

ALGORITMA K-MEANS UNTUK MENGELOMPOKKAN TINGKAT PENGANGGURAN TERBUKA (TPT) MENURUT PROVINSI DI INDONESIA

Agus Bima Saputra¹⁾, Ucta Pradema Sanjaya²⁾, Ita Aristia Sa'ida³⁾

^{1), 2) 3)} Program Studi Teknik Infomatika, Fakultas Sains dan Teknologi, Universitas Nahdlatul Ulama Sunan Giri Bojonegoro Jl Ahmad Yani No. 10, Jambean, Sukorejo, Kec. Bojonegoro Kab. Bojonegoro
Email : agusbimasaputra01@gmail.com¹⁾, uctapradema@unugiri.ac.id²⁾, itaaristia@unugiri.ac.id³⁾

ABSTRACT

Unemployment is one of the main problems faced by many countries, including Indonesia. The Open Unemployment Rate (OER) is an important indicator used to measure the amount of labor force that is not absorbed in the labor market. This research aims to Cluster the provinces in Indonesia based on unemployment rate and school enrollment rate, so as to provide a clearer picture of the distribution of unemployment in different regions: The study identified three main Clusters: Cluster 1: Provinces with high unemployment rates. Cluster 2: Provinces with a medium unemployment rate. Cluster 3: Provinces with low unemployment rates. Distribution: The Clustering results show that 13 provinces are included in Cluster 1, 18 provinces in Cluster 2, and 3 provinces in Cluster 3. This study found that the K-Means algorithm is effective in Clustering provinces based on TPT and school enrollment rates. The Clustering results show significant variation between provinces, with some provinces having higher unemployment rates and lower school enrollment rates. This study successfully Clustered Indonesian provinces based on unemployment and school enrollment rates using the K-Means algorithm. The Clustering results provide valuable insights into the distribution of unemployment in Indonesia and can be used as a basis for more effective policy making.

Keywords : TPT, APS, Data Mining, Clustering, K-means Algorithm

ABSTRAK

Pengangguran merupakan salah satu masalah utama yang dihadapi oleh banyak negara, termasuk Indonesia. Tingkat Pengangguran Terbuka (TPT) adalah indikator penting yang digunakan untuk mengukur jumlah angkatan kerja yang tidak terserap di pasar tenaga kerja. Penelitian ini bertujuan untuk mengelompokkan provinsi-provinsi di Indonesia berdasarkan tingkat pengangguran dan angka partisipasi sekolah, sehingga dapat memberikan gambaran yang lebih jelas tentang distribusi pengangguran di berbagai wilayah. Cluster: Penelitian ini berhasil mengidentifikasi tiga Cluster utama: Cluster 1: Provinsi dengan tingkat pengangguran tinggi. Cluster 2: Provinsi dengan tingkat pengangguran sedang. Cluster 3: Provinsi dengan tingkat pengangguran rendah. Distribusi: Hasil Clustering menunjukkan bahwa 13 provinsi termasuk dalam Cluster 1, 18 provinsi dalam Cluster 2, dan 3 provinsi dalam Cluster 3. Penelitian ini menemukan bahwa algoritma K-Means efektif dalam mengelompokkan provinsi berdasarkan TPT dan angka partisipasi sekolah (APS). Hasil Clustering menunjukkan adanya variasi yang signifikan antar provinsi, dengan beberapa provinsi memiliki tingkat pengangguran yang lebih tinggi dan partisipasi sekolah yang lebih rendah dibandingkan provinsi lainnya. Penelitian ini berhasil mengelompokkan provinsi-provinsi di Indonesia berdasarkan TPT dan angka partisipasi sekolah menggunakan algoritma K-Means. Hasil Clustering memberikan wawasan yang berharga mengenai distribusi pengangguran di Indonesia dan dapat digunakan sebagai dasar dalam pengambilan kebijakan yang lebih efektif.

Kata Kunci : TPT, APS, Data Mining, Clustering, Algoritma K-means

1. Pendahuluan

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).

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). Adapun beberapa fungsi data mining yang diantaranya, Clustering, klasifikasi, asosiasi, estimasi, dan prediksi (Patty, 2010). Clustering merupakan proses mengeksplorasi 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 menyelesaikan permasalahan pengelompokan data. 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.

Penelitian yang menggabungkan antara variabel TPT dan APS juga dilakukan oleh Nugroho dkk (Nugroho, 2024) Pada penelitian ini telah dijelaskan bahwa angka partisipasi sekolah tidak berpengaruh besar terhadap tingkat pengangguran terbuka, hal ini disebabkan karena usaha-usaha yang berada di wilayah Eks Karesidenan Pati hanya membutuhkan tenaga kerja kasar dengan lulusan SMA dan dapat dilatih oleh perusahaan mengenai cara penggunaan alat-alat yang dibutuhkan dalam proses produksi.

Rujukan kedua penelitian yang telah dilakukan oleh (Maliqi & Falgenti, 2023) yang 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.

Rujukan ketiga penelitian yang dilaksanakan oleh (Desdianti et al., 2024) menjelaskan tentang Analisis K-Means dengan menggunakan metode Elbow dengan hasil, Analisis K-Means menggunakan metode Elbow mengidentifikasi empat kluster berdasarkan berbagai indikator, termasuk PDRB, pertumbuhan ekonomi, tingkat partisipasi angkatan kerja, rata-rata lama sekolah, tingkat pengangguran kabupaten/kota, dan indeks pembangunan manusia. Hasil analisis ini menunjukkan empat kluster dengan karakteristik sebagai berikut: Cluster 1 memiliki tingkat pengangguran yang sangat tinggi, Cluster 2 memiliki tingkat pengangguran yang

tinggi, Cluster 3 memiliki tingkat pengangguran yang sedang, dan Cluster 4 memiliki tingkat pengangguran yang rendah.

Rujukan keempat penelitian yang dilakukan oleh (Sita Muharni, 2022) Metode K-Means clustering diterapkan pada data tingkat pengangguran terbuka dari tahun 2016-2018 dan 2019-2021. Hasil analisis ini mengungkapkan pengetahuan baru bahwa hanya satu provinsi yang naik ke Cluster 1, yaitu Provinsi Riau. Pada tahun 2016-2018, provinsi ini termasuk dalam kluster dengan tingkat pengangguran yang tinggi. Namun, pada tahun 2019-2021, Provinsi Riau masuk ke dalam kluster provinsi dengan tingkat pengangguran yang rendah.

Rujukan kelima penelitian dengan metode K-Means (Ansen, 2022) tentang pengelompokan data pengangguran di provinsi Jawa Barat dengan mendapatkan hasil 2 cluster yaitu cluster tinggi (C1) sebanyak 13 kabupaten/kota dan cluster rendah (C2) sebanyak 14 kabupaten/kota.

Rujukan terakhir penelitian dengan menggunakan metode K-Means juga dilakukan oleh (Pastia & Dikananda, 2023) Penelitian ini melakukan clustering terhadap data pengangguran terbuka di Provinsi Jawa Barat. Hasilnya menunjukkan dua kluster, Cluster 0, yang terdiri dari daerah dengan tingkat pengangguran rendah, mencakup Kabupaten Purwakarta, Kabupaten Karawang, Kabupaten Bekasi, Kabupaten Bandung Barat, Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, Kota Tasikmalaya, dan Kota Banjar. Cluster 1, yang terdiri dari daerah dengan tingkat pengangguran tinggi, meliputi Kabupaten Bogor, Kabupaten Sukabumi, Kabupaten Cianjur, Kabupaten Bandung, Kabupaten Garut, Kabupaten Tasikmalaya, Kabupaten Ciamis, Kabupaten Kuningan, Kabupaten Cirebon, Kabupaten Majalengka, Kabupaten Sumedang, Kabupaten Indramayu, dan Kabupaten Subang.

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

2. Metode Penelitian

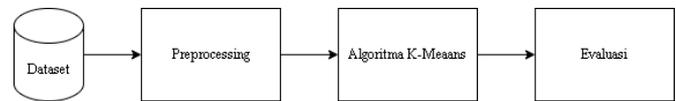
Metode yang digunakan oleh peneliti dalam penelitian ini digunakan untuk Clustering dataset Tingkat Pengangguran Terbuka tahun 2019-2023 dan Angka Partisipasi Sekolah dengan rentang usia 19-23 tahun sebagai data pendamping.

2.1 Algoritma K-Means

K-means Clustering merupakan algoritma unsupervised learning yang biasa digunakan untuk mengelompokan atau mengklasterisasi data dan pengenalan pola. K-means juga merupakan algoritma

untuk Cluster n objek berdasarkan atribut menjadi κ partisi, dimana $\kappa < n$ (Wahidiah, 2021).

Berdasarkan metode penelitian maka algoritma K-Means ditampilkan melalui skema metode pada gambar dibawah. Melalui gambaran dasar pengujian dataset dengan algoritma K-Means.



Gambar 2. 1 Pengujian Algoritma K-Means

Penjelasan skema metode pada gambar adalah sebagai berikut :

a. Dataset

Tahap awal dalam penelitian ini adalah mendapatkan dataset untuk diteliti. Dan pada penelitian ini menggunakan dataset yang diperoleh dari website bps.go.id. dan datanya bersifat public atau bisa diakses oleh masyarakat.

Tabel 2. 1 Dataset Tingkat Pengangguran Terbuka 2019-2023

No	Provinsi	Februari	Agustus	Februari	Agustus
1	Aceh	5,48	6,17	5,75	6,03
2	Sumut	5,57	5,39	5,24	5,89
3	Sumba	5,38	5,38	5,9	5,94
4	Riau	5,36	5,76	4,25	4,23
....
33	Papua Barat	5,81	6,43	5,53	5,38
34	Papua	3,22	3,51	3,49	2,67

Tabel 2. 2 Dataset Angka Partisipasi Sekolah 2019-2023

No	Provinsi	Aps Tahun	Aps Tahun
1	Aceh	34,84	35,24
2	Sumut	27,64	31,08
3	Sumbar	38,79	42,19
4	Riau	28,49	32,32
....
....
33	Papua Barat	35,05	35,74
34	Papua	24,41	26,06

b. Pre-Processing

Pada tahapan ini peneliti menggabungkan antara dua dataset yaitu Tingkat Pengangguran Terbuka dan Angka Partisipasi Sekolah. Juga pada tahapan ini merubah atribut data menjadi X1, X2, X3, X4, X5, Y1,Y2, Y3, Y4, Y5, Z1, Z2, Z3, Z4, dan Z5

Tabel 2. 3 Data setelah di preprocessing

No	Provinsi	X1	X3	X4	Z3	Z4	Z5
1	Aceh	34,8	5,4	6,1		5,7	6,0
		4	8	7	...	35,24	5	3
2	Sumut	27,6	5,5	5,3		5,2	5,8
		4	7	9	...	31,08	4	9
3	Sumbar	38,7	5,3	5,3			5,9
		9	8	8	...	42,19	5,9	4
4	Riau	28,4	5,3	5,7		4,2	4,2
		9	6	6	...	32,32	5	3
.....
33	Papua Barat	35,0	5,8	6,4		5,5	5,3
		5	1	3	...	35,74	3	8
34	Papua	24,4	3,2	3,5		3,4	2,6
		1	2	1	...	26,06	9	7

c. Clustering menggunakan algoritma K-Means

Tahapan ini dataset yang telah melewati proses sebelumnya akan di Clustering menggunakan algoritma K-Means.

d. Evaluasi

Pada tahap ini dilakukan analisis dan evaluasi terhadap kinerja algoritma K-Means terhadap dataset yang digunakan.

Adapun beberapa langkah-langkah algoritma K-means adalah sebagai berikut :

1) Penentuan jumlah kluster

Dalam penentuan jumlah kluster sesuai dengan masalah yang diteliti. Penelitian ini mengacu pada 3 Cluster, meliputi Cluster C1, C2, dan C3.

2) Penentuan kluster awal

Untuk menentukan nilai n suatu pusat kluster awal dilakukan secara acak guna mewakili masukan dan urutan. Pusat awal kluster didapat dari data sendiri bukan dengan menentukan titik baru yaitu dengan acak pusat awal dari data kluster

Tabel 2. 4 Penentuan kluster awal

Cluster	Provinsi	X1	X2	Z4	Z5
C1	SUMATERA					
	UTARA	27,64	5,57	5,24	5,89
C2	JAWA					
	TIMUR	26,33	3,77	4,33	4,88
C3	PAPUA					
	BARAT	35,05	5,81	5,53	5,38

3) Perhitungan antara jarak dan pusat kluster

Untuk menghitung jarak antar data dan pusat kluster dinyatakan dengan rumus Eulidian distance sebagai berikut :

$$D = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2}$$

Dimana :

De merupakan Euclidean Distance

i merupakan banyaknya objek

(x, y) merupakan koordinat objek

(s, t) merupakan koordinat centroid

4) Mengelompokkan data

Jarak yang telah dihitung akan dibandingkan dan dipilih jarak terdekat antara data dengan pusat kluster, hasil dari perhitungan ini menunjukkan bahwa data tersebut berada dalam satu kelompok dengan pusat kluster terdekat. Dalam pengelompokan, pilihlah hasil nilai jarak yang terkecil, kemudian kelompokkan data dengan pusat cluster yang memiliki jarak terkecil.

5) Menentukan pusat kluster baru

Pusat kluster yang baru digunakan untuk melakukan iterasi berikutnya, iterasi ini akan terus berlanjut jika hasil yang diperoleh belum konvergen. Adapun cara menentukan pusat kluster adalah dengan mencari anggota dari masing-masing Cluster, kemudian menghitung pusat baru dinyatakan dengan rumus sebagai berikut:

$$\mu_k = \frac{1}{N_k} \sum_{q=1}^{N_k} xq$$

Dimana :

μ_k = titik pusat dari kelompok ke-K

N_k = banyaknya data pada kelompok ke-K

xq = data ke-q pada kelompok ke-K

6) Menentukan hasil kluster yang sudah konvergen

Dalam K-Means Clustering, data akan mengalami beberapa iterasi. Kluster konvergen adalah data yang dihasilkan iterasi baru telah sama dengan iterasi sebelumnya dan sudah tidak mengalami perubahan. Dalam pencarian kluster konvergen yaitu membagi dengan Within Cluster Variation (WCV) dibagi Between Cluster Variation (BCV). Nilai WCV didapat dari nilai

minimal jarak dengan pusat Cluster dipangkat 2 sedangkan nilai BCV didapat dari jumlah perhitungan jarak antar pusat Cluster.

WCV	4901,46
-----	---------

3. Pembahasan dan Hasil

Hasil pada iterasi 1 ini terdapat 5 provinsi yang termasuk ke dalam C1, 18 provinsi masuk ke dalam C2, dan 11 provinsi yang termasuk ke C3.

Tabel 3. 1 Hasil Perhitungan Iterasi Ke 1

C1	C2	C3	Cluster	Min Jarak	Min Jarak ²
12,74	17,898	2,087	C3	2,09	4,36
0,00	6,008	13,549	C1	0,00	0,00
23,61	28,697	10,885	C3	10,89	118,48
4,34	8,014	11,674	C1	4,34	18,88
9,84	4,391	22,447	C2	4,39	19,28
20,48	15,351	33,576	C2	15,35	235,64
12,15	14,835	9,345	C3	9,34	87,32
15,20	9,841	28,039	C2	9,84	96,85
23,63	18,359	36,730	C2	18,36	337,07
17,49	15,392	29,632	C2	15,39	236,91
8,19	9,179	19,355	C1	8,19	67,12
13,11	12,929	24,407	C2	12,93	167,17
9,82	5,112	23,018	C2	5,11	26,13
61,26	65,796	48,282	C3	48,28	2331,19
6,01	0,000	18,709	C2	0,00	0,00
15,09	14,034	26,680	C2	14,03	196,94
10,51	12,554	11,631	C1	10,51	110,56
9,20	4,978	19,849	C2	4,98	24,79
10,13	12,136	11,067	C1	10,13	102,60
7,74	2,701	20,748	C2	2,70	7,30
10,80	5,152	23,076	C2	5,15	26,54
6,13	2,462	17,694	C2	2,46	6,06
10,18	15,597	4,206	C3	4,21	17,69
5,86	4,874	17,072	C2	4,87	23,76
10,49	7,463	23,444	C2	7,46	55,70
8,00	7,873	14,570	C2	7,87	61,99
15,73	20,513	3,774	C3	3,77	14,24
12,33	15,378	8,406	C3	8,41	70,66
14,52	17,626	8,691	C3	8,69	75,53
13,74	8,256	24,812	C2	8,26	68,16
26,69	31,923	13,717	C3	13,72	188,15
8,24	12,040	7,860	C3	7,86	61,77
13,55	18,709	0,000	C3	0,00	0,00
11,88	6,528	23,575	C2	6,53	42,62

Melakukan Perhitungan jarak antar pusat cluster untuk memperoleh nilai Between Cluster Variation (BCV).

Tabel 3. 2 Hasil Perhitungan antara Jarak dan Pusat Cluster Iterasi Ke 1

C1	C2	6,0084
C1	C3	13,5483
C2	C3	18,7092
BCV		38,27

Menghitung nilai rasio dengan membandingkan nilai rasio dengan membagi nilai Between Cluster Variation (BCV) dengan Within Cluster Variation (WCV).

$$BCV/WCV = 38,27/4901,46$$

$$Rasio = 0,01$$

Setelah dilakukan sebanyak 8 iterasi perhitungan sudah tidak berubah atau sudah konvergen maka hasil perhitungannya adalah sebagai berikut :

Tabel 3. 3 Hasil Perhitungan Iterasi Ke 8

C1	C2	C3	Cluster	Min Jarak	Min Jarak ²
6,42	22,585	24,508	C1	6,42	41,16
9,09	10,198	36,953	C1	9,09	82,70
16,22	33,600	13,665	C3	13,66	186,72
6,37	12,605	34,765	C1	6,37	40,57
16,84	2,866	45,703	C2	2,87	8,21
28,17	10,366	57,046	C2	10,37	107,45
4,16	20,076	28,471	C1	4,16	17,27
22,47	5,445	51,375	C2	5,45	29,65
31,22	13,544	60,090	C2	13,54	183,44
26,20	11,118	52,959	C2	11,12	123,62
16,31	9,219	42,446	C2	9,22	84,99
21,74	10,364	47,621	C2	10,36	107,41
17,83	1,229	46,445	C2	1,23	1,51
52,87	70,948	24,633	C3	24,63	606,81
13,06	5,494	41,815	C2	5,49	30,18
23,77	10,496	50,056	C2	10,50	110,17
5,82	17,846	31,099	C1	5,82	33,92
13,49	7,878	42,270	C2	7,88	62,06
4,46	17,367	31,329	C1	4,46	19,88
15,37	2,995	44,115	C2	2,99	8,97
17,21	4,168	46,107	C2	4,17	17,37
11,69	7,253	40,577	C2	7,25	52,60
6,07	19,940	27,530	C1	6,07	36,85

11,71	8,219	40,143	C2	8,22	67,55
19,09	3,688	47,090	C2	3,69	13,60
7,86	12,730	36,242	C1	7,86	61,78
7,84	25,419	21,599	C1	7,84	61,53
3,82	20,532	27,640	C1	3,82	14,61
5,78	22,779	26,006	C1	5,78	33,46
18,44	8,150	47,043	C2	8,15	66,42
19,60	36,608	11,515	C3	11,52	132,60
2,12	16,937	30,266	C1	2,12	4,49
6,93	23,306	23,886	C1	6,93	48,09
17,45	6,059	46,320	C2	6,06	36,72
				WCV	2534,36

Hasil perhitungan jarak antar pusat Cluster pada iterasi ke 8

Tabel 3. 4 Hasil perhitungan antara jarak dan pusat cluster pada iterasi ke 8

C1	C2	18,1364
C1	C3	28,9595
C2	C3	46,8209
BCV		93,92

Menghitung nilai rasio dengan membandingkan nilai rasio dengan membagi nilai Between Cluster Variation (BCV) dengan Within Cluster Variation (WCV). Setelah membandingkan hasil rasio antara iterasi ke 7 dan ke 8 hasilnya sudah tidak perlu perulangan lagi atau hasil iterasi sudah konvergen.

$$BCV/WCV = 93,92/2534,36$$

$$\text{Rasio} = 0,004$$

4. Kesimpulan

Berdasarkan uraian yang telah dipaparkan dalam bab sebelumnya, maka dapat diambil kesimpulan bahwa, setelah dilakukan proses perhitungan menggunakan algoritma K-Means, terdapat 13 provinsi yang termasuk ke dalam kelompok C1 yang diantaranya adalah provinsi Aceh, Sumatera Utara, Riau, Bengkulu, Bali, Nusa Tenggara Timur, Kalimantan Timur, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Maluku Utara, Papua Barat. 18 provinsi masuk ke dalam kelompok C2 yang diantaranya sebagai berikut Jambi, Sumatera Selatan, Lampung, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, Jawa Barat, Jawa Tengah, Jawa Timur, Banten, Nusa Tenggara Barat, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Utara, Sulawesi Utara, Sulawesi Barat, dan Papua. 3 provinsi yang tergabung dalam kelompok C3 yang diantaranya adalah Provinsi Sumatera Barat, Daerah Istimewa Yogyakarta, dan Maluku.

Penelitian ini diharapkan sebagai referensi bagi instansi terkait untuk mengambil keputusan mengatasi

masalah pengangguran yang masih masif di Indonesia. Untuk penelitian selanjutnya perlu adanya pembaruan dalam penggunaan metode agar ada improvisasi sehingga hasil daripada perhitungan lebih bervariasi.

Daftar Pustaka

- Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics (Switzerland)*, 9(8), 1–12. <https://doi.org/10.3390/electronics9081295>
- Ansen, N. (2022). Penerapan Metode Algoritma K-Means Untuk Clustering. 6(1), 1–5.
- Ardian, R., Syahputra, M., & Dermawan, D. (2022). Pengaruh Pertumbuhan Ekonomi Terhadap Tingkat Pengangguran Terbuka Di Indonesia. *Jurnal Ekonomi, Bisnis Dan Manajemen*, 1(3), 190–198. <https://journal.unimar-amni.ac.id/index.php/EBISMEN/article/view/90>
- Badan Pusat Statistik. (2023). Tingkat Pengangguran Terbuka (TPT) sebesar 5,32 persen dan Rata-rata upah buruh sebesar 3,18 juta rupiah per bulan. 06 November 2023. <https://www.bps.go.id/id/pressrelease/2023/11/06/2002/tingkat-pengangguran-terbuka--tpt--sebesar-5-32-persen-dan-rata-rata-upah-buruh-sebesar-3-18-juta-rupiah-per-bulan.html>
- Desdianti, M., Debatara, N. N., Sulistianingsih, E., & Barat, P. K. (2024). ANALISIS K-MEANS CLUSTERING DENGAN METODE ELBOW PADA PENGELOMPOKAN TINGKAT PENGANGGURAN DI KALIMANTAN BARAT dimulai dengan pemahaman bahwa suatu kumpulan data sebenarnya mempunyai kesamaan antar anggotanya, yang memungkinkan anggota-anggota yang mempunyai. 13(1), 53–62.
- Kantardzic, M. (2020). DATA MINING CONCEPTS, MODELS, METHODS, and ALGORITHMS. In *July 2011*. IEEE PRESS.
- Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining of cluster in K-means. *International Journal of Advance Research in Computer Science and Management Studies*, 1(6), 90–95. <https://www.researchgate.net/publication/313554124>
- Maliqi, R., & Falgenti, K. (2023). Komparasi Kluster Pengangguran Terbuka di Indonesia Sebelum dan Saat Pandemi Covid-19 Menggunakan K-Mean Clustering. *Faktor Exacta*, 16(1), 23–30. <https://doi.org/10.30998/faktorexacta.v16i1.15108>

- Pastia, N. I., & Dikananda, F. N. (2023). Pengelompokan Data Pengangguran Terbuka Menggunakan Algoritma K-Means Berdasarkan Provinsi Jawa Barat. *Jurnal Dinamika Informatika*, 12(1), 59–69.
- Patty, W. (2010). Analisa Sebaran Iluminasi Cahaya Petromaks Dengan Perlakuan Bertudung Dan Tanpa Tudung. *Jurnal Perikanan Dan Kelautan Tropis*, 6(3), 156.
<https://doi.org/10.35800/jpkt.6.3.2010.161>
- Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. *IEEE Access*, 8, 80716–80727.
<https://doi.org/10.1109/ACCESS.2020.2988796>
- Siregar, A. M. (2019). Pengelompokan Bidang Laju Pertumbuhan Ekonomi Indonesia Menggunakan Algoritma K-Means. *Jurnal Accounting Information System (AIMS)*, 2(2), 140–151.
<https://doi.org/10.32627/aims.v2i2.342>
- Sita Muharni, S. A. (2022). Penerapan Metode K-Means Clustering pada Data Tingkat Pengangguran Terbuka Tahun 2016-2018 dan 2019-2021. *Computer Science (CO-SCIENCE)*, 2(2), 109–116.
<https://doi.org/10.31294/coscience.v2i2.1151>
- Sulistiyawati, A., & Supriyanto, E. (2021). Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan. *Jurnal Tekno Kompak*, 15(2), 25. <https://doi.org/10.33365/jtk.v15i2.1162>
- Tarigan, S. U., Saniman, S., & Yetri, M. (2022). Klasterisasi Data Penanganan Dan Pelayanan Kesehatan Masyarakat Menggunakan Algoritma K-Means. *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, 1(3), 193.
<https://doi.org/10.53513/jursi.v1i3.5223>
- Wahidah, N. (2021). Clustering Menggunakan K-Means Algorithm (K-Means Algorithm Clustering). *Fakultas Teknologi Informasi*, 21(1), 70–80.
- Yuan, C., & Yang, H. (2019). Research on K-Value Selection Method of K-Means Clustering Algorithm. *J*, 2(2), 226–235.
<https://doi.org/10.3390/j2020016>