

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Perkembangan internet telah mengubah cara individu berbelanja, berkomunikasi, menyebarkan informasi maupun melakukan transaksi bisnis. Perubahan perilaku transaksi masyarakat dengan menggunakan internet mengakibatkan pertumbuhan e-commerce yang eksponensial. Selain itu, e-commerce tidak hanya untuk membeli produk secara online, tetapi juga untuk membandingkan harga, fitur, dan dapat juga melakukan pembayaran listrik, air dan lainnya [1]. Dengan pertumbuhan transaksi belanja digital masyarakat saat ini, menjadi sulit untuk memprediksi segmen dan pola pelanggan menggunakan model matematika tradisional. Identifikasi tepat waktu dari tren yang baru muncul dari volume data yang besar memainkan peran utama dalam proses bisnis dan pengambilan keputusan. Volume data yang sangat besar ada, tetapi perusahaan haus akan pengetahuan [2].

Segmentasi pelanggan (SP) bertujuan untuk menargetkan pelanggan yang berharga dan mengembangkan kegiatan pemasaran dengan membagi pelanggan menjadi kelompok-kelompok homogen yang lebih kecil sehingga strategi pemasaran dapat menargetkan setiap kelompok secara individual [3]. Segmentasi pelanggan dengan algoritma clustering adalah salah satu algoritma yang banyak digunakan. Teknik clustering meminimalkan jarak intra cluster dan memaksimalkan jarak antar cluster untuk segmentasi data [4]. Selain itu analisis perilaku pelanggan [5]–[7] model RFM (*Recency Frequency Monetary*) merupakan salah satu solusi yang perlu dipertimbangkan dalam analisis pemasaran.

Penerapan algoritma clustering banyak diusulkan dalam banyak karya seperti, [8] menerapkan algoritma clustering untuk klasifikasi otomatis pelanggan dapat mengidentifikasi beberapa tipe pelanggan, dan dinamika perilaku mereka di dalam toko. Selanjutnya [9] menerapkan K-Means++ untuk Identifikasi Nilai Pengguna Berdasarkan Model RFM yang Ditingkatkan dan [10] untuk menganalisis penggunaan sistem CRM (Customer Relationship Management) pada usaha kecil dan menengah (UKM). Penerapan K-Means Clustering pada model RFM dapat secara akurat mengekstrak karakteristik

perilaku pembelian dari setiap jenis pelanggan dan membuat strategi pemasaran yang akurat [5].

Baru-baru ini pendekatan algoritma clustering dengan Association Rules Mining dalam pemasaran, seperti [11] menggunakan model penetapan item berdasarkan asosiasi antar item digunakan untuk menempatkan item di lingkungan retail atau e-commerce untuk meningkatkan penjualan, kemudian [4] mengusulkan model pemasaran silang berdasarkan algoritma penambangan pola sekuensial yang ditingkatkan. Dalam karya [12] mengintegrasikan Association Rules Mining untuk membangun model prediksi hasil pertandingan bola basket dengan tujuan memberikan informasi yang berguna untuk operasi di pasar olahraga.

Terinspirasi dari karya sebelumnya maka dalam penelitian ini kami mengusulkan algoritma clustering yang berbeda (K-Means Clustering, Spectral Clustering, dan Gaussian Mixture Model) pada model RFM untuk memecah profil pelanggan besar menjadi segmen yang jauh lebih kecil dan untuk mendapatkan wawasan tentang perilaku terperinci pelanggan kemudian Association Rules Mining dengan algoritma Apriori digunakan untuk menemukan set item yang sering dibeli bersama menggunakan analisis aturan asosiasi yang diharapkan berguna sebagai sistem rekomendasi.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah apakah algoritma clustering dapat diintegrasikan dengan association rules mining untuk menghasilkan informasi segmentasi pelanggan yang akurat?

1.3 Batasan Masalah

Batasan masalah penelitian ini adalah sebagai berikut:

1. Menggunakan dataset terbuka Online Retail UCI Machine Learning yang dapat di download pada <https://archive.ics.uci.edu/ml/datasets/online+retail>
2. Menggunakan algoritma clustering K-Means Clustering, Spectral Clustering, dan Gaussian Mixture Model.
3. Menggunakan metode Association Rules Mining dengan algoritma Apriori

4. Aplikasi di bangun pada lingkungan kerja Google Colab dengan menggunakan Python 3 menggunakan Jupyter Notebook, library Tensorflow3 dan Keras4 Scikit-image, Sklearn, Pandas dan Numpy dan pengaturan GPU runtime

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk meneliti cara kerja data mining khususnya algoritma clustering dan apriori untuk analisis pemasaran berdasarkan segmentasi pelanggan yang diharapkan dapat berguna sebagai informasi dalam pengambilan kebijakan perusahaan.

1.5 Manfaat Penelitian

Beberapa manfaat yang diharapkan pada penelitian ini adalah sebagai berikut:

- a. Memahami kinerja algoritma clustering yang berbeda untuk segmentasi pelanggan
- b. Memahami cara kerja teknik data mining yang berbeda dan penerapannya
- c. Memberikan bahan referensi bagi perusahaan maupun peneliti lainnya yang tertarik di bidang sistem transportasi cerdas dan visi komputer dengan teknik pembelajaran mendalam

1.6 Keterbaruan

Dalam beberapa tahun terakhir ini penerapan algoritma clustering seperti K-Means, dbscan, K-Means ++, Hierarchical clustering dan lainnya telah terbukti akurat di terapkan untuk segmentasi pelanggan dengan tujuan yang berbeda, namun beberapa karya terbaru melaporkan penerapan kombinasi clustering dengan penambangan aturan asosiasi telah menghasilkan informasi yang berguna pada strategi pemasaran, seperti [13] mengusulkan algoritma baru dan efisien untuk menemukan pengetahuan dasar dalam bentuk aturan asosiasi temporal (TARs) dalam data pembuatan besi Blast furnace, kemudian [11] menerapkan Association Rules Mining untuk mengidentifikasi hubungan yang berguna antara kumpulan data besar dan penambangan aturan asosiasi dapat memberikan wawasan baru kepada pembuat keputusan, secara khusus model yang diusulkan digunakan untuk menempatkan item di lingkungan retail atau e-commerce untuk meningkatkan penjualan

berdasarkan algoritma genetika hibrida dan [14] mengintegrasikan pengelompokan dan penambangan aturan asosiasi untuk menganalisis dan membuat profil pelanggan dalam layanan video on demand (VoD).