

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Pendidikan adalah salah satu elemen kunci dalam pengembangan sumber daya manusia yang berkualitas dan kompetitif. Melalui pendidikan, individu dapat mengeksplorasi potensi diri, meningkatkan kualitas hidup, dan memberikan kontribusi terhadap kemajuan bangsa. Dengan alasan tersebut, pemerintah Indonesia bertekad untuk memastikan akses pendidikan bagi semua segmen masyarakat, terutama bagi anak-anak dari keluarga yang kurang mampu. Salah satu manifestasi nyata dari tekad tersebut adalah Program Indonesia Pintar (PIP) yang bertujuan mendukung pelajar dari keluarga miskin, rentan, atau kelompok prioritas lainnya agar tetap melanjutkan pendidikan dan mendapatkan layanan pendidikan yang layak. Adanya pendidikan yang merata dan memadai di semua kalangan masyarakat akan berimbas terhadap kemajuan sebuah Negara [1].

Perkembangan teknologi telah mendorong kemajuan dalam data mining, yang merupakan salah satu bagian dari penemuan pengetahuan basis data (KDD) dan bertujuan untuk mengekstrak pola, keterkaitan, serta wawasan baru dari kumpulan data yang besar. Metode ini memiliki nilai tinggi karena dapat mendukung proses pengambilan keputusan dengan lebih efisien dan berdasarkan bukti data [2]. Di antara sekian banyak teknik pengelompokan, algoritma *K-Means* muncul sebagai salah satu yang paling diminati karena kemudahannya, kecepatannya serta efektivitas dalam mengelompokkan data dalam skala yang besar. Metode ini beroperasi dengan menetapkan sejumlah pusat kelompok

(*centroid*), lalu menghitung jarak antara data dan *centroid* dengan menggunakan ukuran jarak tertentu. Salah satu ukuran jarak yang sering diaplikasikan adalah *Euclidean Distance*, sebab dapat memberikan perhitungan jarak geometris yang tepat antara data sehingga kelompok yang terbentuk menjadi lebih representatif [3].

SDN 057/IX Tangkit merupakan salah satu sekolah dasar negeri yang berada di wilayah Kecamatan Sungai Gelam, Kabupaten Muaro Jambi, Provinsi Jambi. Sekolah ini berperan penting dalam memberikan layanan pendidikan dasar bagi anak-anak di sekitar Desa Tangkit dan sekitarnya. Sebagai sekolah yang berada di wilayah pedesaan, SDN 057/IX Tangkit memiliki peserta didik dengan latar belakang sosial dan ekonomi yang beragam, sebagian besar berasal dari keluarga dengan tingkat ekonomi menengah ke bawah. Kondisi tersebut menjadikan sekolah ini sebagai salah satu lembaga pendidikan yang aktif mengajukan usulan penerima Program Indonesia Pintar (PIP) agar siswa-siswa dari keluarga kurang mampu tetap dapat melanjutkan pendidikan tanpa terkendala biaya.

Namun, dalam pelaksanaannya, proses penentuan calon penerima Program Indonesia Pintar (PIP) masih dihadapkan pada berbagai tantangan, terutama dalam pengelolaan dan analisis data usulan yang sangat besar volumenya. Proses seleksi yang dilakukan berdasarkan pengamatan subjektif seperti mencatat formulir satu persatu bukan berdasarkan sistem otomatis. Hal ini kerap kali menghasilkan ketidakakuratan dalam penentuan sasaran, disebabkan oleh tidak adanya sistem

yang dapat mengelompokkan data usulan dengan cara yang objektif dan efisien berdasarkan kriteria sosial-ekonomi yang relevan.

Salah satu metode yang dapat digunakan untuk mengatasi permasalahan tersebut adalah algoritma *K-Means Clustering* dengan jarak *Euclidean Distance* sebagai ukuran kemiripan data. Algoritma ini berfungsi untuk mengelompokkan data berdasarkan kesamaan karakteristik antar objek, sehingga data yang mirip akan berada dalam satu kelompok (*cluster*). Dengan menerapkan metode ini pada data usulan penerima PIP, sekolah dapat dengan mudah mengelompokkan siswa berdasarkan kondisi sosial-ekonominya, seperti pendapatan orang tua, pekerjaan, dan jumlah tanggungan. Hasil pengelompokan ini membantu pihak sekolah dalam menilai siswa yang tidak layak, cukup layak dan layak menerima bantuan secara objektif dan efisien.

Untuk menganalisis permasalahan yang terjadi telah dilakukan penelitian terdahulu oleh Rizki Dwi Romadhona dan Arta Ainur Rofiq [4] dengan judul “Algoritma *K-Means Clustering* Sebagai Penentuan Beasiswa PIP pada SMPN 9 Kota Blitar” menunjukkan bahwa penerapan algoritma *K-Means Clustering* dengan jarak *Euclidean* dapat membantu mengelompokkan data calon penerima Program Indonesia Pintar (PIP) secara lebih objektif dan efisien berdasarkan indikator sosial ekonomi seperti pendapatan orang tua, pekerjaan, jumlah tanggungan, serta kepemilikan bantuan sosial. Hasil penelitian tersebut mampu mengelompokkan siswa ke dalam dua klaster, yaitu kelompok layak dan tidak layak menerima bantuan, sehingga proses seleksi menjadi lebih tepat sasaran. Namun, penelitian tersebut masih berfokus pada jenjang sekolah menengah

pertama, sedangkan penerapan metode serupa pada jenjang sekolah dasar belum banyak dikaji.

Kemudian menurut Yana Mulyana, Abdul Fadlil, dan Imam Riadi [5] menyoroti permasalahan penentuan penerima KIP-K di Universitas Muhammadiyah Tasikmalaya yang masih manual sehingga kurang tepat sasaran. Peneliti menggunakan algoritma K-Means dengan Euclidean Distance di RapidMiner untuk mengelompokkan 191 pendaftar berdasarkan atribut sosial ekonomi dan prestasi, menghasilkan tiga kluster: layak, dipertimbangkan, dan tidak layak. Namun, penerapannya masih terbatas pada tingkat perguruan tinggi dan belum dicoba di pendidikan dasar dengan karakteristik sosial ekonomi berbeda.

Selanjutnya penelitian yang dilakukan oleh Rini Gustini dan RZ. Abdul Aziz [6] berfokus pada permasalahan ketidaktepatan sasaran penerima Kartu Indonesia Pintar (KIP) karena proses penentuan penerima masih dilakukan secara manual melalui aplikasi DAPODIK tanpa mempertimbangkan faktor-faktor kelayakan secara komprehensif. Untuk menjawab masalah tersebut, penelitian ini mengembangkan model sistem pendukung keputusan (SPK) menggunakan metode *K-Means* dan *Average Linkage Clustering* guna mengelompokkan siswa berdasarkan kriteria ekonomi dan sosial secara objektif. Namun, penelitian ini masih terbatas pada penerapan dua metode clustering dan belum menggabungkan pendekatan lain yang lebih adaptif terhadap data dinamis, sehingga menimbulkan kebutuhan pengembangan model pengambilan keputusan yang lebih fleksibel,

terintegrasi, dan mampu meningkatkan akurasi dalam penentuan kelayakan penerima bantuan pendidikan.

Dari uraian di atas, dapat disimpulkan bahwa diperlukan suatu metode analisis data yang mampu mengelompokkan data usulan penerima bantuan Program Indonesia Pintar (PIP) berdasarkan kriteria kelayakan siswa, agar pihak sekolah dapat mengambil keputusan yang tepat dan objektif dalam penentuan penerimaan bantuan. Oleh karena itu, penulis mengangkat judul “KLAUSTERISASI DATA USULAN PENERIMA BANTUAN PROGRAM INDONESIA PINTAR MENGGUNAKAN ALGORITMA *K-MEANS* DI SDN 057/IX TANGKIT”

1.2 RUMUSAN MASALAH

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah:

Bagaimana penerapan algoritma *K-Means* dengan *Euclidean Distance* dapat digunakan untuk mengklasterisasi data usulan penerima Program Indonesia Pintar (PIP) di SDN 057/IX Tangkit sehingga diperoleh kelompok siswa yang layak, cukup layak, dan tidak layak menerima bantuan?

1.3 BATASAN MASALAH

Adapun batasan masalah pada penelitian ini adalah:

1. Penelitian ini berfokus pada proses klasterisasi data usulan penerima bantuan Program Indonesia Pintar (PIP) di SDN 057/IX Tangkit.

2. Data yang digunakan merupakan data siswa yang diusulkan sebagai calon penerima PIP di SDN 057/IX Tangkit, yang meliputi beberapa atribut seperti pekerjaan dan penghasilan orang tua.
3. Metode yang digunakan dalam proses pengelompokan data adalah algoritma *K-Means* dengan perhitungan jarak menggunakan *Euclidean Distance*.
4. Penelitian ini hanya menghasilkan kelompok siswa berdasarkan tingkat kelayakan penerimaan Program Indonesia Pintar (PIP), yaitu kategori layak, cukup layak, dan tidak layak.
5. Sistem atau model yang dikembangkan tidak membahas proses verifikasi dan validasi akhir penerima bantuan oleh pihak sekolah atau instansi terkait, melainkan sebatas membantu proses analisis data usulan penerima.
6. Alat bantu (*Tools*) yang digunakan pada penelitian ini adalah *RapidManer* dan *WEKA*

1.4 TUJUAN DAN MANFAAT PENELITIAN

1.4.1 TUJUAN PENELITIAN

- 1 Untuk menerapkan algoritma *K-Means* dengan *Euclidean Distance* dalam proses pengelompokan data usulan penerima Program Indonesia Pintar (PIP) di SDN 057/IX Tangkit.
- 2 Untuk mengelompokkan siswa berdasarkan tingkat kelayakan penerimaan PIP ke dalam beberapa kategori, yaitu layak, cukup layak, dan tidak layak.
- 3 Untuk membantu pihak sekolah dalam pengambilan keputusan yang lebih objektif, cepat, dan akurat dalam menentukan siswa yang berhak menerima bantuan PIP.

- 4 Untuk meningkatkan efektivitas pengelolaan data usulan penerima bantuan pendidikan melalui penerapan metode analisis data berbasis data mining dengan algoritma *K-Means*.

1.4.2 MANFAAT PENELITIAN

1. Hasil penelitian ini dapat membantu pihak sekolah di SDN 057/IX Tangkit dalam mengidentifikasi dan mengelompokkan calon penerima Program Indonesia Pintar (PIP) secara lebih objektif dan efisien, sehingga proses penentuan siswa yang layak, cukup layak dan tidak layak menerima bantuan menjadi lebih tepat sasaran.
2. Penelitian ini dapat menjadi bahan referensi atau dasar pertimbangan dalam pengambilan keputusan terkait penyaluran bantuan pendidikan agar lebih transparan dan sesuai dengan kondisi sosial ekonomi siswa.
3. Penelitian ini dapat menjadi acuan dan bahan pengembangan untuk penelitian berikutnya, khususnya dalam penerapan metode data mining dan algoritma *K-Means* dengan variasi parameter atau metode pembandingan lainnya untuk meningkatkan akurasi hasil klasterisasi.

1.5 SISTEMATIKA PENULISAN

Sistematika penulisan dalam laporan tugas akhir ini disusun untuk memberikan gambaran yang jelas mengenai alur pembahasan penelitian. Adapun susunan penulisan skripsi ini adalah sebagai berikut:

BAB I : PENDAHULUAN

Bab ini berisi uraian mengenai latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan. Bab ini menjadi dasar pemahaman mengenai alasan dan arah penelitian dilakukan.

BAB II : LANDASAN TEORI

Bab ini memaparkan teori-teori pendukung yang relevan dengan penelitian, seperti konsep data mining, klasterisasi, algoritma K-Means, Euclidean Distance, Program Indonesia Pintar (PIP), serta penelitian terdahulu yang berhubungan dengan topik. Bab ini berfungsi sebagai landasan konseptual dari pelaksanaan penelitian.

BAB III : METODOLOGI PENELITIAN

Bab ini menjelaskan metode yang digunakan dalam penelitian, meliputi kerangka kerja penelitian, teknik pengumpulan data, tahapan pra-pemrosesan data, metode analisis data, alat bantu penelitian yang digunakan, serta jadwal pelaksanaan penelitian.

BAB IV : ANALISIS DAN VISUALISASI

Bab ini membahas proses analisis data serta visualisasi hasil klasterisasi penerima Program Indonesia Pintar (PIP) menggunakan algoritma K-Means. Analisis dilakukan untuk mengetahui pola pengelompokan siswa berdasarkan kondisi sosial ekonomi sehingga diperoleh kategori layak, cukup layak, dan tidak layak.

BAB V : PENUTUP

Bab ini memuat kesimpulan yang diperoleh dari hasil penelitian serta saran-saran yang dapat digunakan untuk pengembangan penelitian ke depan atau sebagai pertimbangan bagi pihak sekolah dalam proses seleksi penerima PIP.