

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara berkembang yang memiliki jumlah penduduk terbesar keempat di dunia, setelah China, India, dan Amerika Serikat (Kemnaker RI, 2023). Status tersebut dapat menjadi potensi sekaligus ancaman bagi negara Indonesia. Potensi Indonesia yang terdiri dari aneka ragam kebudayaan, bahasa, hingga sumber daya alam justru menyimpan ancaman dari segi kualitas sumber daya manusianya sendiri. Indonesia memiliki minat baca yang sangat rendah. Survei dari UNESCO mengenai literasi dunia menyatakan minat baca masyarakat Indonesia hanya 0,001% atau hanya satu orang dari 1.000 orang Indonesia yang rajin membaca (Perpusnas, 2021). Berdasarkan PISA 2022 yaitu *Programme for International Student Assessment* sebuah studi internasional yang menilai kualitas sistem pendidikan dengan mengukur hasil belajar yang esensial untuk berhasil di Abad ke-21 menyatakan, Indonesia berada pada peringkat 10 terbawah dalam kategori literasi membaca dengan skor literasi membaca sebesar 359 (EMedia DPR RI, 2024).

Literasi adalah wujud dari keterampilan yang secara nyata, yang secara spesifik adalah keterampilan kognitif dari membaca serta menulis, yang terlepas dari konteks di mana keterampilan itu diperoleh dari siapa serta cara memperolehnya (UNESCO, 2025). Makna literasi dipengaruhi oleh beberapa hal, diantaranya yaitu penelitian akademik, institusi, konteks nasional, nilai-nilai budaya, dan juga pengalaman (Purwati, 2017). Menurut Badan Bahasa (2023), rendahnya tingkat literasi di Indonesia bisa disebabkan oleh berbagai hal, diantaranya yaitu dampak teknologi

seperti *handphone* dan televisi, kurangnya kualitas dan akses bahan bacaan, serta minimnya peran orang tua dalam memperkenalkan dunia baca.

Peran fasilitas literasi seperti perpustakaan, koleksi buku, hingga tenaga perpustakaan memengaruhi kualitas literasi di Indonesia, apabila fasilitasnya tidak memadai maka kesenjangan literasi antar daerah juga dapat terjadi. Indonesia memiliki Indeks Pembangunan Literasi Masyarakat (IPLM) sebagai acuan untuk mengetahui apakah fasilitas, sarana, dan prasarana dalam hal literasi sudah memadai atau belum. Menurut Kemenko PMK RI (2023), IPLM yang dihimpun oleh Perpustakaan Nasional pada tahun 2022 berada pada 64,48 dari skala 100 atau masuk kategori sedang. Pemerintah Indonesia telah menargetkan pada tahun 2024 IPLM harus mencapai angka 71,4. IPLM terdiri dari 7 unsur, yaitu pemerataan layanan perpustakaan, ketercukupan koleksi, ketercukupan tenaga perpustakaan, tingkat kunjungan masyarakat per hari, jumlah perpustakaan ber-Standar Nasional Pendidikan (SNP), keterlibatan masyarakat dalam kegiatan sosialisasi/promosi, dan anggota perpustakaan (Perpustakaan Nasional, 2023).

Pengelompokan kabupaten/kota berdasarkan indikator penyusun IPLM dapat dijadikan pemerintah sebagai acuan dalam mengalokasikan dana untuk meningkatkan fasilitas literasi di tiap daerah. Perbaikan tersebut bisa difokuskan kepada daerah dengan kategori IPLM rendah ataupun sangat rendah, kemudian daerah dengan kategori IPLM sedang dan unggul bisa dijadikan contoh bagi daerah-daerah lainnya, sehingga nilai IPLM dapat merata untuk semua daerah. Pengelompokan kabupaten/kota berdasarkan indikator penyusun IPLM akan dilakukan pada 35 kabupaten/kota di provinsi Jawa Tengah. IPLM di Jawa Tengah tahun 2023 bernilai sebesar 64,4 dan nilai ini masih berada di bawah rata-rata IPLM

Indonesia tahun tersebut yaitu sebesar 68,19. Alasan lainnya yaitu Jawa Tengah menduduki peringkat ketiga sebagai provinsi terpadat di Indonesia, setelah Jawa Barat dan Jawa Timur. Kegiatan pendidikan maupun ekonomi banyak yang terpusat di Jawa Tengah, sehingga diperlukan penelitian lebih lanjut untuk mengetahui daerah mana yang harus difokuskan untuk peningkatan fasilitas literasi masyarakatnya.

Pengelompokan atau analisis *cluster* bertujuan untuk mengelompokan objek data tertentu ke dalam beberapa *cluster* yang memiliki karakteristik sama. Pengelompokan IPLM dilakukan dengan 2 metode, yaitu *K-Medoids (Partitioning Around Medoids)* dan *K-Harmonic Means* karena kedua metode tersebut merupakan bentuk perbaikan dari metode sebelumnya seperti *K-Means*, sehingga kinerjanya akan lebih baik dibandingkan metode *cluster* konvensional lainnya. Kedua metode tersebut tergolong ke dalam *clustering* non hierarki, artinya jumlah *cluster* ditentukan di awal terlebih dahulu.

Algoritma *K-Harmonic Means* merupakan metode *clustering* yang diusulkan oleh Zhang yang kemudian dimodifikasi oleh Hammerly, dimana metode ini berbasis pusat seperti *K-Means* dengan kelebihan mampu menangani masalah inisialisasi pusat *cluster* yang sensitif pada metode *K-Means* (Yang *et al.*, 2009). Proses inisialisasi *centroid* awal sangat memengaruhi hasil *clustering*. Pemilihan *centroid* awal secara random pada *K-Means* cenderung menghasilkan *cluster* yang berbeda pada data yang sama, sehingga inisialisasi *centroid* awal yang kurang optimal menyebabkan hasil *clustering* yang kurang optimal juga.

K-Harmonic Means menggunakan fungsi pembobotan untuk meningkatkan kepentingan data yang nilainya jauh dari titik pusat, sehingga metode ini bersifat tidak sensitif terhadap inisialisasi awal (Salsabila *et al.*, 2022). Perbedaan *K-Means*

dan *K-Harmonic Means* yang lain yaitu terletak pada fungsi tujuan yang dipakai. *K-Harmonic Means* menggunakan fungsi tujuan berupa rata-rata harmonik dari jarak setiap titik data ke pusat *cluster*. Metode *K-Means* diasumsikan tiap titik data mempunyai bobot yang sama sedangkan pada *K-Harmonic Means* tiap titik data diberi bobot yang dinamis berdasarkan suatu rata-rata harmoniknya (Yunistya *et al.*, 2022).

Algoritma *K-Medoids* atau sering disebut juga dengan algoritma PAM (*Partitioning Around Medoids*) dikembangkan oleh Leonard Kaufman dan Peter J. Rousseeuw, yaitu algoritma yang mirip dengan *K-Means* karena kedua algoritma ini *partitional* yang memecah dataset menjadi kelompok-kelompok (Wira *et al.*, 2019). Perbedaan *K-Means* dengan *K-Medoids* terletak pada penentuan pusat *cluster*, dimana algoritma *k-means* menggunakan nilai rata-rata (*mean*) dari setiap *cluster*, sedangkan algoritma *K-Medoids* menggunakan objek data asli sebagai perwakilan (*medoids*) (Kaur *et al.*, 2014). *K-Medoids* cenderung lebih stabil karena hasil akhirnya tidak bergantung pada pemilihan *medoids* awal yang acak. Hal ini dikarenakan *K-Medoids* menggunakan data aslinya sebagai *centroid* dan ada proses perbaikan selama iterasi, sehingga hasil *cluster* tidak akan terjebak pada penentuan *medoids* awal yang acak (Bintoro *et al.*, 2024).

Validasi dilakukan pada tiap *cluster* yang terbentuk untuk mengetahui metode *cluster* dan jumlah *cluster* yang optimal. Validasi *cluster* dapat dilakukan dengan berbagai cara, salah satunya yang paling sering digunakan adalah validasi *Silhouette Coefficients*. Validasi *Silhouette Coefficients* adalah metode untuk melihat kualitas dan kekuatan dari *cluster*. Penelitian sebelumnya telah membahas mengenai perbandingan beberapa metode untuk pengelompokan data. Penelitian yang

dilakukan oleh Widiartha (2011) membandingkan metode *K-Means* dan *K-Harmonic Means*, hasilnya menunjukkan bahwa metode *K-Harmonic Means* lebih baik dibandingkan *K-Means* pada data penelitian yang terdiri dari dataset *Iris*, *Wisconsin Breast Cancer (Cancer)*, *Contraceptive Method Choice (CMC)*, *Glass*, dan *Wine* yang diambil dari *UCI Machine Learning Repository*. Penyebabnya karena *K-Harmonic Means* berhasil pada optimalisasi posisi titik pusat *cluster*. Keunggulan *K-Harmonic Means* lainnya yaitu terletak pada penggunaan parameter p yang berfungsi dalam mengatur pembobotan keanggotaan data dalam *cluster*, sehingga dapat mengurangi pengaruh inisialisasi awal dan meningkatkan kualitas *clustering* secara keseluruhan

Penelitian yang dilakukan Rudianto *et al.* (2024) membandingkan metode *K-Means* dan *K-Medoids*. Kesimpulan yang diperoleh yaitu, pada studi kasus pengelompokan provinsi berdasarkan Indeks Demokrasi Indonesia 2021 diperoleh kesimpulan bahwa metode *K-Medoids* lebih baik daripada *K-Means*. Penelitian lainnya dilakukan Jelita (2023) untuk mengelompokkan Kabupaten/Kota di Indonesia berdasarkan Unsur-Unsur Pembangun Literasi Masyarakat (UPLM) menggunakan *K-Means*. Kesimpulan yang diperoleh yaitu metode *K-Means* menghasilkan *cluster* optimal sebanyak 4 dengan karakteristiknya masing-masing di tiap *cluster*.

Penelitian ini menggunakan algoritma *K-Medoids* dan *K-Harmonic Means* untuk mengelompokkan 35 kabupaten/kota di Provinsi Jawa Tengah berdasarkan indikator penyusun IPLM tahun 2023. *Cluster* optimal dan metode terbaik dipilih menggunakan validasi *Silhouette Coefficients*. Indikator penyusun IPLM yang digunakan yaitu pemerataan layanan perpustakaan, ketercukupan koleksi,

ketercukupan tenaga perpustakaan, tingkat kunjungan masyarakat per hari, jumlah perpustakaan ber-Standar Nasional Pendidikan (SNP), dan keterlibatan masyarakat dalam kegiatan sosialisasi/ promosi. Pengelompokan ini nantinya bermanfaat bagi pemerintah maupun masyarakat setempat dalam pelaksanaan peningkatan tingkat literasi yang merata bagi seluruh daerah.

1.2 Rumusan Masalah

Rumusan masalah dalam penelitian ini dapat dituliskan sebagai berikut:

1. Bagaimana hasil perbandingan nilai *Silhouette Coefficients* pada algoritma *K-Medoids* dan *K-Harmonic Means*?
2. Bagaimana profilisasi karakteristik tiap *cluster* terhadap pengelompokan indikator penyusun IPLM di Jawa Tengah tahun 2023 berdasarkan hasil perbandingan validasi paling optimal?

1.3 Batasan Masalah

Batasan-batasan masalah dalam penelitian ini dapat dituliskan sebagai berikut:

1. Data yang digunakan adalah data indikator penyusun IPLM 1 s.d 6 di Jawa Tengah tahun 2023 menurut Kabupaten/Kota yang diambil dari *website* resmi Badan Pusat Statistik (BPS) Jawa Tengah
2. Penentuan *cluster* dan metode optimal menggunakan validasi *Silhouette Coefficients*
3. Pengolahan data menggunakan *Microsoft Excel* dan R Studio
4. Nilai toleransi yang digunakan sebesar $\varepsilon = 10^{-4}$ dan nilai maksimum iterasi sebesar 100

1.4 Tujuan Penelitian

Tujuan dalam penelitian ini dapat dituliskan sebagai berikut:

1. Menentukan *cluster* optimal berdasarkan hasil perbandingan validasi *Silhouette Coefficients* pada metode *K-Harmonic Means* dan *K-Medoids*
2. Menentukan profilisasi karakteristik tiap *cluster* terhadap pengelompokan indikator penyusun IPLM di Jawa Tengah tahun 2023 berdasarkan hasil validasi *Silhouette Coefficients* paling optimal