

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Kemiskinan tetap menjadi salah satu masalah yang paling mendesak di zaman sekarang. Kemiskinan menjadi permasalahan utama yang sering dihadapi banyak negara di dunia, khususnya negara berkembang seperti Indonesia. Definisi mengenai kemiskinan sangat beragam. Dalam pengertian sempit, kemiskinan dipahami sebagai kondisi keterbatasan uang dan barang untuk memenuhi kebutuhan hidup. Sementara dalam pengertian luas, kemiskinan dipandang sebagai fenomena yang bersifat multidimensi [1]. Badan Pusat Statistik (BPS) mendefinisikan kemiskinan sebagai kondisi dimana seseorang atau sekelompok orang tidak mampu memenuhi hak-hak dasarnya untuk mempertahankan dan mengembangkan kehidupan yang bermartabat, yang diukur berdasarkan pengeluaran untuk memenuhi kebutuhan dasar makanan (mencapai 2.100 kalori per hari) dan kebutuhan dasar non-makanan (seperti perumahan, sandang, pendidikan, dan kesehatan) [1].

Menurut Bank Dunia pada tahun 2004, salah satu penyebab kemiskinan adalah rendahnya pendapatan untuk memenuhi kebutuhan dasar seperti makanan, pakaian, tempat tinggal, tingkat kesehatan dan pendidikan yang layak. Selain itu, kemiskinan juga berkaitan dengan keterbatasan

lapangan pekerjaan dan biasanya mereka yang masuk dalam kategori miskin (*the poor*) tidak memiliki pekerjaan (pengangguran), serta tingkat pendidikan dan kesehatan yang mereka miliki biasanya tidak memadai [2].

Berdasarkan data BPS, jumlah penduduk miskin di Indonesia pada Maret 2024 tercatat sebesar 9,03 persen atau sekitar 25,22 juta jiwa. Angka ini mengalami penurunan dibandingkan Maret 2023 yang mencapai 9,36 persen atau 25,90 juta jiwa [3]. Meskipun terjadi penurunan, tingkat kemiskinan masih relatif tinggi. Tingginya tingkat kemiskinan dapat menyebabkan berbagai masalah sosial yang serius, salah satunya adalah tingginya tingkat pengangguran. Penduduk yang tidak memiliki pekerjaan berisiko mencari alternatif lain untuk memenuhi kebutuhan sehari-hari, termasuk terlibat dalam kegiatan ilegal, yang dapat meningkatkan angka kejahatan. Ini menunjukkan bahwa isu kemiskinan tidak hanya berkaitan dengan faktor ekonomi, tetapi juga berdampak pada stabilitas sosial masyarakat. Oleh sebab itu, penanganan kemiskinan membutuhkan analisis yang lebih mendalam agar kebijakan yang diterapkan dapat lebih tepat sasaran dan efektif.

Setiap provinsi di Indonesia menunjukkan tingkat kemiskinan yang berbeda-beda. Perbedaan ini dipengaruhi oleh banyak faktor pendukung. Dengan menetapkan dan mengukur berbagai indikator kemiskinan, kita dapat lebih mudah memahami tingkat kemiskinan dalam suatu provinsi. Untuk memerangi kemiskinan, pemerintah perlu mengetahui provinsi mana yang memiliki tingkat kemiskinan yang tinggi atau rendah sehingga pemerintah

dapat menetapkan skala prioritas dalam penanganan masalah tersebut. Oleh karena itu, pengelompokan tingkat kemiskinan antar provinsi di Indonesia bisa dilakukan dengan metode yang ada dalam *data mining* yaitu metode *clustering* [4].

*Data Mining* adalah sebuah proses yang bertujuan untuk mengekstrak atau menyaring informasi dari kumpulan data yang memiliki ukuran cukup besar melalui berbagai tahapan untuk memperoleh informasi yang berharga dari data tersebut [5]. *Clustering* atau yang biasa disebut analisis klaster dalam *data mining* adalah teknik yang membagi data ke dalam sejumlah kelompok yang memiliki objek dengan karakteristik serupa [4].

Metode analisis klaster umumnya dibagi menjadi dua kategori, yaitu *Hierarchical Clustering* dan *Non-Hierarchical Clustering*. Pada *Hierarchical Clustering*, setiap data awalnya dianggap sebagai sebuah klaster tersendiri, kemudian secara bertahap klaster digabungkan hingga membentuk kelompok yang lebih besar. Sementara itu, *Non-Hierarchical Clustering* dimulai dengan menentukan jumlah klaster yang diinginkan sejak awal, lalu data dikelompokkan ke dalam klaster tersebut sesuai dengan tingkat kemiripannya [6].

Salah satu metode analisis klaster yang sering digunakan adalah *Average Linkage*. Metode *Average Linkage* merupakan salah satu teknik dalam analisis *Hierarchical Clustering* yang berfungsi untuk mengukur jarak antara klaster. Dalam teknik ini, jarak di antara dua klaster dihitung dengan cara mengambil rata-rata dari semua jarak antara setiap pasangan objek,

yaitu dari setiap objek di klaster pertama ke setiap objek di klaster kedua [7]. Pengelompokan dengan metode *Average Linkage* dapat ditemukan dalam beberapa literatur terdahulu diantaranya Damayanthi, dkk menggunakan *Average Linkage* untuk mengelompokkan provinsi di Indonesia menurut indikator indeks kebahagiaan [8]. Nafisah dan Chandra menggunakan metode *Average Linkage* dalam pengelompokan provinsi Jawa Timur berdasarkan faktor-faktor kemiskinan [9]. Wahyuni dan Jatmiko menggunakan metode *Average Linkage* dalam pengelompokan kabupaten di Pulau Jawa berdasarkan faktor-faktor kemiskinan [10]. Karena metode *Average Linkage* belum banyak diterapkan dalam penelitian terkait kemiskinan pada provinsi-provinsi di Indonesia, maka penulis memilih metode ini untuk mengelompokkan provinsi di Indonesia berdasarkan indikator kemiskinan.

Selain metode *Average Linkage*, metode analisis klaster lain yang sering digunakan adalah metode *K-Medoids*. *K-Medoids* adalah metode klusterisasi yang mengambil objek sebagai pusat klaster pada setiap klaster. Dalam metode ini, pusat klaster dikenal sebagai medoid, sedangkan objek lainnya yang bukan menjadi pusat klaster disebut non medoid. Metode *K-Medoids* sangat berguna untuk mengatasi kumpulan data yang mengandung beberapa data ekstrem (*outlier*) yang berada sangat jauh dari mayoritas data lainnya [11]. Pengelompokan dengan metode *K-Medoids* dapat ditemukan dalam beberapa literatur terdahulu diantaranya Zahra, dkk menggunakan metode *K-Medoids* untuk mengelompokkan tingkat kemiskinan di setiap provinsi di Indonesia [12]. Safitri, dkk menggunakan metode

*K-Medoids* untuk mengelompokkan kemiskinan di provinsi Jawa Barat [13]. Munandar menggunakan metode *K-Medoids* untuk mengelompokkan kemiskinan di provinsi Banten [14]. Karena metode *K-Medoids* telah banyak di terapkan dalam pengelompokan provinsi di Indonesia berdasarkan indikator kemiskinan, maka penulis memilih metode ini untuk mengelompokkan provinsi di Indonesia berdasarkan indikator kemiskinan.

Perbandingan antara metode *Average Linkage* dan *K-Medoids* juga telah dibahas dalam penelitian sebelumnya diantaranya Husna, dkk melakukan pengelompokan terhadap jenis penyakit pasien rawat inap yang memperoleh kesimpulan bahwa metode *Average Linkage* lebih baik dibandingkan metode *K-Medoids* [15]. Rhomaningtias, dkk membandingkan metode *Average Linkage* dan *K-Medoids* dalam mengelompokkan data konsumsi rempah rempah di kabupaten/kota yang memperoleh kesimpulan bahwa metode *Average Linkage* lebih baik dibandingkan metode *K-Medoids* [16]. Karena belum ditemukan perbandingan antara metode *Average Linkage* dan *K-Medoids* dalam penelitian terkait kemiskinan, maka pada penelitian ini penulis akan membandingkan metode *Average Linkage* dan *K-Medoids* dalam mengelompokkan provinsi di Indonesia berdasarkan indikator kemiskinan.

Dengan menggunakan metode *Average Linkage* dan *K-Medoids* pada penelitian ini, dapat diperoleh hasil pengelompokan yang berbeda-beda. Salah satu hal yang menarik adalah mengetahui metode mana yang menghasilkan pengelompokan yang lebih baik berdasarkan dari hasil validitas kluster menggunakan *Silhouette Coefficient* [17].

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang diuraikan sebelumnya, rumusan masalah yang akan dibahas dalam penelitian ini adalah bagaimana hasil pengelompokan provinsi yang terbentuk berdasarkan indikator kemiskinan di Indonesia tahun 2024, dengan menggunakan metode *Average Linkage* dan metode *K-Medoids*, serta bagaimana menentukan hasil pengelompokan yang lebih baik antara metode *Average Linkage* dan metode *K-Medoids*.

## 1.3 Batasan Masalah

Adapun batasan masalah dari penelitian ini, yaitu :

1. Data mengenai kasus kemiskinan yang digunakan adalah data indikator kemiskinan berdasarkan provinsi di Indonesia yang diperoleh dari Badan Pusat Statistik (BPS) pada tahun 2024.
2. Penelitian ini menggunakan metode *Average Linkage* dan *K-Medoids* dalam proses pengelompokan dengan jumlah klaster  $k = 2, 3, 4, \text{ dan } 5$ .
3. Validitas klaster yang digunakan untuk mengetahui hasil klaster yang terbaik adalah *Silhouette Coefficient*.

## 1.4 Tujuan Penelitian

Berdasarkan permasalahan yang telah dijelaskan sebelumnya, tujuan dari penelitian ini adalah untuk mengelompokkan provinsi-provinsi di Indonesia berdasarkan indikator kemiskinan yang ada pada tahun 2024

dengan menggunakan metode *Average Linkage* dan *K-Medoids*, serta untuk mengetahui hasil pengelompokan yang lebih baik antara metode *Average Linkage* dan *K-Medoids*.

## 1.5 Sistematika Penulisan

Penelitian ini terdiri dari lima bab. Bab I mencakup latar belakang, rumusan masalah, batasan masalah, tujuan penelitian dan sistematika penulisan. Pada Bab II terdapat teori-teori yang relevan dengan masalah. Bab III berisikan sumber data, variabel penelitian, dan tahapan analisis data. Pada Bab IV terdapat hasil dan pembahasan yang berisikan uraian analisis serta pembahasan yang akan disimpulkan pada Bab V.

