

BAB I

PENDAHULUAN

1.1. Latar Belakang

Pembangunan merupakan upaya untuk menciptakan perubahan dengan tujuan meningkatkan kondisi yang ada menjadi lebih baik. Setiap negara perlu melakukan pembangunan guna menciptakan kesejahteraan bagi masyarakatnya. Tingkat keberhasilan pembangunan suatu negara dapat dinilai dengan melihat sejauh mana negara mampu mengatasi berbagai persoalan mendasar seperti pengangguran, kemiskinan, ketahanan pangan, buta huruf, serta bagaimana demokrasi ditegakkan. Pembangunan manusia di suatu negara sangat penting untuk mencapai kesejahteraan dan mendorong kemajuan. Kualitas sumber daya manusia yang ada dalam wilayah tersebut perlu dievaluasi terlebih dahulu. *United Nation Development Programme* (UNDP, 1995) menilai keberhasilan pembangunan manusia di suatu negara adalah dengan menggunakan Indeks Pembangunan Manusia (IPM).

Indeks Pembangunan Manusia (IPM) merupakan salah satu indikator penting untuk mengukur tingkat kesejahteraan masyarakat di suatu wilayah. Indeks Pembangunan Manusia (IPM) menjadi salah satu ukuran penting yang menggambarkan keberhasilan pembangunan suatu daerah berdasarkan tiga dimensi dasar, yaitu umur panjang dan hidup sehat, pengetahuan, dan standar hidup layak (UNDP, 1995). Indikator yang digunakan untuk menghitung IPM mencakup umur harapan hidup, harapan lama sekolah, rata-rata lama sekolah, dan pengeluaran per kapita riil. Indikator ini memberikan gambaran komprehensif tentang kualitas hidup

masyarakat dan menjadi acuan bagi pemerintah dalam merumuskan kebijakan pembangunan yang lebih baik (Badan Pusat Statistika, 2023).

Berdasarkan data BPS (Badan Pusat Statistika, 2023) Indonesia tahun 2023, Indeks Pembangunan Manusia (IPM) Indonesia mencapai angka 74,39, mengalami kenaikan sebesar 0,62 dibandingkan tahun sebelumnya yaitu sebesar 73,77. Kota Yogyakarta memiliki IPM tertinggi di Indonesia, yaitu 81,09, sementara Kabupaten Nduga memiliki IPM terendah di Indonesia sebesar 35,19. Perbedaan yang mencolok ini mendorong penulis untuk melakukan pengelompokan berdasarkan indikator IPM. Pengelompokan ini bertujuan untuk mengidentifikasi kemiripan antar kabupaten/kota sehingga dapat memberikan gambaran yang lebih jelas mengenai daerah yang perlu mendapat prioritas dalam upaya peningkatan IPM.

Pengelompokan kabupaten/kota berdasarkan indikator IPM dapat dilakukan dengan menggunakan analisis *cluster* non-hirarki. Analisis klaster dibagi menjadi dua jenis, yaitu *cluster* hirarki dan *cluster* non-hirarki atau *partitioning* (Johnson dan Wichern, 1982). *Cluster* hirarki biasanya digunakan untuk jumlah individu yang tidak terlalu banyak, di mana jumlah *cluster* belum ditentukan. Contoh metode *cluster* hierarki adalah *Agglomerative Hierarchical Clustering* (AHC) dan *Divisive Hierarchical Clustering*. Metode *cluster* non-hirarki atau *partitioning*, jumlah *cluster* yang akan dibentuk sudah ditentukan sebelumnya. Contoh metode *cluster* non-hirarki adalah *K-Means*, *K-Medoids*, *Clustering Large Application* (CLARA), dan DBSCAN. Proses *clustering* non-hierarki dimulai dengan pemilihan banyaknya *cluster* awal yang diinginkan, kemudian objek-objek penelitian digabungkan ke dalam klaster-klaster tersebut (Johnson dan Wichern, 1982). Pada penelitian ini, akan digunakan metode *K-Medoid* dan *Clustering Large Application*

(CLARA) untuk mengelompokan data indikator Indeks Pembangunan Manusia (IPM) di Indonesia.

K-Medoids clustering, yang juga dikenal sebagai *Partioning Around Medoids* (PAM), adalah varian dari *K-Means* yang menggunakan pengukuran jarak dengan memilih salah satu data (*Medoid*) sebagai pusat cluster (Kaufman dan Rousseeuw, 1990). Hal ini menjadikan *K-Medoids* lebih robust terhadap outlier, yaitu data-data ekstrem yang dapat mempengaruhi hasil pengelompokan secara signifikan. Algoritma *K-Medoids* memiliki kesulitan dalam menangani *dataset* yang besar. Untuk mengatasi kesulitan tersebut, digunakan algoritma *Clustering Large Applications* (CLARA), yang merupakan penyempurnaan dari algoritma *K-Medoids*. CLARA dirancang untuk meningkatkan efisiensi komputasi dalam proses pengelompokan data berukuran besar (Kaufman & Rousseeuw, 1990). Algoritma ini bekerja dengan mengambil sampel secara acak dari dataset, kemudian menerapkan algoritma *K-Medoids* pada sampel tersebut untuk menentukan *medoid* yang optimal. Proses ini diulangi dari dataset 1 sampai 5 menggunakan sampel yang berbeda, sehingga diperoleh hasil terbaik yang digunakan sebagai pengelompokan akhir. Kualitas *medoid* akhir dievaluasi berdasarkan jarak rata-rata setiap objek dalam kumpulan data.

Penelitian sebelumnya telah dilakukan oleh Ardini et al. (2023). Penelitian tersebut membahas mengenai perbandingan algoritma *K-Medoids* dan CLARA dalam analisis pengelompokan kemiskinan di Indonesia. Pada penelitian tersebut menunjukkan hasil bahwa dengan menggunakan *K-Medoids* diperoleh 2 *cluster*, sedangkan dengan menggunakan algoritma CLARA diperoleh 3 *cluster*. Berdasarkan hasil pengelompokan kedua algoritma tersebut, diperoleh algoritma

terbaik dengan menggunakan validasi *cluster Silhouette Coefficient* yaitu algoritma *K-Medoids* karena memiliki nilai rata-rata rasio simpangan baku terbesar yaitu 0,667.

Data yang digunakan dalam penelitian ini diketahui memiliki potensi keberadaan *outlier*. Metode *K-Medoids* dan CLARA dalam penelitian ini digunakan untuk mengelompokkan Kabupaten/Kota di Indonesia berdasarkan IPM. Hasil dari penelitian ini diharapkan dapat membantu pemerintah untuk melakukan evaluasi terhadap hasil pembangunan dalam aspek kesehatan, pendidikan, pengeluaran serta membuat kebijakan yang tepat untuk meningkatkan nilai IPM di Indonesia.

1.2. Rumusan Masalah

Sehubungan dengan latar belakang yang telah disampaikan sebelumnya, maka dapat diuraikan rumusan masalah sebagai berikut:

1. Bagaimana hasil pengklasteran pada data Indeks Pembangunan Manusia (IPM) di Indonesia tahun 2023 menggunakan metode *K-Medoids* dan *Clustering Large Application* (CLARA) beserta validasinya?
2. Bagaimana hasil dari jumlah *cluster* optimal berdasarkan Indeks Pembangunan Manusia (IPM) di Indonesia tahun 2023?

1.3. Batasan Masalah

Peneliti membatasi masalah pada sumber data, validasi, dan *software* yang digunakan dalam penelitian ini. Data yang digunakan pada penelitian ini bersumber dari website publikasi Badan Pusat Statistika (BPS) Indonesia tahun 2023. Validasi

yang digunakan yaitu *Silhouette Coefficient* dengan bantuan *software* RStudio dan *Microsoft Excell*.

1.4. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah ditentukan, maka tujuan penulis dalam melakukan penulisan laporan Tugas Akhir adalah sebagai berikut:

1. Menentukan hasil pengklasteran pada data Indeks Pembangunan Manusia (IPM) di Indonesia tahun 2023 menggunakan metode *K-Medoids* dan *Clustering Large Application* (CLARA) beserta validasinya.
2. Menentukan jumlah *cluster* optimal berdasarkan Indeks Pembangunan Manusia (IPM) di Indonesia tahun 2023.