

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Kemiskinan merupakan salah satu masalah mendasar yang menjadi fokus perhatian pemerintah di setiap negara, termasuk Indonesia [1]. Ada berbagai faktor yang menyebabkan kemiskinan di Indonesia, seperti tingginya pertumbuhan penduduk, rendahnya tingkat pendidikan, kurangnya motivasi untuk bekerja, terbatasnya peluang kerja yang tersedia, dan lain sebagainya [2]. Kombinasi dari berbagai faktor ini membuat kemiskinan menjadi isu yang kompleks untuk diatasi.

Menurut Badan Pusat Statistik [3] pada bulan maret 2024, tingkat kemiskinan di Indonesia mencapai 9,03%. Menteri Koordinator bidang Pembangunan Manusia dan Kebudayaan RI, Muhadjir Effendy [4] mengatakan anggotanya terus berupaya mencapai target yang telah ditetapkan oleh Presiden Joko Widodo, yaitu mengurangi tingkat kemiskinan hingga mendekati 7,5%.

Dalam upaya memperbaiki kondisi kemiskinan di Indonesia, pemerintah telah mengeluarkan berbagai program bantuan sosial yang ditujukan untuk membantu masyarakat miskin atau kurang mampu seperti Bantuan Langsung Tunai (BLT), Bantuan Pangan Non Tunai (BPNT), Bantuan Beras Miskin (Raskin), Keluarga Penerima Manfaat (KPM), Program Keluarga Harapan (PKH) dan Program Kartu Prakerja [5]. Program-program ini dirancang untuk membantu masyarakat dalam

memenuhi kebutuhan dasar seperti pangan, kesehatan, pendidikan, hingga mendukung kebutuhan kelompok lansia dan penyandang disabilitas.

Salah satu program yang menjadi andalan pemerintah dalam mengurangi kemiskinan adalah Program Keluarga Harapan (PKH). PKH adalah program untuk mengurangi kemiskinan dan mengembangkan sistem perlindungan sosial bersyarat terhadap masyarakat miskin, yang bertujuan mempercepat pencapaian *Millennium Development Goals* (MDGs) [6].

Meskipun tujuan program ini sangat baik, dalam praktiknya penyaluran bantuan PKH ini sering menghadapi berbagai masalah terkait pengolahan data dan ketidakmerataan dalam penyaluran bantuan [7]. Permasalahan ini juga dialami oleh Desa Talang Belido. Warga mengeluhkan bahwa penyaluran bantuan PKH dianggap tidak merata. Hal ini dikarenakan warga yang merasa berhak mendapatkan bantuan justru tidak terpilih, sementara warga yang dianggap tidak berhak mendapatkan bantuan justru menerima bantuan tersebut.

Permasalahan lainnya adalah jumlah data penduduk yang besar di Desa Talang Belido menyebabkan pengelola data membutuhkan waktu lama untuk menentukan penerima bantuan PKH. Hal ini dikarenakan verifikasi data masih dilakukan secara manual dengan memeriksa data satu per satu. Selain itu, terdapat data yang belum diperbarui atau tidak lengkap. Terkadang, pihak desa hanya menerima usulan dari setiap RT terkait warga yang direkomendasikan untuk mendapatkan bantuan PKH.

Untuk mengatasi masalah pengolahan data tersebut dan penentuan prioritas warga yang berhak mendapat bantuan Program Keluarga Harapan (PKH), bisa menggunakan teknik *data mining*, salah satunya dengan metode *Clustering K-Means* dan algoritma C4.5.

Hal ini bisa dibuktikan dengan melihat penelitian sebelumnya yang berjudul “Penerapan Data Mining Untuk Menentukan Penerima Bantuan BLT Menggunakan Metode *Clustering K-Means* Pada Desa Pamulihan” [8], dengan menggunakan 4 atribut yang terdiri dari tanggungan, pekerjaan, gaji, dan status dengan total 280 data yang menghasilkan 9 *cluster* dengan status penerima bantuan, dalam pertimbangan, dan bukan penerima.

Sementara itu, pada penelitian sebelumnya yang berjudul “Analisis Data Mining Menggunakan Metode Algoritma C4.5 Menentukan Penerima Bantuan Langsung Tunai” [9], dengan menggunakan 4 atribut meliputi penghasilan, status, pekerjaan, dan umur dengan total 121 data yang menghasilkan suatu informasi pengelompokan data warga dengan kategori menerima bantuan dan tidak menerima bantuan.

Serta pada penelitian sebelumnya yang berjudul “Implementasi Algoritma K-Means dan C4.5 Dalam Menentukan Tingkat Penyebaran COVID-19 Di Indonesia” [10] dengan menggunakan 4 atribut yaitu nama provinsi, jumlah kasus positif, jumlah meninggal, dan jumlah sembuh dengan kategori zona hitam, zona merah, zona kuning, dan zona hijau yang dapat memberitahukan informasi mengenai pemetaan jumlah persebaran kasus Covid-19 di Indonesia.

Dari beberapa penelitian di atas, *Data Mining* dipilih karena dapat menggali atau menghasilkan informasi dari data yang besar dan mengidentifikasi pola dari data yang tidak beraturan [5], sehingga bisa memberikan prediksi dalam pengambilan keputusan. *Clustering K-Means* dipilih karena dapat mengelompokkan data berdasarkan *cluster* dengan karakteristik yang sama [11]. Serta algoritma *C4.5* dipilih karena bisa mengklasifikasikan data dan menggambarannya dalam bentuk pohon keputusan [12] sehingga bentuk klasifikasi data mudah dipahami.

Berdasarkan permasalahan dan informasi yang telah diberikan, maka penulis akan melakukan penelitian dengan judul “**Klasifikasi Data Mining Dalam Penentuan Prioritas Bantuan PKH Dengan Metode Clustering K-Means Dan C4.5 Pada Desa Talang Belido**” Dengan menggunakan enam atribut meliputi nama kepala keluarga, pekerjaan, jumlah anak atau tanggungan, pendapatan atau gaji per bulan, kondisi dan status kepemilikan rumah, serta status kesejahteraan penduduk. Dimana penentuan atribut ini ditentukan berdasarkan kriteria penduduk miskin [13] yang memenuhi syarat untuk menerima bantuan PKH [14].

1.2 PERUMUSAN MASALAH

Berdasarkan uraian latar belakang di atas, penulis dapat merumuskan permasalahan sebagai berikut:

1. Bagaimana mengelompokkan data calon penerima bantuan PKH di Desa Talang Belido dengan metode *Clustering K-Means*?

2. Bagaimana membuat model pohon keputusan dalam penentuan prioritas penerima bantuan PKH menggunakan algoritma C4.5 setelah proses *clustering* dilakukan?
3. Bagaimana mengombinasikan metode *Clustering K-Means* dan C4.5 dalam penyusunan strategi penentuan prioritas bantuan PKH yang tepat sasaran di Desa Talang Belido?

1.3 BATASAN MASALAH

Untuk mencegah pembahasan yang meluas dan tetap fokus pada ruang lingkup masalah, penulis menentukan batasan-batasan masalah yang akan dibahas sebagai berikut:

1. Penelitian ini dilakukan pada Desa Talang Belido.
2. Data yang digunakan adalah data penduduk Desa Talang Belido tahun 2023 hingga bulan desember 2024 dengan total 4.775 data penduduk.
3. Penelitian ini menggunakan metode *Clustering* dengan Algoritma *K-Means* dan metode klasifikasi dengan Algoritma C4.5.
4. Tools yang digunakan dalam melakukan pemrosesan data adalah Rapidminer.
5. Atribut yang digunakan yaitu nama kepala keluarga, pekerjaan, jumlah anak atau tanggungan, pendapatan atau gaji per bulan, kondisi dan status kepemilikan rumah, serta status kesejahteraan penduduk.
6. Jumlah tanggungan (Ibu hamil/nifas/menyusui, anak usia dini 0-6 tahun, dan anak usia 6-21 tahun yang belum menyelesaikan wajib belajar).
7. Kategori yang digunakan ada 3 (tiga) yaitu layak, dalam pertimbangan, dan tidak layak.

1.4 TUJUAN PENELITIAN

Penelitian yang penulis lakukan ini mempunyai tujuan yaitu:

1. Mengkluster data calon penerima bantuan PKH di Desa Talang Belido berdasarkan karakteristik yang telah ditentukan menggunakan metode *Clustering K-Means*.
2. Membuat Pohon Keputusan (*Decision Tree*) menggunakan metode C4.5 untuk menentukan prioritas penerima bantuan PKH di Desa Talang Belido berdasarkan hasil *clustering*.
3. Menggabungkan metode *K-Means* dan C4.5 untuk menghasilkan model penentuan prioritas bantuan PKH yang lebih tepat sasaran.

1.5 MANFAAT PENELITIAN

Manfaat yang diharapkan penulis dari penelitian ini adalah sebagai berikut:

1. Mengetahui informasi data penduduk Desa Talang Belido berdasarkan kategori layak, dalam pertimbangan, dan tidak layak dalam penentuan prioritas penerima bantuan PKH.
2. Menghasilkan model pohon keputusan (*Decision Tree*) untuk mempermudah memahami kategori dalam penentuan penerima bantuan PKH.
3. Mengetahui hasil klasifikasi keakurasian data penerima bantuan PKH dengan penggabungan metode *Clustering K-Means* dan algoritma C4.5.

1.6 SISTEMATIKA PENULISAN

Adapun sistematika penulisan laporan tesis dibagi menjadi menjadi beberapa bagian sebagai berikut:

BAB I : PENDAHULUAN

Pada bab ini berisi pendahuluan yang menjelaskan latar belakang, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II : LANDASAN TEORI DAN TINJAUAN PUSTAKA

Pada bab ini dijelaskan mengenai dasar-dasar teori tinjauan pustaka dari para pakar yang bersumber dari buku, jurnal, ataupun internet yang memuat konsep-konsep teoritis yang dijadikan sebagai referensi penulis dalam melakukan penelitian. Teori-teori yang digunakan antara lain mengenai klasifikasi, *data mining*, *Clustering*, *K-Means*, *C4.5*, *PKH*, dan *Rapidminer*.

BAB III : METODE PENELITIAN

Pada bab ini dijelaskan mengenai metode-metode yang digunakan dalam penelitian ini yang meliputi alur penelitian, bahan penelitian, serta alat penelitian.

BAB IV : HASIL PENELITIAN DAN PEMBAHASAN

Pada bab ini membahas tentang hasil analisis dari data yang telah dihitung dengan menggunakan gabungan metode *Clustering K-Means*

dan algoritma C4.5 serta bentuk visualisasi analisis dari tools *Rapidminer* yang digunakan.

BAB V : PENUTUP

Pada bab ini akan membahas mengenai kesimpulan dari hasil penulisan laporan ini disertai saran-saran yang berguna untuk pengembangan penelitian kedepan.