Jurnal Teknologi dan Manajemen Industri Terapan (JTMIT) Vol. 4, No. 4, Desember 2025 pp. 1691 - 1701
P-ISSN: 2829-0232 E-ISSN: 2829-0038

Analisis Pola Pengangguran Menggunakan Metode Clustering Algoritma K-Means Di Wilayah Kabupaten Cirebon

Misbakhul Anam[1], Annisa Maulana Majid[2], Ermanto[3]
[1,2,3)] Jurusan Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa
Jl. Inspeksi Kalimalang No.9 17530 Kabupaten Bekasi Jawa Barat
Email: misbah99@mhs.pelitabangsa.ac.id, annisa.maulanamajid@pelitabangsa.ac.id, ermanto@pelitabangsa.ac.id

ABSTRAK

Pengangguran merupakan salah satu permasalahan sosial dan ekonomi yang dihadapi oleh banyak daerah di Indonesia, termasuk salah satunya di wilayah Kabupaten Cirebon. Penelitian ini bertujuan untuk mengidentifikasi dan menganalisa lebih dalam terkait pola pengangguran di wilayah Kabupaten Cirebon pada tingkat Kecamatan menggunakan teknik clustering algoritma K-Means yang dipadukan dengan aplikasi RapidMiner. Data penelitian terkait pola pengangguran diperoleh dari Dinas Ketenagakerjaan Kabupaten Cirebon periode tahun terbaru yakni tahun 2024, yang meliputi data riwayat pencari kerja seperti Latar Belakang Pendidikan dari terendah sampai tertinggi (SD-S1), Jenis Kelamin, Alamat, serta data jumlah Penyerapan Angkatan Kerja Perusahaan berdasarkan Kecamatan. Hasil 3 cluster dengan kategori Pola Pengangguran Tinggi, Pengangguran Rendah, dan Pengangguran Sedang, dipilih sebagai hasil final dalam penelitian ini dengan nilai evaluasi DBI 0.063 yang menunjukan hasil kualitas clustering mendapatkan nilai yang sangat baik. Visualisasi grafik dan pemetaan wilayah menggunakan ArcGIS pada penelitian ini bertujuan untuk mempermudah dalam memberikan rekomendasi bagi pemerintah daerah, khususnya Dinas Ketenagakerjaan Kabupaten Cirebon, dalam merancang program dan kebijakan yang lebih tepat sasaran dan terstruktur untuk menanggulangi indikasi pola pengangguran yang ada di wilayah Kecamatan Kabupaten Cirebon berbasis data.

Kata kunci: Pengangguran, K-means, RapidMiner, Kabupaten Cirebon, DBI, ArcGIS

ABSTRACT

Unemployment is one of the social and economic problems faced by many regions in Indonesia, including Cirebon Regency. This study aims to identify and analyze in more depth the unemployment pattern in Cirebon Regency at the Sub-district level using the K-Means algorithm clustering technique combined with the RapidMiner application. Research data related to unemployment patterns were obtained from the Cirebon Regency Manpower Office for the latest period, namely 2024, which includes job seeker history data such as Educational Background from lowest to highest (Elementary School - Bachelor Degree), Gender, Address, and data on the number of Company Labor Force Absorption based on Sub-district. The results of three clusters, categorised as Highest Unemployment, Low Unemployment, and Medium Unemployment, were selected as the final results in this study, with a DBI evaluation value of 0.063, indicating that the clustering quality results achieved a very good value. Graphic visualization and regional mapping using ArcGIS in this study aims to facilitate in providing recommendations for local governments, especially the Cirebon Regency Manpower Office, in designing programs and policies that are more targeted and structured to overcome indications of unemployment patterns in the Cirebon Regency District area based on data.

Keywords: *Unemployment, K-means, RapidMiner, DBI, ArcGIS*

Pendahuluan

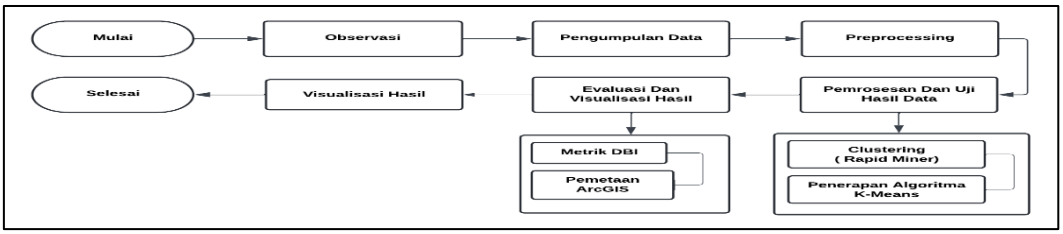
Pekerjaan merupakan salah satu sumber penghasilan yang sangat penting bagi seseorang untuk memenuhi kebutuhan hidup bagi dirinya dan keluarganya, dengan mendapatkan pekerjaan dan bekerja manusia bisa memenuhi keberlangsungan hidupnya dengan layak[1].Kebutuhan hidup yang layak merupakan hak setiap warga negara yang dijamin oleh hukum dasar Negara Kesatuan Republik Indonesia, bahwa setiap warga negara nya berhak mendapatkan pekerjaan dan kehidupan yang layak, hal ini sesuai dalam Pasal 27 Ayat (2) UUD 1945 yang berbunyi, “Tiap-tiap warga negara berhak atas pekerjaan dan penghidupan yang layak”[2]. Namun, saat ini terdapat banyak fenomena masyrakat yang sulit mendapatkan pekerjaan, menyebabkan munculnya pola masyarakat yang tidak atau belum mendapatkan pekerjaan yang disebut pengangguran. Pengangguran merupakan istilah untuk orang yang tidak bekerja sama sekali, sedang mencari pekerjaan, maupun seseorang yang sedang berusaha untuk mendapatkan pekerjaan yang layak[3]. Peneltian ini penulis buat berdasarkan studi kasus ditempat penulis tinggal yakni di Kabupaten Cirebon, dimana Kabupaten Cirebon sendiri memiliki karakteristik wilayah

dengan tingkat pengangguran yang bervariasi di setiap daerahnya. Melalui data pembuatan Kartu AK1 yang dikeluarkan oleh Dinas Tenaga Kerja (Disnaker) dan data penyerapan angkatan kerja di Perusahaan menjadi sumber informasi yang sangat penting untuk menganalisis kondisi pengangguran. Karena sebaran pengangguran sering kali dipengaruhi oleh berbagai faktor yang bervariatif, seperti latar belakang pendidikan yang rendah, keterampilan, ketersediaan lapangan kerja yang tidak seimbang, laju pertumbuhan ekonomi, serta besaran upah di suatu daerah [4].

Penelitian sebelumnya membahas analisis pengangguran mengunakan algoritma K-Means oleh Sintia Kusuma Arum, Rini Astuti, dan Fadhil Muhammad Basyar tahun 2024. Menerapkan pendekatan K-Means Clustering pada Dataset Tingkat Pengangguran Terbuka (TPT) Berdasarkan Tingkat Pendidikan Akhir dapat membantu mengidentifikasi kelompok atau Cluster Tingkat Pendidikan yang lebih rentan mengalami pengangguran[5]. Hasilnya menunjukkan pembagian data ke dalam tiga cluster dengan nilai DBI sebesar 0,487. Penelitian selanjutnya dilakukan oleh Dedy Sutris Martua Simanjuntak, Indra Gunawan, Sumarno, Poningsih, dan Ika Purnama Sari pada tahun 2023. Dengan jumlah cluster (k) = 3, menggunakan RapidMiner yang masing-masing merepresentasikan kategori tinggi, menengah, dan rendah dalam hal tingkat pengangguran. Hasil dari penelitian ini memperkuat efektivitas algoritma K-Medoids/K-Means dalam memetakan wilayah prioritas yang harus mendapat perhatian lebih dari Dinas Ketenagakerjaan[6]. Penelitian oleh Nurul Nurjanah, Nana Suarna, dan Willy Prihartono tahun 2024. Penelitian ini mengimplementasikan tingkat pengangguran di Provinsi Jawa Barat berdasarkan tingkat Pendidikan dengan aplikasi RapidMiner. Nilai DBI terbaik ditemukan pada k = 3 dengan skor 0.214, menandakan kualitas clustering yang cukup baik. Hasil penelitian ini adalah kategori kelompok pengangguran dibagi menjadi 3 yakni cluster rendah, sedang, dan tinggi , yang mengungkap bahwa pengangguran tertinggi terjadi pada tingkat pendidikan SMA, sedangkan pengangguran terendah berasal dari kategori tidak/belum sekolah[7]. Penelitian yang dilakukan oleh Siti Rohaniyah dan Ade Irma Purnamasari tahun 2023. Menggunakan data Dinas Tenaga Kerja dan Transmigrasi melalui situs Open Data Jabar untuk tahun 2021, menghasilkan 3 atribut utama yang digunakan: tingkat pendidikan, jumlah pencari kerja, dan tahun. Algoritma K-Means dijalankan dengan variasi jumlah cluster dari k = 2 hingga k = 10. Hasil terbaik diperoleh pada k = 6 dengan nilai Davies Bouldin Index (DBI) sebesar 0.185.Hasil dari penelitian menerangkan bahwa tingkat pendidikan Diploma I/II/III/Akademi mendominasi jumlah pencari kerja di Cluster 5, dengan jumlah tertinggi sebanyak 1.062 orang[8]. Penelitian selanjutnya dilakukan oleh Pastia dan Dikananda tahun 2023 menjelaskan bahwa penerapan algoritma K-Means efektif untuk mengelompokkan tingkat pengangguran berdasarkan data terbuka Provinsi Jawa Barat tahun 2013–2021. Penelitian ini menghasilkan dua kelompok utama, yaitu daerah dengan tingkat pengangguran tinggi dan rendah, dengan nilai Davies Bouldin Index (DBI) sebesar 0.733. Hasil analisis menunjukkan bahwa wilayah dengan pengangguran tinggi cenderung berada pada kabupaten dengan tingkat urbanisasi dan kepadatan penduduk yang tinggi, sedangkan daerah dengan pengangguran rendah umumnya memiliki aktivitas ekonomi agraris dan industri kecil menengah yang stabil[9].

Penelitian di atas menjadi landasan untuk penelitian yang penulis buat yakni untuk menganalisis pola pengangguran menggunakan metode clustering dengan algoritma K-Means, sebagai referensi dalam memahami distribusi pengangguran bagi pemerintah daerah, khususnya diwilayah Kabupaten Cirebon dalam menyusun langkah-langkah serta pengambilan kebijakan berbasis data dalam upaya penanganan tingkat pengangguran secara tepat, efektif, dan strategis di Tingkat Kecamatan Kabupaten Cirebon.

Metode Penelitian



Gambar 1. Tahapan Penelitian

1. Observasi

Penelitian ini menggunakan metode observasi non-partisipan, yakni metode pengumpulan data dengan observasi dimana peneliti tidak terlibat langsung dan hanya sebagai pengamat independen (Sugiyono, 2012:14)[10]. Dengan cara mengamati lingkungan atau situasi secara langsung, dapat dijadikan referensi untuk pertimbangan pengumpulan data yang diperlukan mengenai permasalahan yang ada.

2. Pengumpulan Data

Setelah memperoleh izin dan menyelesaikan prosedur pengambilan data, selanjutnya data dikumpulkan dalam satu file. Data ini merupakan data jenis kuantitatif. Dimana penelitian ini menggunakan pendekatan kuantitatif dikarenakan mengacu pada perhitungan analisis data penelitian yang berupa angka-angka atau pernyataan-pernyataan yang dinilai dan analisis dengan analisis statistik[11].

3. Preprocessing

Tahap preprocessing merupakan langkah untuk mempersiapkan data agar dapat dianalisis lebih lanjut. Tahap ini mencakup langkah-langkah yang sangat penting seperti:

a) Pembersihan Data:

Data awal yang diperoleh berupa data dengan format Excell, dimana data awal yang diperoleh merupakan data BNBA (*By Name By Address*), yang di dalamnya masih banyak terdapat noise atau data kosong dalam variable yang harus dilakukan pembersihan data terlebih dahulu agar informasi dalam data dapat di proses dengan baik. dalam proesnya pembersihan dilakukan menggunakan pengkodean bahasa phyton yang dipilih karena dilengkapi library seperti Pandas, Plotly, Numpy, scikit-learn dan XGBoost yang memudahkan dalam pengolahan data memungkinkan pengguna untuk melakukan analisis data, machine learning, serta visualisasi geografis dengan lebih mudah dan efisien[12].

b) Seleksi Data:

Pemilahan data merupakan proses seleksi fitur atau variabel yang benar-benar relevan dengan tujuan penelitian. variabel yang tidak mendukung analisis utama, seperti Nama, Nomor Kependudukan, Jenis lokasi pencari kerja (AKAL,AKAD, dan AKAN), dihapus dari dataset, namun tetap digunakan sebagai data pendukung. Hanya variabel yang berkaitan saja yang dipertahankan karena memiliki hubungan erat dengan tujuan analisis pengangguran.

c) Transformasi Data:

Transformasi data dilakukan untuk mengubah struktur atau format data agar lebih sesuai dengan kebutuhan analisis. Dalam penelitian ini, data variable juga di ubah menggunakkan abjad Z-P dari masing-masing nya menandakan urutan data rekap data tahunan kategori Jenis Kelamin, Pendidikan, Penyerapan Angkatan Kerja serta total pendaftar kartu AK1 di tiap Kecamatan tahun 2024. Berikut keterangan indikator abjad pada data:

Tabel 1. Keterangan Data

Z	Jumlah Laki-laki	Tahun 2024
Y	Jumlah Perempuan	Tahun 2024
X	SD/MI	Tahun 2024
W	SMP/MTS	Tahun 2024
V	SMA/MA	Tahun 2024
U	SMK	Tahun 2024
T	D3	Tahun 2024
S	D4/S1	Tahun 2024
P	Penyerapan Angkatan Kerja Lokal	Tahun 2024

d) Normalisasi Data:

Normalisasi dilakukan untuk menyetarakan skala antar variabel sehingga tidak ada satu variabel pun yang mendominasi proses clustering. Metode normalisasi seperti *Min-Max Scaling*, serta agregasi data menjadi bentuk *rasio* yang mengubah nilai data ke dalam rentang 0 hingga 1, sehingga algoritma K-Means dapat mengelompokkan data secara lebih seimbang dan akurat.

Tabel 2. Data Siap Proses

Kecamatan	Z	Y	X	W	…..	S	P
Arjawinangun	0,52471483	0,47528517	0,00382409	0,06309751	…..	0,07456979	0,09695817
Astanajapura	0,48979592	0,51020408	0,00431965	0,07343413	…..	0,05831533	1,24382385
Babakan	0,46803977	0,53196023	0,02733813	0,17769784	…..	0,03669065	0,76704545
Beber	0,46522782	0,53477218	0,00728155	0,0315534	…..	0,0461165	0,28297362
Ciledug	0,4622384	0,5377616	0,03525046	0,13914657	…..	0,04916512	0,83985441
Ciwaringin	0,48275862	0,51724138	0,00424628	0,07643312	…..	0,07430998	0,10526316
Depok	0,47669305	0,52330695	0,0079929	0,05150977	…..	0,04351687	0,11697449
…..	…..	…..	…..	…..	…..	…..	…..

Sebagian besar pada tahap preprocessing ini, penulis menggunakan pemrograman phyton yang di eksekusi pada platform Google Colabs yaitu platform inovatif berbasis cloud yang dirancang oleh Google untuk mempermudah proses penulisan eksekusi, dan berbagi kode program melalui Google Drive langsung melalui browser[13].

4. Pengujian Dan Evaluasi Hasil

RapidMiner sebagai tools utama perangkat lunak yang bersifat terbuka (open source) yang dapat digunakan dalam memberikan sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi (Aprilla Dennis 2013)[14].Algoritma K-Means dipilih sebagai metode untuk melakukan clustering data yang prinsiap kerja nya berusaha untuk mempartisi data yang ada ke dalam bentuk satu atau lebih cluster atau kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lainnya[15]. Inti dari algoritma ini adalah proses iteratif untuk menghitung jarak antara titik data dengan pusat cluster (centroid), dan menempatkan setiap titik data ke cluster dengan jarak terkecil. Berikut adalah langkah proses cara kerja algoritma K-means:

1. **Inisialisasi Jumlah Cluster (k)**
   Tentukan jumlah cluster yang diinginkan. Pemilihan nilai (k) dapat didasarkan pada kebutuhan penelitian (acak) atau melalui metode evaluasi.
2. **Penentuan Centroid Awal**
   Pilih secara acak k buah titik sebagai pusat awal (centroid) dari masing-masing cluster. Titik ini menjadi acuan awal untuk pembentukan kelompok data.
3. **Pengelompokan Data ke Cluster Terdekat**
   Hitung jarak antara setiap data dengan masing-masing centroid menggunakan rumus Euclidean Distance:

$d_e = \sqrt{\{(x_i - c_i)^2 + (y_i - d_i)^2\}}$

Keterangan:

- de = Jarak antara titik data dan centroid
- xi, yi = Koordinat titik data ke-i
- ci, di = Koordinat centroid cluster ke-i
- i = Banyaknya objek

4. **Penghitungan Ulang Centroid**
   Setelah data dikelompokkan, hitung ulang posisi centroid dengan cara mengambil rata-rata dari semua data dalam cluster:

$V_{\{ij\}} = \frac{1}{Ni} = \sum_{k=0}^{Ni} \text{Xkj}$

Keterangan:

- Vij= Centroid rata-rata pada Cluster ke – i untuk variabel ke-j
- Ni = Jumlah anggota Cluster ke-i
- i, k= Indeks dari Cluster
- j = Indeks variabel
- Xkj = Nilai data ke – k variabel ke – j untuk Cluster tersebut

5. **Iterasi hingga Konvergensi**
   Tahap ini merupakan pembukaan awal iterasi baru. Jika anggota cluster tidak berpindah cluster lagi, proses clustering selesai[16].Selain itu, penelitian ini menggunakan metode tambahan Metrik evaluasi DBI (Davies Bouldien Index) untuk mengukur sejauh mana kualitas cluster yang terbentuk saling terpisah dan seberapa rapat cluster tersebut. DBI merupakan metode validasi untuk mengukur kualitas hasil pengelompokan. Sebuah klaster dianggap optimal apabila memiliki nilai DBI yang semakin rendah, karena ini menunjukkan bahwa cluster-cluster tersebut lebih terpisah satu sama lain dan memiliki kedalaman internal yang lebih tinggi[17].

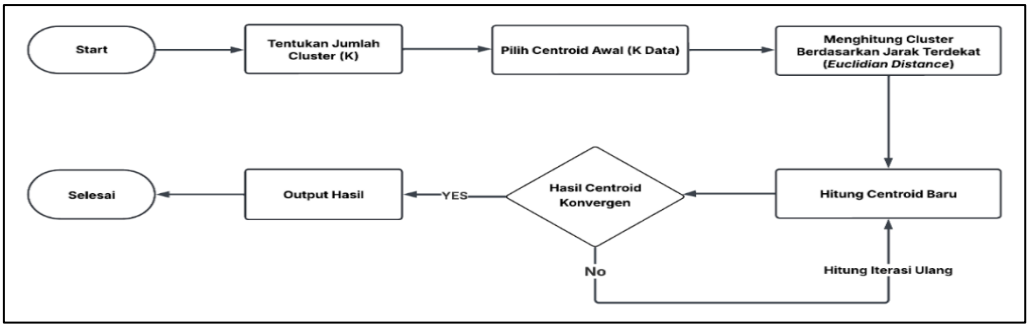
5. Visualisasi Hasil

Mengingat fokus penelitian ini adalah clustering untuk mengelompokan wilayah, selain menggunakan grafik Scatter Plot penulis memilih aplikasi pemetaan ArcGIS karena cocok untuk analisis spasial yang lebih mendalam mengenai pola pengangguran tingkat kecamatan di Kabupaten Cirebon. ArcGIS merupakan sistem informasi geografis yang dikembangkan oleh Esri (*Environmental Systems Research Institute*) yang digunakan untuk mengelola, menganalisis, dan memvisualisasikan data berbasis spasial[18].

Hasil Dan Pembahasan

1. Simulasi Manual Penerapan Algoritma K-Means

Tahap ini adalah tahap untuk mengimplementasikan perhitungan K-Means untuk uji data secara manual menggunakan perhitungan *Euclidian Distance*, simulasi alur penerapan algoritma K-means secara manual dapat di lihat pada diagram flow chart berikut:



Gambar 2. Simulasi Penerapan Algoritma K-Means

Metode perhitungan clustering menggunakan algoritma K-Means secara manual pada penelitian ini menggunakan *Microsoft Excel* untuk memastikan keakuratan proses serta kesesuaian hasil dengan perangkat lunak RapidMiner. Pertama tentukan jumlah cluster (k) secara asumsi awal, yaitu sebanyak 3 (tiga) cluster berdasarkan pertimbangan karakteristik data. Penentuan awal cluster (centroid) sebagai pusat cluster data dilakukan secara acak. Pada konteks penelitian ini penulis menentukan centroid awal dari data utama berupa Kecamatan Arjawinangun (Centroid 1/ C1) Kecamatan Astanajapura (Centroid 2/ C2) Kecamatan Babakan (Centroid 3/C3). Perhitungan cluster berdasarkan jarak terdekat dari setiap data menggunakan rumus *Euclidian Distance*, tujuanya yakni untuk menemukan jarak yang terdekat setiap data dari setiap centroid, penulis melampirkan tiga sampel perhitungan data dari baris dan kolom variabel Z sampai P berdasarkan nama Kecamatan sebagai sampel dalam menghitung jarak terdekat secara manual sesuai rumus *Euclidian Distance*. Berikut ini sampel perhitungan jarak data ke titik pusat cluster:

Cluster 1 (C1):

$d1 = \sqrt{(0{,}524714829 - 0{,}524714829)^2 + (0{,}475285171 - 0{,}475285171)^2 + (0{,}003824092 - 0{,}003824092)^2 + (0{,}063097514 - 0{,}063097514)^2 + (0{,}37667304 - 0{,}37667304)^2 + (0{,}455066922 - 0{,}455066922)^2 + (0{,}026768642 - 0{,}026768642)^2 + (0{,}07456979 - 0{,}07456979)^2 + (0{,}096958175 - 0{,}096958175)^2} = \mathbf{0}$

Cluster 2 (C2):

$d1 = \sqrt{(0{,}524714829 - 0{,}489795918)^2 + (0{,}475285171 - 0{,}510204082)^2 + (0{,}003824092 - 0{,}004319654)^2 + (0{,}063097514 - 0{,}073434125)^2 + (0{,}37667304 - 0{,}326133909)^2 + (0{,}455066922 - 0{,}529157667)^2 + (0{,}026768642 - 0{,}008639309)^2 + (0{,}07456979 - 0{,}058315335)^2 + (0{,}096958175 - 1{,}243823845)^2} = \mathbf{1{,}151730498}$

Cluster 3 (C3):

$d1 = \sqrt{(0{,}524714829 - 0{,}468039773)^2 + (0{,}475285171 - 0{,}531960227)^2 + (0{,}003824092 - 0{,}027338129)^2 + (0{,}063097514 - 0{,}177697842)^2 + (0{,}37667304 - 0{,}294244604)^2 + (0{,}455066922 - 0{,}458992806)^2 + (0{,}026768642 - 0{,}005035971)^2 + (0{,}07456979 - 0{,}036690647)^2 + (0{,}096958175 - 0{,}767045455)^2} =$ **0,691262779**…… seterusnya sampai baris dan kolom terakhir

Setelah dilakukan perhitungan cluster berdasarkan jarak terdekat antar centroid data, berikutnya adalah mengelompokkan data sesuai dengan hasil cluster yang telah di tentukan. Proses peritungan akan berlanjut ke iterasi berikutnya untuk menemukan hasil yang konvergen atau stabil.

2. Menghitung Dan Menentukan Centroid Baru

Perhitungan sebelumnya masih akan dilanjutkan, untuk menentukan hasil perhitungan *Iterasi* tidak mengalami perubahan. Dengan melakukan pengulangan hitungan dari tahap ***Iterasi 1*** sampai dengan ***Iterasi 2*** perhitungan menemukan posisi yang benar-benar stabil, menggunakan table *centroid baru*, diperoleh hasil yang konvergen pada perhitungan ***Iterasi ke 2*** berdasarkan nilai *mean* atau rata-rata dari tiap cluster dalam data. Berikut adalah hasil final ***Iterasi 2*** yang konvergen dan stabil:

Tabel 2. Centroid Baru

Centroid Baru	Z	Y	X	…..	…..	S	P
C1	0,478980928	0,521019072	0,009313579	…..	…..	0,056400496	0,17759919
C2	0,50095484	0,49904516	0,013607598	…..	…..	0,055883458	1,278672309
C3	0,466193702	0,533806298	0,023699667	…..	…..	0,043877484	0,763298232

Tabel 3. Hasil Iterasi 2

Kecamatan	C1	C2	C3	Jarak Terdekat	Cluster	Kecamatan	C1	C2	C3	Jarak Terdekat	Cluster
Arjawinangun	0,2124 16679	1,1855 83611	0,6840 64956	0,2124 16679	1	Lemahabang	0,6259 32062	0,4944 33269	0,1245 2123	0,1245 2123	3
Astanajapura	1,0708 26232	0,0747 9162	0,4869 89084	0,0747 9162	2	Losari	1,3792 32878	0,3216 16656	0,7924 71749	0,3216 16656	2
Babakan	0,6222 4004	0,5189 50023	0,0759 5129	0,0759 5129	3	Mundu	0,2036 13597	0,9360 94274	0,4283 03978	0,2036 13597	1
Beber	0,2205 19022	1,0058 39105	0,5083 07717	0,2205 19022	1	Pabedilan	0,7921 84923	0,5115 88304	0,3132 33595	0,3132 33595	3
Ciledug	0,7201 52628	0,4639 41134	0,1955 45575	0,1955 45575	3	Pabuaran	0,6734 95208	0,4759 4807	0,1088 1793	0,1088 1793	3
Ciwaringin	0,1725 47001	1,1751 38902	0,6655 01165	0,1725 47001	1	Palimanan	0,0866 0621	1,1692 35689	0,6614 32147	0,0866 0621	1
Depok	0,0962 60436	1,1814 48389	0,6706 72917	0,0962 60436	1	Pangenan	0,2600 8491	0,8992 95048	0,4184 15708	0,2600 8491	1
Dukuptuntang	0,1146 78984	1,1604 89991	0,6449 03993	0,1146 78984	1	Pangurangan	0,4637 78668	1,2851 73827	0,7987 04084	0,4637 78668	1
Gebang	0,5748 80688	0,5512 58978	0,0691 09795	0,0691 09795	3	Pasaleman	0,5928 86139	0,5319 28325	0,0358 22471	0,0358 22471	3
Gegesik	0,0344 6612	1,1156 01364	0,6039 1818	0,0344 6612	1	Plered	0,2105 29781	1,1535 12656	0,6710 86484	0,2105 29781	1
Gempol	0,1515 56562	1,0107 30831	0,4930 89957	0,1515 56562	1	Plumbon	0,1340 4152	1,1620 32341	0,6646 05092	0,1340 4152	1
Greged	0,3718 78522	0,7752 5242	0,2712 72577	0,2712 72577	3	Sedong	0,4237 98672	0,8060 37316	0,3864 14215	0,3864 14215	3
Gunungjati	0,1035 33811	1,1378 62939	0,6297 82712	0,1035 33811	1	Sumber	0,0559 28543	1,1152 03634	0,6093 32466	0,0559 28543	1
Jamblang	0,0584 98726	1,1627 78836	0,6492 21255	0,0584 98726	1	Sura Nenggala	0,2211 70602	1,1306 29061	0,6214 17028	0,2211 70602	1
Kaliwedi	0,2105 37981	1,0650 34355	0,5816 75492	0,2105 37981	1	Susukan	0,1325 48888	1,1334 0892	0,6163 83408	0,1325 48888	1
Kapetakan	0,0931 70285	1,1252 26251	0,6178 71365	0,0931 70285	1	Susukan Lebak	0,6254 18409	0,4927 92434	0,1002 76715	0,1002 76715	3
Karang Sembung	0,7169 36371	0,3999 26021	0,1456 63669	0,1456 63669	3	Talun	0,1420 14875	1,1039 08185	0,6058 46831	0,1420 14875	1
Karang Wareng	0,8989 03292	0,3110 33491	0,2456 59136	0,2456 59136	3	Tengah Tani	0,2401 03243	1,1686 99104	0,6903 84705	0,2401 03243	1
Kedawung	0,2123 83089	1,1457 82015	0,6531 00144	0,2123 83089	1	Waled	0,6911 33994	0,4829 38502	0,1710 55177	0,1710 55177	3
Klangenan	0,0732 56705	1,1654 78453	0,6523 51828	0,0732 56705	1	Weru	0,2385 60501	1,1947 79425	0,7137 38165	0,2385 60501	1

Setelah dilakukan perhitungan centroid hasil dari *iterasi 2* diatas, dapat disimpulkan hasil clustering data posisi antar centroid akhirnya menemukan hasil yang stabil, Dengan ini perhitungan manual telah selesai dilakukan dan telah menemukan hasil akhir yang sama. Selanjutnya kita akan mencocokan hasil perhitungan ini dengan tools utama yang digunakan pada penelitian ini yakni aplikasi RapidMiner.

3. Implementasi Tools Rapid Miner

Tahap ini adalah tahap inti dari proses clustering data menggunakan algoritma K-means,dengan menggunakan tools RapidMiner. Secara alur, penggunaan aplikasi ini tidak jauh berbeda dari alur proses clustering K-Means pada umumnya hanya saja penggunaan aplikasi ini lebih praktis dan mudah dalam melakukan clustering data. Berikut ini adalah alur serta tahapan implementasi clustering dengan algoritma K-Means dengan menggunakan tools RapidMiner:

1) Tahapan Input
Setelah membuka workspace pada aplikasi RapidMiner pilih operator **Read CSV** lalu input (impor) file data yang akan kita proses. Karena jenis data awal berbentuk excel untuk menyesuaikan dalam pemrosesan maka data di ubah terleih dahulu pada format csv. lakukan pemeriksaan atau pengecekan terhadap nilai yang hilang (missing values). Penanganan nilai yang hilang ini penting untuk menjaga kualitas data agar hasil analisis lebih akurat dan tidak bias.

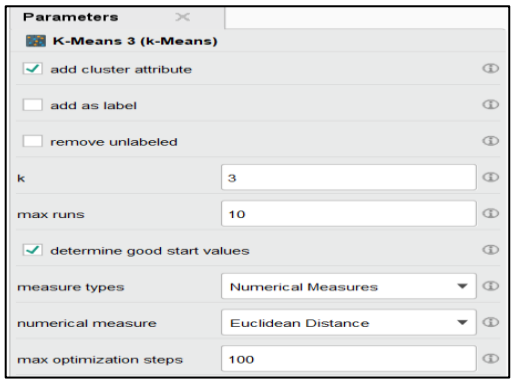
2) Tahapan Transformasi
Pisahkan atribut variabel huruf dan numerik agar proses clustering dapat berjalan dengan baik. Pada konteks penelitian ini variabel Kecamatan adalah *primary key* (variabel unik). Maka dari itu variabel ini akan di edit menjadi tipe data atribut *Id* agar proses clustering nantinya akan optimal. Dengan variabel numerical saja (Interger). Klik **"Change Role"**, lalu pilih atribut *Id* untuk variabel tipe data Kecamatan.Mengingat proses clustering nantinya akan di lakukan di beberapa cabang cluster, maka di perlukan operator yang berperan sebagai duplikasi data. Jadi, setelah seleksi atribut data selesai, selanjutnya tambahkan operator "**Multiply"**. Operator ini berfungsi untuk menduplikasi dataset agar dapat digunakan dalam beberapa proses secara paralel.

3) Tahapan Uji Hasil
Langkah selanjutnya adalah langkah inti dari penelitian ini, yakni proses clustering dan uji hasil pemrosesan dataset. Sebelum proses cluster dan uji hasil dilakukan, penulis terlebih dahulu mengatur parameter fitur clustering algortitma K-Means, seperti centang fitur:

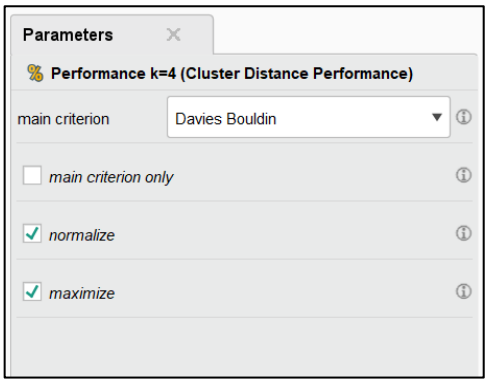
a) ***add cluster attribute***
Atribut ini akan berisi id cluster (misalnya, Cluster 0, Cluster 1, Cluster 2) untuk setiap baris data, menunjukkan ke klaster mana data tersebut ditetapkan, selanjutnya menentukan jumlah cluster (k), pada penelitian ini penulis melakukan uji pada jumlah cluster 3,4, dan 5 sebagai perbandingan hasil terbaik nantinya. Sedangkan parameter (***max run 10)*** yang berarti instruksi perhitungan iterasi akan dilakukan 10 kali untuk menghindari terjebak pada minimum local dengan setiap *run* menggunakan inisialisasi titik awal centroid yang telah ditingkatkan kualitasnya.

b) determine good start values
Yang berfungsi untuk meningkatkan kualitas hasil klastering dan mempercepat konvergensi karena pemilihan centroid awal yang lebih baik. parameter *measure types* ***(numerical measures)*** dipilih karean menyesuaikan tipe data penelitian ini yakni data numerik (kuantitatif), pilih fitur **Euclidian Distance** sebagai pengukuran jarak antara titik data dan centroid klister, selain itu proses optimasi dalam setiap *run* dibatasi hingga maksimum 100 langkah iterasi ***(max optimization steps 100)*** untuk mencapai konvergensi atau mencapai batas komputasi.

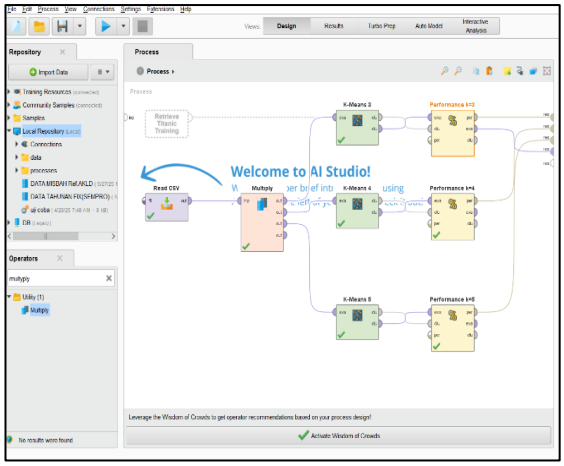
Selanjutnya, pilih operator **Performance** *(Cluster Distance Performance)* yang berfungsi untuk menghitung metrik performa klastering berdasarkan jarak antar klaster dan kepadatan antar klaster. Penamaan jumlah (k) pada operator **Performance** ditujukan agar mempermudah dalam melihat hasil uji dataset. Selanjutnya pilih *main criterion* metrik DBI (Davies Boudin Index) sebagai metrik evaluasi hasil dalam penelitian ini, kemudian centang parameter *normalize* yang berfungsi menskalakan nilai-nilai ke rentang tertentu (misalnya, 0-1), bertujuan untuk memudahkan perbandingan. Karena hasil akhirnya tidak dikalikan dengan minus satu, centang maximize dipilih sehingga menggasilkan nilai DBI yang positif. Terakhir lakukan pengujian hasil dataset dengan tekan tombol "proses" yang teresdia di halaman workspace RapidMiner.



Gambar 3*. Parameter K-Means*



Gambar 4*. Parameter Performance*



Gambar 5. Proses Uji Data Clustering

Row No.	KECAMATAN	cluster	Rasio_Laki	Rasio_Pere..	Rasio_X	Rasio_W	Rasio_V	Rasio_U	Rasio_T
1	ARJAWINAN..	cluster_0	0.525	0.475	0.004	0.063	0.377	0.455	0.027
2	ASTANAJAPU..	cluster_1	0.490	0.510	0.004	0.073	0.326	0.529	0.009
3	BABAKAN	cluster_2	0.468	0.532	0.027	0.178	0.294	0.459	0.005
4	BEBER	cluster_0	0.465	0.535	0.007	0.032	0.420	0.485	0.010
5	CILEDUG	cluster_2	0.462	0.538	0.035	0.139	0.394	0.370	0.012
6	CIWARINGIN	cluster_0	0.483	0.517	0.004	0.076	0.348	0.473	0.023
7	DEPOK	cluster_0	0.477	0.523	0.008	0.052	0.226	0.663	0.009
8	DUKUPUNTA..	cluster_0	0.442	0.558	0.007	0.069	0.310	0.533	0.007
9	GEBANG	cluster_2	0.495	0.505	0.016	0.145	0.304	0.480	0.008
10	GEGESIK	cluster_0	0.478	0.522	0.011	0.031	0.289	0.584	0.025
11	GEMPOL	cluster_0	0.466	0.534	0.051	0.124	0.219	0.556	0.005
12	GREGED	cluster_2	0.393	0.607	0.007	0.072	0.296	0.563	0.013
13	GUNUNGJATI	cluster_0	0.463	0.537	0.006	0.039	0.191	0.681	0.023
14	JAMBLANG	cluster_0	0.468	0.532	0.007	0.049	0.277	0.616	0.016
15	KALIWEDI	cluster_0	0.611	0.389	0	0.035	0.332	0.559	0.026
16	KAPETAKAN	cluster_0	0.507	0.493	0.017	0.063	0.209	0.658	0.009
17	KARANGSEM..	cluster_2	0.484	0.516	0.009	0.088	0.270	0.598	0.006

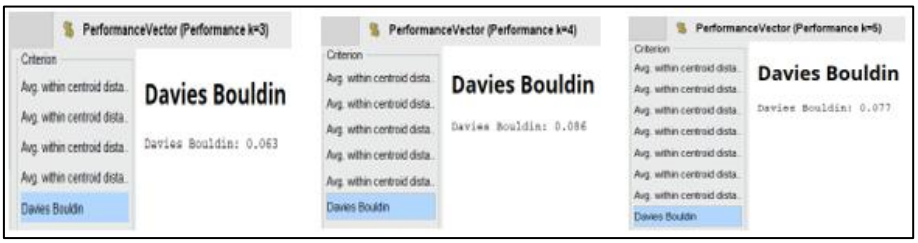
Gambar 6. Hasil Uji Data Clustering

Hasil dari uji hasil di atas menghasilkan 3 clsuter sebagai hasil pengelompokan dengan kulaitas pengelopokan terbaik, yang di indikasikan *cluster 0 , cluster 1 dan cluster 2* atau *cluster 1, cluster 2, dan cluster 3* dalan kategori hitung manual Microsoft Excell.
Keterangan Hasil Jumlah Cluster k=3 :

a) *Cluster 0* : Kecamatan Arjawinangun, Kecamatan Beber, Kecamatan Ciwaringin, Kecamatan Depok, Kecamatan Dukupuntang, Kecamatan Gegesik, Kecamatan Gempol, Kecamatan Gunungjati, Kecamatan Jamblang, Kecamatan Kaliwedi, Kecamatan Kapetakan, Kecamatan Kedawung, Kecamatan Mundu, Kecamatan Palimanan, Kecamatan Pangenan, Kecamatan Panguragan, Kecamatan Plered, Kecamatan Plumbon, Kecamatan Sumber, Kecamatan Suranenggala, Kecamatan Susukan, Kecamataan Talun, Kecamatan Tengah Tani, Kecamatan Weru.
b) *Cluster 1 :* Kecamatan Astanajapura, Kecamatan Losari
c) *Cluster 2 :* Kecamatan Babakan, Kecamatan Ciledug, Kecamatan Gebang, Kecamatan Greged, Kecamatan Karangsembung, Kecamatan Karangwareng, Kecamatan Lemahabang, Kecamtan Pabedilan, Kecamatan Pabuaran, Kecamatan Pasaleman, Kecamatan Sedong, Kecamatan Susukan Lebak, Kecamatan Waled.

4) Tahapan Evaluasi

Setelah diperoleh hasil dan nilai dari masing-masing jumlah cluster yang telah penulis tentukan (Cluster k = 3, k = 4, k = 5). Berdasarkan hasil uji clustering yang telah dilakukan tertera evaluasi metrik DBI dari cluster jumlah k = 3 aalah 0.063, k = 4 adalah 0.086, dan k = 5 adalah 0.077, nilai yang mendekati angka 0, dalam konteks pengukuran metrik DBI menandakan proses uji hasil dari clustering memperoleh hasil yang sangat baik. Oleh karena itu, jumlah (k = 3) dipilih untuk hasil final pada penelitian ini.

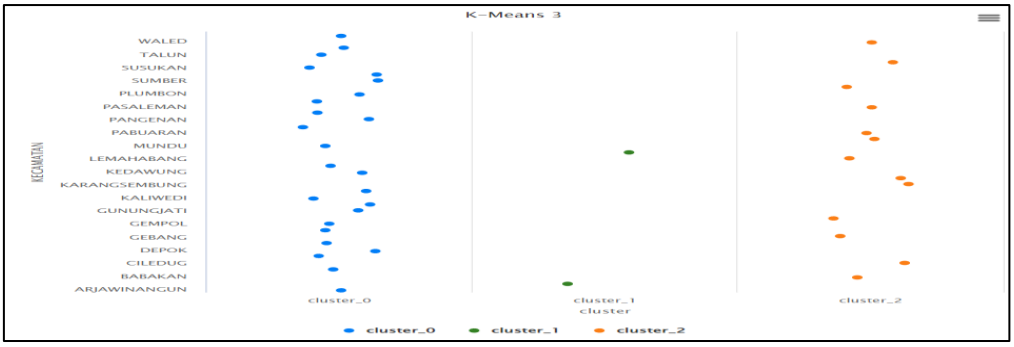


Gambar 7. Hasil Evaluasi Metrik DBI

5) Tahapan Visualisasi

Pemrosesan uji hasil clustering serta evaluasi data yang telah di peroleh akan di visualisasikan dengan menggunakan grafik Scatter Plot dan pemetaan wilayah menggunakan aplikasi geospasial ArcGIS. Tujuan nya yakni untuk menampilkan hasil visual yang lebih menarik dan mudah untuk difahami sebagai sarana untuk analisis data yang mengandung informasi penting terkait gejala dan pola pengangguran.

a) Grafik Scatter Plot

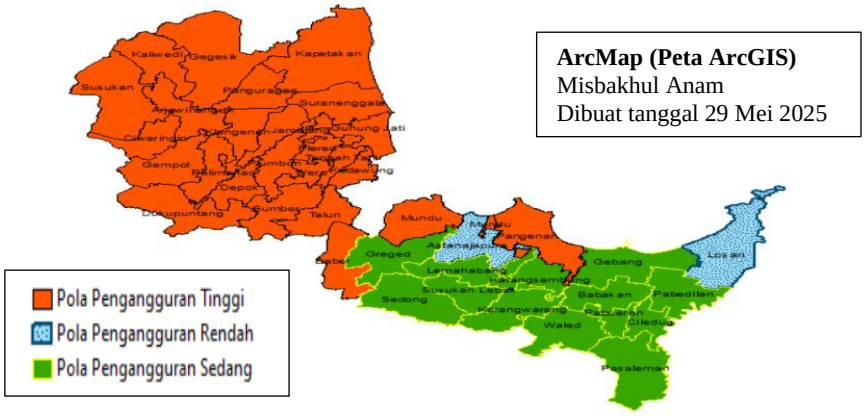


Gambar 8. Grafik Scatter Plot

Grafik ini menampilkan pembagian kecamatan ke dalam tiga cluster berbeda, masing-masing diwakili oleh warna sebagai pembeda antara cluster lainya. Biru untuk Cluster 0, Hijau untuk Cluster 1, dan Orange untuk Cluster 2. Setiap titik merepresentasikan satu Kecamatan di wilayah Kabupaten Cirebon, sehingga pola persebaran antar cluster dapat terlihat jelas secara visual berdasarkan masing-masing kecamatan, seperti yang terlihat dalam grafik scatter plot terlihat jelas terdapat hasil kelompok cluster kecil (Cluster 2), dimana berdasarkan anlisis mendalam data kelompok ini adalah kelompok yang memuat variabel informasi extreme atau sangat berbeda dari kelompok cluster lain, yang mana hanya memuat 2 kelompok Kecamatan.

b) Pemetaan ArcGIS

Proses visualisasi ArcGIS diawali dengan mengupload file dalam format *.shp* yang di dalamnya berisi data-data wilayah Kecamatan yang ada di Kabupaten Cirebon. Sebelum melakukan analisa lebih lanjut, lakukan *setting* variabel dalam data dengan cara memilih opsi *Open Attribut Table*, kemudian sesuaikan format tampilan peta dengan hasil jumlah pengujian clustering algoritma K-Means berdasarkan masing masing kelompok wilayah cluster Kecamatan Kabupaten Cirebon.



Gambar 9. Hasil Pemetaan ArcGIS

Berdasarkan gambar hasil pemetaan ini, pemberian warna sesuai kelompok wilayah cluster dilakukan untuk menandai pemisahaan kelompok masing-masing cluster, dimulai dari Warna Orange (cluster 0) sebagai terindikasi Pola Pengangguran Tinggi, Warna Hijau (cluster 1)Pengangguran Sedang, dan Warna Biru Bercorak (cluster 2) Pengangguran Rendah.

Penjelasan analisis wilayah :

1. **Cluster 0 :** Merupakan wilayah yang memiliki indikasi Pola Pengangguran Tertinggi,dimana wilayah-wilayah kelompok ini, perlu mendapatkan perhatian khusus dari pemerintah daerah khususnya Disnaker Kabupaten Cirebon. Mayoritas wilayah ini ada di bagian Cirebon Barat atau dekat pusat perkotaan. Berdasarkan analisis mendalam dari data jumlah pelamar kerja kartu kuning (AK1) kurang lebih tercatat warga masyarakat di 25 Kecamatan cluster 0, adalah Masyarakat pencari kerja yang mengenyam pendidikan rendah sampai tinggi, berdasarkan data banyak diantaranya memiliki riwayat pendidikan SD maupun SMP sebagai parameter pendidikan rendah dalam data,namun pelamar kerja kelompok ini di dominasi oleh kelompok pendidikan SMA/SMK, bahkan lebih menariknya lagi kelompok cluster ini adalah Kecamatan dengan catatan latar belakang pendidikan tertinggi dibanding dengan kelompok kecamatan cluster lain, seperti Kecamatan Sumber dan Plumbon yang memegang jumlah pola pengangguran terbanyak dengan dominasi pencari kerja berlatar belakang Pendidikan SMA/SMK, maupun D3, dan S1. Selain itu cluster ini, memiliki *demand* atau penyerapan kerja yang sangat rendah yang mana

hanya sekitar kurang dari 20%, bahkan beberapa Kecamatan seperti Kecamatan Arjawinangun, dan Ciwarigin yang hanya menyentuh di bawah presentasi 10%, beberapa faktor sangat mempengaruhi, dari persaingan yang ketat di sekitar sektor industri wilayah yang berada dekat pusat perkotaan yang padat penduduk, ketidaksesuaian kualifikasi pasar dengan lowongan yang dibutuhkan perusahaan (Gap Skill) mengingat angkatan ini adalah angkatan kerja yang sebagian besarnya berfokus di sektor perusahaan . Hal ini juga dibuktikan dari data aktual terjadi lonjakan para pencari kerja (Pendaftar Kartu AK1) yang sangat *extreme* di wilayah Kecamatan kelompok cluster 0. Dibandingkan dengan Kecamatan yang lain, jumlah pencari kerja kelompok wilayah cluster ini sangat tajam berbanding terbalik dengan angka penyerapan angkatan kerja nya baik dari angkatan lokal (AKAL), luar daerah (AKAD), maupun luar negeri (AKAN)

2. **Cluster 1 :** Merupakan kelompok cluster yang terindikasi sebagai daerah dengan indikasi Pola Pengangguran Terendah, artinya kelompok cluster ini merupakan kelompok daerah dengan pemusatan dan penyerapan tenaga kerja yang sangat baik. Kelompok ini hanya terdiri dari dua Kecamatan yakni Kecamatan Astanajapura dan Losari, , kedua Kecamatan ini secara geografis terletak di Kabupaten Cirebon Timur, yang keduanya merupakan wilayah dengan kapasitas pemusatan industri yang sangat besar di berbagai sektor industri yang dikenal dengan wilayah Industri Pantura. Berdasarkan analisis mendalam, data riwayat pendidikan para pencari kerja kelompok ini terbilang merata, tercatat masyarakat pencari kerja masih ada yang berlatar belakang lulusan pendidikan SD dan SMP yang sangat minim, namun secara keseluruhan presentasi *suply* tenaga kerja dengan angka penyerapan kerja di kelompok wilayah ini sangat merata dari SD sampai S1, hal ini di buktikan dengan data aktual yang mecatatkan secara presentasi angka penyerapan tenaga kerja mencapai hingga 85% lebih. Beberapa faktor seperti banyaknya variasi sektor lapangan pekerjaan yang tersedia seperti pergudangan, pelabuhan, pertanian, logistik dan berbagai macam sektor lainya serta tingkat persaingan yang rendah melihat kondisi secara geografis wilayah ini terletak jauh dari pusat kota yang memiliki populasi dan kepadatan penduduk yang sangat tinggi. Minimnya persaingan, persebaran lapangan kerja yang bervariatif merupakan salah satu faktor populasi kedua Kecamatan ini lebih banyak terserap di berbagai bidang pekerjaan.
3. **Cluster 2 :** Kelompok ini adalah wilayah yang memiliki indikasi Pola Pengangguran Sedang, tidak terlalu banyak namun juga tidak begitu sedikit, kelompok wilayah cluster ini beberapa wilayah berada di bagian barat atau pusat kota , sebagian pula berada di pusat industri pantura yakni bagian timur, dalam hal ini, selaku wilayah yang dapat dikatakan sebagai wilayah transisi antara sektor barat dan timur masyarakat pada wilayah cluster ini cenderung memiliki banyak opsi dan peluang mendapatkan akses baik dari wilayah Barat maupun wilayah Timur. Sebagian masyrakat pencari kerja masih banyak terdapat lulusan pendidikan rendah (SD, dan SMP), seperti Kecamatan Pabedilan, dan Waled, namun walaupun demikian pada wilayah-wilayah ini tingkat penyerapan serta pemusatan tenaga kerja cukup baik rata-rata menyentuh presentase 40%. Tetapi dalam hal ini Pemerintah setempat juga perlu melakukan eksplorasi dan pengajian serta langkah awal untuk evaluasi serta melakukan perbaikan khususnya dibidang bursa transfer kerja atau BKK agar perkembangan pada wilayah cluster ini dapat berkembang lebih baik.

Simpulan

Pada penelitian ini dapat disimpulkan bahwasanya, penerapan metode clustering data dengan menggunakan algoritma K-Means berhasil mengelompokkan wilayah Kecamatan di Kabupaten Cirebon dalam beberapa cluster berdasarkan pola tingkat pengangguran. Jumlah hasil 3 cluster dipilih sebagai hasil final karena menhasilkan nilai terbaik dibanding kluster lainya berdasarkan karakteristik masing-masing kelompok wilayah dengan indikasi kategori wilayah dengan Pola Pengangguran Tinggi, Pola Pengangguran Sedang, dan Pola Pengangguran Rendah. Nilai evaluasi metrik DBI 0.063 menunjukan hasil performa clustering yang sangat baik yang menandakan pemisahan antar cluster yang jelas dan optimal. Visualisasi hasil clustering dari grafik Scatter Plot dan ArcGIS memberikan gambaran sebaran wilayah secara geografik dan spasial yang lebih konkret, di mana wilayah-wilayah bagian timur Kabupaten Cirebon cenderung masuk dalam cluster dengan tingkat pola pengangguran yang lebih rendah, menariknya, wilayah di bagian barat dan tengah yang merupakan daerah dekat dengan industri dan pemerintahan justru cenderung menunjukkan tingkat pengangguran yang lebih tinggi. Temuan ini, mengindikasikan bahwa tingginya konsentrasi industri serta tingginya kapasitas sumber daya pendidikan yang tinggi tidak selalu menjadi jaminan rendahnya tingkat pengangguran di suati wilayah, banyak faktor yang harus di imbangi dengan berbagai hal seperti kepadatan penduduk (pendatang dan pribumi) dalam suatu daerah, ketidaksesuaian pasar atau kulafikasi dengan lapangan pekerjaan yang terbatas, masih banyaknya *Gap Skill* atau kurangnya keterampilan, serta kapasitas perusahaan wilayah industri yang belum optimal dalam melakukan pemerataan penyerapan tenaga kerja, khususnya di wilayah *cluster 0*. Oleh karena itu, hasil dari penelitian ini ditujukan agar dapat menjadi referensi serta informasi yang berguna bagi pemerintah daerah, khususnya Dinas Ketenagakerjaan Kabupaten Cirebon, dalam merancang strategi penanggulangan pengangguran yang lebih efektif, terarah, dengan berbasis data. Seperti melakukan melakukan intervensi kebijakan yang lebih spesifik terhadap wilayah yang perlu mendapat perhatian khusus. Bentuk intervensi dapat berupa pelatihan kerja berbasis kebutuhan industri yang merata, peningkatan akses informasi lowongan kerja via website yang lebih efisien di bawah program seperti Bursa Transfer Kerja (BKK), serta kolaborasi antara sektor swasta dan pendidikan vokasi untuk menciptakan program

siap kerja yang lebih terarah baik dari riwayat pendidikan rendah sampai pendidikan tinggi (SD-S1) agar menciptakan sumber daya putra putri daerah yang berpendidikan tinggi dan memiliki kualitas.

Daftar Pustaka

[1] H. Damaryanti, S. A. Alkadrie, and A. Annurdi, "Pemenuhan Upah Minimum Sebagai Upaya Perlindungan Hak Konstitusional," *J. Huk. Media Bhakti*, vol. 1, no. 2, pp. 108–120, 2020, doi: 10.32501/jhmb.v1i2.8.

[2] F. Supriyadi and T. Lesmana, "Jaminan Hukum Atas Pemenuhan Kebutuhan Hidup Layak (KHL) Pada Pekerja di Indonesia Ferry Supriyadi dan Teddy Lesmana Program Studi Ilmu Hukum Universitas Nusa Putra Sukabumi Pendahuluan Pekerja atau buruh merupakan setiap orang yang bekerja un," vol. 2, no. 1, pp. 2961–8754, 2023.

[3] R. M. Sabiq and N. C. Apsari, "Dampak Pengangguran Terhadap Tindakan Kriminal Ditinjau Dari Perspektif Konflik," *J. Kolaborasi Resolusi Konflik*, vol. 3, no. 1, p. 51, 2021, doi: 10.24198/jkrk.v3i1.31973.

[4] S. Aprizkiyandari, N. Satyahadewi, A. N. Pratama, R. Rivaldo, S. I. Nurdiansyah, and S. Helena, "Implementasi K-Means Cluster untuk Menentukan Persebaran Tingkat Pengangguran," *Empiricism J.*, vol. 4, no. 2, pp. 400–406, 2023, doi: 10.36312/ej.v4i2.1518.

[5] S. Kusuma Arum, R. Astuti, and F. Muhammad Basysyar, "Penerapan Algoritma K-Means Pada Dataset Pengangguran Terbuka Berdasarkan Pendidikan Di Provinsi Jawa Barat," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 2, pp. 2221–2226, 2024, doi: 10.36040/jati.v8i2.9440.

[6] D. S. M. Simanjuntak, I. Gunawan, S. Sumarno, P. Poningsih, and I. P. Sari, "Penerapan Algoritma K-Medoids Untuk Pengelompokkan Pengangguran Umur 25 tahun Keatas Di Sumatera Utara," *J. Krisnadana*, vol. 2, no. 2, 2023, doi: 10.58982/krisnadana.v2i2.264.

[7] 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.

[8] S. R. Hani, "Clustering Data Pencari Kerja Menurut Tingkat Pendidikan Menggunakan Algoritma K-Means," *J. Minfo Polgan*, vol. 12, no. 1, pp. 1–14, 2023, doi: 10.33395/jmp.v12i1.12217.

[9] Nok Imas Pastia & Fatiha Nursari Dikananda, "Pengelompokan Data Pengangguran Terbuka Menggunakan Algoritma K-Means Berdasarkan Provinsi Jawa Barat," *J. Din. Inform.*, vol. 12, no. 1, pp. 45–53, 2023, doi: 10.29406/cobit.v1i01.6644.

[10] C. I. D. P. Yanthi and A. Marhaeni, "Pengaruh Pendidikan, Tingkat Upah Dan Pengangguran Terhadap Persentase Penduduk Miskin Di Kabupaten/Kota Provinsi Bali," *J. Piramida*, vol. 11, no. 2, pp. 68–75, 2015, [Online]. Available: https://ojs.unud.ac.id/index.php/piramida/article/download/23280/15301

[11] F. Naomi, G. M. V Kawung, and I. P. F. Rorong, "Pengaruh Inflasi dan Pengangguran Terhadap Kemiskinan di Kota Manado Periode 2007 - 2020," *J. Berk. Ilm. Efisiensi*, vol. 22, no. 6, pp. 97–108, 2022.

[12] R. M. Koretsky, "Python3," *Raspberry Pi OS Syst. Adm. with Syst. Python*, vol. 8, no. 1, pp. 175–305, 2023, doi: 10.1201/b23421-3.

[13] I. Nabilla Audy, T. Nur Padilah, and B. Nurina Sari, "Pengelompokan Daerah Rawan Bencana Alam Di Jawa Barat Menggunakan Algoritma Fuzzy C-Means," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 4, pp. 2799–2803, 2024, doi: 10.36040/jati.v7i4.7205.

[14] H. Rosika *et al.*, "PAKAIAN MENGGUNAKAN METODE K-MEANS Pada era digital perkembangan teknologi semakin berkembang khusunya pengelolaan data penjualan menjadi semakin krusial bagi bisnis untuk memahami perilaku konsumen , mengedentifikasi beberapa kelompok atau cluster memili," vol. 5, pp. 221–231, 2024.

[15] D. D. Darmansah and N. W. Wardani, "Analisis Pesebaran Penularan Virus Corona di Provinsi Jawa Tengah Menggunakan Metode K-Means Clustering," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 8, no. 1, pp. 105–117, 2021, doi: 10.35957/jatisi.v8i1.590.

[16] L. Fimawahib and E. Rouza, "Penerapan K-Means Clustering pada Penentuan Jenis Pembelajaran di Universitas Pasir Pengaraian," *INOVTEK Polbeng - Seri Inform.*, vol. 6, no. 2, p. 234, 2021, doi: 10.35314/isi.v6i2.2096.

[17] M. Minarni, E. I. Sari, A. Syahrani, and P. Mandarani, "Klasterisasi Penyakit Menggunakan Algoritma K-Medoids pada Dinas Kesehatan Kabupaten Agam," *J. Nas. Pendidik. Tek. Inform.*, vol. 10, no. 3, p. 137, 2021, doi: 10.23887/janapati.v10i3.34904.

[18] D. Libora and A. J. Saputra, "Analisa Potensi Genangan Banjir Menggunakan Arc-Gis Di Pulau," *J. Civ. Eng. Plan.*, vol. 4, no. 2, pp. 308–318, 2023, doi: 10.37253/jcep.v4i2.9063.