

BAB I PENDAHULUAN

1.1 Latar Belakang

Pengangguran merupakan salah satu masalah utama dalam pembangunan sosial-ekonomi yang dihadapi oleh banyak negara, termasuk Indonesia.. Tingkat Pengangguran Terbuka (TPT) berfungsi sebagai indikator penting yang mencerminkan proporsi penduduk usia kerja yang aktif mencari pekerjaan namun belum mendapatkan pekerjaan. Selain menjadi permasalahan ekonomi, pengangguran juga berdampak luas pada kesejahteraan sosial, seperti meningkatnya angka kemiskinan, kriminalitas, dan ketimpangan sosial. Pada kenyataannya, masih banyak daerah di Indonesia, termasuk Kota Medan, yang menghadapi masalah kesejahteraan sosial, seperti kemiskinan, anak terlantar, dan pengangguran.[1] Oleh karena itu, memahami pola distribusi pengangguran Di berbagai daerah, hal ini menjadi sangat penting untuk merancang kebijakan yang lebih efektif dan berdasarkan data.

Sebagai salah satu provinsi terbesar dan terpadat di Indonesia, Sumatera Utara memiliki karakteristik sosial-ekonomi yang sangat berbeda. Provinsi ini memiliki banyak kabupaten dan kota yang berbeda dengan kondisi geografis, pembangunan, dan dinamika sosial-ekonomi. Menurut data dari Badan Pusat Statistik (BPS), Tingkat Pengangguran Terbuka (TPT) di Sumatera Utara selama periode 2019-2023 menunjukkan variasi yang signifikan antar wilayah. Kabupaten/kota dengan akses yang lebih baik terhadap pendidikan, layanan kesehatan, dan infrastruktur cenderung memiliki tingkat pengangguran yang lebih rendah dibandingkan dengan wilayah yang kurang berkembang. Selain itu, ketimpangan distribusi lapangan kerja antara daerah perkotaan dan pedesaan turut menjadi faktor yang memperlebar kesenjangan tingkat pengangguran [1]

Sebagai contoh, wilayah perkotaan seperti Medan, yang merupakan pusat ekonomi dan perdagangan, memiliki akses yang lebih luas terhadap peluang kerja formal, tetapi juga menghadapi tantangan dalam menyerap lonjakan angkatan kerja

baru. Di sisi lain, kabupaten yang lebih terpencil seperti Nias atau Mandailing Natal sering kali memiliki keterbatasan dalam menyediakan lapangan kerja yang memadai, sehingga mendorong banyak penduduk usia kerja untuk beralih ke sektor informal atau bahkan migrasi ke kota-kota besar[2]

Melihat kondisi tersebut, penting bagi para pembuat kebijakan untuk memahami pola pengangguran yang ada, termasuk faktor-faktor yang memengaruhi tingkat pengangguran di setiap wilayah. Penggunaan analisis data berbasis klasterisasi dapat membantu mengidentifikasi kelompok-kelompok wilayah dengan karakteristik pengangguran yang serupa. Dengan demikian, pemerintah dapat merancang program dan kebijakan yang lebih terarah, seperti pelatihan keterampilan kerja berbasis kebutuhan lokal atau insentif investasi untuk menciptakan lapangan kerja di daerah yang lebih terpencil.

Oleh karena itu, diperlukan analisis mendalam menggunakan metode yang mampu mengelompokkan wilayah berdasarkan indikator-indikator sosial-ekonomi secara efektif. Metode ini tidak hanya membantu dalam mengidentifikasi disparitas antar wilayah, tetapi juga menjadi dasar bagi pengambilan keputusan yang berbasis bukti, sehingga dapat mendorong pembangunan yang lebih merata dan inklusif di Sumatera Utara. Penelitian ini berfokus pada penerapan metode klasterisasi, seperti K-Medoids, untuk mengeksplorasi pola pengangguran di provinsi ini, serta bagaimana optimasi seperti Particle Swarm Optimization (PSO) dapat meningkatkan kualitas hasil klasterisasi dalam mendukung analisis yang lebih akurat dan bermanfaat.

Data mining adalah bagian penting dari analisis data dan melibatkan kecerdasan buatan, pembelajaran mesin, statistik, dan sistem basis data untuk membuat model yang signifikan dari kumpulan data besar. Clustering (Pengelompokan) dan Classification (Klasifikasi) adalah dua jenis teknik data mining.[3]

Salah satu pendekatan analisis data yang banyak digunakan untuk mengungkap pola tersembunyi dalam data adalah klasterisasi, yang berfungsi untuk mengelompokkan data ke dalam beberapa kelompok berdasarkan kesamaan tertentu. Algoritma K-Medoids menjadi salah satu metode yang sering digunakan karena kemampuannya dalam menghasilkan klaster yang lebih stabil dibandingkan

algoritma lain seperti K-Means. Keunggulan utama K-Medoids terletak pada keandalannya dalam menangani data yang mengandung outlier, yaitu data yang secara signifikan berbeda dari mayoritas data lainnya. Hal ini menjadikan K-Medoids pilihan ideal untuk analisis data sosial-ekonomi yang kompleks dan sering kali mengandung variasi data yang ekstrem.

Namun, meskipun memiliki keunggulan tersebut, K-Medoids tidak luput dari keterbatasan. Salah satu masalah utama adalah sensitivitas algoritma ini terhadap inisialisasi medoid awal. Medoid awal, yang dipilih secara acak, memainkan peran penting dalam menentukan hasil akhir klusterisasi. Jika medoid awal tidak representatif dari dataset, hasil klusterisasi cenderung tidak optimal, menghasilkan kluster yang kurang kohesif atau terlalu tumpang tindih[4]. Keterbatasan ini dapat berdampak pada akurasi analisis, terutama ketika bekerja dengan dataset besar atau yang memiliki distribusi data yang tidak merata.

Untuk mengatasi permasalahan ini, banyak peneliti yang menggabungkan metode optimasi dalam algoritma K-Medoids. Salah satu pendekatan yang populer adalah penerapan algoritma berbasis metaheuristik, seperti Particle Swarm Optimization (PSO). PSO memberikan solusi dengan meniru perilaku sosial organisme, seperti kelompok burung atau ikan, untuk menemukan solusi terbaik dalam ruang pencarian yang kompleks. Dalam konteks K-Medoids, PSO dapat digunakan untuk memilih medoid awal yang lebih representatif, sehingga meningkatkan kualitas klusterisasi secara signifikan. Dengan integrasi ini, K-Medoids tidak hanya mampu menghasilkan kluster yang lebih akurat tetapi juga mengurangi risiko terjebak dalam optima lokal[5]

Pengembangan ini menunjukkan bahwa algoritma K-Medoids yang dioptimasi memiliki potensi besar untuk diterapkan pada berbagai jenis data, terutama data sosial-ekonomi yang sering kali heterogen dan kompleks. Dengan pemilihan medoid yang lebih optimal, proses klusterisasi tidak hanya menjadi lebih efisien tetapi juga memberikan hasil yang lebih bermakna, sehingga dapat digunakan sebagai dasar dalam pengambilan keputusan yang berbasis data.

Dalam konteks klasterisasi, PSO digunakan untuk mengoptimalkan pemilihan medoid awal pada algoritma K-Medoids. Pemilihan medoid awal yang representatif sangat penting karena secara langsung memengaruhi kualitas klasterisasi yang dihasilkan. Dengan menggunakan PSO, medoid awal dapat dipilih secara lebih cerdas berdasarkan evaluasi kohesi intra-klaster dan pemisahan antar-klaster. Hal ini tidak hanya meningkatkan keakuratan klasterisasi tetapi juga mengurangi risiko algoritma terjebak dalam solusi suboptimal atau optima lokal. Integrasi PSO dengan K-Medoids memungkinkan algoritma untuk bekerja lebih baik pada dataset yang kompleks, seperti data sosial-ekonomi, yang sering kali memiliki distribusi data yang tidak merata dan dimensi yang tinggi.

Selain itu, pendekatan ini memberikan efisiensi dalam proses klasterisasi, terutama pada dataset berskala besar. PSO mempercepat proses iterasi dengan memanfaatkan kecerdasan kolektif dari populasi partikel, yang secara simultan mengevaluasi berbagai kemungkinan solusi. Dalam berbagai studi, integrasi PSO dengan K-Medoids telah terbukti memberikan hasil yang superior dibandingkan dengan metode tradisional. Peningkatan ini terlihat dari nilai metrik evaluasi seperti Davies-Bouldin Index (DBI) yang lebih rendah, menunjukkan klasterisasi yang lebih terdefinisi dan kohesif[6]

Pendekatan ini juga membuka peluang untuk pengembangan lebih lanjut dalam bidang klasterisasi, termasuk adaptasi PSO pada variasi algoritma lain atau penerapannya dalam domain yang lebih luas. Dengan kemampuan untuk menangani dataset yang kompleks secara efektif, PSO dan integrasinya dengan algoritma seperti K-Medoids telah menjadi inovasi yang relevan dalam era big data, memberikan solusi untuk berbagai permasalahan analisis data yang semakin kompleks.

Dalam penelitian ini, Metrix Evaluasi DBI digunakan untuk membandingkan performa algoritma K-Medoids dengan dan tanpa integrasi Particle Swarm Optimization (PSO). Pendekatan ini memungkinkan penilaian yang objektif terhadap efektivitas optimasi PSO dalam meningkatkan kualitas klasterisasi. Pada algoritma K-Medoids standar, ketergantungan terhadap medoid awal yang dipilih secara acak sering kali menghasilkan klasterisasi yang kurang

optimal, ditandai dengan nilai DBI yang lebih besar. Sebaliknya, integrasi PSO membantu memilih medoid awal yang lebih representatif, sehingga meningkatkan kohesi intra-klaster dan memperjelas pemisahan antar-klaster, yang secara signifikan menurunkan nilai DBI.

Tujuan dari penelitian ini adalah untuk mengoptimasi Algoritma K-Medoids dengan Partikel Swarm Optimization (PSO) untuk melakukan klusterisasi tingkat pengangguran terbuka di kalangan penduduk Sumatera Utara yang berusia 15 tahun ke atas. Algoritma K-Medoids digunakan sebagai metode dasar untuk klusterisasi karena kemampuan untuk menangani data yang mengandung outlier dan menghasilkan klaster yang stabil. Sementara itu, integrasi PSO digunakan untuk mengatasi keterbatasan pada pemilihan medoid awal, yang seringkali memengaruhi kualitas klusterisasi. Dengan memanfaatkan kemampuan PSO untuk memilih medoid yang lebih representatif, penelitian ini diharapkan dapat menghasilkan klusterisasi yang lebih kohesif dan terdefinisi dengan baik.

Hasil penelitian ini diharapkan akan memberikan wawasan baru tentang teknik klusterisasi untuk menganalisis data sosial-ekonomi, terutama untuk mendapatkan pemahaman yang lebih baik tentang pola distribusi pengangguran di wilayah yang beragam. Hasil penelitian ini dapat menjadi landasan bagi pemerintah daerah untuk membuat program pengentasan pengangguran yang lebih fokus dengan mengidentifikasi kelompok wilayah berdasarkan karakteristik pengangguran mereka. Penelitian ini juga dapat membantu mengembangkan metode klusterisasi berbasis optimasi. Aplikasi metode ini dapat diperluas untuk memeriksa berbagai jenis data sosial lainnya.

Secara keseluruhan, penelitian ini bertujuan untuk tidak hanya memberikan kontribusi akademis dalam hal perbandingan algoritma klusterisasi tetapi juga menyediakan dasar ilmiah yang mendukung pengambilan kebijakan yang lebih efektif dan berbasis data. Dengan hasil klusterisasi yang lebih akurat, penelitian ini diharapkan dapat membantu pemerintah Sumatera Utara mengidentifikasi prioritas wilayah, mengalokasikan sumber daya secara lebih efisien, dan meningkatkan kualitas perencanaan program pembangunan yang inklusif dan berkelanjutan.

1.2 Rumusan Masalah

Rumusan masalah penelitian ini dapat digambarkan sebagai berikut berdasarkan penjelasan latar belakang yang telah diberikan:

1. Bagaimana algoritma K-Medoids dapat digunakan untuk meningkatkan akurasi klasterisasi Pengangguran Terbuka di Sumatera Utara untuk Usia 15 tahun ke atas dengan menggunakan optimasi PSO?
2. Bagaimana Perbandingan performa K-Medoids standar dengan Optimasi PSO berdasarkan metrik Evaluasi Davies-Bouldin Index (DBI)?
3. Bagaimana hasil klasterisasi yang dihasilkan dari kedua metode tersebut dapat digunakan untuk mengidentifikasi pola distribusi pengangguran dan menjadi dasar rekomendasi kebijakan yang lebih terarah di Sumatera Utara?

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Menganalisis Kinerja K-Medoids dengan Optimasi PSO: Untuk mengevaluasi dan membandingkan kinerja algoritma K-Medoids dalam klasterisasi Tingkat Pengangguran Terbuka (TPT) pada penduduk usia 15 tahun ke atas di Sumatera Utara, dengan dan tanpa optimasi menggunakan algoritma Particle Swarm Optimization (PSO), guna meningkatkan akurasi dan efisiensi klasterisasi.
2. Membandingkan dengan K-Medoids dengan Optimasi PSO: Membandingkan kinerja K-Medoids dengan optimasi PSO (Particle Swarm Optimization) terhadap K-Medoids standar, serta menilai peningkatan dalam akurasi klasterisasi dan efisiensi waktu komputasi. Penelitian ini akan mengevaluasi apakah optimasi PSO dapat menghasilkan klaster yang lebih optimal dibandingkan dengan metode standar.
3. Memberikan pemahaman baru tentang pola distribusi pengangguran di Sumatera Utara melalui hasil klasterisasi. Hasil penelitian ini diharapkan dapat membantu pemerintah daerah dalam merancang kebijakan untuk program pengurangan pengangguran yang lebih terfokus dan efektif.

1.4 Manfaat Penelitian

Peneliti mengharapkan beberapa keuntungan dari penelitian ini:

1. Peningkatan Efektivitas Kebijakan Penanggulangan Pengangguran:
Dengan menggunakan metode klasterisasi yang lebih akurat dan optimal, penelitian ini membantu pemerintah daerah dalam mengidentifikasi pola pengangguran yang lebih spesifik di setiap kabupaten/kota di Sumatera Utara. Hal ini memungkinkan penyusunan kebijakan penanggulangan pengangguran yang lebih terfokus dan tanggap terhadap kebutuhan masyarakat. Di wilayah tertentu, serta mengurangi ketidaksesuaian antara kebijakan dan kondisi lokal.
2. Rekomendasi Strategis untuk Pembangunan Ekonomi Lokal:
Hasil klasterisasi dapat menjadi acuan dalam merancang kebijakan pembangunan ekonomi yang lebih inklusif dan berbasis pada kebutuhan masing-masing wilayah. Pemahaman mengenai pola distribusi pengangguran, serta faktor-faktor terkait lainnya (seperti tingkat partisipasi angkatan kerja atau PDRB), dapat menjadi landasan bagi pengembangan strategi pembangunan yang adaptif, tepat sasaran, dan berbasis data.
3. Kontribusi Ilmiah dalam Pengembangan Algoritma Klasterisasi:
Penelitian ini tidak hanya memberikan manfaat praktis dalam konteks kebijakan publik, tetapi juga berkontribusi pada Pengembangan ilmu pengetahuan, terutama dalam bidang klasterisasi data dan optimasi algoritma. Penelitian ini dapat menjadi sumber referensi yang signifikan untuk pengembangan lebih lanjut dalam penelitian yang berfokus pada penerapan algoritma dalam analisis data sosial-ekonomi atau penggunaan algoritma optimasi seperti PSO dalam konteks big data.
4. Manfaat bagi Peneliti dan Mahasiswa:
Penelitian ini memberikan kesempatan bagi peneliti dan mahasiswa untuk mempelajari penerapan algoritma klasterisasi, teknik optimasi seperti PSO, serta metrik evaluasi seperti Davies-Bouldin Index. Hasil penelitian ini dapat digunakan sebagai referensi bagi mahasiswa atau peneliti lain yang tertarik untuk mendalami aplikasi algoritma K-Medoids atau optimasi pada data sosial-ekonomi, serta untuk penelitian terkait klasterisasi di bidang lain.

1.5 Batasan Masalah

Batasan masalah ditentukan dalam penelitian ini untuk memperjelas lingkup dan fokus analisis. Batasan Batasan berikut adalah:

1. Sumber Data

Selama periode 2019-2023, penelitian ini akan menggunakan data dari Tingkat Pengangguran Terbuka (TPT), Tingkat Partisipasi Angkatan Kerja (TPAK), dan Produk Domestik Regional Bruto per Kapita atas Dasar Harga Berlaku (PDRB) menurut Kabupaten/Kota di Sumatera Utara. Data ini dikumpulkan dari Badan Pusat Statistik (BPS) Provinsi Sumatera Utara dan mencakup 33 kabupaten/kota. Data yang digunakan mencakup elemen sosial-ekonomi yang berkaitan dengan perkembangan ekonomi regional dan tingkat pengangguran.

2. Parameter Evaluasi:

Kinerja kedua metode pengelompokan akan dievaluasi menggunakan beberapa parameter. Akurasi pengelompokan diukur dengan indeks Davies-Bouldin (DBI) untuk menilai keterpisahan dan kekompakan kluster. Waktu komputasi akan dianalisis untuk mengetahui seberapa cepat masing-masing metode mencapai konvergensi.

3. Skala Penelitian:

Dalam penelitian ini, penulis hanya akan melihat data dari kabupaten dan kota di Sumatera Utara yang terdaftar dalam data BPS dari tahun 2019 hingga 2023. Penulis tidak akan melihat data dari daerah lain. Fokus utama adalah mengidentifikasi pola pengangguran di Sumatera Utara.

4. Aspek Sosial:

Penelitian ini akan berfokus pada pengelompokan data sosial-ekonomi terkait pengangguran tanpa mempertimbangkan faktor eksternal lainnya yang mungkin memengaruhi kesejahteraan sosial, seperti kebijakan pemerintah dan faktor demografis lainnya.