

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Segmentasi pelanggan bertujuan untuk mengelompokkan pelanggan berdasarkan kesamaan karakteristik. Pengelompokannya didasarkan atas ketertarikan atau permintaan pelanggan yang serupa. Dengan dilakukannya segmentasi terhadap pelanggan, perusahaan dapat memasarkan produknya secara terpersonalisasi sehingga penawaran yang disampaikan sesuai dengan kebutuhan. Dengan teknik segmentasi ini memungkinkan hubungan antara pelanggan dengan perusahaan akan terjalin dengan baik serta dapat mengurangi kemungkinan pelanggan beralih ke kompetitor lain.

Untuk melakukan segmentasi pelanggan dapat dicapai melalui metode *data mining* dengan mengimplementasikan teknik *clustering*. Dalam dunia *clustering*, algoritma *K-Means* sangat populer digunakan dikarenakan kecepatan dan kemudahan dalam pengimplementasiannya. Algoritma *K-Means* sendiri merupakan bagian dari *hard clustering* dimana proses menentukan suatu data masuk ke dalam suatu *cluster* bersifat tegas. Selain *K-Means* algoritma yang belakangan muncul dan mulai sering banyak digunakan ialah *Fuzzy C-Means*, algoritma ini merupakan perluasan dari algoritma *Hard C-Means (K-Means)* dan dikategorikan sebagai metode *soft clustering* dimana penentuan keanggotaan suatu data terhadap suatu *cluster* ditentukan menggunakan derajat keanggotaan, yang memungkinkan suatu data dapat tergolong ke dalam beberapa *cluster*.

Beberapa penelitian yang telah dilakukan mengenai perbandingan kedua algoritma ini menunjukkan bahwa *K-Means* terbukti unggul dalam hal efisiensi, sedangkan *Fuzzy C-Means* unggul dalam hal menghasilkan *cluster* yang berkualitas tanpa memerhatikan metode pengukuran jarak data yang optimal untuk masing-masing algoritma. Penelitian perbandingan kedua algoritma dilakukan dengan melakukan serangkaian uji dengan menerapkan jumlah *cluster* yang berbeda dan metode perhitungan jarak yang digunakan ialah *Euclidean distance* lalu hasil

pengujian akan divalidasi menggunakan indeks validitas. Metode pengukuran jarak memegang peran penting dalam menentukan keanggotaan data terhadap suatu *cluster*, sehingga pemilihan metode pengukuran jarak yang optimal dapat meningkatkan performa dan hasil dari algoritma *clustering*. Maka dari itu, tujuan penelitian ini yaitu untuk mencari dan membandingkan hasil dari penerapan beberapa metode pengukuran jarak data (*Euclidean distance*, *manhattan*, *Chebyshev*, dan *minkowski*) dalam melakukan segmentasi data pelanggan menggunakan algoritma *Fuzzy C-Means clustering* dengan *K-Means* sebagai pembanding yang kemudian dipadukan dengan model analisis RFM (*recency*, *frequency*, dan *monetary*) untuk melakukan analisis terhadap hasil segmentasi.

1.2 Rumusan Masalah

Mengacu pada latar belakang di atas, penulis berhasil mengidentifikasi beberapa masalah diantaranya.

1. Bagaimana melakukan segmentasi pelanggan menggunakan algoritma *Fuzzy C-Means* dengan menerapkan metode pengukuran jarak data *Euclidean*, *Manhattan*, *Chebyshev*, dan *Minkowski*?
2. Bagaimana melakukan segmentasi pelanggan menggunakan algoritma *K-Means* sebagai pembanding terhadap *Fuzzy C-Means* dengan menerapkan metode pengukuran jarak data *Euclidean*, *Manhattan*, *Chebyshev*, dan *Minkowski*?
3. Bagaimana kualitas *cluster* yang dihasilkan dengan menerapkan metode pengukuran jarak data *Euclidean*, *Manhattan*, *Chebyshev*, dan *Minkowski* oleh algoritma *Fuzzy C-Means* dengan *K-Means* sebagai pembanding?
4. Bagaimana melakukan analisis perilaku pelanggan menggunakan model RFM (*Recency*, *Frequency*, dan *Monetary*)?

1.3 Ruang Lingkup

Adapun ruang lingkup pada penelitian ini diantaranya.

1. Studi kasus dilakukan pada Kedai Tepi Jalan.
2. Algoritma yang akan diterapkan ialah *clustering Fuzzy C-Means (FCM)* dengan *clustering K-Means* sebagai pembanding.
3. Analisis perilaku pelanggan menggunakan model RFM (*Recency, Frequency, Monetary*).
4. Metode pengukuran jarak data yang digunakan diantaranya *Euclidean distance, manhattan distance, chebychev distance* dan *minkowski distance*.
5. Data yang digunakan berupa data transaksi periode September 2019 – Oktober 2019.

1.4 Tujuan Penelitian

Tujuan yang akan dicapai melalui penelitian ini yaitu.

1. Mengetahui kualitas *cluster* yang dihasilkan oleh algoritma *Fuzzy C-Means* menggunakan metode pengukuran jarak data *Euclidean, Manhattan, Chebyshev, dan Minkowski* dengan tolak ukur dibandingkan dengan algoritma *K-Means*.
2. Mengetahui kelompok pelanggan pada Kedai Tepi Jalan.

1.5 Manfaat Penelitian

Beberapa manfaat yang akan didapatkan dengan mengacu pada rumusan masalah yang berhasil diidentifikasi serta ruang lingkup penelitian diantaranya.

1. Menghasilkan perbandingan kualitas *cluster* antara metode pengukuran jarak data yang digunakan terhadap algoritma *Fuzzy C-Means*.
2. Menghasilkan formulasi untuk melakukan segmentasi pelanggan bagi Kedai Tepi Jalan.
3. Menyajikan informasi kepada pembaca mengenai teknik dalam melakukan segmentasi pelanggan.
4. Diharapkan dapat menjadi bahan referensi khususnya bagi mahasiswa Informatika dalam menyusun tugas akhir.

1.6 Luaran Penelitian

Adapun luaran dari penelitian ini diantaranya.

1. Hasil perbandingan kualitas *cluster* pada algoritma *Fuzzy C-Means* dan *K-Means* menggunakan metode pengukuran jarak data *Euclidean*, *Manhattan*, *Chebyshev*, dan *Minkowski*.
2. Segmentasi pelanggan pada Kedai Tepi Jalan.

1.7 Sistematika Penulisan

Sistematika penulisan proposal Seminar Teknologi Informasi ini diatur dan disusun ke dalam beberapa bab dengan masing-masing bab berisi sub-sub bab. Adapun gambaran singkat mengenai sistematika penulisan proposal ini dapat dilihat pada daftar berikut.

BAB 1 : Pendahuluan

Pada bab ini dijelaskan secara singkat mengenai latar belakang permasalahan yang diangkat, rumusan masalah, ruang lingkup penelitian, beserta manfaat yang akan didapatkan melalui penelitian ini.

BAB 2 : Tinjauan Pustaka

Bab yang berisi semua landasan-landasan teori yang berkaitan dengan topik penelitian yang diangkat dan akan dijadikan sebagai acuan dalam penulisan penelitian.

BAB 3 : Metodologi Penelitian

Pada bab ini membahas rangkaian cara untuk menyelesaikan masalah penelitian dengan menggunakan metode-metode penelitian guna mengetahui hasil yang akan didapatkan.

BAB 4 : Hasil dan Pembahasan

Bab ini berisi hasil dan pembahasan secara rinci terkait topik penelitian yang diangkat.

BAB 5 : Kesimpulan dan Saran

Merupakan bab akhir pada penelitian ini yang berisi beberapa simpulan terkait penelitian yang telah dilakukan beserta memberikan saran untuk penelitian berikutnya yang mengambil topik serupa.