

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pendidikan menjadi fondasi penting dalam membentuk karakter serta kualitas generasi bangsa. Melalui pendidikan, individu memperoleh kesempatan untuk mengembangkan pengetahuan, sikap, dan keterampilan, terutama pada tingkat sekolah dasar yang menjadi tahap awal pembentukan karakter dan kecerdasan anak[1][2]. Lingkungan sekolah yang nyaman dan didukung oleh fasilitas yang memadai mampu meningkatkan semangat serta motivasi belajar siswa. Sarana seperti ruang kelas, perpustakaan, laboratorium, dan berbagai ruang penunjang lainnya berperan dalam menciptakan suasana belajar yang kondusif. Namun, keterbatasan fasilitas yang masih dialami oleh sebagian sekolah dasar sering kali menjadi hambatan dalam optimalisasi proses pembelajaran [3].

Mengacu pada penjelasan David Eko Wardoyo dari Direktorat Sekolah Dasar Kemendikbudristek dalam tayangan *Advokasi Pengisian Data Sarana dan Prasarana pada Dapodik Sekolah Dasar* tertanggal 27 Maret 2025 yang diakses pada Maret 2025 [4], terungkap bahwa sekitar 120.600 dari total 147.000 sekolah dasar di Indonesia masih belum didukung infrastruktur dan fasilitas yang sesuai dengan ketentuan minimal penyelenggaraan pendidikan dasar. Kesenjangan tersebut turut terlihat pada tingkat daerah, terutama di Kabupaten Bangkalan.

Berdasarkan data Pusat Data dan Teknologi Informasi Kementerian Pendidikan Dasar dan Menengah (Kemendikdasmen) tahun 2024, dari 4.370 ruang kelas sekolah dasar, hanya 809 ruang berada dalam kondisi baik. Selebihnya, 1.317 ruang mengalami kerusakan ringan, 1.061 ruang rusak sedang, dan 762 ruang rusak berat [5]. Kondisi ini diperkuat oleh pernyataan Ketua Komisi D DPRD Bangkalan, Nur Hasan, yang mengungkapkan bahwa sekitar 536 lembaga sekolah dasar di Kabupaten Bangkalan mengalami kerusakan bangunan pada tingkat sedang hingga berat. Ia mendorong pemerintah daerah untuk memberikan perhatian lebih serius terhadap sarana dan prasarana pendidikan melalui alokasi anggaran rehabilitasi yang lebih besar demi mendukung peningkatan mutu pendidikan dasar di Bangkalan [6].

Melihat situasi tersebut, dibutuhkan cara analisis yang tidak hanya terstruktur, tetapi juga mampu menyingkap pola ketimpangan fasilitas sekolah secara lebih tajam. Pemerintah dan pihak terkait memerlukan metode yang sanggup membaca data dalam jumlah besar secara otomatis agar gambaran kondisi sekolah dapat dipahami secara menyeluruh. Dalam hal ini, *data mining* hadir sebagai strategi yang tepat karena berfungsi menggali temuan penting dari kumpulan data besar melalui metode tertentu.

Salah satu teknik di dalamnya adalah *clustering*, yakni metode pengelompokan data berdasarkan kemiripan atribut [7]. Teknik ini cocok digunakan untuk memetakan fasilitas sekolah karena dapat menghimpun sekolah-sekolah dengan karakteristik fasilitas yang sejenis tanpa perlu menetapkan kategori sejak awal. Hasil pengelompokan tersebut membuat perbedaan tingkat kelengkapan fasilitas antarsekolah tampak lebih nyata. Dengan demikian, pihak terkait dapat menyusun langkah peningkatan infrastruktur dan fasilitas pendidikan secara lebih terarah, objektif, dan berbasis bukti.

K-Means, *K-Medoids*, dan *Fuzzy K-Medoids Type-2* dapat dipandang sebagai tiga pendekatan *clustering* yang menawarkan sudut pandang berbeda dalam memetakan distribusi fasilitas sekolah dasar secara lebih presisi. *K-Means* unggul dari sisi kesederhanaan dan efisiensi komputasi, tetapi sensitif terhadap penentuan centroid awal yang berpotensi memengaruhi proses konvergensi maupun hasil akhirnya [8]. Sebaliknya, *K-Medoids* membutuhkan beban komputasi lebih tinggi karena pusat *cluster* dihitung berdasarkan frekuensi data teramati.

Walaupun demikian, metode ini memiliki kelebihan dalam menghadapi outlier karena berpusat pada medoid [9]. Sementara itu, *Fuzzy K-Medoids Type-2* hadir sebagai pengembangan yang menggabungkan kekuatan *K-Medoids* dengan keluwesan logika *Fuzzy Type-2* untuk mengakomodasi data yang tidak pasti dan ambigu. Tidak seperti *K-Medoids* tradisional yang hanya menggunakan titik data sebagai pusat *cluster*, metode ini turut mempertimbangkan derajat keanggotaan *fuzzy* tingkat dua, sehingga dapat menangkap bentuk ketidakpastian yang lebih rumit [10]. Oleh karena itu, penelitian ini menggunakan ketiga algoritma tersebut secara bersamaan untuk menilai performanya pada *dataset* fasilitas pendidikan serta menentukan metode yang paling unggul dalam menangani ketimpangan data

sekolah dasar di Kabupaten Bangkalan.

Penelitian ini mengadopsi kerangka kerja *CRISP-DM* (*Cross Industry Standard Process for Data Mining*) yang mencakup enam tahapan utama: *business understanding*, *data understanding*, *data preparation*, *modelling*, *evaluation*, dan *deployment* [11]. Tahap awal menitikberatkan pada pemahaman persoalan pemerataan fasilitas sekolah dasar di Bangkalan. Selanjutnya, data yang bersumber dari DAPODIK dipelajari dan dipersiapkan melalui proses pengumpulan serta pembersihan agar siap diolah. Pada fase pemodelan, tiga algoritma *clustering* yang telah terbukti efektivitasnya digunakan untuk mengelompokkan data, kemudian performanya dianalisis menggunakan indeks evaluasi yang sesuai.

Sejumlah penelitian terdahulu telah mengkaji perbandingan metode *clustering* untuk pengelompokan fasilitas pendidikan. Salah satunya penelitian oleh Denny Nurdiansyah, Saniyatus Saidah, dan Nita Cahyani yang membandingkan *K-Means*, *K-Medoids*, dan *Random Clustering* dalam mengelompokkan sekolah dasar di Kabupaten Bojonegoro berdasarkan kapasitas dan fasilitas yang tersedia. Hasil penelitian menunjukkan bahwa *K-Means* memberikan performa terbaik dengan lima kelompok kategori, yaitu sangat lengkap (14 sekolah), lengkap (236 sekolah), cukup lengkap (176 sekolah), kurang lengkap (310 sekolah), dan tidak lengkap (177 sekolah). Hasil tersebut dimanfaatkan sebagai dasar rekomendasi bagi Dinas Pendidikan Kabupaten Bojonegoro dalam upaya pemerataan fasilitas sekolah [12].

Studi Akhiril Anwar Harahap, Muhammad Raihan, Nailul Amani, dan Putri Risma Andini membandingkan tiga pendekatan klasterisasi, yaitu *K-Means*, *K-Medoids*, dan *Fuzzy C-Means*, dalam memetakan desa di Indonesia berdasarkan fasilitas sekolah per provinsi. Penelitian ini memanfaatkan data BPS periode 2011–2018 dengan *Davies Bouldin Index (DBI)* sebagai alat evaluasi. Dari hasil pengujian, *K-Means* tampil sebagai metode paling unggul dengan *DBI* 0,451, kemudian *Fuzzy C-Means* 0,491, dan *K-Medoids* 0,638. Berdasarkan hasil tersebut, penelitian ini mendorong pemerintah untuk lebih fokus memperbaiki fasilitas pendidikan di provinsi dengan kondisi terendah, seperti Papua Barat, Maluku, Maluku Utara, dan Bengkulu [13].

Hasil ini sekaligus membuka ruang baru bahwa Fuzzy K-Medoids *Type-2* merupakan pendekatan yang masih jarang dipakai dalam pemetaan fasilitas

pendidikan, sehingga memberi nilai inovatif pada studi ini. Untuk memastikan kualitas pengelompokan yang dihasilkan, diperlukan evaluasi lebih lanjut, termasuk dalam menetapkan jumlah *cluster* yang ideal sesuai karakteristik sekolah dasar. Penelitian ini menggunakan *Silhouette Coefficient* sebagai tolok ukur penentuan jumlah *cluster* dan kualitas klaster. Nilainya berada pada kisaran -1 sampai 1; semakin tinggi mendekati 1, semakin kuat kedekatan data dalam klaster, sedangkan nilai yang mendekati -1 menunjukkan pengelompokan yang lemah [14].

Penelitian ini memadukan algoritma clustering dengan evaluasi *Silhouette Coefficient* sebagai strategi analisis yang utuh dan terarah. Karena tiap algoritma memiliki pola kerja yang berbeda dalam membaca data, maka perbandingan menjadi langkah penting untuk menemukan metode yang paling tepat dalam mengelompokkan fasilitas pendidikan. Walaupun pendekatan *clustering* telah banyak dimanfaatkan dalam studi pendidikan, kebanyakan penelitian terdahulu masih berfokus pada satu atau dua algoritma saja dan belum menyertakan *Fuzzy K-Medoids Type-2* sebagai alternatif pengelompokan. Selain itu, pembahasan mengenai klasterisasi fasilitas sekolah dasar di Kabupaten Bangkalan masih belum banyak dijumpai. Atas dasar tersebut, penelitian ini menempatkan *K-Means*, *K-Medoids*, dan *Fuzzy K-Medoids Type-2* dalam satu kerangka komparatif untuk menilai performa masing-masing, sementara mutu hasil pengelompokan dianalisis melalui *Silhouette Coefficient (SC)*.

Berangkat dari persoalan yang ada, penelitian ini menelaah kinerja *K-Means*, *K-Medoids*, dan *Fuzzy K-Medoids Type-2* untuk memetakan pola sebaran fasilitas sekolah dasar. Pengelompokan tersebut diharapkan mampu menghadirkan gambaran yang lebih terarah bagi Dinas Pendidikan Kabupaten Bangkalan dalam menentukan langkah pemerataan fasilitas, merumuskan kebijakan, dan mengembangkan program pendidikan berdasarkan kebutuhan masing-masing *cluster* sekolah.

1.2 Perumusan Masalah

Perumusan masalah berisi masalah, metode usulan, dan pertanyaan penelitian sebagai berikut:

1.2.1 Permasalahan

Inti persoalan penelitian ini berada pada wajah fasilitas pendidikan sekolah

dasar di Kabupaten Bangkalan yang masih timpang. Ada sekolah yang telah ditunjang fasilitas lengkap, tetapi ada pula yang masih bergulat dengan keterbatasan sarana dan prasarana. Kondisi ini dapat memperlebar jurang mutu pendidikan antarwilayah dan memperlambat terciptanya pemerataan pendidikan dasar.

Sebenarnya, data penting seperti status sekolah, jumlah siswa, guru, ruang kelas, perpustakaan, ruang ibadah, UKS, toilet, hingga akreditasi sudah tersedia di DAPODIK dan Dinas Pendidikan Bangkalan. Namun, data tersebut belum disusun dan dianalisis secara terarah untuk membentuk peta kelompok sekolah berdasarkan kelengkapan fasilitas. Dampaknya, perencanaan kebijakan pemerataan, penguatan infrastruktur pendidikan, dan penentuan prioritas pembangunan menjadi kurang optimal. Karena itu, dibutuhkan pendekatan *data mining* dengan metode *clustering* agar sekolah dapat dikelompokkan sesuai kondisi fasilitasnya, sehingga keputusan yang diambil menjadi lebih akurat dan efektif.

1.2.2 Metode Usulan

Dalam studi ini, pengelompokan sekolah dasar di Kabupaten Bangkalan berdasarkan kelengkapan fasilitas pendidikan dilakukan melalui tiga algoritma, yaitu *K-Means*, *K-Medoids*, dan *Fuzzy K-Medoids Type-2*. Proses penelitian diawali dengan pendekatan *CRISP-DM* sebagai pedoman sistematis dalam pelaksanaan *data mining*. Pada tahap awal, data dipahami melalui proses penelaahan untuk memastikan keterkaitannya dengan fokus masalah. Data yang telah diperoleh kemudian melalui tahap pembersihan dan pengolahan agar layak digunakan pada analisis. Sesudahnya, ketiga metode klasterisasi diterapkan untuk membentuk kelompok sekolah, sementara kualitas hasil pengelompokan diuji menggunakan *Silhouette Coefficient*.

1.2.3 Pertanyaan Penelitian

Dari rangkaian penjelasan dalam latar belakang, lahirlah beberapa pernyataan penelitian yang dirumuskan sebagai berikut:

1. Berapa *cluster* optimal untuk mengelompokkan sekolah dasar di Kabupaten Bangkalan berdasarkan kelengkapan fasilitas pendidikan dengan *K-Means*, *K-Medoids*, dan *Fuzzy K-Medoids Type-2*?
2. Bagaimana hasil *cluster K-Means*, *K-Medoids*, dan *Fuzzy K-Medoids Type-2* berdasarkan *Silhouette Coefficient*?

1.3 Tujuan dan Manfaat

Bagian tujuan dan manfaat memuat sasaran penelitian serta kontribusi yang diharapkan sebagai berikut:

1.3.1 Tujuan Penelitian

Berikut tujuan penelitian yang mengacu pada rumusan masalah di atas:

1. Menentukan klaster optimal sekolah dasar di Kabupaten Bangkalan berdasarkan kelengkapan fasilitas pendidikan dengan metode *K-Means*, *K-Medoids*, dan *Fuzzy K-Medoids Type-2*.
2. Membandingkan hasil klasterisasi *K-Means*, *K-Medoids*, dan *Fuzzy K-Medoids Type-2* berdasarkan nilai Silhouette untuk menentukan metode terbaik.

1.3.2 Manfaat Penelitian

Sesuai rumusan masalah, manfaat penelitian ini adalah sebagai berikut:

1. Manfaat Akademis:

Melalui penerapan dan telaah perbandingan Implementasi algoritma *K-Means*, *K-Medoids*, serta *Fuzzy K-Medoids Tipe-2* di ranah pendidikan, penelitian ini menghadirkan sumbangsih nyata bagi kemajuan teknologi penggalian data. Lebih jauh, hasilnya dapat menjadi pijakan penting bagi penelitian lanjutan yang berfokus pada analisis data pendidikan, terutama dalam mendukung distribusi fasilitas sekolah dasar yang lebih merata dan berkeadilan.

2. Manfaat Praktis:

Studi ini memotret secara lebih tegas kondisi kelengkapan fasilitas sekolah dasar di Kabupaten Bangkalan sebagai pijakan bagi Dinas Pendidikan untuk menata kebijakan pemerataan sarana. Melalui hasil klasterisasi tersebut, diharapkan lahir rancangan pembangunan pendidikan yang lebih presisi, visioner, dan sesuai dengan karakter kebutuhan tiap kelompok sekolah.

1.4 Batasan Masalah

Berikut batasan masalah agar pembahasan tetap fokus dan sesuai dengan tujuan penulis:

1. Penelitian ini berfokus pada sekolah dasar di Kabupaten Bangkalan.
2. Data penelitian menggunakan data sekunder dari DAPODIK dan Dinas

Pendidikan Bangkalan tahun 2025.

3. Variabel penelitian dibatasi pada nama sekolah, status sekolah, peserta didik, guru, ruang kelas, perpustakaan, ruang ibadah, UKS, toilet, dan akreditasi.

1.5 Sistematika Penulisan

Sistematika penulisan proposal ini penulis susun dengan mengikuti aturan penulisan proposal pada program studi S1 Sistem Informasi, Universitas Trunodjoyo Madura. Berikut merupakan metodologi penulisan proposal pada penelitian ini:

BAB I PENDAHULUAN

Menyajikan gambaran menyeluruh tentang alasan lahirnya penelitian, perumusan persoalan utama, hambatan yang ditemui, metode yang dirancang, pertanyaan penelitian, serta uraian tujuan, nilai guna, ruang lingkup, dan alur penulisan karya ilmiah.

BAB II TINJAUAN PUSTAKA

Memuat teori pendukung penelitian, algoritma *K-Means*, *K-Medoids*, *Fuzzy K-Medoids Type-2*, serta penelitian terdahulu sebagai acuan.

BAB III METODOLOGI PENELITIAN

Menjelaskan metodologi penelitian, meliputi tahapan, pemahaman dan persiapan data, penerapan tiga metode, serta pengujian dan evaluasi.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi hasil penelitian, analisis data, evaluasi *Silhouette Coefficient*, dan pembahasan temuan.

BAB V KESIMPULAN DAN SARAN

Bab ini menyajikan inti temuan penelitian sekaligus rekomendasi yang lahir dari rangkaian analisis yang telah dilaksanakan.

DAFTAR PUSTAKA

Menjelaskan sumber rujukan Tugas Akhir. Pada penelitian ini, daftar pustaka dibuat otomatis menggunakan Mendeley.

DAFTAR LAMPIRAN

Bagian ini mencakup lampiran yang berisi data dan informasi pendukung yang relevan dengan penelitian.