnya.
6		<i>Include</i> <i>/uses</i>	Menunjukkan hubungan <i>use case</i> tambahan ke sebuah <i>use case</i> yang ditambahkan fungsionalitas dari <i>use case</i> lain.

Sumber: [15]

## 2. Class Diagram

*Class Diagram* adalah diagram yang menggambarkan hubungan antar kelas dalam suatu sistem, termasuk atribut dan metode yang dimiliki oleh masing-masing objek. Berikut adalah simbol dari *Class Diagram* [12]:

**Tabel 2.3 Simbol Class Diagram**


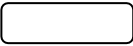
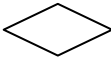


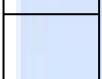
No	Simbol	Nama	Keterangan
1		Kelas ( <i>Class</i> )	Kumpulan objek-objek yang berbagi atribut dan operasi yang sama.
2		Antarmuka ( <i>Interface</i> )	Berorientasi objek.
3		Asosiasi Berarah ( <i>Directed Association</i> )	Hubungan antarkelas dengan makna kelas yang satu digunakan oleh kelas yang lain.
4		Generalisasi	Hubungan dimana objek anak berbagi struktur data dan berbagi perilaku dari objek yang ada di atasnya objek induk.
5		Kebergantungan ( <i>Dependency</i> )	Hubungan antarkelas dengan makna yang bergantung antarkelas.
6		Agregasi ( <i>Aggregation</i> )	Hubungan antarkelas dengan makna semua-bagian.

Sumber: [15].

## 3. Activity Diagram

*Activity Diagram* adalah diagram yang menggambarkan aliran kegiatan, aksi, dan kontrol secara terstruktur dalam suatu sistem, sehingga memudahkan pemahaman proses yang terjadi. Berikut adalah simbol pada *Activity Diagram* [12].

Tabel 2.4 Simbol *Activity Diagram*

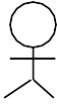
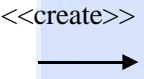
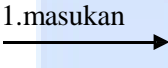
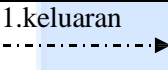
No	Simbol	Nama	Keterangan
1		Status Awal ( <i>Start Point</i> )	Menggambarkan aktivitas awal pada sistem.
2		Aktivitas ( <i>Activity</i> )	Menggambarkan aktivitas yang dilakukan sistem.
3		Percabangan ( <i>Decision</i> )	Menggambarkan pilihan atau keputusan.
4		Penggabungan ( <i>Join</i> )	Penggabungan dimana lebih dari satu aktivitas digabungkan menjadi satu.
5		Status Akhir ( <i>End Point</i> )	Menggambarkan aktivitas akhir yang dilakukan sistem.
6		<i>Swimlane</i>	Memisahkan aktivitas berdasarkan aktor.

Sumber: [15]

#### 4. *Sequence Diagram*

*Sequence Diagram* adalah diagram yang menggambarkan bagaimana objek atau elemen dalam sebuah kelas saling berinteraksi secara berurutan dari awal hingga akhir proses. Diagram ini termasuk jenis dalam UML, yang digunakan untuk memodelkan dinamika sistem dan memperlihatkan alur komunikasi antar objek dalam suatu skenario atau proses. Berikut ini adalah simbol dari *Sequence Diagram* [12].

Tabel 2.5 Simbol *Sequence Diagram*


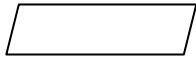

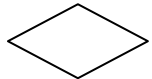
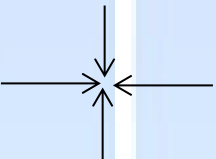
No	Simbol	Nama	Keterangan
1		Aktor ( <i>Actor</i> )	Deskripsi dari orang, proses, atau sistem lain yang berinteraksi dengan sistem informasi yang akan dibuat di luar sistem informasi yang akan dibuat itu sendiri.
2		Garis Hidup ( <i>Lifeline</i> )	Menggambarkan kehidupan dari suatu objek.
3		Objek ( <i>Object</i> )	Menggambarkan objek yang berinteraksi pesan.
4		Waktu aktif	Objek dalam keadaan aktif dan berinteraksi.
5		Pesan tipe <i>create</i>	Suatu objek membuat objek yang lain, arah panah mengarah pada objek yang dibuat.
6		Pesan tipe <i>call</i>	Suatu objek memanggil operasi yang ada pada objek lain atau dirinya sendiri.
7		Pesan tipe <i>send</i>	Suatu objek mengirimkan data atau informasi ke objek lainnya, arah panah mengarah pada objek yang dikirim.
8		Pesan tipe <i>return</i>	Suatu objek yang telah menjalankan suatu operasi kembalian ke objek tertentu, arah panah mengarah pada objek yang menerima kembalian.
9		Pesan tipe <i>destroy</i>	Sebuah objek mengakhiri hidup objek yang lain, arah panah mengarah pada objek yang diakhiri.

Sumber: [15].

### 1.1.6.3 Flowchart

*Flowchart* adalah diagram yang memperlihatkan alur logis dari sebuah program atau prosedur dalam suatu sistem, sehingga memudahkan pemahaman langkah-langkah prosesnya. Berikut adalah simbol dari *Flowchart* [13].

Tabel 2.6 Simbol *Flowchart*

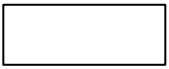
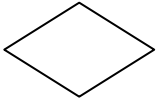

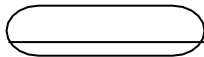
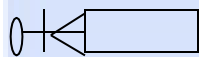
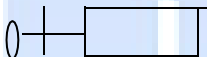
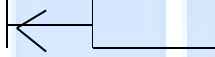
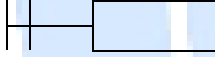
No	Gambar	Keterangan
1		Terminal untuk memulai dan mengakhiri suatu program.
2		Proses, suatu simbol yang menunjukkan sikap pengolahan yang dilakukan oleh komputer.
3		Input-Output, untuk memasukkan data maupun menunjukkan hasil dari suatu proses.
4		<i>Predefined Process</i> , suatu simbol untuk menyediakan tempat-tempat pengolahan data dalam <i>storage</i> .
5		<i>Decision</i> , suatu kondisi yang akan menghasilkan beberapa kemungkinan jawaban atau pilihan.
6		<i>Connector</i> , suatu prosedur akan masuk dan keluar melalui simbol ini dalam lembar yang sama.
7		<i>Off line connector</i> , merupakan simbol untuk masuk dan keluarnya suatu prosedur pada lembar kertas yang lain
8		Arus atau <i>flow</i> , prosedur yang dapat dilakukan dari atas kebawah, bawah keatas, dari kekanan, atau dari kanan kekiri.
9		<i>Document</i> , merupakan simbol untuk data yang berbentuk informasi.
10		Untuk menyatakan sekumpulan langkah proses yang ditulis sebagai prosedur.

Sumber: [16].

#### 1.1.6.4 Entity Relationship Diagram (ERD)

*Entity Relationship Diagram* (ERD) adalah diagram dengan notasi grafis yang digunakan dalam perancangan *Database* untuk menunjukkan hubungan antara satu data dengan data lainnya. Berikut adalah simbol ERD [12]:

Tabel 2.7 Simbol *Entity Relationship Diagram*

No	Simbol	Keterangan
1		<i>Entity</i> : digunakan untuk menggambarkan obyek yang di indentifikasi ke dalam lingkungan
2		Relasi/aktivitas dalam <i>entity</i>
3		<i>Simple atribut</i>
4		<i>Field</i> atau <i>primary key attribute</i>
5		Hubungan antar <i>entity</i> dengan derajat kardinalitas relasi <i>optional many</i>
6		Hubungan antar <i>entity</i> dengan derajat kardinalitas relasi <i>optional one</i>
7		Hubungan antar <i>entity</i> dengan kardinalitas relasi <i>mandatory many</i>
8		Hubungan antar <i>entity</i> dengan derajat kardinalita relasi <i>mandatory one</i>

Sumber: [15].

### 1.1.7 Perangkat Lunak yang Digunakan

Dalam pembuatan skripsi ini, penulis membutuhkan beberapa perangkat lunak pendukung, dalam memecahkan masalah yang ada. Beberapa penjelasan dari perangkat lunak pendukung yaitu sebagai berikut:

## 1. *Hypertext Preprocessor (PHP)*

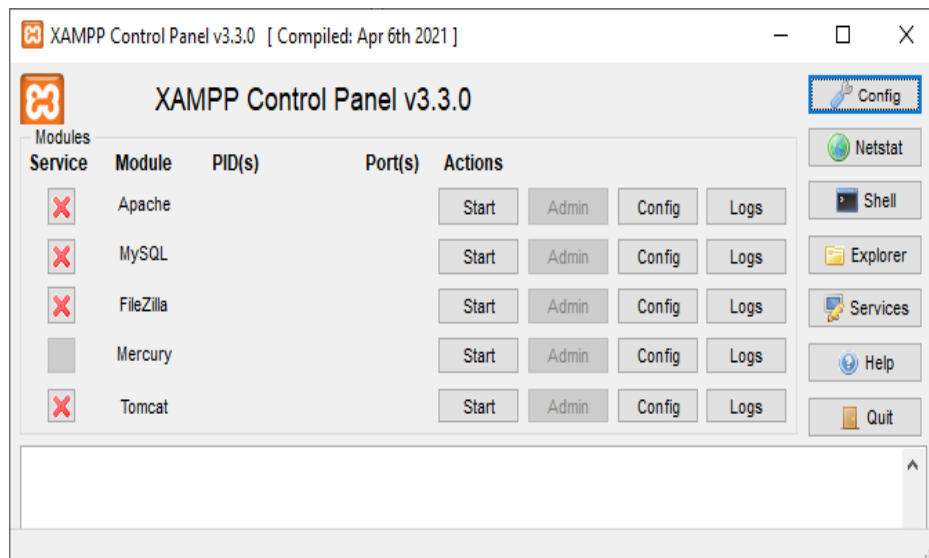
Penelitian ini menggunakan beberapa perangkat lunak utama yang saling mendukung. Bahasa pemrograman yang digunakan adalah PHP versi 8.2, yang dipilih karena merupakan versi stabil dan *modern* yang menyediakan peningkatan performa, keamanan, serta kompatibilitas dengan berbagai pustaka dan *Framework* yang digunakan dalam pengembangan aplikasi *web*. PHP bertindak sebagai "otak" yang memproses logika algoritma K-Means sebelum hasilnya ditampilkan di browser.



**Gambar 2.3 PHP**

## 2. XAMPP

Untuk menjalankan *Server* lokal selama proses pengembangan, digunakan XAMPP versi 8.2.4, yang di dalamnya telah terintegrasi *Apache Web Server*, PHP 8.2, dan MySQL. Penggunaan XAMPP memudahkan peneliti dalam menyiapkan lingkungan pengembangan tanpa perlu melakukan instalasi terpisah untuk *Server* maupun *database*, sekaligus mempermudah proses konfigurasi dan pengujian aplikasi sebelum diunggah ke *server* produksi. XAMPP adalah perangkat lunak *open-source* yang menyediakan lingkungan pengembangan *server web* secara lokal (*Localhost*) sehingga tidak perlu secara manual meng-*install* dan mengonfigurasi *Server Apache*, MySQL/MariaDB, PHP, namun XAMPP sudah menyatukan semuanya [14].



**Gambar 2.4 XAMPP**

### 3. MySQL

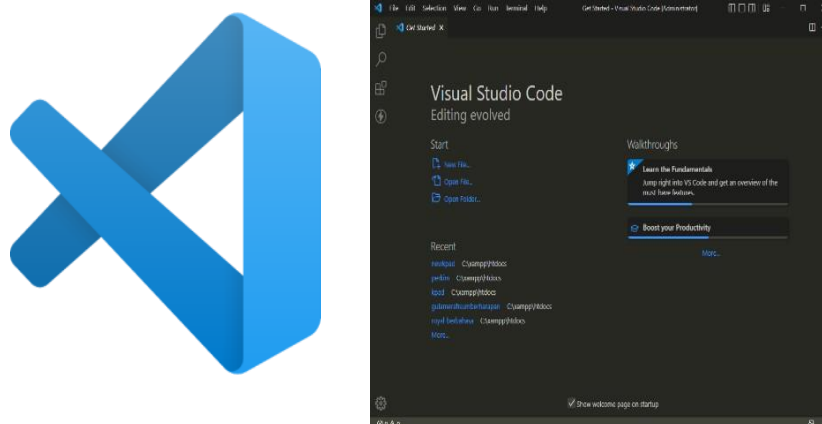
Basis data sistem ini dikelola menggunakan MySQL versi 8.0, yang dikenal sebagai sistem manajemen basis data relasional (RDBMS) dengan performa unggul, mendukung *query* kompleks, serta mampu menangani pertumbuhan data siswa dan lainnya dalam jangka panjang. MySQL adalah sebuah sistem manajemen basis data relasional (*Relational Database Management System*) yang menggunakan bahasa SQL sebagai standar untuk mengakses datanya. MySQL versi ini juga kompatibel penuh dengan PHP 8.2, sehingga proses integrasi berjalan stabi. MySQL merupakan perangkat lunak sistem manajemen basis data yang fleksibel dan mendukung berbagai aktivitas penting, seperti membuat, menyimpan, mengubah, menghapus, serta mengambil data dalam tabel, sekaligus memungkinkan pemilihan data sesuai kebutuhan pengguna. MySQL disebut relasional karena data disimpan dalam tabel-tabel terpisah yang saling terhubung (berelasi). Fungsi utamanya meliputi pembuatan dan penghapusan basis data, pembuatan/modifikasi tabel, kolom dan relasi [15].



**Gambar 2.5 MySQL**

#### **4. *Visual Studio Code***

Penelitian ini menggunakan *Visual Studio Code* (VS Code) versi 1.92 sebagai *code editor* utama. VS Code versi ini mendukung berbagai ekstensi seperti PHP *IntelliSense*, MySQL, dan *Live Server*, yang sangat membantu dalam proses penulisan kode, *debugging*, serta integrasi sistem. Stabilitas dan fleksibilitasnya menjadikan VS Code sebagai pilihan editor yang umum digunakan dalam pengembangan aplikasi web modern. *Visual Studio Code* adalah editor kode sumber yang dikembangkan oleh *Microsoft* dan bersifat gratis serta *open-source*. *Visual Studio Code* digunakan sebagai lingkungan pengembangan utama (*Integrated Development Environment* atau IDE) untuk menulis, menyunting, dan melakukan *debugging* pada skrip PHP, file CSS untuk tampilan dashboard, pengelolaan basis data SQL dan perangkat lunak ini juga dirancang untuk membantu pengembang menulis dan mengedit kode dari berbagai bahasa pemrograman, seperti HTML, CSS, dan *JavaScript*. Selain fitur pengeditan, *Visual Studio Code* juga menyediakan alat untuk menemukan kesalahan kode, menyelesaikan kode secara otomatis, mengelola kontrol versi, serta mendukung berbagai ekstensi yang memudahkan setiap tahap pengembangan perangkat lunak [19].



**Gambar 2.6** Tampilan *Visual Studio Code*

## 5. *Google Chrome*

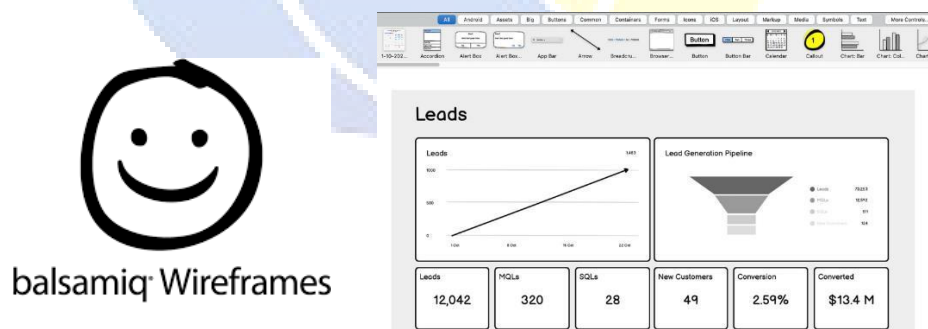
Pengujian tampilan antarmuka serta fungsionalitas aplikasi dilakukan menggunakan *Google Chrome* versi 129, karena *Browser* ini menyediakan fitur *Developer Tools* yang lengkap seperti *Console*, *Network Monitor*, dan *Performance Insights* yang sangat membantu dalam mendeteksi kesalahan script, memantau respons *Server*, mengoptimalkan tampilan antarmuka pada berbagai perangkat, serta *google chrome* juga perangkat lunak penjelajah web (*web browser*) yang digunakan sebagai media antarmuka untuk menampilkan sistem informasi yang dikembangkan. Selain itu, fitur *Developer Tools* pada *Google Chrome* digunakan secara intensif dalam tahap pengujian untuk melakukan *debugging* pada tampilan dashboard dan memastikan algoritma K-Means menampilkan data yang akurat pada sisi pengguna. Fitur *Chrome Developer Tools* digunakan secara ekstensif untuk memantau *network traffic* antara klien dan server *Chrome* juga memungkinkan pengguna masuk dengan akun *Google*, sehingga data seperti *bookmark* dan riwayat penelusuran dapat diselaraskan di berbagai perangkat secara mudah [16].



**Gambar 2.7 Google Chrome**

## 6. *Balsamiq Wireframes*

Dalam tahap perancangan antarmuka pengguna (UI), penelitian ini juga menggunakan *Balsamiq Wireframes* versi 4.7.5 untuk membuat rancangan sketsa tampilan aplikasi. *Balsamiq* membantu peneliti membuat desain antarmuka yang sederhana namun terstruktur dengan baik, sehingga memudahkan proses validasi desain sebelum diimplementasikan pada tahap pengembangan. Dalam konteks pengembangan sistem informasi berbasis *web*, *prototipe Balsamiq* sering diuji pada *Browser* modern seperti *Google Chrome*, sehingga tim dapat memverifikasi tampilan, navigasi, dan responsivitas desain di lingkungan yang mendekati kondisi nyata pengguna [17].



**Gambar 2.8 Balsamiq Wireframes**

## 1.2 Tinjauan Penelitian

Penelitian ini menggunakan beberapa referensi sebagai dasar pertimbangan dalam penyusunan skripsi, di mana tinjauan tersebut diperoleh dari berbagai karya penelitian sebelumnya yang membahas permasalahan serupa, antara lain:

1. Yusuf Durachman & Abdul Wahab bin Abdul Rahman (2025), *Analyzing Students' Interest in Mathematics Through the Implementation of the K-Means Clustering Algorithm*. Dalam penelitian ini, data nilai matematika dan keaktifan siswa dianalisis menggunakan algoritma *K-Means* untuk mengelompokkan tingkat minat belajar siswa menjadi tiga kategori, yaitu minat tinggi, sedang, dan rendah. Hasil *clustering* membantu guru dalam memahami pola minat siswa dan menentukan metode pembelajaran yang lebih tepat. Penelitian ini relevan sebagai acuan penerapan *K-Means* untuk analisis minat belajar siswa [18].
2. Dede Wintana, Siti Komalasari, & Nurhayati (2022), *Analyzing Students' Interest in Mathematics Through the Implementation of the K-Means Clustering Algorithm*. Dalam penelitian ini, data nilai matematika, hasil kuesioner minat belajar, dan tingkat keaktifan siswa dianalisis menggunakan algoritma *K-Means* untuk mengelompokkan siswa ke dalam tiga kategori minat belajar, yaitu tinggi, sedang, dan rendah. Hasil *clustering* membantu guru dalam memahami perbedaan minat belajar siswa dan menentukan strategi pembelajaran yang lebih tepat. Penelitian ini relevan karena menunjukkan penerapan *K-Means* secara langsung untuk menganalisis minat belajar siswa. [19]

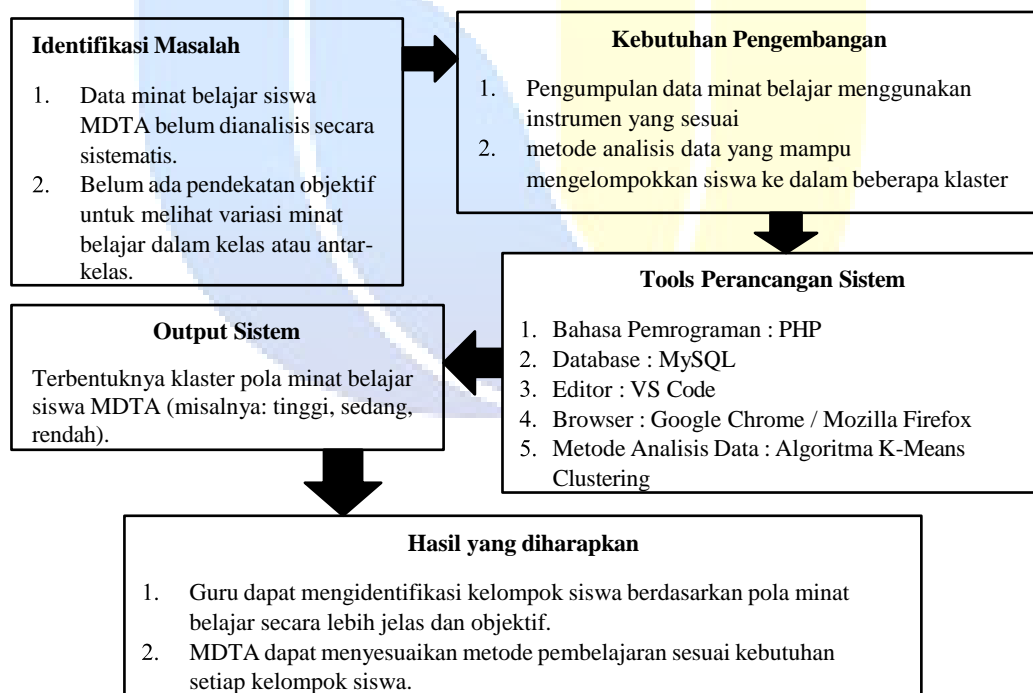
3. Dewi Rengganis, Nur Istiqomah, & Ema Rahmawati (2024), *K-Means Clustering to Improve Interest Grouping Model for High School Students*  
Penelitian ini menggunakan data kuesioner minat akademik siswa untuk membentuk *cluster* minat belajar menggunakan algoritma *K-Means*. Hasil *clustering* membantu sekolah dalam menentukan program pembinaan sesuai kecenderungan minat siswa. Penelitian ini relevan dengan studi ini karena membahas pengelompokan minat belajar melalui pendekatan *data mining*. [20]
4. Elvira D. N. Lestari, Syarifah F. Rezky & Afdal Alhafiz (2023) Penerapan *Data Mining* Untuk Pengelompokan Data Stok Barang Dengan *Metode K-Means Clustering*  
Penelitian ini fokus pada pengelompokan stok barang pada sebuah percetakan/gudang dengan tujuan mengidentifikasi barang “sangat laris”, “cukup laris”, dan “kurang laris”. Hasil *clustering* diharapkan membantu mengoptimalkan stok dan meminimalkan penumpukan barang. PDF artikel tersedia secara *public* [21].
5. Tatang Sujana, Siti Rahmawati, & Andri Hermawan (2023), *Implementation of Data Mining Using the K-Means Algorithm to Group Students Based on Academic Performance*  
Dalam penelitian ini, nilai akademik siswa dari sekolah berbasis keagamaan dianalisis menggunakan *K-Means* untuk mengelompokkan siswa ke dalam tiga kategori performa. Hasil menunjukkan bahwa *K-Means* efektif dalam memetakan performa belajar berdasarkan kesamaan nilai. Penelitian ini sangat relevan karena metode dan objek penelitiannya serupa dengan kondisi penelitian di MDTA. [22]

Dari berbagai penelitian sebelumnya, diketahui bahwa penggunaan sistem informasi pendidikan dan algoritma *data mining* seperti *K-Means* dapat membantu

mengelola data siswa serta menganalisis pola belajar. Namun, sebagian besar penelitian masih berfokus pada sekolah formal dan belum diterapkan secara khusus untuk menilai minat belajar di lembaga non-formal seperti MDTA. Oleh karena itu, penelitian ini menggunakan algoritma *K-Means* untuk mengelompokkan minat belajar siswa MDTA Qur'an Kisaran berdasarkan ketiga data tersebut.

### 1.3 Kerangka Pemikiran

Kerangka pemikiran penelitian ini disusun berdasarkan permasalahan utama yang terjadi pada Madrasah Diniyah Takmiliah Awaliyah (MDTA), yaitu belum tersedianya analisis terstruktur untuk mengetahui pola minat belajar siswa. Ketiadaan analisis tersebut menyebabkan guru kesulitan dalam mengidentifikasi kelompok siswa dengan minat belajar tinggi, sedang, maupun rendah. Akibatnya, strategi pembelajaran tidak dapat disesuaikan berdasarkan kebutuhan dan karakteristik siswa. (Lihat Gambar 2.7).



**Gambar 2.9 Kerangka Pemikiran**

## **1.4 Tinjauan Umum Instansi**

### **1.4.1 Sejarah Singkat Madrasah Diniyah Takmiliyah Awaliyah (MDTA)**

Madrasah Diniyah Takmiliyah Awaliyah (MDTA) Qur'an Kisaran bermula dari cita-cita mulia seorang ibu bernama Hj. Salbiah. Beliau dikenal sebagai sosok dermawan yang memiliki kepedulian tinggi terhadap perkembangan pendidikan agama Islam di lingkungannya. Keinginannya adalah agar anak-anak di Kelurahan Sentang, Kabupaten Asahan, memperoleh pendidikan tambahan berupa pembelajaran Al-Qur'an, ibadah shalat, serta ilmu-ilmu agama Islam lainnya pada waktu sore hari. Harapan tersebut ia sampaikan kepada anak-anaknya sebelum wafat sebagai amanah yang ingin diwujudkan untuk memberikan manfaat kepada masyarakat sekitar.

Setelah beliau meninggal dunia, anak-anaknya berupaya mewujudkan harapan tersebut dengan mendirikan lembaga pendidikan yang tidak hanya berbentuk sekolah arab atau MDTA, tetapi berkembang menjadi Yayasan Pendidikan Qur'an (YPQ). Yayasan ini kemudian menaungi berbagai jenjang pendidikan, mulai dari TK, SD, MTs, MA, hingga Pondok Pesantren, yang seluruhnya berlandaskan nilai-nilai Al-Qur'an.

MDTA Qur'an Kisaran secara resmi berdiri pada 16 Januari 2023. Lembaga ini memperoleh pengakuan dari Kementerian Agama Kabupaten Asahan pada 16 Oktober 2023 dengan Nomor Sertifikat B.3300/Kk.02.06/5PP.00.8/X/2023 dan Nomor Statistik Diniyah Takmiliyah 311212090130. MDTA ini berlokasi di Jl. Jahe Lingkungan IV, Kelurahan Sentang, Kecamatan Kisaran Timur, Kabupaten Asahan, Provinsi Sumatera Utara.

Pada awal berdirinya, MDTA Qur'an Kisaran membuka tiga kelas dengan jumlah 90 santri. Proses pembelajaran didukung oleh tenaga pendidik berpengalaman yaitu Julfirman, S.Pd.I sebagai Kepala MDTA, serta Siti Aisyah Samosir, S.Pd, Erpa Lestari, dan Mica Adella sebagai guru. Seluruh tenaga pendidik berasal dari lulusan sarjana dan pesantren sehingga memiliki kualitas yang memadai dalam memberikan pendidikan keagamaan.

MDTA Qur'an Kisaran hadir sebagai jawaban atas kebutuhan masyarakat yang menginginkan pendidikan agama Islam tambahan bagi anak-anak sekolah dasar. Kurikulum yang digunakan mencakup mata pelajaran seperti Al-Qur'an, Hadits, Aqidah, Akhlak, Fiqih, Tarikh Islam, Bahasa Arab, dan praktik ibadah. Selain pembelajaran, MDTA juga menyediakan kegiatan ekstrakurikuler, termasuk pembinaan olahraga untuk mengikuti kompetisi seperti Pekan Olahraga Santri Diniyah (PORSADIN).



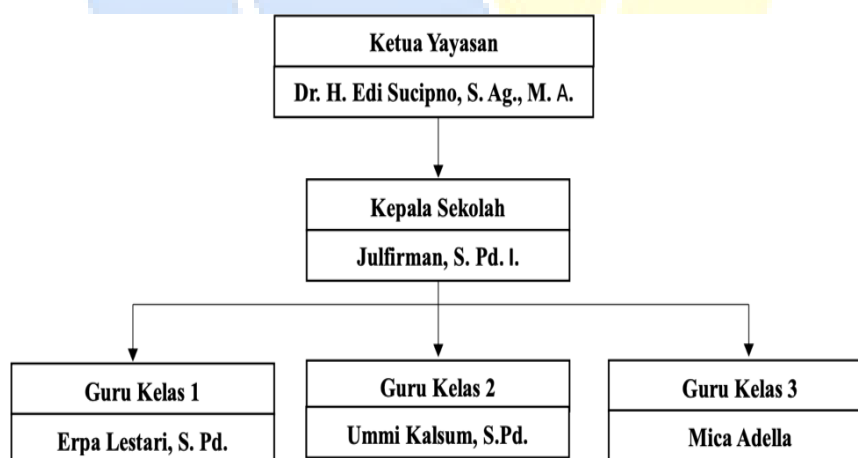
**Gambar 2.10** Lokasi Penelitian

MDTA Qur'an Kisaran berlokasi di Jl. Jahe Lingkungan IV, Kelurahan Sentang, Kecamatan Kisaran Timur, tepat di dalam kawasan utama Yayasan

Pendidikan Qur'an (YPQ) Kisaran. Lokasi ini merupakan pusat kegiatan pendidikan Islam terpadu yang menaungi berbagai jenjang pendidikan, mulai dari TK, SD, MTs, MA, hingga Pondok Pesantren. Dengan berada dalam satu kompleks yang terintegrasi, MDTA Qur'an Kisaran memiliki lingkungan belajar yang kondusif, religius, dan kaya akan aktivitas pendidikan. Keberadaannya di tengah ekosistem pendidikan yang lengkap tersebut menjadikan MDTA tidak hanya sebagai tempat belajar sore hari, tetapi juga sebagai bagian dari pusat pembinaan karakter Qur'ani yang terus berkesinambungan dari tingkat dasar hingga pendidikan menengah dan pesantren.

#### 1.4.2 Struktur Organisasi

Struktur organisasi MDTA (Madrasah Diniyah Takmiliah Awaliyah) merupakan susunan atau pembagian tugas yang menggambarkan hubungan kerja, wewenang, dan tanggung jawab antara setiap unsur yang ada di dalam lembaga MDTA. Struktur ini berfungsi untuk memastikan proses pendidikan berjalan terarah, terkoordinasi, dan sesuai dengan tujuan lembaga



**Gambar 2.11 Struktur Organisasi MDTA Qur'an**

### 1.4.3 Tugas dan Wewenang

#### 1. Ketua Yayasan

Dalam rangka mendukung terselenggaranya pengelolaan lembaga yang terarah dan sesuai dengan visi Yayasan Pendidikan Qur'an Kisaran, diperlukan penegasan mengenai peran serta tanggung jawab setiap unsur yang terlibat. Sehubungan dengan itu, tugas dan wewenang Ketua Yayasan sebagai pimpinan tertinggi dalam struktur kelembagaan dapat dijelaskan sebagai berikut:

- a. Menetapkan arah kebijakan dan visi lembaga.
- b. Mengawasi dan memantau seluruh kegiatan pendidikan MDTA.
- c. Mengambil keputusan strategis terkait pengembangan madrasah.
- d. Mengangkat dan memberhentikan Kepala MDTA.
- e. Menyetujui program kerja, laporan, dan rencana pengembangan madrasah.
- f. Menjamin ketersediaan sarana dan prasarana pendukung pembelajaran.

#### 2. Kepala Sekolah

Kepala Sekolah Bertugas:

- a. Mengelola seluruh kegiatan pendidikan, pembelajaran, dan administrasi madrasah.
- b. Menyusun kurikulum, jadwal, dan kalender pendidikan.
- c. Membina dan mengkoordinasikan seluruh guru serta tenaga pendidik.
- d. Mengambil keputusan operasional untuk kelancaran kegiatan belajar mengajar.
- e. Membagi tugas kepada guru dan staf sesuai kebutuhan madrasah.
- f. Mengawasi perkembangan akademik, kedisiplinan, dan akhlak santri.
- g. Menjalin komunikasi dengan wali santri, masyarakat, dan Yayasan.

- h. Menyusun laporan kegiatan dan hasil evaluasi kepada Ketua Yayasan.
- i. Mengatur suasana belajar yang tertib, aman, dan religius.

### 3. Guru Kelas

Guru Kelas Bertugas:

- a. Melaksanakan pembelajaran sesuai kurikulum Kemenag.
- b. Menilai kemampuan santri meliputi pengetahuan, keterampilan, dan akhlak.
- c. Mendidik dan membimbing santri dalam ibadah, akhlakul karimah, serta kedisiplinan.
- d. Memberikan pembinaan sesuai karakter dan kebutuhan santri.
- e. Menentukan metode pembelajaran yang sesuai dengan kondisi kelas.
- f. Menjaga ketertiban, kenyamanan, dan keamanan kelas.
- g. Berkomunikasi dengan wali santri mengenai perkembangan anak.
- h. Membantu pelaksanaan kegiatan MDTA seperti ujian, dan ekstrakurikuler.

### i. **Visi dan Misi MDTA Qur'an**

#### 1. **Visi MDTA Qur'an**

“Insan Qur’ani dan Madani”

#### 2. **Misi MDTA Qur'an**

- 1. Mewujudkan Peserta Didik Yang Memiliki Kemampuan Membaca, Menghafal, Memahami, Dan Mengamalkan Al- Qur'an.
- 2. Memiliki Dan Menguasai Ilmu Tanzilyah Dan Kaunyah.
- 3. Memiliki Kemampuan Berbahaa, Berpidato Dan Skill Dalam IPTEK.
- 4. Memiliki Kebiasaan Berfikir, Bersikap, Dan Berbuat Sesua Dengan Al- Qur'an.

## b. Perumusan Hipotesis

Hipotesis merupakan jawaban sementara terhadap rumusan masalah yang disusun berdasarkan landasan teori dan hasil penelitian terdahulu. Hipotesis ini masih bersifat dugaan dan akan dibuktikan kebenarannya melalui proses pengolahan dan analisis data dalam penelitian. Berdasarkan kajian pustaka dan penelitian sebelumnya yang membahas penerapan *data mining* dalam bidang pendidikan, khususnya penggunaan algoritma *K-Means* untuk pengelompokan data siswa, diketahui bahwa metode *clustering* mampu mengidentifikasi pola dan karakteristik siswa berdasarkan kesamaan data akademik dan perilaku belajar. Penelitian-penelitian di bidang *Educational Data Mining* menunjukkan bahwa algoritma *K-Means* efektif digunakan untuk mengelompokkan siswa ke dalam beberapa kategori berdasarkan tingkat minat, keterlibatan, dan performa belajar siswa [3]. Oleh karena itu, hipotesis dalam penelitian ini dirumuskan sebagai berikut:

1. Penerapan algoritma *K-Means* mampu mengelompokkan siswa MDTA kedalam kluster minat belajar yang berbeda berdasarkan indikator – indikator minat belajar yang diukur .
2. Hasil *clustering* yang diperoleh mampu menunjukkan pola minat belajar siswa secara jelas dan terstruktur, yaitu terbentuknya kluster minat belajar tinggi, sedang, dan rendah. Pola ini memberikan gambaran objektif mengenai variasi tingkat minat belajar siswa yang sebelumnya sulit diidentifikasi hanya melalui pengamatan subjektif guru.
3. Informasi mengenai kluster minat belajar yang dihasilkan dapat dimanfaatkan oleh guru dan pihak madrasah untuk memahami karakteristik masing-masing

kelompok siswa. Informasi ini mendukung penentuan strategi pembelajaran yang lebih tepat sasaran, pemberian pendampingan yang sesuai, serta membantu proses seleksi siswa dengan minat belajar tinggi untuk mengikuti kegiatan pengembangan prestasi seperti PORSADIN.

