

Pengelompokan Wilayah di Indonesia Berdasarkan Tingkat Pengangguran Menggunakan Metode *Gaussian Mixture Model* (GMM)

Clustering of Unemployment Rates in Indonesia Using the Gaussian Mixture Model (GMM)

Devina Nadifa Nur Aulia¹, M. Al Haris¹, Saeful Amri², Rochdi Wasono¹

¹ Program Studi Statistika, Universitas Muhammadiyah Semarang, Kota Semarang

² Program Studi Sains Data, Universitas Muhammadiyah Semarang, Kota Semarang

Corresponding author : devinaaulia427@gmail.com

Abstrak

Latar belakang : Tingginya tingkat pengangguran merupakan permasalahan serius yang berdampak pada aspek sosial dan ekonomi suatu negara, termasuk Indonesia. Pada tahun 2024, Indonesia mencatatkan tingkat pengangguran tertinggi di Asia Tenggara sebesar 5,2%. Kondisi ini menuntut adanya strategi yang tepat dalam menangani masalah tersebut, salah satunya melalui pengelompokan wilayah berdasarkan karakteristik pengangguran. **Metode :** *Gaussian Mixture Model* (GMM) dipilih karena memiliki keunggulan dalam mengidentifikasi pola distribusi data yang kompleks, mampu menangkap keragaman karakteristik antar wilayah, serta lebih fleksibel dibandingkan metode *clustering* lainnya. GMM tidak hanya membagi data ke dalam kelompok secara tegas, tetapi juga memberikan probabilitas keanggotaan setiap provinsi terhadap masing-masing cluster. Dengan demikian, GMM lebih sesuai digunakan untuk data sosial-ekonomi yang seringkali memiliki karakteristik yang saling tumpang tindih. Data yang digunakan adalah data indikator pengangguran di Indonesia. **Hasil penelitian :** Hasil analisis menunjukkan bahwa model terbaik diperoleh dengan jumlah cluster optimal sebanyak dua berdasarkan nilai *Bayesian Information Criterion* (BIC) tertinggi sebesar -624,5664. **Kesimpulan :** Cluster 1 beranggotakan 24 provinsi yang umumnya memiliki tingkat pengangguran dan kemiskinan lebih tinggi dan cluster 2 beranggotakan 14 provinsi dengan partisipasi kerja laki-laki lebih tinggi namun tingkat pengangguran lebih rendah.

Kata Kunci : *bayesian information crietrion* (BIC), *expectation maximization*, *gaussian mixture model* (GMM), pengangguran, pengelompokan

Abstract

Background : The high unemployment rate represents a serious issue that affects both the social and economic aspects of a country, including Indonesia. In 2024, Indonesia recorded the highest unemployment rate in Southeast Asia at 5.2%. This condition necessitates the implementation of appropriate strategies to address the problem, one of which is regional clustering based on unemployment characteristics. **Method :** The *Gaussian Mixture Model* (GMM) was selected due to its advantages in identifying complex data distribution patterns, capturing the heterogeneity of regional characteristics, and offering greater flexibility compared to other clustering methods. Unlike traditional clustering techniques that assign observations strictly to a single group, GMM provides the probability of each province's

membership across clusters. Consequently, GMM is more suitable for socio-economic data, which often exhibit overlapping characteristics. The dataset employed in this study consists of unemployment indicators in Indonesia. **Result :** The analysis revealed that the best-fitting model was obtained with an optimal number of two clusters, as indicated by the highest Bayesian Information Criterion (BIC) value of -624.5664. **Conclusion :** Cluster 1 comprises 24 provinces that generally exhibit higher unemployment and poverty rates, while Cluster 2 consists of 14 provinces characterized by higher male labor force participation but lower unemployment rates.

Keywords : bayesian information crietrion (BIC), clustering, expectation maximization, gaussian mixture model (GMM), unemployment

PENDAHULUAN

Pengangguran merupakan masalah yang dihadapi hampir setiap negara dan berdampak pada aspek sosial maupun pertumbuhan ekonomi, khususnya di negara berkembang seperti Indonesia [1]. Di Indonesia, pengangguran dipengaruhi oleh berbagai faktor, antara lain tingginya laju pertumbuhan penduduk, ketidakseimbangan antara jumlah tenaga kerja dan ketersediaan lapangan pekerjaan, perkembangan teknologi yang belum sepenuhnya diadaptasi oleh tenaga kerja usia produktif, tingkat pendidikan, perubahan struktur ekonomi, serta tingkat upah [2]. Ketiadaan pendapatan pada masyarakat yang menganggur menurunkan daya beli dan tingkat konsumsi, sementara peningkatan partisipasi angkatan kerja yang tidak diimbangi dengan perluasan lapangan kerja memperparah masalah ini [3], [4], [5].

Pandemi COVID-19 telah memberikan dampak signifikan terhadap kondisi ketenagakerjaan di Indonesia, terutama melalui penurunan tajam dalam permintaan dan penawaran tenaga kerja di berbagai sektor ekonomi. Situasi ini menyebabkan lonjakan tingkat pengangguran, yang meningkat dari 5,23% pada Agustus 2019 menjadi 7,07% pada Agustus 2020. Kenaikan ini mencerminkan semakin sulitnya masyarakat dalam memperoleh pekerjaan, serta meningkatnya jumlah kasus Pemutusan Hubungan Kerja (PHK) yang terjadi secara masif di berbagai bidang industri. Penurunan aktivitas ekonomi yang terjadi selama masa pandemi memaksa banyak perusahaan untuk melakukan efisiensi, termasuk dengan mengurangi jumlah tenaga kerja guna mempertahankan kelangsungan operasional mereka. Dampak dari kebijakan tersebut terasa cukup luas dan mendalam, terutama dalam kurun waktu antara awal tahun 2020 hingga akhir tahun 2022 [6].

Berdasarkan data *International Monetary Fund* (IMF) pada April 2024, Indonesia mencatat tingkat pengangguran tertinggi di Asia Tenggara sebesar 5,2%, diikuti Filipina, Brunei Darussalam, Malaysia, Vietnam, Singapura, dan Thailand [7], sementara Badan Pusat Statistik (BPS) melaporkan TPT Februari 2024 sebesar 4,28% atau turun 0,63 persen poin dari tahun sebelumnya [8]. Untuk menekan angka tersebut, pemerintah daerah menggalakkan program pemberdayaan UMKM, pembentukan Badan Usaha Milik Desa (BUMDes) sesuai UU No. 6 Tahun 2014,

penguatan kelompok tani, dan inisiatif ekonomi lainnya, sejalan dengan ketentuan UU No. 20 Tahun 2008 tentang Usaha Mikro [9], [10], [11], [12]. Program-program ini diharapkan efektif jika disertai prioritas pada wilayah dengan pengangguran tertinggi melalui strategi segmentasi atau pengelompokan wilayah berdasarkan analisis pola provinsi di Indonesia [13].

Pengelompokan provinsi di Indonesia berdasarkan tingkat pengangguran dapat dilakukan melalui analisis *clustering* yang merupakan bagian analisis multivariat untuk mengelompokkan objek berdasarkan kesamaan karakteristik sekaligus membedakannya dari objek yang berbeda [14], [15]. *Clustering* teknik ini terbagi menjadi dua metode utama, yakni *hierarchical clustering* dan *non-hierarchical clustering* [16]. Metode *hierarchical clustering* dimulai dengan menggabungkan objek yang memiliki tingkat kesamaan tertinggi, kemudian mengelompokkan objek lain berdasarkan kedekatan berikutnya [17]. Sementara itu, *non-hierarchical clustering* mengelompokkan objek ke dalam jumlah *cluster* yang telah ditentukan sebelumnya [18].

Metode *non-hierarchical clustering* tidak melalui proses hierarki. Contoh yang populer adalah K-Means yang termasuk kategori *unsupervised learning* karena data tidak memiliki label kelas, sehingga keanggotaannya dapat berubah antar-*cluster* [19]. Metode lain adalah *Gaussian Mixture Models* (GMM) yang memodelkan data dengan distribusi Gaussian menggunakan parameter seperti rata-rata dan varians. GMM efektif untuk data besar, fleksibel dalam bentuk cluster, serta memungkinkan data memiliki keanggotaan ganda [20]. Estimasi parameter GMM dilakukan dengan algoritma *Expectation Maximization* (EM) melalui dua tahap iteratif, yaitu *Expectation step* (E-step) yang berfungsi mengestimasi nilai serta variabel yang hilang atau belum tersedia, sedangkan tahap *Maximization* atau M-Step bertujuan untuk memperbaiki parameter berdasarkan nilai yang diperoleh dari tahap E-Step. Kedua tahapan tadi dilakukan secara berulang hingga mencapai titik konvergensi [21]. Setelah didapatkan nilai estimasi langkah berikutnya yaitu menentukan jumlah *cluster* dengan memperhatikan nilai *Bayesian Information Criterion* (BIC) terbesar [22], [23].

Penelitian sebelumnya menunjukkan keunggulan metode GMM dalam berbagai bidang penerapannya. [24] menggunakan GMM untuk mengelompokkan provinsi di Indonesia berdasarkan penyebaran COVID-19 dan membentuk enam klaster dengan nilai BIC tertinggi sebesar -467,3299. [16] menerapkannya pada pengelompokan kabupaten/kota di Pulau Sulawesi berdasarkan indikator penanganan stunting yang menghasilkan empat klaster dengan BIC tertinggi -6278,888. [25] membandingkan GMM dengan DBSCAN dan *hierarchical clustering* pada data rumah sakit di Jakarta, dan GMM terbukti memiliki performa terbaik dengan nilai Davies-Bouldin terendah, menghasilkan empat klaster. [26] membandingkan GMM dan K-Means untuk pengelompokan *cloud workloads*, dan hasilnya GMM mampu mengidentifikasi pola penggunaan sumber daya yang lebih spesifik. Berdasarkan temuan tersebut, penelitian ini bertujuan mengelompokkan

provinsi di Indonesia berdasarkan tingkat pengangguran menggunakan metode GMM.

METODE

Sumber Data dan Variabel Penelitian

Data yang digunakan dalam penelitian ini adalah data kuantitatif yang merupakan data sekunder yang diperoleh melalui website <https://bps.go.id/>. Variabel yang digunakan dalam penelitian ini meliputi Tingkat Pengangguran Terbuka (TPT) (X_1), presentase penduduk miskin (X_2), angka partisipasi angkatan kerja laki-laki (X_3), angka partisipasi angkatan kerja perempuan (X_4), rata-rata laju pertumbuhan penduduk (X_5), dan persentase penduduk berusia 25 tahun ke atas dengan pendidikan SMA ke atas (X_6) pada tiap-tiap provinsi di Indonesia pada tahun 2024 yang berjumlah 38 observasi.

Tahapan-tahapan penelitian

Tahapan yang dilakukan dalam mengelompokkan wilayah di Indonesia berdasarkan tingkat pengangguran menggunakan metode GMM adalah sebagai berikut [16], [26]:

1. Mendeskripsikan data indikator tingkat pengangguran wilayah-wilayah di Indonesia
2. Melakukan *preprocessing* data dengan standarisasi data menggunakan Z-score dengan persamaan sebagai berikut:

$$z = \frac{x_i - \bar{x}}{S} \quad (1)$$

dengan

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (2)$$

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (3)$$

3. Pengelompokan menggunakan Model GMM dengan tahapan sebagai berikut:
 - a. Menetapkan jumlah cluster sebanyak $k = 1, 2, \dots, 9$.
 - b. Menginisiasi parameter model GMM yaitu rata-rata (μ), covarians (Σ), dan bobot campuran dari semua komponen GMM (λ). Model GMM dirumuskan dengan persamaan berikut:

$$P(X) = \sum_{k=1}^K \pi_k N(X|\mu_k, \Sigma_k) \quad (4)$$

- c. Menghitung estimasi parameter dengan Expectation Step (E-Step). Pada tahap E-Step dalam algoritma Expectation-Maximization (EM) digunakan untuk menghitung probabilitas bahwa setiap data termasuk dalam salah satu distribusi Gaussian dalam model.
 - d. Menghitung estimasi parameter dengan Maximization Step (M-Step). Perbarui parameter model, yaitu mean, kovarian, dan bobot prior berdasarkan nilai responsibility yang dihitung pada tahap E-Step.
 - e. Jika memenuhi syarat konvergen atau nilai log-likelihood antara dua iterasi berturut-turut sangat kecil (di bawah ambang batas tertentu), maka proses dianggap konvergen, maka lanjut ke langkah selanjutnya. Jika tidak maka berulang ke langkah 6.
4. Menentukan jumlah *cluster* optimal berdasarkan nilai BIC terkecil dalam model GMM dengan persamaan berikut:

$$BIC = -2 \log(\hat{L}) - k \log(n) \quad (5)$$

5. Melakukan pemetaan dan karakterisasi berdasarkan cluster optimal.

HASIL DAN PEMBAHASAN

Deskriptif statistik dari indikator yang digunakan untuk mengelompokkan wilayah di Indonesia disajikan pada Tabel 1.

Tabel 1. Statistik Deskriptif Indikator yang Digunakan untuk Mengelompokkan Wilayah di Indonesia

Variabel	Mean	Median	Min	Max
X ₁	4,39	4,19	1,32	6,75
X ₂	11,15	10,22	4	32,97
X ₃	84,07	84,70	76,62	90,37
X ₄	56,82	55,41	47,05	85,71
X ₅	1,29	1,36	0,31	1,93
X ₆	42,88	41,93	19,96	69,27

Statistik deskriptif pada Tabel 1 menunjukkan bahwa setiap variabel memiliki tingkat variasi yang berbeda antar provinsi. Tingkat Pengangguran Terbuka (X₁) menunjukkan rentang dari 1,32 di Provinsi Papua Pegunungan hingga 6,75 di Provinsi

Jawa Barat. Hal ini mencerminkan adanya ketidaksamaan kondisi pasar tenaga kerja antar wilayah. Persentase Penduduk Miskin (X_2) memperlihatkan variasi yang lebih besar, yaitu antara 4 di Bali hingga 32,97 di Provinsi Papua Pegunungan yang mengindikasikan terdapat kesenjangan kesejahteraan antar provinsi di Indonesia. Angka Partisipasi Angkatan Kerja Laki-Laki (X_3) dan Angka Partisipasi Angkatan Kerja Perempuan (X_4) juga menunjukkan perbedaan antar provinsi, khususnya pada partisipasi perempuan yang berkisar antara 47,05 di Provinsi Jambi hingga 85,71 di Provinsi Papua Pegunungan. Hal ini menandakan masih adanya ketidakmerataan keterlibatan perempuan dalam pasar kerja di berbagai wilayah. Sementara itu, Presentase Penduduk Usia 25 Tahun ke Atas dengan Pendidikan SMA ke Atas (X_6) menunjukkan ketimpangan yang cukup tinggi, mulai dari 19,96 di Provinsi Papua Pegunungan hingga 69,27 di Provinsi DKI Jakarta, sehingga mencerminkan adanya ketidakmerataan capaian pendidikan antar provinsi.

Secara keseluruhan, perbedaan rentang yang cukup besar pada beberapa variabel, khususnya pada Tingkat Pengangguran Terbuka (TPT) (X_1), Presentase Penduduk Miskin (X_2), Angka Partisipasi Angkatan Kerja Laki-Laki (X_3), Angka Partisipasi Angkatan Kerja Perempuan (X_4), dan Presentase Penduduk Usia 25 Tahun ke Atas Dengan Pendidikan SMA ke Atas (X_6), menunjukkan bahwa kondisi antar provinsi tidak bersifat homogen. Variasi yang tinggi ini menjadi dasar perlunya dilakukan analisis pengelompokan (*clustering*) untuk mengidentifikasi provinsi-provinsi yang memiliki karakteristik serupa.

Preprocessing Data

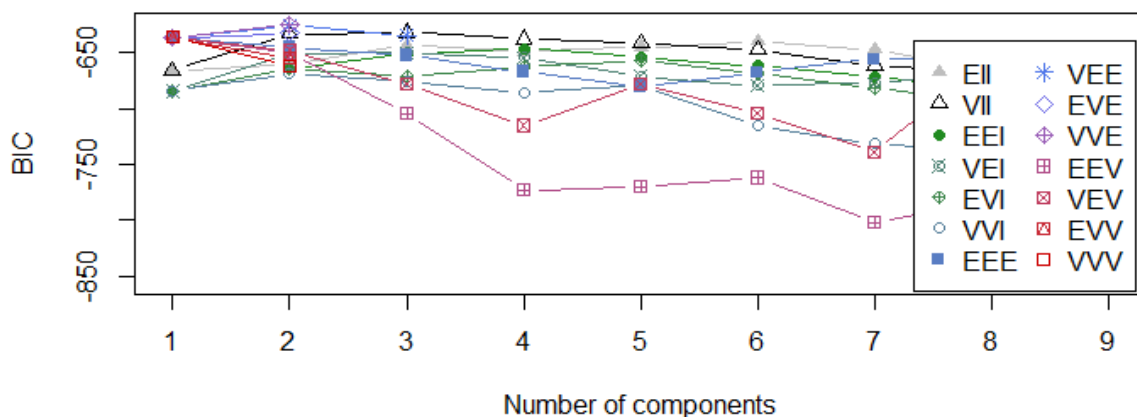
Preprocessing data dilakukan untuk menstandarisasi data menggunakan metode Z-score dengan persamaan 1. Standarisasi dilakukan untuk merubah nilai masing-masing atribut ke skala yang sama untuk memastikan semua variabel berkontribusi setara dalam analisis. Hasil dari standarsiasi data pada setiap variabel dari 38 Provinsi di Indonesia ditunjukkan pada Tabel 2.

Tabel 2. Hasil Standarisasi Data

Provinsi	X_1	X_2	X_3	X_4	X_5	X_6
Aceh	0,9656	0,4565	-1,3058	-0,9500	0,2938	0,4744
Sumatera Utara	0,8591	-0,4685	0,0657	0,2269	0,3250	1,1319
Sumatera Barat	0,9656	-0,7679	-0,5105	0,1226	0,4155	0,5853
Riau	-0,4890	-0,6642	0,2770	-1,2891	0,2314	0,2362
⋮	⋮	⋮	⋮	⋮	⋮	⋮
Papua Pegunungan	-2,1779	3,2346	2,4169	3,8118	0,4185	-2,2986

Pengelompokan Menggunakan Model *Gaussian Mixture Model* (GMM)

Pengelompokan tingkat pengaggaran menggunakan metode GMM dilakukan dengan mencobakan *cluster* awal (k) dari 2 hingga 9 *cluster*. Langkah ini dilakukan agar seluruh kemungkinan jumlah *cluster* yang relevan dapat dievaluasi secara menyeluruh, sehingga hasil pengelompokan yang diperoleh mampu menggambarkan struktur data secara lebih akurat. Pemilihan rentang ini didasarkan pada keunggulan metode GMM yang memungkinkan dan evaluasi seluruh hasil pengelompokan pada 14 model karakteristik geometrik gaussian (EII, VII, EEI, VEI, EVI, VVI, EEE, VEE, EVE, VVE, EEV, VEV, EVV, VVV) untuk setiap nilai k . Setiap model diuji dengan variasi jumlah *cluster* mulai dari $k = 2$ hingga $k = 9$ berdasarkan nilai BIC. BIC digunakan karena mampu memberikan penilaian yang seimbang antara kecocokan model terhadap data dan kompleksitas model. BIC akan dihitung untuk setiap kemungkinan jumlah *cluster*, kemudian dibandingkan. Model dengan nilai BIC terbesar (terlepas dari negatif atau positif) akan dianggap sebagai model yang terbaik karena memberikan keseimbangan paling baik antara akurasi dan kesederhanaan model. Gambar 1 menunjukkan grafik nilai BIC terhadap jumlah komponen (*cluster*) dan struktur dari model GMM.



Gambar 1. Grafik Hasil Perhitungan *Ratio BIC Changes GMM*

Gambar 1 menunjukkan 14 model yang terbentuk dari komponen 2 hingga 9. Model tersebut menggambarkan karakteristik geometrik gaussian dalam GMM. Hasil tersebut menunjukkan bahwa 14 model pada *cluster* 2 cenderung menghasilkan nilai BIC tinggi dibandingkan *cluster* lainnya serta model terbaik yang didapat adalah model VVE pada *cluster* 2. Nilai BIC dari 14 model yang terbentuk pada *cluster* 2 ditunjukkan pada Tabel 3.

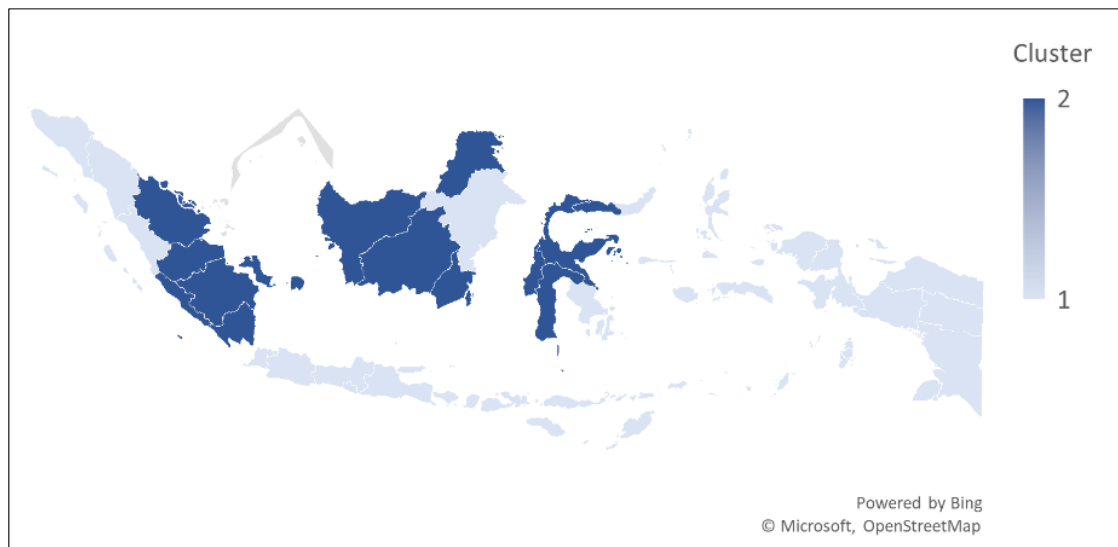
Tabel 3. Nilai BIC dari 14 Model yang Terbentuk pada *Cluster* 2

Model	Kluster	Nilai BIC
EII	2	-660,5613

VII	2	-633,3573
EEI	2	-664,4011
VEI	2	-651,1020
EVI	2	-664,2476
VVI	2	-668,7450
EEE	2	-645,5284
VEE	2	-625,8566
EVE	2	-632,6664
VVE	2	-624,5664
EEV	2	-650,5559
VEV	2	-648,2102
EVV	2	-655,0783
VVV	2	-661,6880

Hasil *Clustering* Menggunakan *Gaussian Mixture Model (GMM)*

Hasil *clustering* wilayah di Indonesia berdasarkan tingkat pengangguran menggunakan GMM menghasilkan 2 *cluster*. Provinsi-provinsi yang tergolong ke dalam cluster 2 (warna biru tua) umumnya tersebar di wilayah barat Indonesia, seperti DKI Jakarta, Kalimantan Timur, dan beberapa provinsi di Pulau Sumatera. Provinsi-provinsi ini cenderung memiliki tingkat pengangguran yang lebih tinggi serta karakteristik sosial-ekonomi tertentu yang membedakan mereka dari wilayah lainnya. Sebaliknya, cluster 1 (warna biru muda) mencakup sebagian besar provinsi di wilayah timur dan tengah Indonesia, termasuk Papua, Nusa Tenggara Timur, Sulawesi Tengah, dan lainnya, yang secara umum memiliki tingkat pengangguran dan variabel sosial ekonomi yang lebih rendah atau berbeda polanya. Berikut ini adalah hasil *clustering* provinsi yang disajikan dalam bentuk visual dan tabel.



Gambar 2. Grafik *Clustering* Wilayah di Indonesia Berdasarkan Tingkat Pengangguran Menggunakan GMM

Tabel 4. Hasil *Clustering* setiap Provinsi

Cluster	Provinsi
<i>Cluster 1</i>	Aceh, Sumatra Utara, Sumatra Barat, Kep.Riau, DKI Jakarta, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, Banten, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Timur, Maluku Utara, Papua Barat, Papua Barat Daya, Papua, Papua Selatan, Papua Tengah, dan Papua Pegunungan.
<i>Cluster 2</i>	Riau, Jambi, Sumatera Selatan, Bengkulu, Lampung, Kep. Bangka Belitung, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Utara.

Untuk mengetahui karakteristik dari masing-masing *cluster* yang terbentuk, dilakukan perhitungan rata-rata dari semua variabel pembentuk *cluster*. Nilai rata-rata variabel pembentuk masing-masing *cluster* disajikan pada Tabel 5.

Tabel 5. Nilai Rata-rata setiap Variabel bagi Kedua *Cluster*

Cluster	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆
1	4,583	12,799	83,450	59,064	1,307	45,007
2	4,056	8,325	85,156	52,973	1,276	39,245

Berdasarkan Tabel 5, dapat dilihat bahwa *cluster 1* memiliki nilai rata-rata Tingkat Pengangguran Terbuka (X₁) lebih tinggi dibandingkan dengan *cluster 2*. Selain itu, *cluster 1* juga menunjukkan dominasi nilai rata-rata Presentase Penduduk

Miskin (X_2), Angka Partisipasi Angkatan Kerja Perempuan (X_4), Rata-rata Laju Pertumbuhan Penduduk (X_5), dan Persentase Penduduk Usia 25 Tahun ke Atas dengan Pendidikan SMA ke Atas (X_6). Sementara itu, *cluster* 2 hanya unggul pada variabel Angka Partisipasi Angkatan Kerja Laki-Laki (X_3) dengan rata-rata sebesar 85,15643, lebih tinggi dibandingkan *cluster* 1 yang sebesar 83,45042. Hal ini menunjukkan bahwa *cluster* 1 cenderung mencerminkan wilayah dengan Tingkat Pengangguran dan Kemiskinan yang lebih tinggi, serta pertumbuhan penduduk dan tingkat pendidikan yang lebih baik. Sebaliknya, *cluster* 2 menggambarkan wilayah dengan tingkat pengangguran dan kemiskinan yang lebih rendah, namun dengan partisipasi kerja laki-laki yang lebih tinggi dan beberapa indikator sosial yang relatif lebih rendah dibandingkan *cluster* 1.

KESIMPULAN

Wilayah-wilayah di Indonesia jika dilihat dari indikator tingkat pengangguran memberikan variasi yang cukup beragam. Hasil pengelompokan wilayah-wilayah di Indonesia dengan GMM menghasilkan 2 *cluster*. Cluster 1 cenderung beranggotakan provinsi di wilayah timur dan tengah Indonesia, sedangkan provinsi-provinsi yang tergolong ke dalam cluster 2 umumnya tersebar di wilayah barat Indonesia. Cluster 1 mencerminkan wilayah dengan tingkat pengangguran dan Kemiskinan yang lebih tinggi dari cluster 2, akan tetapi pertumbuhan penduduk dan tingkat pendidikannya lebih baik dari cluster 2.

DAFTAR PUSTAKA

- [1] A. Octamaya Tenri Awaru, Nurulmiati Nurulmiati, Nita Talia, Mustika Mustika, Melfin Dwi Sofiana, and Yulianti Yulianti, "Dampak Sosial Ekonomi Pemberdayaan Masyarakat Kelompok Tani Swadaya Oleh Industri PT. Manakkara Unggul Lestari," *Harmon. J. Ilmu Komun. dan Sos.*, vol. 1, no. 3, pp. 196-203, Jul. 2023, doi: 10.59581/harmoni-widyakarya.v1i3.1013.
- [2] Akramunnisa and Fajriani, "K-Means Clustering Analysis pada PersebaranTingkat Pengangguran Kabupaten/Kota di Sulawesi Selatan," *J. Varian*, vol. 3, no. 2, pp. 103-112, Apr. 2020, doi: 10.30812/varian.v3i2.652.
- [3] N. A. Rahman and N. Z. Riani, "Analisis Pengangguran Terbuka di Indonesia," *J. Kaji. Ekon. dan Pembang.*, vol. 5, no. 3, p. 43, Sep. 2023, doi: 10.24036/jkep.v5i3.15285.
- [4] D. Purwanto, R. A. Pratama, R. B. Lein, I. Prastyo, and M. Al Haris, "Clustering of Districts in Central Java According to People's Welfare Indicators Using Ward's Method," *Var. J. Stat. Its Appl.*, vol. 7, no. 1, pp. 73-82, 2025, doi: <https://doi.org/10.30598/variancevol7iss1page73-82>.
- [5] C. N. Rianda, "Analisis Dampak Pengangguran Berpengaruh Terhadap Individual," *At-Tasyri' J. Ilm. Prodi Huk. Ekon. Syariah*, vol. 2, no. 1, p. 17, Jul. 2020, doi: 10.47498/tasyri.v12i01.358.
- [6] S. Frisnoiry, H. M. W. Sihotang, N. Indri, and T. Munthe, "Analisis Permasalahan Pengangguran di Indonesia," *Kompak J. Ilm. Komputerisasi Akunt.*, vol. 17, no. 1, pp. 365-374, 2024, doi: <https://doi.org/10.51903/kompak.v17i1.1866>.

- [7] International Monetary Fund, "The World Economic Outlook (WEO) database," International Monetary Fund. [Online]. Available: <https://www.imf.org/en/Countries/IDN>
- [8] Badan Pusat Statistik, "Tingkat Pengangguran Terbuka (TPT) sebesar 4,91 persen dan Rata-rata upah buruh sebesar 3,27 juta rupiah per bulan," Badan Pusat Statistik. [Online]. Available: <https://www.bps.go.id/id/pressrelease/2024/11/05/2373/tingkat-pengangguran-terbuka--tpt--sebesar-4-91-persen-.html>
- [9] A. Sidik and M. Nanang, "Upaya Pemerintah Desa untuk Mengurangi Angka Pengangguran di Desa Darmawangi Kecamatan Tomo Kabupaten Sumedang Provinsi Jawa Barat," *eJournal Pembang. Sos.*, vol. 11, no. 1, pp. 579-589, 2023, doi: <https://ejournal.ps.fisip-unmul.ac.id/site/?p=1661>.
- [10] F. D. Kurniawan and L. Fauziah, "Pemberdayaan Usaha Mikro Kecil dan Menengah (UMKM) dalam Penanggulangan Kemiskinan," *JKMP (Jurnal Kebijak. dan Manaj. Publik)*, vol. 2, no. 2, pp. 165-176, Sep. 2014, doi: 10.21070/jkmp.v2i2.436.
- [11] Sufi, "Sistem Pengelolaan Badan Usaha Milik Desa (BUMDes) dalam Pembangunan Tambak Milik Desa (Studi Desa Lhok Euncien Kecamatan Baktiya Barat)," *Negot. J. Ilmu Adm. Bisnis*, vol. 4, no. 1, pp. 46-58, 2021, [Online]. Available: <https://ojs.unimal.ac.id/na/article/view/4923/pdf>
- [12] L. G. Pramono and Yuliawati, "Peran Kelompok Tani Terhadap Pendapatan Petani Padi Sawah di Kelurahan Kauman Kidul Kecamatan Sidorejo Kota Salatiga," *Agritech J. Fak. Pertan. Univ. Muhammadiyah Purwokerto*, vol. 21, no. 2, p. 129, Jan. 2020, doi: 10.30595/agritech.v21i2.5064.
- [13] R. Rahmati and A. W. Wijayanto, "Analisis Cluster dengan Algoritma K-Means, Fuzzy C-Means dan Hierarchical Clustering (Studi Kasus: Indeks Pembangunan Manusia Tahun 2019)," *JIKO (Jurnal Inform. dan Komputer)*, vol. 5, no. 2, pp. 73-80, 2021, doi: <http://dx.doi.org/10.26798/jiko.v5i2.422>.
- [14] A. M. Dyfa, M. Al Haris, and A. Imron, "Pengelompokan Kabupaten/Kota di Jawa Tengah Berdasarkan Rasio Tenaga Kesehatan Menggunakan Metode Fuzzy Gustafson Kessel," in *Seminar Nasional Sains Data 2025 (SENADA 2025)*, 2025, pp. 347-362. doi: <https://doi.org/10.33005/senada.v5i1.488>.
- [15] M. Musfiani, "Analisis Cluster dengan Menggunakan Metode Partisi pada Pengguna Alat Kontrasepsi di Kalimantan Barat," *Bimaster Bul. Ilm. Mat. Stat. dan Ter.*, vol. 8, no. 4, Oct. 2019, doi: 10.26418/bbimst.v8i4.36584.
- [16] R. Rahman, J. Junaidi, and N. F. Gamayanti, "Pengelompokan Kabupaten/Kota di Pulau Sulawesi Berdasarkan Indikator Indeks Khusus Penanganan Stunting Menggunakan Gaussian Mixture Model," *Maj. Ilm. Mat. dan Stat.*, vol. 23, no. 2, p. 148, Sep. 2023, doi: 10.19184/mims.v23i2.36389.
- [17] Musdalifah M. Ramly, S. Sudarmin, and B. Poerwanto, "Analisis Hierarchical Clustering Multiscale Bootstrap (Kasus: Indikator Kemiskinan di Provinsi Sulawesi Selatan Tahun 2020)," *VARIANSI J. Stat. Its Appl. Teach. Res.*, vol. 4, no. 3, pp. 142-152, Dec. 2022, doi: 10.35580/variasiunm26.
- [18] U. Syafiyah, D. P. Puspitasari, I. Asrafi, B. Wicaksono, and F. M. Sirait, "Analisis Perbandingan Hierarchical dan Non-Hierarchical Clustering pada Data Indikator Ketenagakerjaan di Jawa Barat Tahun 2020," in *Seminar Nasional Official Statistics*, Nov. 2022, pp. 803-812. doi: 10.34123/semnasoffstat.v2022i1.1221.
- [19] M. P. Lawuna, Y. Zega, and R. N. Mendrofa, "Analisis Cluster Mahasiswa Pendidikan Matematika Universitas Nias," *J. Cendekia J. Pendidik. Mat.*, vol. 8,

-
- no. 2, pp. 946-961, Apr. 2024, doi: 10.31004/cendekia.v8i2.2940.
- [20] A. Triayudi *et al.*, *Clustering*. Sleman: PT Penamudamedia, 2024.
- [21] A. Milson, D. E. Herwindiati, and N. J. Perdana, "Penerapan Klasifikasi Suara Sebagai Autentikasi Keamanan Sistem Login Menggunakan Gaussian Mixture Models," *Comput. J. Comput. Sci. Inf. Syst.*, vol. 8, no. 1, pp. 104-109, Apr. 2024, doi: 10.24912/computatio.v8i1.16229.
- [22] F. Yessica, D. Kusnandar, and N. Imro'ah, "Implementasi Metode Latent Class Cluster Analysis dalam Pengelompokan Wilayah Berdasarkan Indikator Indeks Pembangunan Manusia," *Bul. Ilm. Math. Stat dan Ter.*, vol. 11, no. 2, pp. 213-220, 2022, [Online]. Available: <https://jurnal.untan.ac.id/index.php/jbmstr/article/view/53269/75676592495>
- [23] R. Narulita and A. Kudus, "Penerapan Gaussian Mixture Model pada Klasterisasi Desa di Kalimantan Barat berdasarkan Indeks Desa Membangun Tahun 2021," in *Bandung Conference Series: Statistics*, Jul. 2022, pp. 226-233. doi: 10.29313/bcss.v2i2.3877.
- [24] J. Riyono, S. D. Puspa, and C. E. Pujiastuti, "Simulasi Clustering Provinsi di Indonesia dalam Penyebaran Covid-19 berdasarkan Indikator Kesehatan Masyarakat Menggunakan Algoritma Gaussian Mixture Model," *Majamath J. Mat. dan Pendidik. Mat.*, vol. 5, no. 1, pp. 43-60, 2022, doi: MAJAMATH: Jurnal Matematika Dan Pendidikan Matematika.
- [25] K. E. Setiawan and A. Kurniawan, "Pengelompokan Rumah Sakit di Jakarta Menggunakan Model DBSCAN, Gaussian Mixture, dan Hierarchical Clustering," *J. Inform. Terpadu*, vol. 9, no. 2, pp. 149-156, Sep. 2023, doi: 10.54914/jit.v9i2.995.
- [26] E. Patel and D. S. Kushwaha, "Clustering Cloud Workloads: K-Means vs Gaussian Mixture Model," in *Procedia Computer Science*, 2020, pp. 158-167. doi: 10.1016/j.procs.2020.04.017.