

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Knowledge Discovery in Database (KDD) merupakan cara pengekstrasian pengetahuan dari data yang berukuran besar untuk menemukan pola-pola baru sehingga mendapatkan sebuah pengetahuan dan informasi yang baru. Teknologi *data mining* digunakan untuk menggali pengetahuan yang ada didalam *database* (Elda et al., 2021). Istilah KDD dan *data mining* biasanya dimanfaatkan secara bergantian dalam memperjelas tahapan penggalian informasi tertanam pada *database* besar.

Data mining merupakan inti dari proses KDD (Nurhachita & Negara, 2021). *Data mining* adalah proses yang dapat mengidentifikasi pengetahuan potensial berguna dan bermanfaat yang ada dalam sebuah *database* besar menggunakan bantuan dari teknik *statistic*, matematik, kecerdasan buatan dan *machine learning*. *Data mining* juga merupakan kegiatan berupa pengumpulan dan pengolahan data yang tidak digunakan untuk menganalisis dan menemukan sebuah nilai atau pengetahuan yang dapat membantu dalam penyelesaian sebuah masalah (Dwita Elisa Sinaga et al., 2022).

Beberapa teknik yang ada dalam *data mining* salah satunya yaitu *classification*. Salah satu algoritma *data mining* yang digunakan untuk klasifikasi data adalah algoritma C4.5. Menurut Situmorang et al. (2022), algoritma C4.5 adalah algoritma klasifikasi data dengan teknik pohon keputusan (*Decision Tree*) yang populer dan disukai karena memiliki kelebihan-kelebihan. Kelebihan ini yaitu, dapat mengolah data numerik dan diskret, dapat menangani nilai atribut yang hilang, menghasilkan aturan-aturan yang mudah diinterpretasikan, tercepat diantara algoritma-algoritma yang lain, keakuratan prediksi, dan skalabilitas (kemampuan untuk membangun model secara efisien untuk data berjumlah besar).

Penerapan *data mining* dengan metode *C4.5 Decision Tree* dapat membantu menggali pengetahuan yang digunakan untuk analisis pengadaan peralatan dan mesin kantor. Dalam pengadaan peralatan dan mesin kantor, terdapat banyak faktor yang harus dipertimbangkan, seperti ketersediaan anggaran, kebutuhan, dan jenis peralatan yang dibutuhkan. Oleh karena itu, penerapan *data mining* dapat membantu dalam melakukan analisis terhadap data yang telah terkumpul, sehingga dapat memberikan informasi yang lebih akurat dan efektif.

Penelitian sebelumnya penerapan algoritma *C4.5* untuk *data mining* sudah banyak dilakukan, yaitu pada penelitian yang dilakukan oleh Lestari, *et al.* (2022) mengusulkan algoritma *C4.5* untuk mengklasifikasikan Usaha Mikro, Kecil, dan menengah (UMKM) sebanyak 3.808 data UMKM menghasilkan 91% UMKM termasuk dalam kategori mikro, 8% kategori kecil, dan 1% kategori sedang dengan akurasi algoritma *C4.5* memiliki akurasi lebih baik dari *naïve bayes* dengan nilai 99,2% untuk *C4.5* dan 95,41% untuk *naïve bayes*. Selanjutnya Sumiati, *et al.* (2023) melakukan penelitian untuk memberikan solusi alternatif dalam mengatasi keterbatasan tenaga medis dengan mengklasifikasikan gangguan jantung berdasarkan data rekam medis *electrocardiogram* menggunakan algoritma *C4.5 Decision Tree* dengan akurasi 75,33%.

Penelitian lain yang dilakukan oleh Tohir, *et al.* (2023) bertujuan untuk mengetahui perbedaan algoritma *C4.5* dan *random forest* untuk memprediksi proses persalinan ibu hamil menggunakan 1.000 data latih dan 200 data uji dengan akurasi penerapan algoritma *C4.5* dan *random forest* menghasilkan 96% dan 95%. Andesti, *et al.* (2022) melakukan penelitian untuk mengetahui potensi peningkatan kecakapan berbahasa Inggris dengan 7 faktor yang ditetapkan sebagai parameter dengan tingkat akurasi 90% dari 90 responden dengan faktor yang mempengaruhi potensi peningkatan kecakapan berbahasa Inggris adalah *Practice*.

Pada penelitian ini, algoritma *C4.5* diterapkan untuk analisis pengusulan pengadaan peralatan dan mesin kantor karena akurasi algoritma ini seperti yang diuraikan pada penelitian terdahulu memiliki akurasi yang cukup tinggi yaitu di atas 70%. Data dalam penelitian ini bersumber dari data KIB dan *database* aplikasi SIMAK BMN pada Satker Rektorat Universitas Andalas. Penerapan teknik *data mining* diharapkan dapat mempercepat proses pengambilan keputusan.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, permasalahan penelitian ini dirumuskan sebagai berikut :

1. Bagaimana penerapan metode *C4.5 Decision Tree* digunakan untuk analisis pengadaan peralatan dan mesin kantor?
2. Bagaimana mengukur akurasi hasil analisis pengadaan peralatan dan mesin kantor dari metode *C4.5 Decision Tree* ?

1.3 Batasan masalah

Untuk mencapai tujuan penelitian ini, peneliti memfokuskan batasan penelitian sebagai berikut :

1. Data yang digunakan merupakan data dari KIB dan *database* aplikasi SIMAK BMN Satuan Kerja (Satker) Rektorat Universitas Andalas.
2. Atribut yang digunakan yaitu, nama barang, kondisi barang, bahan, tahun perolehan, nilai perolehan dan keterangan pengusulan pengadaan.
3. Metode yang digunakan adalah algoritma *C4.5 Decision Tree*.
4. Aplikasi yang digunakan untuk melakukan pengujian adalah *rapidminer 10.2*.

1.4 Tujuan Penelitian

Berdasarkan batasan masalah yang telah diuraikan, maka dapat ditentukan tujuan dari penelitian sebagai berikut :

1. Menerapkan algoritma *C4.5 Decision Tree* untuk analisis pengadaan peralatan dan mesin kantor.
2. Untuk mengetahui akurasi hasil prediksi pengadaan peralatan dan mesin kantor.

1.5 Manfaat Penelitian

Adapun manfaat yang ingin dicapai oleh penulis dalam penelitian ini adalah sebagai berikut :

1. Memberikan kontribusi bagi pengembangan ilmu pengetahuan dan teknologi khususnya dalam bidang *data mining*.
2. Dapat memberikan informasi analisis pengadaan peralatan dan mesin kantor baik yang diusulkan maupun yang tidak diusulkan.
3. Memberikan inspirasi bagi penelitian selanjutnya yang terkait dengan penggunaan teknologi *data mining* dalam analisis pengadaan peralatan dan mesin kantor.

1.6 Sistematika Penulisan

Adapun sistematika penulisan tugas akhir ini dibagi ke dalam enam bab yang dapat dijelaskan sebagai berikut :

Bab I: Pendahuluan

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

Bab II : Landasan Teori

Pada bab ini dijelaskan teori-teori mengenai teknik untuk menunjang penulisan tesis. Adapun dibahas dalam bab ini adalah teori yang berkaitan tentang *data mining* dan metode algoritma *C4.5 Decision Tree*.

Bab III : Metodologi Penelitian

Bab ini membahas tentang analisis dan penggunaan secara matematis metode *C4.5 Decision Tree*.

Bab IV : Analisis dan Perancangan

Bab ini membahas hasil implementasi metode algoritma *C4.5 Decision Tree* pada data peralatan dan mesin kantor (aset tetap) Rektorat Universitas Andalas dari data KIB dan *database* SIMAK BMN

Bab V : Implementasi dan Hasil

Bab ini membahas hasil penerapan algoritma C4.5 *Decision Tree* pada data peralatan dan mesin kantor (aset tetap) Rektorat Universitas Andalas dari *database* SIMAK BMN.

Bab VI : Kesimpulan dan Saran

Bab ini membuat kesimpulan dan hasil penelitian keunggulan metode algoritma C4.5 *Decision Tree* dalam melakukan analisis pengadaan peralatan dan mesin kantor pada Satker Rektorat Universitas Andalas.