

BAB I

PENDAHULUAN

1.1. Latar Belakang

Tanaman padi menjadi sumber bahan pangan utama hampir dari setengah penduduk dunia. Tidak terkecuali Indonesia, hampir seluruh penduduk Indonesia memenuhi kebutuhan bahan pangannya dari tanaman padi. Dengan demikian tanaman padi merupakan tanaman yang mempunyai nilai spiritual, budaya, ekonomi, dan politik yang penting bagi bangsa Indonesia karena memengaruhi hajat hidup orang banyak (Ishaq, M. Rumiati, A.T. dan Permatasari, E.O., 2017).

Produksi padi di Jawa Barat dari Januari hingga September 2018 sebesar 8.108 ton Gabah Kering Giling (GKG). Produksi tertinggi terjadi pada bulan Maret yaitu sebesar 302 ton. Sementara itu, potensi produksi padi pada bulan Oktober, November, dan Desember masing-masing sebesar 521 ton, 446 ton, dan 445 ton. Dengan demikian, perkiraan total produksi padi 2018 adalah 9.539 ton. Tiga kabupaten dengan produksi padi (GKG) tertinggi terjadi di kabupaten Indramayu, Karawang, dan Subang dengan produksi masing-masing sebesar 1.392 ton, 1.124 ton, dan 991 ton (Metodologi, Perbaikan, Perhitungan Data, dan Produksi Beras., 2018).

Dalam dunia ilmu komputer, dikenal luas *data mining* sebagai teknik penggalian data untuk mencari sebuah pola tersembunyi demi menghasilkan sebuah pengetahuan baru di dalam sekumpulan data. Secara khusus *data mining* memiliki metode tersendiri berdasarkan tujuan dari pemanfaatan himpunan data yaitu estimasi, prediksi, klasifikasi, klasterisasi dan asosiasi.

Salah satu teknik yang bisa digunakan untuk tujuan pemetaan sebuah data adalah teknik klasterisasi (*clustering*). *Clustering* merupakan sebuah teknik dalam *data mining* yang berfungsi untuk mengelompokan data (*grouping*) berdasarkan kemiripannya ke dalam klaster. Ada beberapa algoritma yang dapat digunakan, salah satu yang populer digunakan untuk klasterisasi suatu *dataset* adalah *k-means* (Ridlo, M.r. Defiyanti, S. dan Primajaya, A., 2017). Adapun penelitian sebelumnya mengenai pengelompokan, penggunaan algoritma *fuzzy c-means* dan *k-means* dalam jurnal yang berjudul “Deteksi Penyakit Diabetes Dengan Metode *Fuzzy C-Means Clustering* dan *K-Means Clustering*”. Hasil dari penelitian tersebut adalah

pengelompokan terdiri dari kelompok diabetes dan non-diabetes. Pengujian untuk masing-masing metode dilakukan terhadap 9 data. Hasil pengujian terbaik metode *k-means* adalah 73,438% dan metode *fuzzy c-means* adalah 82,812% (Praja, A. Lubis, C. dan Herdiwindiati, D.E., 2017). Dilihat dari penelitian terdahulu pengelompokan biasanya menggunakan teknik *Clustering*. Untuk itu metode yang akan digunakan yaitu algoritma *k-means* dan *fuzzy c-means* untuk mengelompokkan produksi padi di provinsi Jawa Barat.

1.2. Rumusan Masalah

Berdasarkan latar belakang diatas, dapat dirumuskan suatu permasalahan yaitu:

1. Bagaimana akurasi dari algoritma *k-means* dan *fuzzy c-means* dalam mengelompokkan hasil produksi padi di provinsi Jawa Barat?
2. Bagaimana mengetahui akurasi produksi padi yang cukup baik, sangat baik dan baik?

1.3. Tujuan Penelitian

Tujuan dari penelitian ini yaitu:

1. Menggunakan algoritma *k-means* dan *fuzzy c-means* dalam mengelompokkan hasil produksi padi di provinsi Jawa Barat.
2. Mengetahui daerah yang memiliki produksi padi yang cukup baik, sangat baik dan baik.

1.4. Manfaat

Manfaat dari penelitian ini yaitu:

1. Penelitian ini dapat dijadikan sebagai studi literatur yang akan dilakukan selanjutnya.
2. Mengetahui daerah mana saja yang produksinya cukup baik, sangat baik dan baik.