

BAB I

PENDAHULUAN

1.1 Latar Belakang

Salah satu indikator untuk melihat kemajuan suatu negara ialah faktor ekonomi negara karena kemiskinan merupakan salah satu permasalahan yang dialami oleh hampir semua negara. Pertumbuhan ekonomi yang stabil kerap kali dianggap sebagai prasyarat pembangunan negara dan kenaikan kesejahteraan masyarakat (Todaro & Smith, 2015). Negara dengan pertumbuhan ekonomi yang tinggi cenderung memiliki lebih banyak sumber daya untuk mengalokasikan anggaran pembangun infrastruktur, pemberian layanan publik, dan program dalam rangka mengurangi kemiskinan dan meningkatkan kesejahteraan masyarakat. Selama keberjalananya, pertumbuhan ekonomi yang tinggi belum tentu berarti penurunan angka kemiskinan secara signifikan karena aspek kemiskinan sendiri merupakan persoalan yang cukup kompleks. Kemiskinan merupakan masalah serius terutama untuk negara berkembang seperti Indonesia karena kemiskinan sebuah wilayah dalam jangka panjang akan berdampak pada pembangunan nasional.

Secara umum, kemiskinan merupakan suatu keadaan seseorang yang tidak dapat memenuhi kebutuhan dasarnya. Badan Pusat Statistik menggunakan konsep kemampuan memenuhi kebutuhan dasar dalam mengukur kemiskinan, sehingga kemiskinan dapat diartikan sebagai ketidakmampuan seseorang dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur menurut garis kemiskinan (BPS, 2023). Garis kemiskinan merupakan suatu nilai

pengeluaran minimum kebutuhan makanan dan bukan makanan yang harus dipenuhi agar tidak dikategorikan miskin.

Menurut Badan Pusat Statistik (BPS) pada publikasi Profil Kemiskinan di Indonesia bulan Maret 2023, upaya negara Indonesia dalam mengatasi masalah kemiskinan telah dilakukan secara berkala dan berkelanjutan melalui program dan kebijakan seperti pembagian bantuan sosial yang sudah dijalani. Akan tetapi, masih ada ketidaksamaan kondisi penduduk miskin yang cukup signifikan antar wilayah. Oleh karena itu, diperlukan adanya analisis yang mendalam terhadap data penduduk miskin di setiap wilayah Indonesia karena hal ini menjadi fundamental dalam merumuskan strategi dan program yang akan dijalankan nantinya sebagai upaya mengatasi keparahan kondisi penduduk miskin yang lebih tepat sasaran dan efektif.

Salah satu pendekatan yang dapat digunakan untuk menganalisis kondisi penduduk miskin di Indonesia adalah dengan melakukan pengelompokan (*clustering*) kabupaten/kota yang ada di Indonesia berdasarkan indikator kemiskinan. Analisis *cluster* merupakan sebuah analisis yang digunakan untuk mencari kemiripan dalam data dan menempatkan data yang mirip tersebut ke dalam suatu kelompok. *Clustering* membagi kumpulan data tersebut ke dalam kelompok dengan anggota kelompok yang memiliki kesamaan lebih besar dibandingkan kelompok lainnya (Xu & Wunsch, 2008). Pengelompokan ini bertujuan mengidentifikasi beberapa wilayah yang memiliki karakteristik kondisi penduduk miskin serupa sehingga pemerintah dapat menyiapkan kebijakan dan program yang sesuai dan terfokus dengan kondisi masing-masing kelompok wilayah.

Saat menganalisis dengan metode *clustering*, agar metode yang digunakan dapat *robust* atau tidak sensitif terhadap data *outlier*, ada beberapa metode yang

dapat digunakan, diantaranya adalah algoritma *Partitioning Around Medoid* (PAM) dan *Clustering Large Application* (CLARA). Algoritma PAM merupakan salah satu metode *clustering* yang *robust* terhadap *outlier* dalam data. Algoritma PAM yang dikembangkan oleh Kauffman & Rousseeuw pada tahun 1990 menggunakan objek representatif yang disebut dengan *medoid* sebagai pusat *cluster*. *Medoid* merupakan objek yang letaknya terpusat di dalam suatu *cluster* (Kaufman & Rousseeuw, 1990). Algoritma ini digunakan untuk menangani kekurangan dari algoritma *K-Means* yang sensitif terhadap *outlier* karena objek tersebut memiliki karakteristik sangat jauh terhadap mayoritas data lainnya, sehingga jika dimasukkan ke dalam sebuah *cluster* objek tersebut dapat mendistorsi nilai rata-rata *cluster* tersebut. Penggunaan *medoid* membuat PAM lebih tahan terhadap *outlier* karena *medoid* cenderung lebih mewakili sebagian besar objek dalam *cluster* dibandingkan dengan rata-rata. Algoritma PAM memiliki kelebihan dapat *robust* terhadap *outlier*, tetapi algoritma PAM memiliki kesulitan dalam menangani *dataset* yang besar. Untuk mengatasi kesulitan tersebut, digunakan algoritma *Clustering Large Application* (CLARA) yang merupakan pengembangan dari Algoritma PAM. CLARA digunakan untuk meningkatkan efisiensi komputasi pada *clustering* data berukuran besar (Kaufman & Rousseeuw, 1990). Hal ini terjadi dikarenakan CLARA bekerja dengan memilih sampel secara acak dari *dataset* dan menerapkan algoritma PAM pada sampel tersebut hingga mendapatkan *medoid* yang optimal. Proses tersebut dilakukan secara berulang kali dengan menggunakan sampel yang berbeda sehingga mendapatkan hasil terbaik yang diambil sebagai hasil pengelompokan akhir. Kualitas *medoid* akhir diukur dengan jarak rata-rata setiap objek dalam kumpulan data.

Penelitian serupa yang sebelumnya pernah dilakukan adalah penelitian Alfiah *et al.*, (2022). Penelitian tersebut membahas mengenai *clustering* menggunakan metode K-Medoids berdasarkan indikator kemiskinan di Jawa Timur tahun 2020. Pada penelitian tersebut menunjukkan *clustering* menggunakan metode K-Medoids memperoleh 2 *cluster* optimal setelah melihat nilai *silhouette coefficient*-nya. Terdapat 30 kabupaten/kota pada *cluster* 1 dan 8 kabupaten/kota pada *cluster* 2. Penelitian yang telah dilakukan selanjutnya dilakukan oleh Ayuni *et al.*, (2024) mengenai penerapan metode K-Medoids dan CLARA dengan validasi nilai *silhouette coefficient* untuk data Indeks Pembangunan Manusia berdasarkan kabupaten/kota di Indonesia. Pada penelitian tersebut menunjukkan *clustering* menggunakan metode K-Medoids dan CLARA memperoleh *cluster* optimal yang terbentuk adalah sama-sama 2 *cluster*. Nilai *silhouette coefficient* tertinggi berada pada metode CLARA yaitu sebesar 0,438 sedangkan pada metode K-Medoids sebesar 0,340. Hal ini menunjukkan bahwa metode CLARA memberikan hasil *clustering* yang lebih optimal dibandingkan dengan metode K-medoids dalam mengelompokkan kabupaten/kota di Indonesia berdasarkan nilai Indeks Pembangunan Manusia.

Berdasarkan latar belakang tersebut, Penelitian ini akan melakukan *clustering* menggunakan algoritma PAM dan CLARA. Jarak yang digunakan pada penelitian ini adalah jarak *Euclidean*. Nilai *Davies Bouldin Index* digunakan untuk membandingkan kedua algoritma tersebut sehingga mendapatkan hasil terbaik. Hasil dari penelitian ini diharapkan memperoleh hasil *cluster* kabupaten/kota di Indonesia berdasarkan data penduduk miskin yang lebih akurat dan berguna dalam membuat skala prioritas wilayah serta strategi

penanganan kondisi penduduk miskin yang lebih efektif. Hasil *clustering* ini juga dapat memberikan informasi yang berguna untuk pemerintah dan pemegang kepentingan mengambil keputusan dalam upaya mengurangi keparahan kondisi penduduk miskin masing-masing kelompok wilayah dan meningkatkan kesejahteraan penduduk miskin secara merata.

1.2 Rumusan Masalah

Berdasarkan latar belakang maka dapat dirumuskan permasalahan sebagai berikut:

1. Bagaimana *clustering* pada data penduduk miskin di Indonesia menggunakan algoritma *Partitioning Around Medoids* (PAM) dan *Clustering Large Application* (CLARA)?
2. Bagaimana mendapatkan *cluster* terbaik dari algoritma *Partitioning Around Medoids* (PAM) dan *Clustering Large Application* (CLARA) dengan Nilai *Davies Bouldin Index*?
3. Bagaimana analisis perbandingan Nilai *Davies Bouldin Index* dari algoritma *Partitioning Around Medoids* (PAM) dan *Clustering Large Application* (CLARA) untuk data penduduk miskin?

1.3 Batasan Masalah

Pada penelitian ini, batasan masalah yang digunakan sebagai berikut:

1. Data yang digunakan pada penelitian kali ini yaitu data Penduduk Miskin Tahun 2023 pada kabupaten/kota di seluruh Indonesia.
2. Pengelompokan pada penelitian ini menggunakan algoritma *Partitioning Around Medoids* (PAM) dan *Clustering Large Application* (CLARA).

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah disebutkan sebelumnya, maka tujuan penelitian sebagai berikut:

1. Mendapatkan hasil *clustering* pada data di Indonesia menggunakan algoritma *Partitioning Around Medoids* (PAM) dan *Clustering Large Application* (CLARA).
2. Mendapatkan hasil *cluster* terbaik dari algoritma *Partitioning Around Medoids* (PAM) dan *Clustering Large Application* (CLARA) menggunakan nilai *Davies Bouldin Index*.
3. Mendapatkan hasil analisis perbandingan dari algoritma *Partitioning Around Medoids* (PAM) dan *Clustering Large Application* (CLARA) untuk data indikator kemiskinan.