

BAB I

PENDAHULUAN

1.1. Latar Belakang

Konsep Data Mining adalah *tools* penting dalam manajemen informasi karena jumlah informasi yang semakin besar. *Data Mining* sering disebut sebagai *knowledge discovery in database* (KDD) yaitu kegiatan yang meliputi pengumpulan, pemakaian data historis untuk menemukan keteraturan, pola hubungan dalam set data berukuran besar. *Output* dari *Data Mining* ini dapat digunakan untuk pengambilan keputusan di masa depan (Alfina, et al. 2012).

Salah satu teknik dalam *Data Mining* adalah *clustering*. Pengertian *clustering* dalam *Data Mining* adalah pengelompokan sejumlah data atau objek ke dalam *cluster* (*group*) sehingga setiap dalam *cluster* tersebut akan berisi data yang semirip mungkin dan berbeda dengan *objek* dalam *cluster* yang lainnya. Adapun tujuan dari *clustering* ini adalah untuk meminimalisasikan *objective function* yang diatur dalam proses *clustering*, yang pada umumnya berusaha meminimalisasikan variasi antar *cluster* (Agusta, 2007). K-Means merupakan salah satu metode data *clustering non hirarki* yang dalam prosesnya mempartisi data yang ada ke dalam bentuk satu atau lebih *cluster*/kelompok.

Kabupaten Bondowoso merupakan daerah yang tergolong memiliki sumber daya manusia yang rendah di bidang pendidikan seperti buta huruf. Sangat perlu dilakukan pemerataan pendidikan namun sampai saat ini pendidikan di Bondowoso cenderung terpusat hanya di kecamatan kota saja.

Oleh karena itu perlu dilakukan pengelompokkan kecamatan berdasarkan bidang pendidikan. Salah satu metode yang dapat menjawab permasalahan ini adalah *clustering* atau pengelompokan. Dengan begitu dapat dilakukan penelitian lebih mendalam untuk mengetahui kepadatan sekolah maupun guru sehingga pemerataan pendidikan di Kabupaten Bondowoso dapat dilakukan lebih maksimal. Berdasarkan hal tersebut peneliti memilih topik penelitian yang berjudul “Pengelompokan Kecamatan Menggunakan Metode K-Means Cluster”.

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, maka perumusan masalah dalam penulisan ini adalah:

Bagaimana mengimplementasikan sebuah sistem yang dapat digunakan untuk mengelompokkan kecamatan di Kabupaten Bondowoso berdasarkan bidang pendidikan dengan menggunakan algoritma *K-Means Clustering*.

1.3. Batasan Masalah

Luasnya cakupan masalah yang berkaitan dengan penentuan pusat cluster, maka pembatasan masalah yang dilakukan dalam penelitian ini adalah:

1. Metode *K-Means Clustering* digunakan untuk menghitung jarak antara data dan centroid menggunakan rumus *Euclidean Distance Space*. Untuk pusat cluster (*Centroid*) diambil secara *random*.
2. Data yang digunakan dalam penelitian hanya data Kecamatan di Kabupaten Bondowoso sebanyak 23 kecamatan dan hanya menggunakan kriteria berdasarkan bidang pendidikan seperti berikut “Jumlah Sekolah, Jumlah Guru, dan Jumlah Siswa”.

1.4. Tujuan Penelitian

Berdasarkan dari rumusan masalah diatas maka tujuan penelitian adalah:

Mengelompokkan kecamatan berdasarkan bidang pendidikan yang ada di Kabupaten Bondowoso dengan menggunakan algoritma *k-means cluster*.

1.5. Manfaat Penelitian

Manfaat dari penelitian ini adalah:

Membantu pemerintah mengambil keputusan terkait masalah pemberdayaan di bidang pendidikan dalam suatu wilayah.