

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Data mining merupakan proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan metode tertentu. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses *Knowledge Discovery in Database* (KDD) secara keseluruhan [1]. KDD merupakan kegiatan yang meliputi pengumpulan, pemakaian data, historis untuk menemukan keteraturan, pola atau hubungan dalam dataset yang berukuran besar [2]. Dalam tahap ini, *Data mining* bisa diterapkan dengan berbagai teknik diantaranya *classification*, *association* dan *clustering* [3]. Dengan data mining dimana serangkaian prosesnya akan menghasilkan suatu nilai tambah berupa informasi atau pengetahuan baru yang selama ini tidak diketahui secara manual dari sekumpulan data.

Clustering merupakan sebuah teknik dari analisis multivariable yang digunakan untuk mengelompokkan objek-objek (variabel atau data) sehingga dapat menghasilkan suatu informasi untuk membantu pelaksanaan pengujian terhadap objek dan pada akhirnya dapat menyajikan suatu hipotesis berdasarkan relasi yang terjadi [4]. Klasterisasi dapat digunakan untuk mengklasifikasikan daerah-daerah yang padat, menemukan pola-pola distribusi, dan menemukan keterkaitan antara setiap data [5]. Algoritma *Clustering* yang paling banyak digunakan dalam *Data Mining* yaitu Algoritma *K-Means*, dimana metode ini akan

mengelompokkan berdasarkan jarak setiap data dengan titik pusat pada masing-masing *cluster* [6]. Algoritma *K-means* memiliki potensi sebagai teknik yang dapat dipakai untuk melakukan pengelompokan data.

Algoritma *K-means* merupakan salah satu cara pengelompokkan data sekatan (nonhierarki) yang berusaha membagi data yang ada kedalam bentuk dua atau lebih kelompok [7]. Untuk proses pengelompokkan dilakukan dengan melihat setiap data yang memiliki karakteristik yang sama maka akan dikelompokkan dalam 1 kluster dan untuk data yang berbeda maka akan dikelompokkan dalam kluster yang lain [8]. Penggunaan klasterisasi dapat mengidentifikasi suatu wilayah dengan menemukan pola-pola karakteristik hingga menemukan keterkaitan yang menarik antara atribut-atribut data [9]. Algoritma *K-Means* digunakan dalam hal ini akan selesai dan memiliki solusi terbaik. Sehingga nantinya dapat membantu dalam menentukan kategori prioritas untuk mendapatkan Program BPNT [10]. Pada penelitian ini penulis menggunakan Data Mining metode *Clustering K-means* untuk menentukan prioritas penerima BPNT.

Program Bantuan Pangan Non Tunai (BPNT) merupakan bantuan sosial yang disalurkan secara non tunai dari pemerintah yang diberikan kepada Keluarga Penerima Manfaat (KPM) setiap bulannya melalui uang elektronik selanjutnya digunakan untuk membeli bahan pangan di e-warong yang bekerja sama dengan Bank Himbara [11]. Tujuan adanya program ini dapat mengurangi beban pengeluaran KPM melalui pemenuhan kebutuhan sebagai kebutuhan pangan dan mendorong pencapaian tujuan pembangunan berkelanjutan [12]. Program yang dikeluarkan dari pemerintah diharapkan dapat memenuhi hak masyarakat miskin

dalam meningkatkan ketahanan pangan serta sekaligus sebagai penanggulangan kemiskinan di Indonesia serta mendapatkan hasil yang maksimal untuk kedepannya [13]. Mengingat program BPNT merupakan program dari pemerintah pusat serta melihat angka persentase penduduk miskin mengalami kenaikan, sehingga Kecamatan Koto Baru merupakan salah satu wilayah sasaran diterapkannya program tersebut. Pelaksanaan program BPNT dilakukan secara bertahap terhadap komoditi tambahan yang disesuaikan dengan kemampuan potensi lokal daerah.

Penelitian ini juga pernah dilakukan oleh beberapa peneliti seperti Said Nanda Saputra, dkk [14] yaitu Penerapan *K-Means* dilakukan melalui *Google Colab* dan metode *clustering*, ditemukan 2 kluster dengan 666 data dalam kluster 1 dan 16 data dalam kluster 2. Evaluasi menggunakan *Silhouette Score* menunjukkan hasil yang baik dengan nilai 0.91697965. Dengan demikian, algoritma *K-Means Clustering* dapat mengidentifikasi pola karakteristik penerima BPNT dan membedakan kelompok yang layak dan tidak layak menerima bantuan. Kemudian Penelitian yang dilakukan oleh Djuniardi Suhardinata, dkk [15] yaitu Hasil nilai dari evaluasi dengan menggunakan metode *Silhouette Coefficient* untuk 3 kluster yaitu 0.440104 sehingga hasil kluster yang terbentuk dapat dimasukkan kedalam kategori *cluster weak structure* dengan range $0.25 < SC \leq 0.5$ dan merupakan kluster yang baik pada tingkat itu. Dari hasil dapat ditarik sebuah kesimpulan bahwa algoritma *K-Means* sangat baik dalam mengelompokkan penerima bantuan. Dari hasil penelitian ini bahwa kemiripan dari pola yang ditemukan itu berdasarkan pada nilai *attribute* yang dimiliki setiap datanya.

Kecamatan Koto baru yang beralamat di Kota Sungai Penuh Provinsi Jambi, dimana penyaluran program BPNT di kecamatan tersebut masih belum merata dan tidak tepat sasaran dikarenakan jumlah penerima bantuan di setiap desa jumlahnya tidak selalu sama setiap tahunnya, ada yang meningkat dan ada juga yang menurun, ekonomi setiap keluarga juga berbeda-beda. Maka dari itu diperlukan penerapan teknik data mining terhadap penentuan penerima bantuan dengan cara mengklaster data penduduk miskin menggunakan Algoritma *k-means* agar penerima bantuan tepat sasaran dan berguna bagi yang membutuhkan. Pada penelitian ini data yang akan digunakan sebanyak 1523 dengan 10 atribut.

Dengan jumlah populasi penerima yang besar memunculkan kendala sulitnya menentukan data penerima bantuan yang benar-benar layak menerima bantuan. Dengan menerapkan teknik data mining dan menggunakan Algoritma *K-Means* dapat memudahkan proses pengelompokan data penduduk miskin yang berhak menerima program BPNT.

Berdasarkan permasalahan yang telah di uraikan tersebut maka penulis tertarik melakukan penelitian dengan judul **“KLAUSTERISASI KELAYAKAN PENERIMA BANTUAN PANGAN NON TUNAI (BPNT) DENGAN MENGGUNAKAN ALGORITMA K-MEANS (STUDI KASUS : KECAMATAN KOTO BARU KOTA SUNGAI PENUH)”**.

1.2 RUMUSAN MASALAH

Berdasarkan latar belakang masalah yang telah dibahas sebelumnya, maka rumusan masalah yang akan dibahas pada penelitian ini adalah :

1. Bagaimana menerapkan Algoritma K-Means dalam melakukan *clustering* untuk menentukan kelayakan penerima BPNT ?
2. Bagaimana mengevaluasi hasil penerapan metode K-Means clustering dalam memprediksi kelayakan penerima BPNT ?

1.3 BATASAN MASALAH

Untuk menghindari terjadinya pembahasan yang meluas, maka penulis membatasi pembahasan permasalahan berdasarkan tema dan judul sebagai berikut:

1. Penelitian ini berfokus pada faktor-faktor yang mempengaruhi kelayakan penerima bantuan BPNT yaitu jumlah keluarga, kondisi rumah, status kepemilikan, penghasilan, dan status sosial.
2. Data yang digunakan dalam penelitian ini yaitu data penerima bantuan sosial tahun 2022/2023.
3. Penelitian hanya dilakukan di Kecamatan Koto Baru.
4. Penelitian ini menggunakan Algoritma K-Means dalam mengklaster data penerima BPNT.
5. Pengujian hasil analisis menggunakan tools WEKA.

1.4 TUJUAN DAN MANFAAT PENELITIAN

1.4.1 Tujuan Penelitian

Adapun tujuan yang akan disampaikan penulis dalam penelitian ini adalah:

1. Untuk menerapkan Algoritma *K-Means* dalam mengklaster data penduduk untuk menentukan kelayakan penerima BPNT.
2. Untuk mengevaluasi hasil penerapan Algoritma *K-Means* dalam memprediksi kelayakan penerima BPNT.

1.4.2 Manfaat Penelitian

Adapun beberapa manfaat yang akan disampaikan penulis pada penelitian ini adalah :

1. Penelitian ini dapat memberikan kontribusi bagi pengembangan metode penentuan kelayakan penerima bantuan BPNT.
2. Penelitian ini dapat memberikan rekomendasi bagi pemerintah dalam meningkatkan akurasi penentuan kelayakan penerima bantuan BPNT.
3. Penelitian ini dapat mempermudah pihak staff dalam mengklaster data penduduk miskin yang berhak menerima bantuan.

1.5 SISTEMATIKA PENULISAN

Adapun sistematika penulisan tugas akhir ini, penulis menguraikan dalam beberapa bab yaitu :

BAB I : PENDAHULUAN

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

BAB II : LANDASAN TEORI

Pada bab landasan teori ini membahas tentang teori-teori dan pendapat para ahli yang berhubungan dengan permasalahan yang dianalisis, teori-teori yang digunakan antara lain mengenai *data mining*, *clustering*, algoritma *k-means*, Bantuan Pangan Non Tunai (BPNT), dan WEKA.

BAB III: METODOLOGI PENELITIAN

Pada bab ini menjelaskan tentang kerangka kerja penelitian, metode pengumpulan data, metode *clustering*, serta alat bantu yang digunakan pada penelitian ini.

BAB IV: ANALISIS DAN HASIL

Bab ini dilakukan pengelompokan data menggunakan metode *k-means* dan menampilkan praproses data sebelum diolah ke dalam tools dan bentuk visualisasi analisis dari tools WEKA yang digunakan.

BAB V : PENUTUP

Bab ini yang berisikan kesimpulan-kesimpulan yang diambil dari hasil analisis serta saran-saran yang mencakup keseluruhan dari hasil penelitian.