

BAB I

PENDAHULUAN

A. Latar Belakang

Pengangguran adalah salah satu masalah ekonomi yang dihadapi oleh banyak negara, termasuk Indonesia. Definisi pengangguran mencakup individu dalam angkatan kerja yang menganggur tetapi aktif mencari peluang kerja atau yang belum pernah bekerja sebelumnya (BPS Indonesia, 2006).

Provinsi Sumatera Barat merupakan salah satu daerah di Indonesia dengan tingkat pengangguran yang relatif tinggi. Badan Pusat Statistik mengukur tingkat pengangguran dengan menggunakan indikator seperti Tingkat Pengangguran Terbuka di Sumatera Barat yang merupakan persentase jumlah pengangguran terhadap angkatan kerja pada tahun 2019 sebanyak 5,38%, di tahun 2020 naik menjadi 6,88%, kemudian di tahun 2021 turun menjadi 6,52%, dan pada tahun 2022 menurun menjadi 6,28% (BPS Sumbar, 2023). Berdasarkan data tersebut tingkat pengangguran terbuka di Provinsi Sumatera Barat dari tahun 2019 hingga 2022 tidak menurun secara signifikan.

Tingkat pengangguran yang tinggi atau rendah memiliki dampak yang signifikan terhadap suatu wilayah (Itang, 2015). Oleh sebab itu, diperlukan penelitian yang lebih mendalam mengenai keterkaitan antara tingkat pengangguran dan faktor-faktor yang memengaruhi perubahan tingkat tersebut. Metode pengelompokan yang diterapkan dalam penelitian ini adalah analisis kluster atau *Cluster Analysis* (Akramunnisa dan Fajriani, 2020).

Analisis Kluster adalah metode multivariat yang dimaksudkan untuk membagi objek menjadi beberapa kategori, sehingga perbedaan antara satu kelompok dan kelompok lainnya dapat terlihat dengan jelas (Nafisah dan Chandra 2017). Pengelompokan merupakan suatu proses yang bertujuan untuk mengelompokkan data ke dalam beberapa kluster, dimana data dalam satu

klaster memiliki tingkat kesamaan yang tinggi, sedangkan data antar klaster mempunyai tingkat kesamaan rendah (Nurmin, Hayati, dan Goejantoro, 2022).

Terdapat berbagai metode dalam pengelompokan yang dapat digunakan untuk mengelompokkan data berdasarkan kemiripan atau karakteristik tertentu, yaitu *K-Means*, *K-Medoids*, *Gaussian Mixture Model* (GMM) dan *Fuzzy C-Means* (FCM). Metode *K-Means* dan *K-Medoids* adalah metode yang cenderung sensitif terhadap outlier dan memerlukan penentuan jumlah klaster sebelumnya. Sedangkan *Gaussian Mixture Model* sulit menangani data yang tidak terdistribusi normal (Syukron dkk, 2022). Kekurangan dari metode ini dapat diatasi dengan metode *Fuzzy C-Means*.

Fuzzy C-Means adalah metode pengelompokan yang memberikan fleksibilitas dalam menentukan tingkat keanggotaan dalam suatu klaster dan memungkinkan satu titik data untuk tergabung dalam beberapa klaster dengan tingkat keanggotaan yang spesifik (Bezdek, 1981). Konsep dasar *Fuzzy C-Means* (FCM) adalah penentuan awal pusat klaster, yang berfungsi untuk mengelompokkan setiap titik data ke dalam kategori tertentu. Penentuan pusat klaster dan nilai keanggotaan akan dilakukan secara berulang hingga didapatkan pusat klaster yang paling optimal. Pusat klaster akan dipindahkan ke lokasi yang lebih tepat, karena keadaan di pusat klaster serta setiap titik data berdasarkan tingkat keanggotaannya masih belum mencapai akurasi yang memadai (Nurmin, Hayati, dan Goejantoro, 2022).

Berdasarkan penjelasan di atas, penelitian ini bertujuan untuk menganalisis penerapan metode *Fuzzy C-Means* dalam mengelompokkan tingkat pengangguran di Provinsi Sumatera Barat. Sehingga, judul penelitian ini adalah **“Analisis Klaster Menggunakan Algoritma *Fuzzy C-Means* untuk Tingkat Pengangguran di Provinsi Sumatera Barat”**.

B. Rumusan Masalah

Berdasarkan latar belakang yang penulis paparkan, maka perlu diberikan rumusan masalah yaitu bagaimana tingkat pengangguran di Provinsi Sumatera

Barat dapat dikelompokkan menggunakan algoritma *Fuzzy C-Means*.

C. Batasan Masalah

Batasan masalah sangat penting untuk mencegah bias dalam penelitian. Batasan yang ditentukan dalam studi ini adalah sebagai berikut:

1. Data pengangguran yang dianalisis berasal dari data sekunder yang didapatkan melalui Badan Pusat Statistik Sumatera Barat tahun 2023.
2. Variabel yang dianalisis meliputi Tingkat Partisipasi Angkatan Kerja (TPAK), Produk Domestik Regional Bruto (PDRB), Indeks Pembangunan Manusia (IPM), dan Penanaman Modal Dalam Negeri (PMDN).
3. Metode yang digunakan adalah algoritma *Fuzzy C-Means*.

D. Tujuan Masalah

Berdasarkan isu yang telah diuraikan sebelumnya, sasaran dari penelitian ini adalah untuk mengkategorikan lokasi di Provinsi Sumatera Barat berdasarkan tingkat pengangguran dengan menggunakan algoritma *Fuzzy C-Means*.

E. Manfaat Penelitian

Manfaat yang didapatkan dari penelitian ini adalah sebagai berikut:

1. Menyajikan pemahaman yang lebih mendalam mengenai penerapan algoritma *Fuzzy C-Means* dalam analisis tingkat pengangguran.
2. Menambah literatur dalam bidang analisis data, khususnya metode klusterisasi untuk data sosial-ekonomi.
3. Mempermudah proses pengambilan keputusan dalam program penanganan pengangguran.