

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Perkembangan, efektifitas dan kemajuan teknologi yang pesat mempengaruhi dan menghasilkan manfaat luar biasa di berbagai aspek kehidupan, penggunaan teknologi sangat membantu pada instansi pendidikan, pemerintahan maupun swasta di dalam memenuhi kebutuhan dan menyelesaikan suatu pekerjaan. Hal ini menyebabkan perubahan di beberapa aspek kehidupan, termasuk dalam dunia usaha. Salah satu yang bisa dimanfaatkan adalah data mining, yaitu alat atau proses yang mengekstraksi informasi penting dari big data dengan merubahnya menjadi pola informasi yang menarik. Informasi dapat ditemukan dalam database menggunakan teknik tertentu. K-Means dan C4.5 saat ini merupakan algoritma pengelompokan dan klasifikasi yang banyak digunakan. Algoritma K-Means digunakan untuk iterative clustering dengan cara membagi dataset menjadi beberapa kategori yang telah ditentukan. Algoritma C4.5 dimanfaatkan untuk membangun sebuah *decision tree* (pohon keputusan) yang dapat digunakan untuk memprediksi keputusan dengan menggunakan beberapa aturan keputusan.

CV.Togu Togu Cabang Medan merupakan perusahaan yang bergerak dan memulai usaha di bidang *E-Commerce* / penjualannya di bidang online. Keberadaan perusahaan ini kompatibel dengan situasi di Indonesia sekarang ini sedang membangun. Dalam menentukan proses pendataan penjualan terdapat kendala-kendala antara lain efisiensi waktu, banyaknya perbandingan variabel yang di uji, mengklarifikasi data penjualan membutuhkan perhitungan yang valid sehingga dibutuhkan perbandingan metode untuk mendapatkan sebuah informasi yang benar-benar terjamin keasliannya. Berdasarkan hal tersebut diharapkan dapat mempermudah pihak CV.Togu Togu on Cabang Medan dalam mendata penjualan dari tumpukan-tumpukan data yang telah ada.

Berdasarkan pernyataan tersebut maka penulis tertarik untuk membuat sistem untuk mempercepat proses *Clustering* data penjualan maka dari itu diambil judul “**PERBANDINGAN K-MEANS DAN C4.5 TERHADAP DATA PENJUALAN PADA CV.TOGU – TOGU ON CABANG MEDAN**”.

1.2. Rumusan Masalah

Dari latar belakang yang dijelaskan masalah dapat dirumuskan sebagai:

1. Data penjualan belum dimanfaatkan secara maksimal.

2. Ditemukan kemungkinan beberapa item barang yang dijual mengalami kelebihan dan kekurangan stok.
3. Perusahaan belum menemukan cara yang efektif dan efisien untuk dipraktikkan.
4. menerapkan teknik data mining klastering agar dapat membantu perusahaan CV.Togu Togu on Cabang Medan dalam mengklarifikasikan data penjualan dengan menggunakan perbandingan algoritma K-Means dan C4.5.

1.3. Batasan Masalah

Batasan masalah dalam penelitian ini adalah:

1. Membahas Pengklarifikasian data penjualan di CV.Togu Togu on Cabang Medan
2. Bandingkan algoritma K-Means dan C4.5 dengan menggunakan data riwayat penjualan yang diambil sebagai data testing per bulan Juni 2021 – Juli 2022 dari CV.Togu Togu on Cabang Medan.

1.4. Tujuan dan Manfaat Masalah

1.4.1. Tujuan Penelitian

Tujuan penelitian dari penulisan penelitian ini adalah sebagai berikut :

1. Mengetahui penerapan K-Means dan Algoritma C4.5 dalam mengoptimalkan kasus kelebihan dan kekurangan stok item yang akan dijual.
2. Untuk meningkatkan akurasi algoritma C4.5 dengan menerapkan K-Means di pengklarifikasian data penjualan di CV.Togu Togu on Cabang Medan

1.4.2. Manfaat Penelitian

Adapun manfaatnya antara lain:

1. Mengetahui hasil peningkatan akurasi dari pengolahan data penjualan sebelum dan sesudah diterapkan K-Means dan Algoritma C4.5.
2. **Menjaga stok barang agar tidak yang kosong, karena hal tersebut adalah salah satu cara agar pelanggan tetap senang.**

1.5. Keterbaruan

Penelitian (Fani Mulyana Nasution, 2019) menerapkan algoritma *K-Means* untuk mengelompokkan tingkat ketahanan tanaman pangan dengan tujuan untuk meningkatkan produksi hasil dari tanaman pangan di masyarakat lokal dan perkotaan di Provinsi Sumatera Utara, teknik ini digunakan untuk menentukan daerah dengan potensi ketahanan pangan untuk membantu kebutuhan pangan dengan memanfaatkan metode *K-Means* ditinjau berdasarkan luas panen, produksi, dan luas tanam, penggunaan metode *K-Means* ditujukan untuk penggolongan daerah dengan hasil tanaman pangan tinggi, sedang, dan rendah. Penelitian ini menghasilkan pengelompokan dari beberapa jenis tanaman pangan dengan jumlah kluster sebanyak 3 cluster dimana cluster pertama kelompok dengan potensi ketahanan pangan yang tinggi, cluster kedua kelompok dengan potensi ketahanan pangan yang sedang, dan cluster ketiga dengan potensi ketahanan yang rendah[12].

Penelitian (Prasetyo & Santosa, 2019) menjelaskan keakuratan prediksi algoritma C4.5 dan k-Means untuk memprediksi IPK mahasiswa FTI UKDW semester pertama. Data yang dimanfaatkan ialah data mahasiswa FTI UKDW tahun 2008 hingga 2016, sebagai data pelatihan dan tahun 2017 sebagai data ujian. Atribut yang digunakan diantaranya adalah kategori, negara bagian, lokasi, dan tingkat ICE, sedangkan jalur yang tidak melewati menggunakan kategori, status, lokasi, tingkat ICE, angka, bahasa, ruang, dan analogi. Akurasi akan dihitung dengan tabel silang. Algoritma C4.5 memperoleh hasil yaitu 77,45% dan algoritma k-Means mencapai hasil terbaik sebesar 60,78%.. Tingkat akurasi rata-rata adegan dengan jalur pencapaian adalah 55,27%, dan tingkat akurasi rata-rata adegan tanpa jalur pencapaian adalah 38,95%[13].

Penelitian (Salsabila & Intani, 2021) mengimplementasikan metode K-Mean dan algoritma C4.5 dalam tujuan menentukan tingkat penularan Covid-19 di Indonesia. Tujuan dari penelitian ini adalah untuk mengurutkan tingkat penularan kasus Covid-19 menurut provinsi di Indonesia. Pengelompokan menggunakan algoritma clustering K-Means dan algoritma klasifikasi C4.5 secara kombinasi. Algoritma KMeans digunakan untuk mengelompokkan data berdasarkan provinsi ke dalam cluster regional di Indonesia. Hasil pengelompokan menggunakan algoritma C4.5 untuk melihat aturan dalam bentuk pohon keputusan[14].

Penelitian (Girsang, at all, 2022) Penerapan Algoritma C4.5 untuk menentukan penerima manfaat program subsidi perkotaan Hal ini dikarenakan sistem yang dimanfaatkan masihlah secara manual, sehingga sering terjadi kesalahan saat memasukkan informasi calon penerima manfaat. Ini memengaruhi keakuratan permintaan bantuan. Oleh karena itu diperlukan suatu sistem yang dapat mengatasi kesalahan kinerja dan meminimalisir kesalahan pada kantor perangkat desa dengan menggunakan data mining. Dari hasil penelitian, dibuatlah sebuah aplikasi yang diharapkan dapat memprediksi penerimaan utilitas secara akurat dan terpercaya, yang kemudian digunakan untuk mendukung kantor desa. menyelesaikan masalah terkait persetujuan program pengabdian kepada masyarakat[15].