

BAB I

PENDAHULUAN

1.1 Latar Belakang

CV. Cipta Usaha Selaras merupakan perusahaan yang berlokasi di Kota Bekasi dan bergerak di bidang produksi *jumbo bag* untuk kebutuhan industri bahan kimia, pasir silika, batu bara, dan *wood pellet*. Perusahaan ini mengalami fluktuasi penjualan yang cukup signifikan setiap tahunnya, dengan peningkatan permintaan menjelang akhir tahun dan penurunan drastis di awal tahun, sehingga menyulitkan perencanaan produksi. Selain itu, karakteristik perilaku pelanggan yang bervariasi, mulai dari pelanggan yang bertransaksi rutin hingga pelanggan tidak tetap, menyebabkan perusahaan kesulitan dalam memahami pola pembelian secara menyeluruh. Hal ini berdampak pada ketidaktepatan dalam penentuan strategi pemasaran dan pengelolaan permintaan pasar.

Untuk mengatasi tantangan dalam memahami pola pembelian pelanggan dan ketidakpastian permintaan, berbagai metode analisis telah digunakan. Algoritma *K-Means* sering diterapkan untuk mengelompokkan pelanggan berdasarkan kebiasaan belanja mereka[1], sementara metode *Simple Linear Regression* digunakan untuk meramalkan tren penjualan di masa mendatang[2]. Penelitian terdahulu menunjukkan bahwa teknik data mining efektif dalam menganalisis pola pembelian dan penjualan. Kombinasi *K-Means* dan *Linear Regression* dalam analisis *bandwidth* meningkatkan akurasi prediksi tren data[3]. Teknik klasterisasi dan prediksi berbasis data membantu perusahaan mengelola bisnis lebih efisien.

Selain *K-Means*, metode seperti *Decision Tree* dan *Naïve Bayes* juga digunakan dalam analisis pola pelanggan. Namun, *Decision Tree* rentan terhadap *overfitting* saat data terlalu kompleks, sehingga hasilnya kurang akurat [4]. *Naïve Bayes* mengasumsikan variabel independen, padahal dalam bisnis sering kali variabel saling berhubungan. Oleh karena itu, *K-Means* lebih unggul karena mampu mengelompokkan pelanggan tanpa asumsi ketat terkait pola pembelian [5].

Meskipun *K-Means* efektif untuk mengelompokkan pelanggan, penentuan jumlah klaster yang kurang tepat dapat menghasilkan segmentasi yang tidak akurat, membuat strategi pemasaran kurang efektif. Selain itu, pola pembelian yang bervariasi di CV. Cipta Usaha Selaras dapat menyebabkan hasil klasterisasi kurang stabil. Sementara itu, *Simple Linear Regression* hanya mampu melihat hubungan sederhana dalam data, sehingga kurang akurat dalam memprediksi lonjakan atau penurunan permintaan akibat faktor eksternal seperti kebijakan industri atau perubahan ekonomi. Untuk mengatasi hal ini, penelitian ini akan menggunakan *Silhouette Score* dalam menentukan jumlah klaster optimal. Selain itu, *preprocessing* data akan dilakukan untuk mengurangi pengaruh outlier yang dapat mengganggu hasil analisis. Untuk meningkatkan akurasi prediksi tren penjualan, penelitian ini akan menerapkan teknik validasi model agar hasilnya lebih tepat, sehingga pola musiman dapat terdeteksi dengan lebih baik.

Penelitian ini bertujuan untuk mengelompokkan pelanggan berdasarkan pola pembelian menggunakan algoritma *K-Means*, serta memprediksi total nilai pembelian pelanggan berdasarkan frekuensi pembeliannya dengan Regresi Linear Sederhana. Diharapkan metode ini membantu CV. Cipta Usaha Selaras meningkatkan efisiensi operasional dan menyusun strategi pemasaran yang lebih

efektif. Dengan demikian, penelitian ini berjudul **“Klasterisasi dan Prediksi Pola Penjualan di CV. Cipta Usaha Selaras dengan Algoritma K-Means dan Regresi Linear Sederhana”**.

1.2 Identifikasi Masalah

Berdasarkan latar belakang penelitian yang telah di uraikan, terdapat beberapa identifikasi masalah dalam penelitian ini sebagai berikut:

1. CV. Cipta Usaha Selaras masih menggunakan *paperbased* dalam pencatatan, dan belum adanya pengelompokan data pelanggan berdasarkan pola pembelian.
2. Belum diterapkannya metode prediktif yang tepat untuk memperkirakan total pembelian pelanggan, sehingga menyulitkan perusahaan dalam merumuskan strategi pemasaran dan operasional secara efektif dan berbasis data.

1.3 Rumusan Masalah

Berdasarkan identifikasi masalah dan latar belakang penelitian, terdapat beberapa rumusan masalah dalam penelitian ini sebagai berikut:

1. Bagaimana penerapan algoritma *K-Means* untuk mengelompokkan pelanggan berdasarkan pola pembelian mereka di CV. Cipta Usaha Selaras?
2. Bagaimana penerapan Regresi Linear Sederhana untuk memprediksi total pembelian pelanggan berdasarkan frekuensi pembelian di CV. Cipta Usaha Selaras?

1.4 Tujuan Penelitian

Tujuan yang terdapat pada penelitian ini sebagai berikut:

1. Menerapkan algoritma *K-Means* dalam sistem pengolahan data dapat memberikan kemudahan dalam pengelompokan data pelanggan perusahaan berdasarkan pola pembelian secara digital.
2. Penelitian ini bertujuan untuk membangun model prediksi total pembelian pelanggan berdasarkan frekuensi pembelian untuk mendukung perencanaan operasional dan strategi pemasaran yang lebih efektif, menggunakan metode Regresi Linear Sederhana.

1.5 Manfaat Penelitian

Beberapa manfaat yang didapat dari penelitian ini adalah:

1. Membantu perusahaan memahami pola pembelian pelanggan untuk menyusun strategi pemasaran yang lebih tepat.
2. Memberikan prediksi nilai total pembelian pelanggan berdasarkan frekuensi transaksi, sehingga dapat digunakan sebagai dasar dalam perencanaan produksi dan pengelolaan permintaan.
3. Meningkatkan efisiensi operasional melalui sistem pencatatan dan analisis data yang lebih terstruktur.

1.6 Batasan Masalah

Agar penelitian ini jelas dan terarah maka diberikan batasan terhadap permasalahan yang akan diteliti, yang berfokus pada:

1. Data yang digunakan merupakan data penjualan tahun 2024 dari CV. Cipta Usaha Selaras yang mencakup kolom: Customer, Nama Barang, Qty PO, Harga/pcs, Total Harga, dan Tanggal Transaksi, dengan jumlah 358 data.

2. Data sekunder yang digunakan diperoleh dari arsip internal CV. Cipta Usaha Selaras berupa laporan penjualan, dengan periode pengambilan data dilakukan selama satu bulan, yaitu pada Maret hingga April 2024.

1.7 Sistematika Penelitian

Sistematika penulisan menjelaskan susunan isi skripsi secara terstruktur untuk memberikan gambaran menyeluruh mengenai isi penelitian. Adapun sistematika penulisan dalam skripsi ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi tentang latar belakang, identifikasi masalah, tujuan, penelitian, manfaat penelitian, rumusan masalah, batasan masalah.

BAB II TINJAUAN PUSTAKA

Berisi tentang berbagai teori yang berkaitan dengan permasalahan yang akan dibahas seperti *Machine Learning*, Data Mining, Klusterisasi, Algoritma *K-Means*, *Silhouette Score*, Prediksi Penjualan, Regresi Linear Sederhana, Evaluasi Model Prediksi, *CRISP-DM*.

BAB III METODOLOGI PENELITIAN

Merupakan bab yang menjelaskan tentang lokasi penelitian, metode, pengumpulan data, metode analisis data, dan metode teknik lain yang digunakan.

BAB IV HASIL DAN PEMBAHASAN

Menyajikan proses pemodelan, analisis hasil klusterisasi dan prediksi, serta evaluasi model.

BAB V PENUTUP

Menyimpulkan hasil penelitian dan memberikan saran untuk pengembangan lebih lanjut.

