

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Indonesia Digital Printing merupakan salah satu perusahaan yang bergerak pada bidang Merchandise dan Souvenir. Data transaksi penjualan yang ada pada Indonesia Digital Printing semakin lama semakin banyak jumlahnya, hal tersebut sangat disayangkan jika tidak dimanfaatkan dengan baik karena hanya akan menjadi kuburan data alangkah lebih baik jika data tersebut dapat dimanfaatkan untuk mencari informasi yang bermanfaat untuk strategi penjualan sehingga dapat bersaing dengan competitor. Pemahaman yang baik terhadap pelanggan dapat digunakan perusahaan untuk berinvestasi mengelompokkan pelanggan dalam beberapa cluster.

Hubungan antara Perusahaan dengan pelanggan sangat penting untuk menunjang perkembangan dan kelangsungan perusahaan. Pihak manajemen perusahaan harus mampu untuk mengenali beberapa tipe pelanggan dan mempercayainya dengan meningkatkan pemahaman perusahaan akan kebutuhan mereka sebagai individu sehingga dapat mempertahankan loyalitasnya terhadap perusahaan.

Melihat semakin banyaknya pelanggan pada perusahaan tersebut, maka terdapat data-data pelanggan yang harus diolah agar dapat meningkatkan kualitas perusahaan. Salah satu hal yang penting dalam pengelolaan pelanggan adalah

bagaimana suatu perusahaan dapat mengelola dan mengelompokkan beberapa data agar mudah untuk pengambilan keputusan. Sehingga perusahaan dapat lebih mudah untuk memilih pelanggan yang akan diberikan hadiah yang berupa cinderamata dan dilakukan secara kontinu di setiap tahunnya.

K-Means Clustering adalah suatu metode penganalisaan data atau metode Data Mining yang melakukan proses pemodelan tanpa supervisi (unsupervised) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Terdapat dua jenis data clustering yang sering dipergunakan dalam proses pengelompokan data yaitu Hierarchical dan Non-Hierarchical.

K-Means merupakan salah satu metode data clustering non-hierarchical atau Partitional Clustering. Sehingga dengan menggunakan metode K-Means Clustering dapat mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain.

Berdasarkan latar belakang tersebut penelitian ini dilakukan untuk mengimplementasikan data mining menggunakan K-Means Clustering dalam pengelompokan pelanggan berdasarkan jumlah transaksi masing-masing cluster untuk pengambilan keputusan. Data dikumpulkan melalui observasi dan interview yang dilakukan kepada pemilik perusahaan yang bersangkutan. Dimulai dari mengumpulkan data transaksi pelanggan, selanjutnya data preprocessing dengan

memilih data yang dibutuhkan, kemudian menentukan nilai dari pelanggan dengan menggunakan parameter Recency, Frequency dan Monetary (RFM). Setelah nilai-nilai RFM didapatkan berdasarkan data transaksi yang sudah diolah. Hasil RFM tersebut diolah kembali untuk menentukan Cluster menggunakan aplikasi Python sebagai penerapan K-Means Clustering.

1.2 Rumusan Masalah

Rumusan masalah yang akan dibahas pada penelitian ini adalah Bagaimana cara mengelompokkan data pelanggan dengan menentukan nilai RFM berdasarkan data hasil pre proses sebelumnya, ada 3 atribut yang akan digunakan dalam penentuan nilai RFM nya, yaitu; tanggal terakhir pelanggan transaksi, frekuensi transaksi pelanggan, dan total belanja pelanggan. Setelah mendapat nilai-nilai dari setiap atribut RFM, tahap berikutnya adalah melakukan clustering menggunakan algoritma K-Means ke dalam k kelompok.

1.3 Ruang Lingkup

Berdasarkan rumusan masalah di atas, dapat dijabarkan ruang lingkup penelitian adalah sebagai berikut :

- a. Sumber data utama yang digunakan dalam penelitian ini adalah dataset dari perusahaan Indonesia Digital Printing dari bulan Maret-Mei 2020, karena pada data untuk tahun sebelumnya tidak lengkap datanya.
- b. Menggunakan analisis penyelidikan data untuk mengenali data lebih lanjut dan pencarian pengetahuan awal. Pengolahan data untuk menentukan nilai

pelanggan menggunakan atribut sesuai dengan yang dibutuhkan oleh model

RFM, terdiri dari:

- Waktu transaksi terakhir (Recency)
- Jumlah frekuensi transaksi (Frequency)
- Total Belanja Pelanggan (Monetary)

c. Data mentah yang diperoleh dari data pelanggan pada perusahaan Indonesia

Digital Printing akan melalui praproses data yang meliputi :

- Proses pembersihan data untuk menghilangkan baris data dengan nilai kosong atau data yang tidak valid (data cleaning and integration)
- Pemilihan data berdasarkan atribut yang disesuaikan dengan model RFM yaitu: atribut waktu transaksi, jumlah frekuensi transaksi dan jumlah nominal transaksi untuk setiap pelanggan (data selection)
- Persiapan data dengan cara mereduksi kolom data yang tidak sesuai dengan kebutuhan dalam melakukan proses klasterisasi (data preprocessing)
- Transformasi data kedalam bentuk yang terukur sehingga dapat digunakan sebagai atribut untuk proses klasterisasi (transformation)

d. Aplikasi yang dibangun menggunakan bahasa pemrograman Python.

e. Hasil penelitian berupa pengelompokan data pelanggan yang menghasilkan 4 cluster.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini yaitu perusahaan dapat melakukan pengelompokan data pelanggan secara otomatis dengan hasil pengelompokan yang diharapkan sesuai dengan karakteristik masing-masing data.

1.5 Tujuan Penelitian

Perusahaan Indonesia Digital Printing memiliki data transaksi penjualan semakin lama semakin banyak sehingga akan terjadi penumpukan, penumpukan data tersebut dapat dimanfaatkan untuk mencari informasi yang bermanfaat dan tersembunyi sehingga hal tersebut dapat digunakan untuk melakukan strategi penjualan. Pengolahan penumpukan data yang demikian dapat dilakukan dengan penerapan data mining. Penelitian ini bertujuan untuk melakukan pengelompokan data dengan menggunakan teknik clustering menggunakan algoritma K-Means. Sehingga akan diketahui beberapa cluster dengan karakteristik yang berbeda.

BAB I PENDAHULUAN

Bab ini berisikan secara umum mengenai latar belakang, rumusan masalah, ruang lingkup, manfaat penelitian dan tujuan penelitian.

BAB II LANDASAN TEORI

Bab ini berisikan mengenai konsep dan dasar-dasar teori yang digunakan dalam data mining pengelompokkan pelanggan menggunakan RFM dan K-Means Clustering, serta teori-teori yang berhubungan dengan bahasa pemrograman yang akan digunakan dalam skripsi ini. Sebagai referensi dalam pembuatan sistem.

BAB III METODE PENELITIAN

Bab ini berisikan analisis seputar permasalahan yang mencakup berbagai hal penelitian ini. Kemudian dilanjutkan dengan perancangan sistem yang mencakup UML, perancangan interface dan perancangan lainnya yang berkaitan dengan sistem.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi mengenai hasil analisis penelitian dan pembahasan mengenai penerapan metode yang digunakan, hasil penelitian yang dicapai dari proses penyeleksian data sampai pengolahan data serta implementasi program menggunakan bahasa pemrograman Python.

BAB V PENUTUP

Bab ini berisi kesimpulan dari hasil semua tahap yang telah dilalui selama penulisan skripsi ini, serta saran-saran yang berkaitan dengan pembuatan skripsi.

DAFTAR PUSTAKA

Bagian ini berisikan daftar pustaka yang menjadi sumber referensi dalam penelitian dan penyusunan naskah skripsi.