

BAB I

PENDAHULUAN

1.1 Latar Belakang

Data mining adalah proses menemukan pola-pola baru dari sebuah data yang sangat besar sehingga mendapatkan sebuah pengetahuan dan informasi baru. *Data mining* juga disebut sebagai kegiatan pengumpulan dan pemakaian data dalam menentukan pola, keteraturan dan hubungan dalam data set yang sangat besar. (Rafiska R., Defit S. dan Nurcahyo G.W., 2018). *Data mining* juga sering disebut *Knowledge Discovery In Database* (KDD). KDD merupakan proses pencarian pengetahuan yang baru dari sebuah data yang tersimpan dalam sebuah basis data sehingga menemukan pengetahuan yang baru. Penggunaan *Knowledge Discovery In Database* (KDD) membutuhkan sebuah tipe analisis sehingga dapat menemukan hubungan antara atribut masukan dan target, oleh karena itu tipe analisis yang biasa digunakan dalam proses *Knowledge Discovery In Database* (KDD) adalah klasifikasi. Tipe analisis klasifikasi bertujuan untuk meningkatkan kehandalan hasil yang di peroleh dari data. Algoritma C4.5 merupakan salah satu metode klasifikasi yang melibatkan koleksi dan kontruksi pohon keputusan yaitu menggunakan atribut masukan sehingga mendapatkan target yang diinginkan. (Hendrian S., 2018).

Berdasarkan penelitian terdahulu yang dilakukan oleh (Mirqotussa'adah, Muslim M.A., Sugiharti E., Prasetyo B. Dan Alimah S.,2018). Dengan judul "Penerepan *Dizcretization* dan Teknik *Bagging* Untuk Meningkatkan Akurasi Klasifikasi Berbasis *Ensemble* pada Algoritma C4.5 dalam Mendiagnosa Diabetes". Pada penelitian ini Algoritma C4.5 menggunakan *discretization* dan teknik *bagging* menunjukkan peningkatan 6,26%. Dengan akurasi awal 68,61%, setelah diterapkan *discretization* dan teknik *bagging* menjadi 74,87%. Sehingga penerapan *discretization* dan teknik *bagging* dapat meningkatkan akurasi algoritma C4.5 pada klasifikasi *dataset* diabetes.

Penelitian serupa juga dilakukan oleh (Darmawan E., 2018), dengan judul “*C4.5 Algorithm Application for Prediction of Self Candidate New Students in Higher Education*” pada penelitian ini algoritma C4.5 bisadi implementasi berdasarkan data siswa untuk mengetahui jumlah kandidat siswa baru yang akan melakukan registrasi dan tidak melakukan registrasi dan aplikasi algoritma C 4.5 juga bisa mengklasifikasi data dari siswa baru sehingga bisa menghasilkan prediksi siapa kandidat siswa yang mengundurkan diri menjadi siswa baru pada perguruan tinggi .

Penelitian sejenis juga dilakukan oleh (Rianto R. dan Iswari N.M.S., 2018), dengan judul “Rancang Bangun Aplikasi Pendeteksi Penyakit Ginjal Kronis dengan Menggunakan Algoritma C4.5”. Pada penelitian ini algoritma C4.5 telah berhasil dibuat, dengan bantuan dari weka library dalam pembuatan algoritma decision tree yang dihasilkan dapat melakukan prediksi terhadap penyakit ginjal kronis. Proses pengujian dilakukan dengan menggunakan cross-validation dan berdasarkan hasil yang sudah dihitung aplikasi weka memiliki akurasi 91.50% pada saat decision tree dibuat tanpa menggunakan preprocess menu.

Penelitian sejenis juga dilakukan oleh (Hermanto B. dan Azhari SN., 2017), dengan judul “Klasifikasi Nilai Kelayakan Calon Debitur Baru Menggunakan Decision Tree C4.5”. Pada penelitian ini algoritma C4.5 untuk kasus data latihan nasabah debitur pembiayaan kendaraan sepeda motor memberikan hasil akurasi dan performa yang cukup baik. Terbukti sistem telah mampu menghasilkan keputusan bahwa data valid di atas 70% untuk semua skenario kasus.

Ceria Swalayan merupakan sebuah Toko Retail yang menjual berbagai produk kebutuhan Sehari-hari. Ceria Swalayan harus teliti dalam menggunakan modal penjualan yang dikeluarkan setiap hari sehingga mendapat keuntungan. Oleh karena itu pihak manajemen harus mempertimbangkan hal yang pembeli minati terhadap produk yang dijual dan hanya membelanjakan uang modal sesuai dengan kebutuhan penjualan. Produk yang diminati oleh pembeli sangatlah sulit diketahui mengingat banyaknya produk dijual serta banyaknya pembeli yang membeli. Pengelola Ceria Swalayan hanya mengandalkan penyimpanan data perbelanjaan untuk mengetahui produk yang banyak terjual namun manajemen Ceria Swalayan tidak mengetahui secara pasti produk paling diminati setiap harinya.

Berdasarkan Permasalahan diatas dan beberapa contoh penerapan data

mining, maka peneliti mengusulkan menggunakan *data mining* yang dapat digunakan untuk menggali data-data pembelanjaan sehinggadapat mengetahui minat pembeli dalam pembelian produk.Namun untuk menggunakan *data mining* dibutuhkan sebuah metode yang tepat untuk dapat memberikan hasil yang tepat. Oleh karena itu peneliti menggunakan salah satu metode data mining yaitu algoritma C4.5. Sehingga dengan adanya teknik *data mining* menggunakan algoritma C4.5 maka pihak manajemen Ceria Swalayan dapat mengetahui dengan mudah minat pembeli terhadap produk yang dijual sehingga pihak manajemen Ceria Swalayan bisa menyediakan produk sesuai dengan kebutuhan nya. Berdasarkan latar belakang yang telah diuraikan maka peneliti menyimpulkan judul **“Penentuan Minat Pembeli Terhadap Produk Menggunakan Algoritma C4.5 Pada Ceria Swalayan”**.

1.2 Perumusan Masalah

Berdasarkan uraian latar belakang di atas, maka dapat dirumuskan masalah sebagai berikut:

1. Bagaimana menentukan minat pembeli terhadap produk yang dijual oleh Ceria Swalayan?
2. Bagaimana menerapkan algoritma C4.5 dalam menentukan minat pembeli terhadap produk yang dijual oleh Ceria Swalayan?
3. Bagaimana menentukan atribut yang akan digunakan dalam menentukan minat pembeli terhadap produk yang dijual Ceria Swalayan menggunakan algoritma C4.5 ?

1.3 Batasan Masalah

Agar pembahasan tidak melebar dan tidak melebihi dari yang disajikan sesuai dengan pembahasan maka peneliti membatasi masalah penelitian sebagai berikut :

1. Penelitian dilakukan di Ceria Swalayan, Kota Pasir Pengaraian, Kabupaten Rokan Hulu, Provinsi Riau.
2. Metode data mining yang digunakan adalah Klasifikasi Algoritma C4.5.

3. Data yang digunakan berdasarkan data penjualan detail produk di Ceria Swalayan.
4. Atribut yang digunakan yaitu Jumlah, Harga, Satuan

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut :

1. Menentukan minat pembeli terhadap produk yang dijual oleh Ceria Swalayan.
2. Menerapkan algoritma C4.5 untuk menentukan minat pembeli terhadap produk yang dijual oleh Ceria Swalayan.
3. Menentukan atribut yang akan digunakan dalam menentukan minat pembeli terhadap produk yang dijual Ceria Swalayan menggunakan algoritma C4.5.

1.5 Manfaat Penelitian

Manfaat penelitian ini adalah sebagai berikut :

1. Ceria Swalayan dapat dengan mudah menentukan minat pembeli terhadap produk yang dijual sehingga Ceria Swalayan dapat menyediakan produk yang sesuai dengan kebutuhan.
2. Mengetahui dan memahami penggunaan Algoritma C4.5 dalam menentukan minat pembeli terhadap produk yang dijual oleh Ceria Swalayan.
3. Mengetahui dan memahami penentuan atribut yang akan digunakan dalam menentukan minat pembeli terhadap produk yang dijual Ceria Swalayan menggunakan algoritma C4.5.

1.6 Sistematika Penulisan

Untuk mempermudah penulisan penelitian ini, sistematika disesuaikan dengan template yang diatur dalam tata penulisan program studi masing-masing sebagai berikut :

BAB I : PENDAHULUAN

Dalam bab ini dibahas mengenai latar belakang masalah, perumusan masalah, batasan masalah, tujuan dan manfaat dan sistematika penelitian.

BAB II : LANDASAN TEORI

Bab ini membahas tentang teori *data mining* dan Algoritma C4.5 yang berhubungan dengan permasalahan yang di angkat pada penelitian ini.

BAB III : METODOLOGI PENELITIAN

Bab ini membahas tentang analisa dan penggunaan secara matematis Algoritma C4.5 dalam menyelesaikan permasalahan yang diangkat dalam penelitian ini.

BAB IV : ANALISIS DAN PERANCANGAN

Bab ini membahas tentang analisis dan perancangan dalam menentukan minat pembeli terhadap produk yang dijual oleh Ceria Swalayan

BAB V : IMPLEMENTASI DAN HASIL

Bab ini membahas hasil implementasi Algoritma C4.5 dalam menentukan minat pembeli terhadap produk yang dijual oleh Ceria Swalayan

BAB VI : PENUTUP

Bab ini akan disajikan hasil kesimpulan dan saran dari peneliti dari hasil penelitian yang telah dibuat sehingga dapat dikembangkan oleh peneliti berikutnya dan menjadi lebih baik lagi.