

BAB I

PENDAHULUAN

1.1 Latar Belakang

Knowledge Discovery in Database (KDD) adalah suatu proses *komprehensif* yang melibatkan ekstraksi atau identifikasi pola, pengetahuan, dan informasi yang berpotensi berguna dari kumpulan data yang besar. Hasil dari proses ini adalah pengetahuan dan informasi yang baru, sah, mudah dipahami serta bermanfaat (Munir *et al*, 2023). *Data Mining* merupakan bagian dari sebuah inti proses dari KDD meliputi perkiraan dari sebuah algoritma dalam pencarian data, membangun model dan menemukan pola yang belum diketahui sebelumnya (Alkhairi & Situmorang, 2022).

Data Mining adalah suatu teknik yang digunakan untuk mengungkap sebuah informasi yang tersembunyi dalam kumpulan data yang besar yang bertujuan untuk menemukan pola yang menarik yang tidak diketahui sebelumnya (Alfarizi & Andri, 2021). *Data Mining* adalah istilah yang digunakan untuk menggambarkan proses penggalian pengetahuan dari sekumpulan informasi (Lestari,Agustiansyah, 2023). Penerapan data mining tidak hanya bertujuan untuk memperoleh pengetahuan atau informasi baru, tetapi juga dapat meningkatkan pemahaman baru tentang hal yang sudah diketahui sebelumnya (Muttaqin & Defriani, 2020).

Klasifikasi merupakan suatu model atau fungsi yang digunakan untuk menggambarkan dan membedakan kelas atau konsep. Tujuannya adalah untuk melakukan prediksi kelas dari objek yang tidak memiliki label kelas yang diketahui (Andarista, Jananto 2022). Klasifikasi adalah salah satu metoda Algoritma *Data mining* yang digunakan untuk mengali informasi. Algoritma C4.5 termasuk salah satu Algoritma dalam metode klasifikasi yang memiliki fungsi untuk membangun pohon keputusan (Alkhairi & Situmorang, 2022).

Algoritma C4.5 adalah suatu metode yang dapat digunakan untuk melakukan klasifikasi atau pengelompokan dalam sebuah dataset yang berfokus pada pembentukan pohon keputusan (*Decision Tree*) (Putri, Irawan, Rizky, 2021). Klasifikasi Algoritma C4.5 menggunakan pohon keputusan dimulai dari atas (akar) lalu kebawah (akhir), dimana atribut paling atas adalah akar yang disebut node yang akan mewakili atribut dan dibawah dinamakan daun untuk mewakili kelas dan perhitungan yang disebut nilai Entropy dan nilai Gain (Sari & Purwanto, 2023).

Algoritma C4.5 tidak hanya dikembangkan pada bidang teknologi saja, banyak bidang telah mengimplementasikan seperti di bidang Perbankan pada topik *Penerapan Data Mining Untuk Memprediksi Transaksi Nasabah Dengan Algoritma C4.5* (Marlina & Bakri, 2021). Pada Bidang Pertanian juga diterapkan dengan topik *Penerapan Data Mining Algoritma C4.5 Terhadap Prediksi Faktor Menurunnya Hasil Panen Padi* (Siahaan et.al, 2022). Pada Bidang Pemerintahan juga menerapkan Algoritma C4.5 dengan topik *Penerapan Data Mining Algoritma C4.5 Untuk Klasifikasi Hasil Pengujian Kendaraan Bermotor* (Andarista & Jananto, 2022) Pada Dinas perhubungan Kabupaten Rembang.

Berdasarkan penelitian sebelumnya, Penulis melakukan penelitian untuk optimalisasi pendistribusian bantuan sosial dengan menggunakan Algoritma C4.5. Penelitian ini dilakukan di Nagari Limo Koto Kabupaten Pasaman sehingga pendistribusian bantuan sosial lebih tepat sasaran. Oleh karena itu diangkatlah judul “Penerapan *Data mining* Menggunakan Algoritma C.45 Untuk Menentukan Jenis dan Kelayakan Penerima Bantuan Sosial (Studi Kasus Di Nagari Limo Koto Kabupaten Pasaman)”

1.2 Perumusan Masalah

Berdasarkan latar belakang tersebut diatas, maka rumusan masalah yang diajukan dalam penelitian ini adalah sebagai berikut :

1. Bagaimana Penerapan *Data Mining* dengan metode Algoritma C4.5 dapat mengklasifikasikan kelayakan penerima bantuan sosial di Nagari Limo Koto Kabupaten Pasaman.
2. Bagaimana Penerapan metode algoritma C.45 dapat membantu penyelesaian pendistribusian bantuan yang tepat sasaran.

1.3 Batasan Masalah

Agar penelitian Tesis ini terarah dan tujuan tercapai sesuai dengan yang diharapkan, maka perlunya pembatasan masalah yaitu :

1. Data yang digunakan adalah data dari Dinas Sosial Kabupaten Pasaman untuk wilayah kenagarian Limo Koto Kabupaten Pasaman.
2. Metode yang digunakan yaitu metode Algoritma C.45

1.4 Tujuan Penelitian

Tujuan dari Penelitian ini adalah :

1. Memahami dan menganalisa pengembangan Sistem pendukung keputusan dengan metode C4.5 dapat mengklasifikasikan kelayakan penerima bantuan sosial di Nagari limo koto Kabupaten pasaman
2. Menganalisa sejauh mana penerapan metode Algoritma C4.5 dapat membantu penyelesaian pendistribusian bantuan yang tepat sasaran.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah :

1. Dapat digunakan sebagai sistem pendukung keputusan dalam pengelompokan kelayakan bantuan sosial di Nagari Limo Koto Kabupaten Pasaman sehingga tidak terjadi lagi penyaluran bantuan sosial yang tidak tepat sasaran.
2. Hasil dari penelitian ini dapat digunakan untuk diterapkan dalam pemberian bantuan sosial di Nagari Limo Koto Kabupaten pasaman, dan dapat dijadikan acuan untuk penelitian selanjutnya.

1.6 Sistematika Penulisan

Sistematika penulisan disesuaikan dengan template yang diatur dalam tata penulisan program studi masing-masing. Seperti :

BAB I PENDAHULUAN

Berisi Latar Belakang, Perumusan Masalah, Batasan Masalah, Tujuan Penelitian, Manfaat Penelitian, dan Sistematika Penelitian.

BAB II LANDASAN TEORI

Pada bab ini membahas tentang landasan teori yang meliputi konsep dan penerapan Knowledge Discovery in Database (KDD), Data Mining, Algoritma C45.

BAB III METODE PENELITIAN

Bab ini membahas tentang analisa dan Penggunaan secara matematis algoritma C45.

BAB IV ANALISA DAN PEMBAHASAN

Pada bagian ini akan dibahas tentang analisis sistem dan pembahasan sistem yang sedang berjalan.

BAB V IMPLEMENTASI DAN PENGUJIAN

Bab ini membahas hasil dan implementasi algoritma C45 Pada Pemberian Bantuan sosial di Nagari Limo Koto.

BAB VI KESIMPULAN DAN SARAN

Bab ini membahas kesimpulan dan hasil penelitian penerapan data mining menggunakan algoritma C45 untuk menentukan jenis dan kelayakan bantuan sosial studi kasus di Nagari Limo Koto Kabupaten Pasaman.