

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pengangguran terbuka merupakan bagian dari angkatan kerja menganggur atau sedang mencari pekerjaan (Ardian et al., 2022). Tingkat pengangguran terbuka (TPT) merupakan indikator untuk mengukur angkatan kerja yang tidak terserap di pasar tenaga kerja dan menggambarkan kurang dimanfaatkannya pasokan tenaga kerja. Hasil TPT Sakernas sebesar 5,32% pada Agustus 2023. Artinya, dari 100 orang yang bekerja, ada sekitar 5 orang yang menganggur. Pada Agustus 2023, TPT turun sebesar 0,54% dibandingkan Agustus 2022. Pada Agustus 2023, TPT laki-laki sebesar 5,42%, menunjukkan trend meningkat dibandingkan TPT perempuan yang sebesar 5,15%. TPT perempuan dan laki-laki sama dengan TPT nasional turun 0,51% dibandingkan Agustus 2022 dan 0,60%. Berdasarkan wilayah pemukiman, TPT di perkotaan (6,40%) jauh lebih tinggi dibandingkan TPT di perdesaan (3,88%). Dibandingkan Agustus 2022, TPT DKI Jakarta tercatat mengalami penurunan sebesar 1,34%. Sedangkan TPT perdesaan mencatatkan kenaikan sebesar 0,45%. Pada bulan Agustus 2023, TPT kelompok usia remaja (15-24 tahun) tertinggi adalah sebesar 19,40%. Sedangkan TPT penduduk lanjut usia (60 tahun ke atas) paling rendah yaitu 1,28%. Struktur TPT menurut kelompok umur ini juga sesuai dengan tahun sebelumnya. Dibandingkan Agustus 2022, terjadi penurunan TPT pada semua kelompok umur, dengan penurunan terbesar pada kelompok umur diatas 60 tahun sebesar 1,57% (Badan Pusat Statistik, 2023).

Pendidikan memegang peranan penting dalam pembangunan negara. Hal ini sesuai dengan Pasal 27 UUD 1945 yang menyatakan bahwa pendidikan adalah penyiapan warga negara yang baik, yaitu warga negara yang mengetahui hak dan kewajibannya (Nazarpour et al., 2022). Angka Partisipasi Sekolah (APS) merupakan ukuran kemampuan sistem pendidikan dalam mengakomodasi penduduk usia sekolah. Angka ini memperhitungkan perubahan demografi, khususnya di kalangan generasi muda. Langkah-langkah yang banyak digunakan dalam pendidikan. Pertumbuhan siswa menunjukkan perubahan jumlah siswa yang dapat ditampung pada setiap tingkat kelas (Budianto, 2023). Angka Partisipasi

Sekolah (APS) merupakan ukuran daya serap penduduk terhadap sistem pendidikan, tahun kehidupan sekolah, dan merupakan indikator yang menentukan kemajuan pendidikan di suatu bidang tertentu(Nazarpour et al., 2022).

Penelitian yang dilakukan oleh(Rahmayani & Andriyani, 2022) menjelaskan antara keterkaitan variabel angka partisipasi sekolah dan variabel tingkat pengangguran terbuka, dengan kesimpulan bahwa variabel angka partisipasi sekolah tidak berpengaruh secara signifikan terhadap tingkat kemiskinan di provinsi Sumatera, akan tetapi variabel tingkat pengangguran terbuka berpengaruh secara signifikan terhadap tingkat kemiskinan terhadap tingkat kemiskinan di provinsi Sumatera.

Pada penelitian yang dilakukan juga oleh(Nugroho, 2024) juga menjelaskan tentang keterkaitan antara variabel angka partisipasi sekolah dan variabel tingkat pengangguran terbuka, dengan kesimpulan angka partisipasi sekolah tidak berpengaruh besar terhadap tingkat pengangguran terbuka yang disebabkan oleh lapangan usaha yang berada pada wilayah Se- Eks Karesidenan Pati hanya membutuhkan tenaga kerja kasar dengan lulusan SMA dan dapat dilatih oleh perusahaan dalam tata cara penggunaan alat-alat yang diperlukan dalam proses produksi.

Data mining merupakan proses berulang dimana kemajuannya didasarkan pada penemuan, baik melalui metode otomatis ataupun manual. *Data mining* sangat berguna dalam skenario analisis eksplorasi ketika tidak ada gagasan pasti. *Data mining* merupakan proses pencarian informasi baru, berharga, dan tidak sepele ini merupakan proses kerja sama antara manusia dengan komputer(Kantardzic, 2020). *Data mining* juga disebut penemuan pengetahuan dalam basis data, mengacu pada proses tidak sepele dalam proses ekstraksi informasi yang tersembunyi, yang sebelumnya tidak diketahui dan berpotensi berguna(seperti aturan pengetahuan, batasan, keteraturan) dari basis data(Chen et al., 1996). Adapun beberapa fungsi *data mining* yang diantaranya, *Clustering*, klasifikasi, asosiasi, estimasi, dan prediksi(Patty, 2010).

Clustering merupakan proses mengeksploitasi struktur yang mendasari distribusi data dan menentukan aturan untuk mengelompokkan data dengan karakteristik yang serupa(Ahmed et al., 2020). *Clustering* merupakan alat yang

berguna dalam ilmu data. metode untuk menemukan struktur kluster dalam kumpulan data yang dicirikan oleh kesamaan terbesar dalam *Cluster* yang sama dan ketidaksamaan terbesar di antara *Cluster* yang berbeda. Hirarki merupakan metode pengelompokan paling awal yang digunakan oleh ahli biologi dan ilmuwan sosial, sedangkan analisis kluster menjadi cabang dari analisis multivariat statistik (Sinaga & Yang, 2020). adapun beberapa jenis metode dalam *Clustering* yang dapat digunakan *d-based Clustering*, *density-based Clustering*, *distribution-based Clustering*, *hierarchical Clustering* atau *connectivity-based Cluster*.

K-means merupakan salah satu algoritma *Unsupervised learning* sederhana yang dapat memecahkan masalah pengelompokan. Metode ini secara sederhana dan mudah mengklasifikasikan kumpulan data tertentu berdasarkan jumlah *Cluster* (dengan asumsi k *Cluster*) yang ditentukan secara spekulatif. (Kodinariya & Makwana, 2013). Algoritma *K-means* merupakan algoritma pengelompokan berulang yang sederhana. Dengan menggunakan jarak sebagai metrik dan diberikan K kelas dalam kumpulan data, hitung jarak rata-rata, memberikan *centroid* awal, dengan setiap kelas dijelaskan oleh *centroid*. Untuk set data X berisi titik n data multidimensi dan kategori K akan dibagi dengan jarak, sebagai indeks kemiripan dengan target pengelompokan untuk meminimalkan jumlah kuadrat dari berbagai jenis (Yuan & Yang, 2019). Sejumlah penelitian telah menerapkan metode *K-Means Clustering* telah diterapkan di berbagai bidang. Dalam bidang kesehatan digunakan untuk klusterisasi data penanganan dan pelayanan kesehatan masyarakat (Tarigan et al., 2022). Di bidang ekonomi digunakan untuk pengelompokan bidang laju pertumbuhan ekonomi di Indonesia (Siregar, 2019). Di bidang pendidikan digunakan untuk penentuan siswa kelas unggulan (Sulistiyawati & Supriyanto, 2021). Oleh karena itu, metode *K-Means Clustering* digunakan dalam penelitian ini.

Pada penelitian (Maliqi & Falgenti, 2023) menjelaskan tentang *Clustering* menggunakan metode *K-Means* dengan hasil, klusterisasi menggunakan *k-mean Clustering* menunjukkan bahwa pandemi Covid-19 sangat mempengaruhi kenaikan tingkat pengangguran terbuka pada provinsi di Indonesia.

Penelitian lain yang dilakukan oleh (Desdianti et al., 2024) menjelaskan tentang Analisis *K-Means* menggunakan metode Elbow dengan hasil, Hasil yang

diperoleh dengan mengelompokkan PDRB, pertumbuhan ekonomi, tingkat partisipasi angkatan kerja, rata-rata lama sekolah, tingkat pengangguran Kabupaten/Kota dan indeks pembangunan manusia menghasilkan 4 *Cluster* yang terdiri dari *Cluster* 1 memiliki tingkat pengangguran sangat tinggi, *Cluster* 2 dengan karakteristik tingkat pengangguran tinggi, *Cluster* 3 dengan karakteristik tingkat pengangguran sedang dan *Cluster* 4 memiliki karakteristik tingkat pengangguran rendah.

Penelitian yang juga dilakukan oleh (Sita Muharni, 2022) menjelaskan tentang penerapan metode *K-Means Clustering* pada data tingkat pengangguran terbuka Tahun 2016-2018 dan 2019-2021 dengan hasil, didapatkan pengetahuan baru hanya satu provinsi yang naik *Cluster* 1 yaitu Provinsi Riau, Pada tahun 2016- 2018 provinsi ini masuk *Cluster* provinsi dengan tingkat pengangguran yang tinggi, pada tahun 2019-2021 provinsi Riau naik peringkat menjadi provinsi yang memiliki tingkat pengangguran yang rendah.

Berdasar pada latar belakang tersebut, maka pada penelitian ini menggunakan Algoritma *K-Means* dengan *dataset* gabungan antara *dataset* Tingkat Pengangguran Terbuka dan *dataset* Angka Partisipasi Sekolah dengan rentang waktu 2019-2023.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka dapat dirumuskan bagaimana mengelompokkan Tingkat Pengangguran Terbuka dan Angka Partisipasi Sekolah menurut provinsi dengan menggunakan Algoritma *K-Means*?

1.3 Tujuan Penelitian

Berdasarkan pemaparan latar belakang diatas maka tujuan dari penelitian ini untuk mengimplementasikan Algoritma *K-Means Clustering* dalam pengelompokan Tingkat Pengangguran Terbuka dan Angka Partisipasi Sekolah menurut provinsi.

1.4 Batasan Masalah

Dalam menyelesaikan masalah perlu dibuat suatu batasan masalah agar tidak menyimpang dari pokok permasalahan. Oleh karena itu, batasan yang dibuat dalam permasalahan ini adalah sebagai berikut :

1. Menggunakan Algoritma *K-Means Clustering*.
2. Menggunakan gabungan 2 *dataset* yaitu Tingkat Pengangguran Terbuka dan Angka Partisipasi Sekolah sebagai data pendamping.
3. Data diperoleh dari website bps.go.id.
4. Penelitian ini menggunakan rentang usia antara 19-23 tahun.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut :

1. Dapat digunakan sebagai bahan evaluasi dan acuan dalam mengatasi permasalahan Tingkat Pengangguran Terbuka dan Angka Partisipasi Sekolah guna mempersiapkan generasi emas.
2. Pihak atau Instansi terkait dapat melakukan langkah-langkah untuk mengurangi Tingkat Pengangguran Terbuka di setiap provinsi di Indonesia.

UNUGIRI