

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG MASALAH

Data mining merupakan suatu proses yang digunakan untuk menemukan pola dan keterhubungan dalam kumpulan data yang besar dan kompleks [1]. Dalam proses ini, teknik statistik, matematika digunakan untuk mengekstrasi informasi yang bernilai dari data tersebut. Penerapan data mining dapat diterapkan dalam berbagai sektor seperti bisnis, pemerintahan, kesehatan, dan ilmu pengetahuan. Misalnya, dalam konteks sosial, data mining bisa digunakan untuk mengelompokkan populasi miskin dengan menganalisis data demografis, ekonomi, dan sosial [2]. Dalam Data Mining, metode yang beroperasi dengan mencari dan mengelompokkan data yang memiliki karakteristik serupa antara satu entitas data dengan yang lainnya disebut dengan *Clustering* [3]. Salah satu metode data mining yang digunakan pada penelitian ini adalah *clustering*, karena data yang di gunakan tidak memiliki kelas.

Clustering adalah suatu metode dalam data mining yang berfokus pada pencarian dan pengelompokkan data-data yang memiliki karakteristik serupa antara satu data dengan data lainnya yang telah dikumpulkan [3]. Objek akan disusun menjadi satu atau lebih kelompok, sehingga objek yang terdapat dalam kelompok yang sama memiliki tingkat kesamaan yang tinggi antara satu dengan yang lain [4]. Tujuannya adalah agar objek-objek yang tergabung dalam suatu kelompok memiliki kesamaan (atau keterkaitan) satu sama lain dan secara

berbeda (atau tidak berkaitan) dengan objek-objek dalam kelompok lainnya [5]. Dalam *clustering*, terdapat beberapa metode yang digunakan, salah satunya adalah metode *K-Means* [6]. Pada penelitian ini metode *K-Means* digunakan untuk mengelompokkan data penduduk miskin.

K-Means adalah metode analisis data atau teknik data mining yang secara otomatis melakukan proses pemodelan tanpa supervisi (*unsupervised*) dan merupakan salah satu metode pengelompokkan data melalui sistem partisi [7]. Algoritma *K-Means* memiliki kemampuan yang baik dalam menilai ukuran objek yang sangat akurat [8]. Oleh karena itu, algoritma ini cenderung lebih efisien ketika digunakan untuk memproses sejumlah besar objek [9]. Salah satu penerapan metode *K-Means* adalah pengelompokkan penduduk miskin.

Salah satu permasalahan yang masih dihadapi oleh negara Indonesia yaitu masalah kemiskinan. Masalah kemiskinan merupakan masalah yang kompleks dan bersifat multidimensional sehingga menjadi prioritas pembangunan. Selama ini, pemerintah Indonesia telah banyak memiliki program-program untuk mengentaskan kemiskinan yang ada [10]. Dalam rangka percepatan penanggulangan kemiskinan sekaligus pengembangan kebijakan dibidang sosial, dimulai pada tahun 2007 pemerintah melaksanakan Program Keluarga Harapan (PKH). PKH juga sebagai sarana untuk mengembangkan sistem jaminan sosial bagi masyarakat sangat miskin [11]. Bisa disimpulkan bahwa mayoritas penduduk Indonesia masih mengalami kondisi ekonomi yang kurang memadai, atau dengan kata lain, berada dalam keadaan tidak sejahtera atau miskin [12]. Berdasarkan

Penjelasan di atas, perlu adanya upaya untuk mengatasi masalah kemiskinan yang ada di Indonesia.

Upaya pengentasan kemiskinan diharapkan mampu mengatasi kesenjangan perekonomian dan membantu perekonomian masyarakat miskin. Akan tetapi pelaksanaan program PKH ini masih memiliki kendala, yaitu penentuan keluarga yang benar-benar berhak untuk menerima bantuan tersebut. Maka dari itu, untuk membantu pemerintah kecamatan Mendahara Ulu dalam menentukan warga yang layak menerima bantuan, penulis bermaksud untuk membantu pemerintah Kecamatan Mendahara Ulu dalam pengelompokan (mengklaster) data penduduk yang layak menerima bantuan PKH secara tepat dengan teknik data mining menggunakan metode *K-Means Clustering*. Adapun data yang digunakan pada penelitian ini adalah data penduduk miskin Kecamatan Mendahara Ulu tahun 2023. dengan jumlah penduduk miskin sebanyak 10736 orang dengan 2654 KK. Mayoritas warga miskin di kecamatan Mendahara Ulu bekerja sebagai petani yang berpenghasilan tidak menentu.

Pada penelitian [13] membahas tentang metode *K-Means clustering* untuk pengelompokan data penduduk yang layak menerima Bantuan Langsung Tunai Dana Desa secara tepat menggunakan *algoritma k-means* sebagai sistem pendukung keputusan yang tepat. Metode *K-Means* pada hakekatnya merupakan metode untuk pengklasterisasi (*clustering*). Kelebihan metode *K-Means clustering* ini sangat sederhana di implementasikan dan paling banyak digunakan dalam pengklasteran. Dari hasil pengklasteran data penduduk yang memenuhi kriteria

penerima BLT-DD tersebut, maka akan didapatkan hasil perhitungan kelayakan calon penerima dana bantuan yang tepat.

Pada penelitian [14] penulis dalam penelitian ini menjelaskan bahwa metode *K-Means Cluster Analysis* cukup efektif diterapkan dalam proses pengklasifikasian karakteristik terhadap objek penelitian. *Algoritma K-Means* juga tidak terpengaruh terhadap urutan objek yang digunakan, hal ini dibuktikan ketika penulis mencoba menentukan secara acak titik awal pusat *cluster* dari salah satu objek pada permulaan perhitungan.

Berdasarkan pemaparan di atas penulis akan mengangkat judul penelitian **“Penerapan *Algoritma K-Means* Untuk Menentukan Prioritas Penerima Bantuan Sosial PKH Di Kecamatan Mendahara Ulu Tanjung Jabung Timur”**.

1.2 RUMUSAN MASALAH

Dari latar belakang masalah yang telah diuraikan sebelumnya, maka perumusan masalah yang dapat diambil dalam penelitian ini adalah:

1. Bagaimana menerapkan *Algoritma Clustering K-Means* dalam Menentukan Prioritas Penerima Bantuan Sosial PKH Di Kecamatan Mendahara Ulu?
2. Bagaimana mengevaluasi *algoritma K-means clustering* dalam mengelompokkan kelayakan calon penerima Program Keluarga Harapan (PKH) di Kecamatan Mendahara Ulu?

1.3 BATASAN MASALAH

Untuk menghindari perluasan lingkup masalah, dan agar penelitian ini berjalan dengan baik, maka dibuatlah beberapa batasan penelitian sebagai berikut:

1. Penelitian ini hanya berfokus pada faktor-faktor yang mempengaruhi kelayakan calon penerima bantuan Program Keluarga Harapan (PKH) di Kecamatan Mendahara Ulu, yaitu Tunjangan Keluarga, Penghasilan Keluarga, Status Sosial Keluarga, Kepemilikan Rumah, dan Kondisi Rumah.
2. Data yang digunakan dalam penelitian ini merupakan data primer yang diperoleh dari Kecamatan Mendahara Ulu Tanjung Jabung Timur.
3. Penelitian ini hanya dilakukan di kecamatan Mendahara Ulu Kabupaten Tanjung Jabung Timur.
4. Penelitian ini menggunakan Algoritma *K-Means* dalam mengkluster data penerima bantuan sosial PKH.
5. Pengujian hasil analisis menggunakan *tools* SPSS

1.4 TUJUAN DAN MANFAAT

1.4.1 Tujuan Penelitian

Adapun tujuan penelitian adalah sebagai berikut:

1. Untuk menerapkan Algoritma *K-Means Clustering* dalam menentukan kelayakan calon penerima bantuan Program Keluarga Harapan (PKH) di Kecamatan Mendahara Ulu.

2. Untuk mengevaluasi Algoritma *K-Means Clustering* dalam mengelompokkan kelayakan calon penerima bantuan Program Keluarga Harapan (PKH) di Kecamatan Mendahara Ulu.

1.4.2 Manfaat Penelitian

Adapun manfaat penelitian adalah sebagai berikut:

1. Penelitian ini diharapkan dapat memberikan kontribusi bagi pengembangan metode menentukan kelayakan calon penerima bantuan Program Keluarga Harapan (PKH) di Kecamatan Mendahara Ulu.
2. Penelitian ini di harapkan bisa memberikan saran kepada pemerintah untuk meningkatkan tingkat akurasi menentukan kelayakan calon penerima bantuan Program Keluarga Harapan (PKH) di Kecamatan Mendahara Ulu.
3. Penelitian ini diharapkan bisa membantu staf dengan lebih mudah mengklaster data penduduk kurang mampu yang memenuhi syarat untuk menerima bantuan.

1.5 SISTEMATIKA PENULISAN

Sistematika penelitian ini, sebagai berikut:

BAB I : PENDAHULUAN

Pada bab ini menguraikan tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB II : LANDASAN TEORI

Pada bab ini terdiri dari penjelasan mengenai studi sebelumnya yang mendukung tugas akhir dan konsep atau teori-teori yang memiliki keterkaitan terhadap topik penelitian, yaitu bantuan sosial PKH, data mining, metode *k-means clustering*, dan memprediksi penerima bantuan dengan mengolah data penduduk menggunakan aplikasi IBM SPSS Statistics.

BAB III : METODOLAGI PENELITIAN

Pada bab ini menjelaskan mengenai kerangka kerja penelitian, metode atau teknik pengumpulan data serta alat-alat dan bahan-bahan pendukung selama melakukan penelitian.

BAB IV : ANALISIS DAN HASIL

Pada bab ini menjelaskan tentang gambaran umum objek organisasi penelitian, serta menganalisis dan menghitung data penduduk menggunakan metode Algoritma *k-means clustering*.

BAB V : PENUTUP

Pada bab ini berisi kesimpulan dan saran yang diperoleh dari hasil penelitian yang telah dilakukan