

BAB III

ANALISIS DAN PERANCANGAN

3.1. Metode Penelitian

Metode penelitian yang digunakan dalam penelitian ini adalah penelitian kuantitatif dengan pendekatan deskriptif yang bertujuan untuk mengelompokkan siswa berdasarkan tingkat minat baca menggunakan algoritma K-Means Clustering. Penelitian kuantitatif dipilih karena data yang digunakan bersifat numerik dan dapat diolah secara statistik untuk menghasilkan informasi yang objektif. Data dikumpulkan melalui observasi langsung di perpustakaan SD Negeri 01 Kotapinang, wawancara dengan guru, serta pencatatan aktivitas siswa seperti frekuensi kunjungan ke perpustakaan, jenis buku yang dipinjam, durasi membaca, dan partisipasi dalam kegiatan literasi. Data yang terkumpul kemudian diolah melalui tahapan preprocessing, meliputi pembersihan data (data cleaning), transformasi data, dan normalisasi agar siap dianalisis menggunakan algoritma K-Means.

Tahapan analisis dengan algoritma K-Means dimulai dengan menentukan jumlah cluster yang akan digunakan, dalam hal ini tiga cluster yang merepresentasikan tingkat minat baca tinggi, sedang, dan rendah. Algoritma K-Means kemudian mengelompokkan data siswa berdasarkan kesamaan atribut perilaku membaca yang dimiliki. Hasil clustering divisualisasikan untuk memudahkan interpretasi, sehingga pihak sekolah dapat memahami distribusi minat baca siswa secara lebih jelas. Metode ini dipilih karena kemampuannya dalam mengidentifikasi pola dan kelompok yang tersembunyi di dalam data tanpa

memerlukan label awal. Dengan demikian, hasil penelitian diharapkan dapat menjadi dasar perancangan program literasi yang lebih efektif dan tepat sasaran sesuai dengan karakteristik setiap kelompok siswa.

3.2. Lokasi dan Waktu Penelitian

3.3.1. Lokasi Penelitian

Penelitian ini dilakukan di UPTD SD Negeri 01 Kotapinang, yang berlokasi di Kecamatan Kotapinang, Kabupaten Labuhanbatu Selatan, Provinsi Sumatera Utara. Sekolah ini merupakan salah satu sekolah dasar negeri yang memiliki peran penting dalam memberikan pendidikan dasar bagi anak-anak usia 7–12 tahun di wilayahnya. Sebagai sekolah yang menerapkan Kurikulum Merdeka, SD Negeri 01 Kotapinang berkomitmen untuk mengembangkan kompetensi literasi dan numerasi siswa, sesuai dengan tujuan Profil Pelajar Pancasila. Salah satu fasilitas yang menjadi fokus penelitian ini adalah perpustakaan sekolah, yang memiliki koleksi buku pelajaran, buku cerita anak, hingga referensi umum. Namun, berdasarkan hasil observasi awal, pemanfaatan perpustakaan oleh siswa masih rendah, terlihat dari minimnya frekuensi kunjungan dan kurangnya antusiasme siswa untuk membaca secara mandiri.

Secara fisik, SD Negeri 01 Kotapinang memiliki bangunan yang cukup memadai untuk kegiatan belajar mengajar, termasuk ruang kelas, ruang guru, dan ruang perpustakaan. Akan tetapi, perpustakaan sekolah masih memerlukan penataan yang lebih baik, penambahan fasilitas pendukung seperti sudut baca yang nyaman, serta program literasi yang lebih terstruktur. Faktor-faktor inilah yang mendorong peneliti untuk memilih lokasi ini sebagai objek penelitian, dengan harapan hasil analisis minat baca menggunakan algoritma K-Means Clustering

dapat memberikan rekomendasi konkret untuk mengoptimalkan peran perpustakaan sebagai pusat literasi. Pemilihan lokasi ini juga mempertimbangkan dukungan dari pihak sekolah dan guru yang memiliki semangat untuk mengembangkan budaya membaca di kalangan siswa.

3.3.2. Waktu Penelitian

Penelitian ini dilaksanakan pada bulan **Juli 2025** di UPTD SD Negeri 01 Kotapinang. Pemilihan waktu tersebut didasarkan pada ketersediaan data aktivitas perpustakaan yang cukup lengkap serta kondisi siswa yang sedang aktif mengikuti kegiatan belajar mengajar setelah libur semester. Selama periode penelitian, dilakukan pengumpulan data melalui observasi langsung di perpustakaan, wawancara dengan guru, serta pencatatan perilaku membaca siswa seperti frekuensi kunjungan, jenis buku yang dipinjam, dan durasi membaca. Waktu pelaksanaan yang relatif singkat namun terfokus ini memungkinkan peneliti untuk memperoleh gambaran yang jelas mengenai pola minat baca siswa, sehingga data yang terkumpul dapat dianalisis secara optimal menggunakan algoritma K-Means Clustering.

3.3. Populasi dan Sampel

Populasi dalam penelitian ini adalah seluruh siswa yang terdaftar di UPTD SD Negeri 01 Kotapinang pada tahun ajaran 2024/2025, yang terdiri dari siswa kelas I hingga kelas VI. Untuk kepentingan penelitian, sampel diambil dengan menggunakan teknik purposive sampling, yaitu pemilihan sampel berdasarkan kriteria tertentu yang relevan dengan tujuan penelitian. Kriteria yang digunakan antara lain siswa yang aktif tercatat dalam data kunjungan perpustakaan selama periode pengumpulan data pada bulan Juli 2025. Dengan kriteria tersebut, diperoleh

sejumlah sampel yang dianggap mewakili variasi perilaku membaca di sekolah, baik dari segi frekuensi kunjungan, jenis buku yang dipinjam, maupun durasi membaca. Sampel inilah yang kemudian dianalisis menggunakan algoritma K-Means Clustering untuk mengelompokkan siswa berdasarkan tingkat minat bacanya.

3.4. Alat dan Bahan

Alat dan bahan yang digunakan dalam penelitian ini meliputi perangkat keras dan perangkat lunak untuk mendukung proses pengumpulan serta analisis data. Perangkat keras yang digunakan adalah satu unit laptop yang berfungsi untuk pengolahan data, serta alat tulis untuk pencatatan hasil observasi di lapangan. Perangkat lunak yang digunakan meliputi Microsoft Excel untuk pengolahan awal dan pembersihan data, serta Python dengan pustaka scikit-learn sebagai media implementasi algoritma K-Means Clustering. Selain itu, digunakan formulir observasi dan daftar wawancara sebagai instrumen pengumpulan data perilaku siswa di perpustakaan, seperti frekuensi kunjungan, jenis buku yang dipinjam, dan durasi membaca. Bahan penelitian berupa data primer yang diperoleh langsung dari aktivitas siswa di perpustakaan SD Negeri 01 Kotapinang selama bulan Juli 2025, yang kemudian diolah untuk dianalisis lebih lanjut.

3.5. Teknik Pengumpulan Data

Teknik pengumpulan data dalam penelitian ini dilakukan melalui kombinasi observasi, wawancara, dan dokumentasi untuk memperoleh informasi yang akurat mengenai perilaku membaca siswa di perpustakaan SD Negeri 01 Kotapinang. Observasi dilakukan secara langsung dengan mencatat frekuensi kunjungan siswa, jenis buku yang dipinjam, serta durasi membaca selama berada di perpustakaan.

Wawancara dilakukan dengan guru dan pustakawan untuk memperoleh gambaran mendalam tentang kebiasaan membaca siswa dan program literasi yang telah dijalankan sekolah. Dokumentasi digunakan untuk mengumpulkan data sekunder berupa catatan kunjungan perpustakaan, daftar peminjaman buku, dan arsip kegiatan literasi. Data yang diperoleh dari ketiga teknik ini kemudian diolah menjadi dataset yang siap dianalisis menggunakan algoritma K-Means Clustering.

3.6. Arsitektur Sistem

Arsitektur sistem dalam penelitian ini dirancang untuk memfasilitasi proses pengumpulan, pengolahan, dan analisis data minat baca siswa di perpustakaan SD Negeri 01 Kotapinang menggunakan algoritma K-Means Clustering. Sistem ini terdiri dari tiga lapisan utama, yaitu lapisan input, lapisan pemrosesan, dan lapisan output. Pada lapisan input, data perilaku siswa seperti frekuensi kunjungan, jenis buku yang dipinjam, durasi membaca, dan partisipasi dalam kegiatan literasi dimasukkan ke dalam sistem melalui formulir observasi dan dokumentasi perpustakaan. Lapisan pemrosesan bertanggung jawab untuk melakukan preprocessing data, termasuk pembersihan, transformasi, dan normalisasi, sebelum dilakukan pengelompokan menggunakan algoritma K-Means. Sementara itu, lapisan output menghasilkan hasil clustering berupa pengelompokan siswa ke dalam kategori minat baca tinggi, sedang, dan rendah, yang kemudian divisualisasikan dalam bentuk grafik atau tabel untuk memudahkan interpretasi dan pengambilan keputusan oleh pihak sekolah.

3.7. Desain Aktivitas Sistem

Desain aktivitas sistem pada penelitian ini menggambarkan alur kerja mulai dari pengumpulan data perilaku membaca siswa di perpustakaan hingga proses

pengelompokan menggunakan algoritma K-Means Clustering. Aktivitas ini mencakup tahap input data, preprocessing, pemrosesan clustering, dan penyajian hasil dalam bentuk visual yang dapat digunakan pihak sekolah untuk merancang program peningkatan literasi yang tepat sasaran.

3.7.1. Langkah-langkah Metode Backpropagation

Untuk melakukan pengolahan pada metode K-Means, tahap pertama yaitu menentukan data yang akan digunakan pada penelitian ini.

1. Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan untuk memperoleh informasi mengenai perilaku membaca siswa di perpustakaan SD Negeri 01 Kotapinang yang akan dianalisis menggunakan algoritma K-Means Clustering. Data dikumpulkan melalui observasi langsung terhadap aktivitas siswa di perpustakaan, seperti frekuensi kunjungan, jenis buku yang dipinjam, durasi membaca, serta partisipasi dalam kegiatan literasi. Selain itu, dilakukan wawancara dengan guru dan pustakawan untuk mendapatkan informasi tambahan mengenai kebiasaan membaca siswa dan program literasi yang telah diterapkan sekolah. Dokumentasi data sekunder berupa catatan kunjungan perpustakaan dan daftar peminjaman buku juga digunakan untuk melengkapi dataset penelitian. Seluruh data yang terkumpul kemudian diolah menjadi format terstruktur agar dapat dianalisis secara optimal dengan metode clustering.

Tabel 3. 1. Data Sampel Penelitian

No.	Nama Siswa	Frekuensi Membaca	Durasi Membaca	Jenis Bacaan	Motivasi Membaca	Partisipasi
1	Abdullah Rambe	6	0.5	Buku Pelajaran	Hiburan	Klub Buku
2	Abel Monautu Siregar	4	2.1	Buku Pelajaran	Hiburan	Klub Buku
3	Adelia Marbun	2	2.6	Bacaan Digital	Hiburan	Kunjungan Perpustakaan

4	Aldi Manosor Rajagukguk	4	1.8	Majalah	Hiburan	Kunjungan Perpustakaan
5	Amanda Putra Manurung	1	2.7	Buku Pelajaran	Tugas	Lomba
6	Anggy Sri Mahesti Br Purba	5	2.5	Majalah	Hiburan	Kunjungan Perpustakaan
7	Arifin Sitorus Pane	2	1.5	Majalah	Pengetahuan	Lomba
8	Arini Anastasya Marbun	2	1.9	Buku Cerita	Pengetahuan	Lomba
9	Arsita Ayu Lestari	2	1.0	Bacaan Digital	Hiburan	Klub Buku
10	Arya Maulana	6	2.7	Majalah	Pengetahuan	Lomba
11	Arya Sadewa	6	2.0	Buku Pelajaran	Tugas	Kunjungan Perpustakaan
12	Aulia Rahmadani Harahap	7	1.6	Bacaan Digital	Tugas	Klub Buku
13	Berni Nduru	5	1.8	Bacaan Digital	Menambah Pengetahuan	Lomba
14	Boni Asian Panjaitan	1	2.2	Majalah	Tugas	Kunjungan Perpustakaan
15	Candra Putra Marbun	5	2.1	Majalah	Hiburan	Klub Buku
16	Candy Mayasari Br Silalahi	7	0.6	Buku Pelajaran	Hiburan	Klub Buku
17	Celvin Lusiana Waruwu	1	1.7	Majalah	Tugas	Kunjungan Perpustakaan
18	Cenli Indriani Naibaho	5	2.3	Buku Cerita	Tugas	Lomba
19	Charli Billvans Sihite	6	1.5	Majalah	Hiburan	Kunjungan Perpustakaan
20	Chokky Parlindungan Sitohang	4	1.4	Majalah	Tugas	Kunjungan Perpustakaan

Data yang telah diperoleh dari hasil pengumpulan data ini berjumlah **20 entri** yang merepresentasikan perilaku membaca siswa di SD Negeri 01 Kotapinang. Setiap data mencakup lima variabel utama, yaitu frekuensi membaca (jumlah kunjungan atau aktivitas membaca dalam periode tertentu), durasi membaca (lamanya waktu yang dihabiskan untuk membaca dalam satuan jam), jenis bacaan (kategori bahan bacaan yang dipilih siswa seperti buku pelajaran, majalah, bacaan digital, atau buku cerita), motivasi membaca (alasan utama siswa membaca, seperti hiburan, tugas, pengetahuan, atau menambah pengetahuan), serta partisipasi (keterlibatan siswa dalam kegiatan literasi seperti klub buku, kunjungan perpustakaan, atau lomba). Dataset ini bersifat heterogen karena menunjukkan variasi perilaku dan preferensi membaca antar siswa, yang nantinya akan dianalisis

menggunakan algoritma K-Means Clustering untuk mengidentifikasi kelompok siswa dengan tingkat minat baca yang berbeda.

2. Pembersihan Data

Pembersihan data pada penelitian ini dilakukan untuk mengubah variabel kategorikal menjadi bentuk numerik agar dapat diolah oleh algoritma K-Means Clustering. Proses ini dilakukan dengan memberikan kode numerik pada setiap kategori sesuai dengan tabel partisi yang telah ditetapkan. Untuk variabel Jenis Bacaan, kategori Bacaan Digital diberi nilai 0, Buku Cerita bernilai 1, Buku Pelajaran bernilai 2, dan Majalah bernilai 3. Variabel Motivasi Membaca dikodekan dengan nilai 0 untuk Hiburan, 1 untuk Menambah Pengetahuan, dan 2 untuk Tugas. Sementara itu, variabel Partisipasi diberi nilai 0 untuk Klub Buku, 1 untuk Kunjungan Perpustakaan, dan 2 untuk Lomba. Transformasi ini bertujuan untuk menyederhanakan data sehingga setiap atribut memiliki format numerik yang konsisten, memudahkan proses perhitungan jarak dan pengelompokan pada tahap analisis clustering.

Tabel 3. 2. Tabel Konservi

Variabel	Partisi	Nilai
Jenis Bacaan	Bacaan Digital	0
	Buku Cerita	1
	Buku Pelajaran	2
	Majalah	3
Motivasi Membaca	Hiburan	0
	Menambah Pengetahuan	1
	Tugas	2
Partisipasi	Klub Buku	0
	Kunjungan Perpustakaan	1
	Lomba	2

Tabel 3. 3. Data Hasil Konversi

No.	Nama Siswa	Frekuensi Membaca	Durasi Membaca	Jenis Bacaan	Motivasi Membaca	Partisipasi
1	Abdullah Rambe	6	0.5	2	0	0
2	Abel Monautu Siregar	4	2.1	2	0	0

3	Adelia Marbun	2	2.6	0	0	1
4	Aldi Manosor Rajagukguk	4	1.8	3	0	1
5	Amanda Putra Manurung	1	2.7	2	2	2
6	Anggy Sri Mahesti Br Purba	5	2.5	3	0	1
7	Arifin Sitorus Pane	2	1.5	3	1	2
8	Arini Anastasya Marbun	2	1.9	1	1	2
9	Arsita Ayu Lestari	2	1	0	0	0
10	Arya Maulana	6	2.7	3	1	2
11	Arya Sadewa	6	2	2	2	1
12	Aulia Rahmadani Harahap	7	1.6	0	2	0
13	Berni Nduru	5	1.8	0	1	2
14	Boni Asian Panjaitan	1	2.2	3	2	1
15	Candra Putra Marbun	5	2.1	3	0	0
16	Candy Mayasari Br Silalahi	7	0.6	2	0	0
17	Celvin Lusiana Waruwu	1	1.7	3	2	1
18	Cenli Indriani Naibaho	5	2.3	1	2	2
19	Charli Billvans Sihite	6	1.5	3	0	1
20	Chokky Parlindungan Sitohang	4	1.4	3	2	1

Pada tabel di atas merupakan hasil konversi data perilaku membaca siswa di SD Negeri 01 Kotapinang ke dalam bentuk numerik untuk memudahkan proses analisis menggunakan algoritma K-Means Clustering. Variabel Jenis Bacaan, Motivasi Membaca, dan Partisipasi yang semula berbentuk kategori telah diubah menjadi kode angka sesuai dengan partisi yang telah ditentukan, sedangkan variabel Frekuensi Membaca dan Durasi Membaca tetap dipertahankan dalam bentuk angka asli karena sudah bersifat numerik. Sebagai contoh, kategori “Buku Pelajaran” dikodekan menjadi angka 2, “Hiburan” menjadi 0, dan “Klub Buku” menjadi 0. Proses konversi ini membuat seluruh data bersifat homogen dalam format numerik, sehingga dapat langsung digunakan untuk perhitungan jarak antar data dan proses pengelompokan pada tahap clustering. Dengan bentuk data ini, setiap atribut dapat dianalisis secara kuantitatif untuk mengidentifikasi pola minat baca siswa secara lebih akurat.

3. Perhitungan Data

Perhitungan data pada penelitian ini dilakukan dengan menggunakan algoritma K-Means Clustering untuk mengelompokkan siswa berdasarkan kesamaan karakteristik pada variabel frekuensi membaca, durasi membaca, jenis bacaan, motivasi membaca, dan partisipasi. Setelah data dikonversi seluruhnya ke dalam bentuk numerik, proses perhitungan dimulai dengan menentukan jumlah cluster yang akan dibentuk, diikuti dengan penentuan centroid awal secara acak. Selanjutnya, jarak setiap data siswa terhadap centroid dihitung menggunakan rumus Euclidean Distance, dan data akan dikelompokkan ke dalam cluster dengan jarak terdekat. Proses ini berulang dengan memperbarui posisi centroid hingga posisi centroid stabil atau tidak mengalami perubahan signifikan. Hasil perhitungan ini menghasilkan pengelompokan siswa yang dapat digunakan untuk mengidentifikasi pola minat baca berdasarkan kesamaan perilaku dan motivasi membaca mereka.

$$\text{Centroid 1} = 6, 0.5, 2, 0, 0$$

$$\text{Centroid 2} = 4, 2.1, 2, 0, 0$$

$$\text{Centroid 3} = 2, 2.6, 0, 0, 1$$

$$\begin{aligned} d1, 1 &= \sqrt{(6 - 6)^2 + (0.5 - 0.5)^2 + (2 - 2)^2 + (0 - 0)^2 + (0 - 0)^2} \\ &= 0 \end{aligned}$$

$$\begin{aligned} d1, 2 &= \sqrt{(6 - 4)^2 + (0.5 - 2.1)^2 + (2 - 2)^2 + (0 - 0)^2 + (0 - 0)^2} \\ &= 2.56 \end{aligned}$$

$$\begin{aligned} d1, 3 &= \sqrt{(6 - 2)^2 + (0.5 - 2.6)^2 + (2 - 0)^2 + (0 - 0)^2 + (0 - 1)^2} \\ &= 5.04 \end{aligned}$$

Hasil selanjutnya akan dipaparkan dalam bentuk tabel yaitu sebagai berikut.

Tabel 3. 4. Hasil Clustering

No.	Nama Siswa	X1	X2	X3	Centroid
1	Abdullah Rambe	0	2.56	5.04	1
2	Abel Monautu Siregar	2.56	0.00	3.04	2

3	Adelia Marbun	5.04	3.04	0.00	3
4	Aldi Manosor Rajagukguk	2.77	1.45	3.69	2
5	Amanda Putra Manurung	6.15	4.17	3.16	3
6	Anggy Sri Mahesti Br Purba	2.65	1.78	4.24	2
7	Arifin Sitorus Pane	4.80	3.22	3.49	2
8	Arini Anastasya Marbun	4.89	3.17	1.87	3
9	Arsita Ayu Lestari	4.50	3.03	1.89	3
10	Arya Maulana	3.29	3.22	5.20	2
11	Arya Sadewa	2.69	3.00	4.94	1
12	Aulia Rahmadani Harahap	3.20	4.15	5.57	1
13	Berni Nduru	3.42	3.18	3.41	2
14	Boni Asian Panjaitan	5.82	3.87	3.76	3
15	Candra Putra Marbun	2.14	1.41	4.39	2
16	Candy Mayasari Br Silalahi	1.00	3.35	5.83	1
17	Celvin Lusiana Waruwu	5.70	3.89	3.85	3
18	Cenli Indriani Naibaho	3.64	3.17	3.88	2
19	Charli Billvans Sihite	1.73	2.52	5.12	1
20	Chokky Parlindungan Sitohang	3.29	2.55	4.29	2

Hasil perhitungan pada tabel di atas menunjukkan bahwa setiap data siswa dikelompokkan berdasarkan jarak terdekatnya ke salah satu dari tiga centroid, yaitu Centroid 1 (C1), Centroid 2 (C2), dan Centroid 3 (C3). Nilai pada kolom X1, X2, dan X3 merupakan jarak dari data ke masing-masing centroid. Data siswa akan masuk ke cluster yang memiliki nilai jarak terkecil. Contohnya, pada baris pertama, nilai X1 adalah 0 sehingga paling kecil dibanding X2 dan X3, maka siswa tersebut masuk ke C1.

C1 (Centroid 1) adalah titik pusat cluster pertama yang menjadi acuan bagi siswa-siswa dengan jarak terdekat ke C1. Siswa yang masuk ke C1 berarti karakteristik data mereka paling mirip atau paling dekat dengan nilai rata-rata (centroid) dari cluster pertama. C2 (Centroid 2) adalah titik pusat cluster kedua, di mana siswa-siswa yang masuk kelompok ini memiliki data yang jaraknya paling

dekat ke nilai rata-rata cluster kedua. Sementara C3 (Centroid 3) merupakan titik pusat cluster ketiga, yang menjadi acuan bagi siswa dengan jarak terkecil ke nilai rata-rata cluster ketiga.

Dengan pengelompokan ini, C1, C2, dan C3 tidak hanya menunjukkan urutan kelompok, tetapi juga menggambarkan kemiripan pola data antar anggota di dalamnya. Siswa dalam satu cluster memiliki kesamaan karakteristik membaca, motivasi, atau partisipasi yang serupa dengan centroid yang mewakili cluster tersebut. Perbedaan nilai jarak pada X1, X2, dan X3 membantu menentukan seberapa erat hubungan data siswa terhadap masing-masing centroid.