

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Pada era globalisasi saat ini, kemajuan teknologi sangat mempengaruhi berbagai segi dalam kehidupan. Segala macam kegiatan yang dilakukan dalam kehidupan sehari-hari ini menghasilkan berbagai informasi dimana informasi ini akan diolah menjadi suatu data dan disimpan dalam suatu *database*. Data sangat penting digunakan untuk menegaskan suatu permasalahan. Data terdiri dari dua jenis yaitu kategorik dan numerik. Data kategorik merupakan data yang disajikan dalam bentuk bukan angka sedangkan data numerik merupakan data yang disajikan dalam bentuk angka. Data yang sering digunakan dan dianalisis merupakan data yang berjenis numerik karena data numerik sudah berupa angka sehingga lebih mudah.

Salah satu kegiatan dalam menganalisis data yaitu mengelompokkan data menjadi beberapa kelompok. Pengelompokan biasanya menggunakan analisis kelompok. Analisis ini digunakan untuk mengklasifikasikan objek ke dalam kelompok yang relatif homogen didasarkan pada suatu set variabel yang dipertimbangkan untuk diteliti. Potensi *clustering* adalah dapat digunakan untuk mengetahui struktur dalam data yang dapat dipakai lebih lanjut dalam berbagai aplikasi secara luas seperti klasifikasi, pengolahan gambar, dan pengenalan pola. *Clustering* dapat diterapkan ke dalam data yang kuantitatif (numerik), kualitatif (kategorikal), atau kombinasi dari keduanya (Supranto, 2004).

Cluster secara umum merupakan wujud himpunan bagian dari suatu himpunan data dan metode *clustering* dapat diklasifikasikan berdasarkan himpunan bagian yang dihasilkan apakah *fuzzy* atau *crisp (hard)*. Metode *hard clustering* merupakan model yang berdasar pada teori himpunan klasik, yang mana suatu objek menjadi anggota atau tidak menjadi anggota secara penuh ke dalam suatu kelompok. *Hard clustering* membagi fakta ke dalam sejumlah himpunan bagian secara eksklusif. Metode *fuzzy clustering* mengizinkan objek untuk menjadi bagian

dari beberapa kelompok secara bersamaan dengan perbedaan level keanggotaan (Kusrini & Luthfi, 2009). Terdapat banyak algoritma yang digunakan untuk *clustering*. Salah satunya *Fuzzy C-Means* (FCM). *Fuzzy C-Means Clustering* (FCM) merupakan algoritma *clustering* data yang setiap datanya menjadi anggota dari suatu kluster dengan derajat didefinisikan dengan level keanggotaan (Jang, 2004).

Beberapa tahun terakhir, sudah dieksplorasi berbagai pengelompokan data dengan menggunakan *Fuzzy C-Means* (FCM) namun, dalam penelitian-penelitian yang sudah ada pengelompokan data dengan menggunakan *Fuzzy C-Means* (FCM) ini menggunakan kriteria jarak tanpa pembobot.

Berdasarkan hasil penelitian yang ditulis oleh Dik (2014) dalam jurnal yang berjudul “*Weighted Distances for Fuzzy Clustering*” menunjukkan bahwa algoritma *clustering* baru atau pengelompokan dengan menggunakan jarak terbobot yang diusulkan lebih baik dari pada mengelompokkan dengan menggunakan jarak tanpa pembobot. Untuk mengetahui bagaimana pengelompokan data dengan menggunakan jarak terbobot seperti yang diusulkan oleh Amina dkk, maka penulis ingin membahas sebagian beserta penerapan tentang jarak terbobot baru untuk *Fuzzy C-Means*.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang sudah diuraikan sebelumnya, maka permasalahan dalam penelitian ini yaitu bagaimana penerapan tentang jarak terbobot untuk *Fuzzy C-Means*

1.3 Batasan Masalah

Sesuai perumusan masalah yang sudah diuraikan sebelumnya, maka penulis memberi batasan masalah yaitu adanya jarak terbobot dan jarak tanpa pembobot dalam metode *Fuzzy C-Means Clustering*. Jarak yang digunakan dalam *Fuzzy C-Means* yaitu *euclidean*, *euclidean weighted*, *manhattan*, *manhattan weighted*. Untuk membuktikan teori yang akan dibahas, dalam studi kasus menggunakan data produksi pertanian jenis tanaman pangan per provinsi tahun 2015.

1.4 Jenis Penelitian dan Metode Analisis

Jenis penelitian dalam tugas akhir ini adalah penelitian teoritis, yang mengacu pada jurnal yang berjudul “*Weighted Distances for Fuzzy Clustering*” (Amina dkk, 2014).

1.5 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengetahui bagaimana penerapan jarak terbobot dalam pengelompokan suatu objek dengan metode *Fuzzy C-Means*.

1.6 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah mempermudah penyelesaian kasus untuk mengelompokkan sekumpulan data dengan metode *Fuzzy C-Means* berbasis jarak terbobot. Karena pada penelitian ini, peneliti membuat *syntax* untuk mencari nilai pembobot.