

BAB I

PENDAHULUAN

1.1 Latar Belakang

Terjadinya Revolusi Industri 4.0 memberikan dampak yang sangat besar terhadap dunia pendidikan di Indonesia, diawali dengan munculnya digitalisasi sistem pendidikan yang mengharuskan seluruh bidang pendidikan harus beradaptasi dengan perubahan yang terjadi. Revolusi Industri 4.0 memberikan dampak positif bagi dunia pendidikan, dimana sistem pembelajaran yang semakin maju dan berkembang memberikan dampak yang sangat besar bagi sekolah. Jarak dan batas wilayah bukanlah hambatan untuk memahami dan berinteraksi dengan dunia luar. Sejak munculnya Revolusi Industri 4.0 dalam dunia pendidikan telah membawa dampak positif bagi kemajuan dan perkembangan sistem pembelajaran, namun jika disalahgunakan dapat berdampak buruk bagi dunia pendidikan (Natalia Shefia Sembiring et al., 2022).

Saat ini, sektor pendidikan perlu berdaya saing tinggi dengan memanfaatkan kualitas dan kuantitas sumber daya manusia (SDM) yang ada, seperti siswa yang berprestasi. Namun pada kenyataannya, kemampuan menguasai dan memahami orang berbeda-beda. Hal ini dipicu oleh perbedaan pengalaman dan pengetahuan. (Ismaya & Zyen, 2022). Sekolah dan berbagai pihak di bidang pendidikan telah melakukan banyak upaya untuk meningkatkan prestasi akademik siswa, sehingga membantu mencapai standar pendidikan nasional. Tidak hanya itu, sistem informasi kini menjadi sumber daya yang dapat digunakan untuk meningkatkan kualitas dan prestasi siswa. Teknologi data mining merupakan teknologi pengolahan data yang dapat menemukan dan mengidentifikasi pola data yang tersembunyi. Hasil pengolahan data berguna untuk pengambilan keputusan di masa depan (Dacwanda & Nataliani, 2021).

Berdasarkan hasil wawancara dengan salah satu guru SMP Kedungadem mengungkapkan bahwa pendidik masih menggunakan perhitungan manual dalam menilai dan mengidentifikasi siswa berprestasi berdasarkan nilai karakter akademik. Hal ini dapat menyebabkan hasil yang tidak akurat, dan proses seleksi dapat memakan waktu lama, kemungkinan 1-2 minggu atau lebih, sehingga

menyulitkan pendidik untuk melakukan analisis lebih lanjut terhadap status siswa. Selain itu, penilaian yang dilakukan lebih fokus pada bidang akademik. Penilaian pengembangan diri dilakukan berdasarkan standar penilaian minimal dan hanya didasarkan pada penilaian subyektif guru yang bersangkutan. Sistem evaluasi seperti ini tentu saja kurang memadai karena tidak sejalan dengan visi dan misi SMP Kedungadem yang berharap dapat menghasilkan siswa yang tidak hanya unggul secara akademis namun juga berakhlak mulia.

Dengan terus meningkatnya volume data siswa SMP Kedungadem setiap tahunnya mengakibatkan penumpukan data yang tidak terolah secara maksimal. Volume data yang terus meningkat memerlukan beragam teknik atau metode untuk mengolahnya menjadi informasi yang dapat menjadi bahan pertimbangan para pendidik dalam proses pengambilan kebijakan, serta untuk memberikan peringatan dini (*early warning*) terhadap siswa tertentu dari hasil *clustering* prestasi rendah yang dapat mengakibatkan siswa -siswi tidak lulus atau gagal (Desi et al., 2022).

Data mining dapat diterapkan untuk mengidentifikasi siswa berprestasi. Data mining juga sering digunakan dalam pengelolaan data untuk menghasilkan pengetahuan yang tidak diketahui secara manual (Sarah et al., 2022). Data mining adalah metode pemrosesan data yang umum digunakan. Ada banyak metode dalam data mining, salah satunya adalah pengelompokan data yang telah dilakukan pada penelitian-penelitian sebelumnya, yaitu membandingkan algoritma *K-means* dengan *K-medoids* untuk mengelompokkan “Penyakit Pasien Di Puskesmas Cigugur Tengah”. Tujuannya adalah mengklasifikasikan penyakit mana yang paling banyak diderita pasien dan membantu pemerintah, khususnya Departemen Kesehatan dengan memberikan edukasi kesehatan kepada masyarakat. Pada penelitian ini dilakukan perbandingan penerapan algoritma *K-Means* dan *K-Medoids*, serta dilakukan pengujian validitas terhadap hasil cluster yang terbentuk. Nilai efektivitas Davies Bouldin Index sebagai metode analisis cluster sebesar -0,453 dan algoritma *k-medoids* sebesar -1,276. Dari hasil tersebut dapat disimpulkan bahwa algoritma *K-means* dianggap sebagai algoritma yang lebih baik. Dan algoritma *K-means* lebih unggul dari algoritma *K-medoids* yang menghasilkan nilai rata – rata Davies Bouldin Index sebesar -1,276 (Sugianto et al., 2020).

Pada penelitian sebelumnya, peneliti juga berhasil melakukan teknik *clustering* menggunakan algoritma K-Means untuk “Pengelompokan Topik Skripsi Mahasiswa”. Hal ini dapat menunjukkan kemampuan mahasiswa terhadap setiap kelompok bidang keahlian dan kemampuan. Hasil pengelompokan dianalisis berdasarkan jumlah matakuliah pada masing-masing kelompok bidang keahlian. Jumlah maksimum matakuliah dalam satu kelompok keahlian dan kemampuan menunjukkan bahwa mahasiswa memiliki kemampuan yang lebih baik dalam bidang tersebut. Oleh karena itu, topik skripsi yang sesuai untuk kelompok bidang keahlian tersebut direkomendasikan. Dengan memilih topik skripsi yang sesuai dengan bidangnya, diharapkan mahasiswa dapat menyelesaikan skripsinya dengan lebih baik dan tepat waktu (Rafi & Defriani, 2020).

Untuk mengatasi permasalahan tersebut dengan cepat, tepat dan akurat diperlukan suatu metode data mining *clustering* yaitu K-means. *Clustering* adalah metode untuk mencari dan mengelompokkan data dengan karakteristik serupa di antara setiap kumpulan data. Ada beberapa algoritma *clustering*, salah satunya adalah K-means. Algoritma K-Means adalah algoritma *clustering* yang sederhana dengan membagi database menjadi beberapa *cluster* (Ismaya & Zyen, 2022). Berdasarkan latar belakang tersebut, penelitian ini menggunakan metode k-means untuk mengelompokkan nilai-nilai karakter akademik siswa.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka permasalahan yang dapat dirumuskan sebagai berikut.

1. Bagaimana cara menerapkan algoritma K-Means clustering dalam mengelompokkan nilai-nilai karakter akademik siswa?
2. Bagaimana hasil evaluasi perhitungan jarak menggunakan Davies Bouldin Index?

1.3 Batasan Masalah

Batasan masalah dalam penelitian adalah sebagai berikut:

1. Algoritma *K-Means clustering* yang digunakan dalam penelitian ini tidak boleh dibandingkan dengan algoritma lainnya.

2. Data diperoleh dari hasil kuesioner yang disebar di sekolah.
3. Data yang diklasterisasi hanya berupa angka, tidak menguji data berupa gambar.
4. Informasi yang dihasilkan berupa kelompok siswa yang memiliki nilai karakter akademik baik dan kurang baik.

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut.

1. Menerapkan algoritma K-Means clustering dalam mengelompokkan nilai-nilai karakter akademik siswa untuk mengetahui kelompok siswa yang mempunyai nilai karakter akademik baik dan kurang baik.
2. Mengevaluasi hasil perhitungan dengan Davies Bouldin Index.

1.5 Manfaat Penelitian

Adapun manfaat yang diharapkan dari penelitian ini adalah:

1. Menambah wawasan dan pengetahuan untuk penerapan algoritma *K-Means* pada data siswa yang berprestasi.
2. Menemukan literatur yang relevan dan memahami proses penggunaan algoritma k-means untuk menentukan siswa berprestasi berdasarkan karakter akademik.
3. Bagi institusi terkait, hal ini membantu institusi pendidikan mengidentifikasi siswa dengan prestasi rendah yang kemungkinan besar tidak lulus.
4. Bagi peneliti, dapat memberikan wawasan untuk menentukan siswa berprestasi berdasarkan karakter akademik dengan algoritma *K-Means*.