Jurnal Ilmiah Sistem Informasi dan Ilmu Komputer
Volume. 6 Nomor. 2 Juli 2026
e-ISSN: 2827-7953; p-ISSN: 2827-8135; Hal 128-139
DOI: https://doi.org/10.55606/juisik.v6i2.2414
Tersedia: https://journal.sinov.id/index.php/juisik

Penerapan Metode *K-Means Clustering* dalam Pengelompokkan Tingkat Pengangguran di Provinsi Lampung

Imelia Puspita Sari[1*], Muhammad Rijal Firmansyah[2], Arzella Meilantika[3], Ade Herlina[4], Nuari Anisa Sivi[5]
[1-5] Sistem Informasi, Universitas Nahdlatul Ulama Lampung, Indonesia
Email: Imeliapuspitastr11@gmail.com[1], kgrijalfirman@gmail.com[2], arzellameylantika@gmail.com[3], adeherlina985@gmail.com[4], nuarianisasivi@gmail.com[5]*

[*]*Penulis Korespondensi: Imeliapuspitastr11@gmail.com*

Abstract. *Unemployment is one of the main problems in economic development in Indonesia. One of the indicators used to measure the economic condition of a region is the Open Unemployment Rate. The purpose of this study is to provide information regarding a more detailed grouping of the number of unemployment rates in each district/city in Lampung Province, through the application of the K-Means clustering algorithm method. The data used in the study were taken from the official website of the Central Statistics Agency (BPS) of Lampung Province in 2025, which includes data on the Open Unemployment Rate (TPT) and the Poverty Rate. The application of K-Means clustering was used to break down the data into groups. The research method includes four steps: data collection, data preparation, modeling, and evaluation (using the Davies Bouldin Index theory), using the RapidMiner application as a support to facilitate data mining processing. The results of this study are to provide information about the unemployment rate groups in Lampung Province into 3 groups, namely low unemployment rate, medium unemployment rate, and high unemployment rate, which will divide all districts in Lampung Province.*

Keywords: *Data Mining;K-Means Clustering; Lampung; Open Unemployment Rate; Unemployment.*

Abstrak. Pengangguran adalah salah satu permasalahan pokok pada pembangunan perekonomian di Indonesia. Salah satu indikator yang yang dipakai untuk mengukur kondisi ekonomi suatu wilayah adalah dengan Tingkat Pengangguran Terbuka. Tujuan dari penelitian ini adalah untuk memberikan informasi mengenai pengelompokkan lebih detail jumlah tingkat pengangguran tiap kapupaten/kota yang ada di Provinsi Lampung, dengan melalui penerapan metode algoritma K-Means clustering. Data yang digunakan dalam penelitian, diambil dari situs resmi *website* Badan Pusat Statistik (BPS) Provinsi Lampung pada tahun 2025, yang mencakup data Tingkat Pengangguran Terbuka (TPT) dan Tingkat Kemiskinan. Penerapan K-Means clustering digunakan untuk memecahkan data secara berkelompok. Metode penelitian mencakup empat langkah yaitu pengumpulan data, persiapan data, modeling, dan evaluasi (menggunakan teori *Davies Bouldin Index*), dengan menggunakan aplikasi RapidMiner sebagai pendukung untuk mempermudah pengolahan data mining. Hasil dari penelitian ini adalah memberikan informasi tentang kelompok tingkat pengangguran yang ada di Provinsi Lampung ke dalam 3 kelompok, yaitu tingkat pengangggguran rendah, tingkat pengangguran sedang, dan tingkat pengangguran tinggi, yang akan membagi seluruh kabupaten yang ada di Provinsi Lampung

Kata kunci: Data Mining;*K-Means Clustering;* Lampung; Pengangguran; Tingkat Pengangguran Terbuka.

1. LATAR BELAKANG

Pada pembangunan suatu wilayah, masalahan perekonomian tidak dapat dilepaskan dari kondisi ketersediaan lapangan pekerjaan. Tidak tersedianya lapangan pekerjaan akan menjadi salah satu faktor naiknya tingkat pengangguran disuatu wilayah. Pengangguran adalah salah satu permasalahan utama dalam perekonomian yang mencerminkan ketidaksesuainan antara jumlah tenaga kerja dengan ketersediaan lapangan pekerjaan [1]. Permasalahan tingkat pengangguran sampai saat ini masih menjadi perhatian di berbagai wilayah Indonesia. Indikator utama yang digunakan untuk menakar kondisi tenaga kerja yang ada suatu wilayah

Naskah Masuk: 18 April 2026; Revisi:16 Mei 2026; Diterima: 22 Juni 2026; Tersedia: 24 Juni 2026

umumnya menggunakan Tingkat Pengangguran Terbuka (TPT) [1]. Tingkat penganguran merupakan urutan dari prevalensi pengangguran dan dihitung sebagai presentase dengan membagi jumlah individu yang menganggur dengan semua jumlah individu yang bekerja [2]. Tingginya angka tingkat pengangguran akan menunjukkan kurang optimalnya penyerapan tanaga kerja, tidak seimbangnya pertumbuhan angka kerja dengan lapangan pekerjaan. Oleh karena itu, tingkat pengangguran dapat menjadi salah satu indikator dalam menilai pertumbuhan perekonomian wilayah tersebut.

Provinsi Lampung merupakan provinsi yang letaknya ada di ujung selatan pulau Sumatra juga menghadapi tantangan dalam mengelola ketenagakerjaan. Di tinjau dari Badan Pusat Statistik, provinsi Lampung memiliki jumlah penduduk sekitar 9.623,79 ribu jiwa dengan jumlah kemiskinan ada di 229,16 ribu jiwa dan jumlah pengangguran terbuka 206,80 ribu jiwa. Jumlah tersebut diakulasikan dari seluruh kabupaten/kota yang ada di Provinsi Lampung. [3]

Pengelompokan tingkat pengangguran ditiap kabupaten/kota dapat dilakukan dengan menerapkan metode K-means Clustering. Metode k-means mempunyai beberapa keunggulan seperti mudah untuk digunakan, cepat dijalankan, dengan hasil yang akurat [4]. Melalui permasahan di atas, mengenai penglompokan tingkat pengangguran kapupaten/kota di Provinsi Lampung, sehingga dapat diketahui dengan mudah wilayah yang memiliki tingkat pengangguran paling tinggi maupun tingkat pengangguran paling rendah. Penelitian ini diharapkan akan memberikan tambahan informasi dan kemudahan dalam melakukan optimalisasi penanggulagan tingkat pengangguran yang ada di wilayah Lampung.

Pengelompokan tingkat pengangguran ditiap kabupaten/kota dapat dilakukan dengan menggunakan metode K-means Clustering. Metode k-means memiliki beberapa keunggulan seperti mudah untuk digunakan, cepat dijalankan, dengan hasil yang akurat [4]. Melalui permasahan di atas, mengenai penglompokan tingkat pengangguran kapupaten/kota di Provinsi Lampung, sehingga dapat diketahui dengan mudah wilayah yang memiliki tingkat pengangguran paling tinggi maupun tingkat pengangguran paling rendah. Penelitian ini diharapkan akan memberikan tambahan informasi dan kemudahan dalam melakukan optimalisasi penanggulagan tingkat pengangguran yang ada di wilayah Lampung.

2. KAJIAN TEORITIS

Penelitian Terdahulu

Penelitian "Implementasi K-Means Clustering untuk mengelompokan Tingkat Pengangguran", penelitian ini bertujuan untuk mengklasifikasikan dengan lebih detail lagi dalam menganalisis data pengangguran yang ada di Jawa Barat melalui algoritma K-Means Clustering. [5]

"Penerapan Algoritma K-Means Clustering dan Hierarchical Clustering dalam Mengelompokkan Data Pengangguran di Karawang". Penelitian ini mengangkat kota Karang sebagai objek yang merupakan kawasan industri. Hal ini memicu penduduk luar untuk melakukan migrasi, mempertimbangkan dengan naikknya jumlah penduduk maka besar kemungkinan akan ikut naiknya jumlah pengangguran yang ada di daerah tesebut. Tujuannya dengan memanfaatkan algoritma K-Means Clustering dan Hierarchical Clustering dalam pengemlompokannya untuk memperoleh informasi. [6]

Penelitian "Penerapan Metode K-Means Pada Pengelompokkan Pengangguran di Indonesia". Bertujuan untuk menganalisis tingkat pengangguran yang ada di Indonesia. Diharapkan akan memberikan masukkan terhadap pemerintah untuk memperluas lapangan pekerjaan. Khusunya provinsi dengan minim lapangan pekerjaan berdampak pada pengangguran. [7]

Penelitian "Penerapan Algoritma K-Means Pada Data Pengangguran Di Jawa Barat". Penelitian ini dilakukan pada tahun 2023, dengan tujuan untuk memberikan gambaran pola dan dampak pengangguran di Provinsi Jawa Barat, dengan menggunakan metode K-means Clustering. [8]

K-Means

Metode yang digunakan analisis data untuk pengelompokan adalah dengan algoritma K-Means Clustering. K-Means adalah algoritma yang dapat digunakan untuk mengolah data dalam jumlah besar atau kecil [5]. K-Means merupakan metode pengolahan data, dengan tujuan untuk membagi data dalam satu atau beberapa kelompok, sehingga data dengan karakteristik yang serupa dapat berada dalam kelompok yang sama. Metode K-means memiliki proses yang sederhana untuk diterapkan dan dioperasikan, kecepatan yang baik, mudah untuk digunakan, serta mudah beradaptasi, sehingga banyak dimanfaatkan dalam berbagai aplikasi berskala kecil sampai dengan menengah [7]. Algoritma K-Means dapat bekerja dengan menghitung jumlah cluster (k) yang telah ditentukan berdarkan pusat cluster dengan menggunakan persamaan ukuran jarak geometris [5].

Clustering

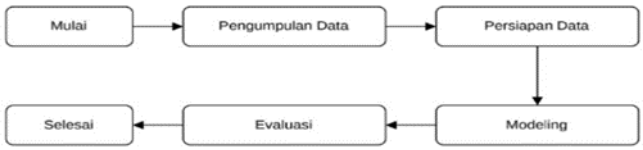
Clustering adalah metode data mining yang memiliki sifat *unsupervised,* yaitu proses pengelompokkan data, dilakukan tanpa arahan, intuksi, maupun target output tertentu [5]. Clustering mengarah pada pengelompokkan rekaman yang mirip, mengamati, serta akan membuat kelas objek yang mempunyai kesamaan [9]. Clustering akan coba mambagi beberapa informasi ke dalam kelompok-kelompok yang mempunyai tingkat kemiripian relatif tinggi, sehungga kesamaan dalam kelompok menjadi lebih maksimal [10].

Pengangggguran

Sukirno (1994) menyatakan bahwa pengangguran bukan merupakan seseorang yang tidak memiliki pekerjaan, tetapi tidak secara aktif melakukan pencarian pekerjaan. Pengangguran merupakan kondisi ketika seseorang yang telah memasuki usia angkatan kerja dan ingin memperoleh pekerjaan, namun belum berhasil mendapatkan mendapatkan pekerjaan [5]. Tingginya angka pengangguran akan mencerminkan kondisi ketersediaan lapangan pekerjaan.

3. METODE PENELITIAN

Penelitian ini menamfaatkan algoritma K-Means untuk mengelompokkan data dalam tiga cluster berdasarkan tinggi rendahnya data yang ada. Proses penelitian meliputi dilakukan proses pengumpulan data, proses persiapan data, modeling, dan evaluasi. Berikut tahapan proses sesuai dengan gambar di bawah ini.



Gambar 1. Alur Penelitian.

Pengumpulan Data

Pada penelitian ini, data diperoleh dari sumber resmi yaitu Badan Pusat Statistik Provinsi Lampung (BPS). Jenis data yang dikumpulkan meliputi data tingkat pengangguran terbuka, serta tingkat kemiskinan yang ada di provinsi Lampung. Proses pengambilan data dilakukan dengan mengakses dan mengunduh melalui situs publikasi resmi website BPS.

Persiapan Data

Pada tahapan persiapan data, dilakukan dengan tujuan memastikan data sudah benar dan siap untuk diolah dan dilakukan analisis. Dilakukan untuk menidentifikasi dan mengatatasi

ketidaksesuaian data. Data *selection* merupakan proses pemilihan data dari kumpulan data operasional yang dilakukan sebelum memulai tahapan selanjutnya. Proses pembersihan data dilakukan dengan menghapus data yang sama, membersihkan data yang tidak konsisten, serta membenahi kesalahan pada data [11].

Modeling

Pada tahap modeling akan dilakukan pengolahan data dengan metode K-Means, untuk menggolongkan data sesuai kemiripan karakter akan dikelompokkan dalam satu cluster. Proses pengelompokkan menggunakan bantuan aplikasi RapidMiner untuk mempermudah perhitungan dan visualisasi data. Berikut lngkah-langkah metode K-Means Clustering. A.) Tetapkan jumlah *cluster* (k), yang hendak digunakan pada pengelompokkan data;b.) Inisialisasi centroid awal, yang dipilih secara acak oleh sistem pada aplikasi RapidMiner;c.) Menghitung jarak antara setiap data dengan centroid menggunakan metode Euclidean Distance untuk menentukan kedekatan antara data. $d(x,y)=\sqrt{\sum_{i=1}^{n}(xi-yi)^2}$*;d.)* Mengklasifikasikan setiap data dalam cluster yang memliki jarak paling dekat pada centroid;e.) Melakukan pembaharuan centroid dengan menghitung nilai rata-rata dari semua data, pada masing-masing cluster.;f.) Ulangi proses perhitungan jarak dan pembaharuan centroid hingga tidak terjadi perubahan anggota clustet atau telah melalui kondisi konvergen.

Evaluasi

Pada tahapan ini, dilakukan pengujian untuk evaluasi hasil kerja perhitungan *maching learning* yang digunakan sebagai informasi yang mudah dipahami berlandaskan data yang sudah diolah[6]. Informasi yang diperoleh dari proses data mining, disajikan dalam bentuk mudah untuk dipahami. Pendekatan *David Bouldin Index* (DBI) digunakan dalam tahapan ini untuk menetukan cluster yang paling ideal dan akan menunjukkan hasil yang diperoleh, sudah tercapai atau belum[12].

4. HASIL DAN PEMBAHASAN

Bersadarkan hasil penelitian yang telah, pada pembahasan ini akan dipaparkan tahapan pengelompokan data tingkat pengangguran terbuka yang ada di Provinsi Lampung memanfaatkan algotitma K-Means. Proses pengelompokkan dilakukan dengan menggunakan *maching learning* yaitu menggunakan *software*RapidMiner studio. Tahap-tahap analisis pengelompokkan data sebagai berikut.

Data

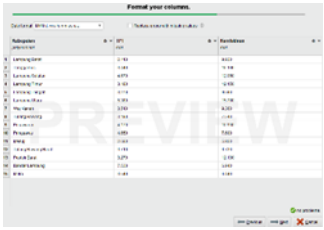
Data yang digunakan berasal dari *website* resmi Badan Pusat Statistik (BPS) Provinsi Lampung pada tahun 2025. Data yang diambil meliputi data tingkat pengangguran terbuka (TPT) tahun 2025 dan data tingkat kemiskinan pada tahun 2025. Data tersebut dapat dilihat pada gambar 2 berikut :

Kabupaten	TPT	Kemiskinan
Lampung Barat	2.11	9,93
Tanggamus	3.44	10,1
Lampung Selatan	4.67	12,05
Lampung Timur	3.18	12,15
Lampung Tengah	3.27	9,59
Lampung Utara	5.38	15,78
Way Kanan	3.51	9,36
Tulang Bawang	3.19	7,54
Pesawaran	4.17	10,93
Pringsewu	4.65	7,6
Mesuji	2.68	5,92
Tulang Bawang Barat	3.71	6,72
Pesisir Barat	3.27	12,13
Bandar Lampung	7.53	6,94
Metro	3.54	6,44

Gambar 2. Data Tingkat Pengangguran&Kemiskinan.

Berdasarkan data di atas, tiap daerah yang ada di Lampung memiliki perbedaan jumlah tingkat pengangguran terbuka dan tingkat kemiskinannya. Dimana terdapat daerah yang memiliki tingkat pengangguran tinggi dan tingkat penangguran rendah. Hal ini, meemungkinkan adanya potensi untuk dilakukan pengelompokan data..

Data *Preparation*

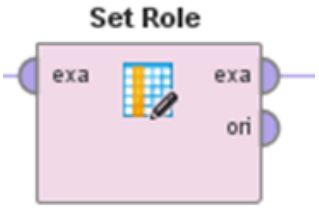


Gambar 3. Penyesianan format data

Pada tahapan ini,dilakukan pengecekan data dan juga penyesuaian format-format data, agar data yang ada dapat diolah.

Pada gambar di atas, menunjukkan proses pengecekan dan juga penyesuaian format atribut pada dataset, seblum dilakukan pengolahan lebih lanjut menggunakan K-Means clustering. Setiap atribut akan diperiksa tipe datanya, seperti atribut kategorikal (*polynominal)* merupakan jenis data yang mewakili kerakteristik (data ini tidak dapat dihitung secara matematis), dan numerik (*real*) data ini dinyatakan dengan angka dan dapat diukur secara matematis.

Selanjutnya, akan dilakukan pengaturan atribut yang menggunakan operator set role dan juga proses normalisasi data. hal ini dilakukan agar data dapat digunakan secara optimal dalam proses clustering



Gambar 4. *Set Role*

Gambar di atas, menunjukkan proses pengaturan atribut menggunakan set role. Pada tahapan ini atribut Kabupaten akan ditetapkan menjadi atribut *id,* sehingga tidak akan terhitung dalam proses clustering. Sedangkan atribut lainnya akan ditetapkan sebagai variabel, dalam pengelompokan data algoritma K-Means Clustering.

Setelah dilakukan pengaturan atribut dengan menggunakan operator Set Role, setiap atribut akan mulai menjalankan perannya masing-masing. Atribut kabupaten akan menjalankan perannya sebagai identitas data dan atribu tingkat pengangguran dan tingkat kemiskinan akan digunakan sebagai variabel dalam proses clustering.



Gambar 5. Hasil pengaturan artibut.

Maka selanjutnya akan dilakukan proses normalisasi pada data. Hal ini diterapkan untuk menyamakan skala antar tiap variabel, sehingga tidak ada dominasi pada salah satu variabel dalam proses clustering [13]. Proses normalisasi ini membantu algoritma untuk bekerja dengan lebih efisien dan menghasilkan model yang baik.



Gambar 1. Normalize

Dan, pada gambar 6 yang ditampilkan di atas, merupakan tampilan yang dilakukan sebelum tahap clustering menggunakan algoritma K-Means Clustering, dinamakan normalize. Lalu setelah dilakukan normalisasi, variabel-variabel dalam dataset yang awalnya mempunyai rentan nilai yang tidak sama akan diseagamkan menjadi serupa..

Row No.	Kabupaten	TPT	Kemiskinan
1	Lampung Bar...	-1.368	0.140
2	Tanggamus	-0.344	0.201
3	Lampung Sel...	0.603	0.909
4	Lampung Ti...	-0.544	0.945
5	Lampung Te...	-0.475	0.016
6	Lampung Uta...	1.149	2.262
7	Way Kanan	-0.290	-0.067
8	Tulang Bawa...	-0.536	-0.727
9	Pesawaran	0.218	0.502
10	Pringsewu	0.588	-0.706
11	Mesuji	-0.929	-1.315
12	Tulang Bawa...	-0.136	-1.025
13	Pesisir Barat	-0.475	0.938
14	Bandar Lamp...	2.804	-0.945
15	Metro	-0.267	-1.127

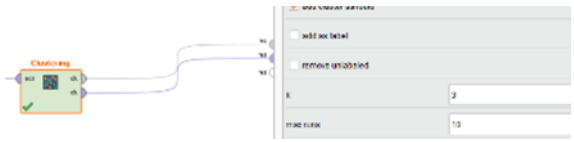
Gambar 7. Hasil Normalize

Proses (Modeling)

Pada tahap modeling, merupakan tahapan paling penting, yaitu proses analisis data dengan *software* RapidMiner menggunakan algoritma K-Means clustering. Digunakan untuk menetapkan jumlah dan kelompok cluster pada data. cluster yang dipakai yaitu cluster-3 atau (k-3).

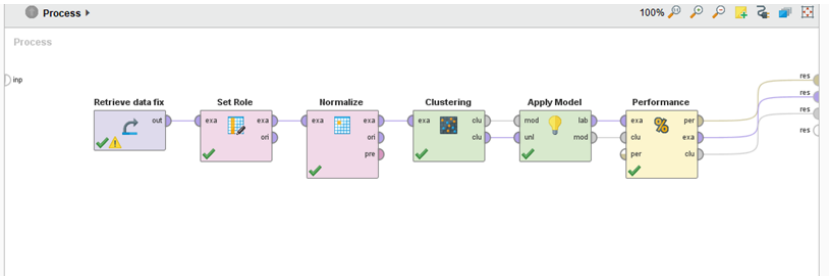
Cluster 0	Tingkat Pengangguran Rendah
Cluster 1	Tingkat Pengangguran Sedang
Cluster 2	Tingkat Penagngguran inggi

Gambar 8. Jenis Cluster



Gambar 2. *Clustering*

Untuk menentukan jumlah cluster, diperlukan operator algotitma K-Means dengan parameter k = 3, max run = 10, seperti yang diampilkan pada gambar 9 di atas. Dan berdasarkan hasil perhitungan dekat, setiap data akan dikelompokkan ke dalam cluster dengan jarak terdekat terhadap centroid, melalui proses iteratif hingga diperoleh hasil yang stabil. Adapun rangkaian proses clustering yang diterapkan pada penelitian, dapat ditinjau pada gambar berikut.



Gambar 3. Alur algoritma K-means Clustering

Gambar di atas menunjukkan rangkaian proses yang dilakukan untuk mengelompokkan data tingkat pengangguran yag ada di Lampung, dengan menerapkan algoritma K-Means Clustering pada RapidMiner. Proses yang terjadi meliputi pengaturan

artibut hingga tahap terakhir yaitu evaluasi yang mengaplikasikan teori *Davies Bouldin Index* (DBI).

Hasil Clstering

Berdasarkan hasil proses clustering yang dilakukan dengan menerapkan algoritma K-Means Clustering, memperoleh hasil pengelompokan tingkat pengangguran di Provinsi Lampung ke dalam 3 cluster.

Row No.	Kabupaten	cluster	TPT	Kemiskinan
1	Lampung Bar...	cluster_0	-1.368	0.140
2	Tanggamus	cluster_0	-0.344	0.201
3	Lampung Sel...	cluster_1	0.603	0.909
4	Lampung Ti...	cluster_1	-0.544	0.945
5	Lampung Te...	cluster_0	-0.475	0.016
6	Lampung Uta...	cluster_1	1.149	2.262
7	Way Kanan	cluster_0	-0.290	-0.067
8	Tulang Bawa...	cluster_0	-0.536	-0.727
9	Pesawaran	cluster_1	0.218	0.502
10	Pringsewu	cluster_0	0.588	-0.706
11	Mesuji	cluster_0	-0.929	-1.315
12	Tulang Bawa...	cluster_0	-0.136	-1.025
13	Pesisir Barat	cluster_1	-0.475	0.938
14	Bandar Lamp...	cluster_2	2.804	-0.945
15	Metro	cluster_0	-0.267	-1.127

Gambar 11. Tabel hasil cluster

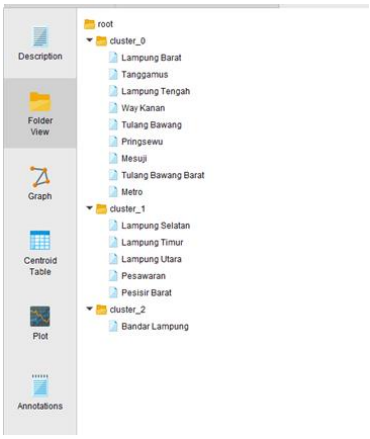
Pada gambar di atas, menunjukkan 15 kabupaten yang ada di Lampung telah dikelompokkan dalam 3 cluster, yaitu cluster 0 mempresentasikan sebagai pemilik tingkat pengangguran rendah, cluster 1 akan mempresentasikan pemilik tingkat pengangguran sedang, serta cluster 2 mempresentasikan pemilik tingkat pengangguran tinggi. Berikut merupakan cluster model yang diperoleh dari 15 item (kabupaten) yang ada.

Cluster Model

```
Cluster 0: 9 items
Cluster 1: 5 items
Cluster 2: 1 items
Total number of items: 15
```

Gambar 12. Cluster model.

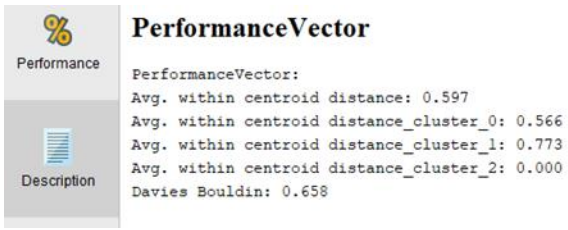
Hasil tersebut menunjukkan adanya perbedaan kondisi tingkat pengangguran yang ada di Provinsi Lampung. Sehingga dapat digunakan untuk membantu memahami pola pengelompokkan data berdasarkan tingkat kemiripan karakteristik. Dari seluruh kabupaten yang ada di Lampung terdapat 9 kabupaten yang termasuk di dalam cluster tingkat pengangguran rendah, 5 kabupaten masuk dalam cluster sedang, dan 1 kabupaten dalam cluuster tinggi. Berikut dibawah ini, detail kelompok cluster.



Gambar 13. Hasil Cluster.

Evaluasi

Teori *Davies Bouldin Index* (DBI) merupakan teori evaluasi yang digunakan dalam penelitian ini. Metode yang digunakan untuk mengukur kualitas hasil clustering yang telah dilakukan dengan menerapkan algoritma K-Means Clustering. DBI berusaha untuk memperbesar jarak antar *cluster* dan juga memperkecil jarak di dalam cluster [14]. Dan, pada penelitian yang telah dilakukan, ditentukan jumalah cluster (k) sebanyak 3 cluster. Ketiga cluster tersebut mewakili dari kategori tingkat pengangguran rendah, sedan, dan tinggi. Berikut di bawah ini nilai DBI yang diperoleh untuk mengevaluasi hasil clustering yang telah dilakukan sebelumnya.



Gambar 14. Nilai DBI.

Berdasarkan gambar 14, di atas didapatkan nilai *Davies Bouldin Index (*DBI) senilai 0.658. Nilai tersebut menunjukkan bahwa nilai hasil clustering diperoleh cukup baik, dikarenakan nilai DBI yang dihasilkan mendekati angka 0. Semakin kecil nilai DBI atau semakin kecil nilai positifnya yang diperoleh maka akan semakin ideal [15]. Dan oleh karenanya, nilai DBI yang mendekali angka 0 menandakan bahwa *cluster* yang terbentuk memeiliki tingkat kecocokan yang tinggi dalam satu cluster, serta perbedaan yang cukup jelas dengan cluster lainnya. Dengan itu, hasil clustering dapat dikatakan mampu untuk mengelompokkan data tingkat pengangguran yang ada di Provinsi Lampung dengan cukup baik.

5. KESIMPULAN

Penelitian ini dilakukan dengan menggunakan data yang berasal dari data Tingkat Pengangguran Terbuka (TPT) dan tingkat kemiskinan tahun 2025 yang diambil dari *website* resmi Badan Pusat Statistik (BPS) Provinsi Lampung. Pengolahan dilakukan menggunakan RapidMiner dengan jumlah cluster mencapai 3 kelompok, yaitu tinggi, rendah, dan sedang.

Hasilnya menunjukan bahwa dari 15 kabupaten/kota di Provinsi Lampung, ada 9 kabupaten/kota yang termasuk ke dalam cluster tingkat rendah, 5 kabupaten/kota dalam cluster sedang, dan 1 kabupaten/kota yang masuk dalam kategori cluster tingkat tinggi.

Selain itu, dari evaluasi menggunakan teori *Davies Bouldin Indeks* (DBI), Provinsi Lampung memperoleh nilai DBI sebesar 0.658, yang mana nilai ini sudah menunjukkan bahwa nilai hasil cluster sudah cukup baik karena nilai DBI-nya yang mendekati 0, yang berarti cluster yang terbentuk tersebut memiliki kemiripan yang tinggi dalam satu cluster.

Maka dari itu, metode K-Means Clustering bisa diterapkan menjadi salah satu metode *Data Mining* untuk membantu proses pengelompokkan tingkat pengangguran di Provinsi Lampung supaya informasi yang dihasilkan dapat lebih mudah dipahami dan dianalisis.

DAFTAR REFERENSI

R. D. Pratiwi, M. N. Nurhadiyah, and S. Ningsih, “ANALISIS TREN TINGKAT PENGANGGURAN TERBUKA DAN HUBUNGANNYA DENGAN KONDISI EKONOMI PROVINSI LAMPUNG PERIODE 2014-2024,” vol. 3, pp. 413–424, 2026.

P. Dan and P. Ekonomi, “INVESTASI MODAL MANUSIA BIDANG PENDIDIKAN : DAMPAK,” vol. 6, no. 1, pp. 21–37, 2020.

B. P. Statistik, “No Title,” Badan Pusat Statistik Provinsi Lampung. [Online]. Available: https://lampung.bps.go.id/id

W. Syaharani *et al.*, “Penerapan Algoritma k-Means pada Clustering Data Penduduk Miskin untuk Penghapusan Kemiskinan Ekstrem Application of k-Means Algorithm on Clustering Poor Population Data,” vol. 13, pp. 1732–1747, 2024.

N. Nurjanah, N. Suarna, W. Prihartono, T. Informatika, R. P. Lunak, and G. Tasikmalaya, “IMPLEMENTASI K-MEANS CLUSTERING UNTUK MENGELOMPOKAN,” vol. 8, no. 2, pp. 2462–2468, 2024.

A. Alif, R. Mulyana, A. R. Juwita, A. M. Siregar, and A. Fauzi, “Penerapan Algoritma K-means Clustering dan Hierarchical Clustering dalam Mengelompokkan Data Pengangguran di Karawang,” no. 2021, 2024, doi: 10.33364/algoritma/v.21-2.2155.

P. Studi and S. Informasi, “Penerapan Metode K-Means Pada Pengelompokkan Pengangguran Di Indonesia,” vol. 6, pp. 61–74.

N. P. Dharshinni, G. Singh, J. April, T. Naibaho, R. D. L. Tobing, and A. Simamora, “Penerapan Algoritma K-Means Pada Data Pengangguran Di Jawa Barat,” vol. 3, no. 1, pp. 22–33, 2023.

N. Sari, H. H. Handayani, and A. M. Siregar, "Implementasi Clustering Data Kasus Covid 19 Di Indonesia Menggunakan Algoritma K-Means," vol. 11, no. 1, pp. 7–12, 2023.

Z. Nabila, A. R. Isnain, P. Permata, and Z. Abidin, "ANALISIS DATA MINING UNTUK CLUSTERING KASUS COVID-19 DI PROVINSI LAMPUNG DENGAN ALGORITMA K-MEANS," 2021. [Online]. Available: https://api.semanticscholar.org/CorpusID:237684110

I. Sumiati, A. Rasyid, F. Adinata, M. H. A.- Musa, R. Herdiana, and C. Lukman, "Penerapan Metode Two-Step Cluster Untuk Pengelompokan Desa Berdasarkan Kepadatan Penduduk," vol. 05, no. 02, pp. 36–41, 2021.

N. I. Pastia and F. N. Dikananda, "Pengelompokan Data Pengangguran Terbuka Menggunakan Algoritma K-Means Berdasarkan Provinsi Jawa Barat Program Studi Manajemen Informatika , STMIK IKMI Cirebon," vol. 12, no. 1, pp. 59–69, 2023.

M. Anam and A. M. Majid, "Analisis Pola Pengangguran Menggunakan Metode Clustering Algoritma K-Means Di Wilayah Kabupaten Cirebon," vol. 4, no. 4, pp. 1691–1701, 2025.

I. Kamila *et al.*, "Perbandingan Algoritma K-Means dan K-Medoids untuk Pengelompokan Data Transaksi Bongkar Muat di Provinsi Riau," vol. 5, no. 1, pp. 119–125, 2019.

D. Ananda, S. Rohimah, B. Susilo, D. Wulan, and A. Mustofa, "Implementation of K-Means in Grouping Birth Certificate Data in Indonesia Implementasi K-Means Dalam Pengelompokkan Data Akta Kelahiran di Indonesia," pp. 66–71, 2022.