

BAB I

PENDAHULUAN

1.1. Latar Belakang

Kabupaten Banyumas adalah salah satu kabupaten yang ada di Provinsi Jawa Tengah dengan 27 kecamatan, 331 desa, dan jumlah penduduk lebih dari 1,8 juta jiwa [1]. Data tersebut menunjukkan bahwa Kabupaten Banyumas memiliki penduduk yang heterogen. Namun, di tengah keragaman ini terdapat aspek sosial ekonomi yang perlu menjadi perhatian. Sosial ekonomi berarti posisi atau kedudukan seseorang pada suatu kelompok di masyarakat yang dipengaruhi oleh beberapa faktor seperti jumlah pendapatan, pendidikan dan juga aktivitas ekonominya [2]. Indikator ini mempengaruhi banyak aspek di kehidupan masyarakat, seperti peluang untuk mendapat pekerjaan, pendapatan, akses kesehatan, dan juga pendidikan [2].

Klasterisasi merupakan teknik untuk mengelompokkan sekumpulan data kedalam kelompok-kelompok sehingga suatu objek pada kelompok memiliki banyak kemiripan, dan memiliki banyak perbedaan dengan suatu objek di kelompok yang lain [3]. Penerapan teknik klasterisasi untuk mengelompokkan desa di Kabupaten Banyumas berdasarkan kondisi sosial ekonomi telah dilakukan pada penelitian [4] menggunakan Latent class cluster yang menghasilkan 3 cluster data.

Salah satu algoritma pengelompokan yang sering digunakan yaitu K-Means, metode ini membagi data ke dalam sejumlah cluster [5]. K-Means bekerja menggunakan nilai *centroid* yang menjadi pusat dari data dan dipilih secara random [5]. Proses perhitungan jarak pada K-Means umumnya menggunakan Euclidean Distance [6]. Namun, terdapat metode lain seperti Chebyshev, Canberra, Cosine Similarity, Minkowski, dan Manhattan Distance. Berdasarkan penelitian [7] yang membandingkan 6 metode perhitungan jarak pada k-means. Metode pengukuran jarak yang digunakan termasuk jarak Euclidean, jarak Manhattan, jarak Chebyshev, jarak Minkowski, jarak rata-rata, dan perbedaan rata-rata karakter. Penelitian ini menemukan bahwa metode jarak Chebyshev menunjukkan performa yang unggul pada data yang memiliki banyak atribut [7]. Sedangkan dalam penelitian [8] diperoleh hasil bahwa jarak Manhattan memiliki performa yang baik dengan data

dengan outlier. Perbedaan metode perhitungan jarak akan mempengaruhi performa algoritma K-Means dalam mengelompokkan [7].

Perbandingan metode untuk menghitung jarak, yaitu jarak eculidean, manhattan, dan chebyshev pada K-Means diharapkan dapat meningkatkan kualitas hasil pengelompokkan desa di kabupaten Banyumas menjadi beberapa klaster berdasarkan kondisi sosial ekonominya. Hal ini dapat mempermudah pemangku kebijakan dalam pengambilan keputusan, seperti pemetaan pembangunan akses kesehatan, dan pendidikan yang akan meningkatkan kesejahteraan masyarakat.

1.2. Rumusan Masalah

Keadaan sosial ekonomi setiap desa di Kabupaten Banyumas ada yang sama ada yang tidak. Untuk mengetahui kemiripan setiap desa dengan desa lainnya, perlu dilakukan pengelompokkan data berdasarkan aspek aspek sosial ekonomi. Selain itu, terdapat beberapa metode perhitungan jarak yang dapat mempengaruhi kualitas kelompok yang dihasilkan oleh K-Means.

1.3. Pertanyaan Penelitian

Berdasarkan perumusan masalah yang telah dijelaskan sebelumnya, pertanyaan penelitian yang diajukan adalah sebagai berikut:

1. Bagaimana hasil perbandingan perhitungan jarak K-Means dengan jarak eucledian, jarak manhattan, dan jarak chebyshev untuk klasterisasi desa berdasarkan kondisi sosial ekonomi di Kabupaten Banyumas?
2. Metode pengukuran jarak mana yang paling baik dalam mengelompokkan desa berdasarkan kondisi sosial ekonomi di Kabupaten Banyumas?

1.4. Batasan Masalah

Batasan masalah dalam penelitian ini adalah:

1. Dataset yang digunakan pada penelitian ini berasal dari Badan Pusat Statistik, yaitu Data Potensi Desa 2018
2. Variabel yang digunakan adalah variable yang berhubungan dengan indikator sosial ekonomi sebanyak 22 variabel.
3. Menggunakan K-means untuk pengelompokkan

4. Menggunakan euclidean, chebyshev dan manhattan distance untuk dibandingkan sebagai metode perhitungan jarak K-Means
5. Evaluasi hasil pengelompokkan dilakukan dengan melihat nilai silhouette score, Davies Bouldin Index, Waktu eksekusi, dan Ruang yang digunakan.

1.5. Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Membandingkan metode pengukuran jarak, yaitu metode perhitungan jarak euclidean, jarak manhattan, dan jarak chebyshev untuk pengelompokkan desa di Kabupaten Banyumas berdasarkan kondisi Sosial Ekonomi
2. Menentukan metode pengukuran jarak yang paling baik untuk mengelompokkan desa di Kabupaten Banyumas

1.6. Manfaat Penelitian

Membandingkan metode perhitungan jarak pada K-Means diharapkan dapat menghasilkan pengelompokkan dengan kualitas yang lebih baik dalam mengelompokkan desa di kabupaten Banyumas menjadi beberapa klaster berdasarkan kondisi sosial ekonominya. Selain itu, penelitian ini mendukung tujuan pembangunan berkelanjutan (*Sustainable Development Goals/SDGs*), khususnya SDG 1 (*End Poverty*), SDG 8 (*Decent Work and Economic Growth*), dan SDG 11 (*Sustainable Cities and Communities*). Dengan pengelompokkan desa yang akurat, pemerintah dapat lebih efektif mengentaskan kemiskinan, mendorong pertumbuhan ekonomi lokal, mengurangi ketimpangan sosial, dan memastikan pembangunan desa yang berkelanjutan.