

# PENERAPAN ALGORITMA K-MEANS UNTUK MENGELOMPOKKAN KABUPATEN/KOTA DALAM UPAYA PEMETAAN LAPANGAN PEKERJAAN BARU

Bella Nurseptia, Apriade Voutama, Nono Haeryana

*Program Studi Sistem Informasi, Universitas Singaperbangsa Karawang*

Jl. HS.Ronggo Waluyo, Karawang

bella.nurseptia19007@student.unsika.ac.id, apriade.voutama@staff.unsika.ac.id,

nono@staff.unsika.ac.id

**Abstract** - Being a developing nation, Indonesia struggles with unemployment. Growing the number of communities in an area can have a lot of negative effects. Optimizing the creation of new jobs is the best approach to raise awareness. notes West Java is one of the Indonesian provinces with a high unemployment rate, according to the Indonesian Central Statistics Agency (BPS). In reality, West Java Province has just become the first province in Indonesia with significant volumes of foreign investment (PMA). This study will classify districts/cities in West Java based on the attributes of TPT, TPAK, and PMA. From the data processed using the K-Means algorithm, it produces 3 clusters, where the third cluster has different characteristics so that it can help the government in West Java optimize the mapping of new jobs.

**Keywords** - Data Mining, Clustering, KMeans, New Jobs

**Abstrak** - Pengangguran menjadi suatu permasalahan yang dihadapi Indonesia sebagai negara berkembang. Banyak dampak negatif yang dapat muncul akibat tingginya jumlah pengangguran di suatu wilayah. Cara untuk mengatasi tingginya tingkat pengangguran adalah dengan mengoptimalkan pembangunan lapangan pekerjaan baru. Berdasarkan catatan Badan Pusat Statistik (BPS) Indonesia, Jawa Barat merupakan salah satu provinsi di Indonesia dengan tingkat pengangguran yang tinggi. Padahal sejak beberapa tahun terakhir Provinsi Jawa Barat menjadi peringkat pertama Provinsi di Indonesia dengan aliran dana Penanaman Modal Asing (PMA) yang besar. Penelitian ini akan mengelompokkan Kabupaten/Kota di Jawa Barat berdasarkan atribut TPT, TPAK dan PMA. Dari data yang diolah dengan menggunakan algoritma *K-Means* menghasilkan 3 cluster, dimana ketiga cluster tersebut memiliki karakteristik yang berbeda sehingga dapat membantu pemerintah di Jawa Barat mengoptimalkan pemetaan lapangan pekerjaan baru.

**Kata Kunci** - Data Mining, Clustering, KMeans, Lapangan Pekerjaan Baru

## I. PENDAHULUAN

Salah satu hal yang selalu menjadi perbincangan sekaligus permasalahan yang dihadapi Indonesia adalah masalah pengangguran. Dampak negatif yang dapat muncul akibat tingginya jumlah pengangguran di suatu wilayah diantaranya ketidakstabilan ekonomi, munculnya hambatan pada pertumbuhan ekonomi, menurunnya tingkat kesejahteraan masyarakat dan permasalahan sosial ekonomi lainnya yang berujung pada kemiskinan[1]. Pengangguran menjadi salah satu penyumbang dari peningkatan kemiskinan[2]. Permasalahan kemiskinan ini juga dipicu dengan kurangnya perhatian pemerintah pada kebutuhan masyarakat[3], misalnya kurang meratanya lapangan pekerjaan pada tiap daerah,

Cara untuk mengatasi tingginya tingkat pengangguran adalah dengan membuat dan menyusun strategi untuk mengoptimalkan pembangunan lapangan pekerjaan baru. tingkat pengangguran tiap wilayah pastinya berbeda-beda, karenanya pemerintah harus mengetahui daerah mana saja yang memiliki tingkat pengangguran tinggi dan rendah[4].

Berdasarkan catatan Badan Pusat Statistik (BPS) Indonesia, Provinsi Jawa Barat menjadi salah satu Provinsi yang ada di Indonesia dengan TPT atau Tingkat Pengangguran Tinggi yang tergolong tinggi. Dapat dilihat pada tahun 2019 per bulan agustus persentase tingkat pengangguran yang awalnya sempat turun menjadi 8.04% dari tahun sebelumnya, pada tahun 2020 tingkat pengangguran meningkat menjadi 10.48% dan pada 2021 kembali turun menjadi 9.82% [5]. Padahal sejak beberapa tahun terakhir Provinsi Jawa Barat menjadi peringkat pertama provinsi di Indonesia yang memiliki aliran dana Penanaman Modal Asing (PMA) dan Penanaman Modal Dalam Negeri (PMDN) yang besar. PMA merupakan kegiatan penanaman modal oleh investor asing guna keberlangsungan usaha suatu negara, sedangkan PMDN adalah kegiatan ekspansi usaha atau penanaman modal pada suatu negara menggunakan modal dalam negeri. Penelitian sebelumnya[6] menyebutkan bahwa penanaman modal atau investasi asing berpengaruh dengan tingkat pengangguran terbuka.

Dari permasalahan di atas diperlukan analisa mendalam mengenai tingkat pengangguran

menggunakan atribut atribut yang dapat mempengaruhi perubahan TPT tersebut. Metode pengelompokan dapat dimanfaatkan untuk penelitian seperti ini[7]. K-Means merupakan algoritma yang dapat digunakan untuk melakukan pengelompokan data[8]. Kegiatan mengelompokkan atribut-atribut berdasarkan karakteristik yang dimiliki disebut dengan *cluster analysis*. *cluster analysis* menekankan konsep pencarian pada pusat cluster secara iteratif, dimana pusat dari cluster dipilih berdasarkan jarak minimum setiap data pusat cluster (centroid)[9].

Penelitian yang sudah pernah menggunakan metode K-Means untuk melakukan clustering yaitu penelitian[7], penelitian ini menggunakan algoritma kmeans untuk mengelompokkan Kabupaten/Kota di Jawa Barat berdasarkan indikator jumlah pengangguran, jumlah penduduk, seta besaran UMK atau upah minimum kabupaten/kota. Data yang digunakan pada penelitian ini bersumber dari website Badan Pusat Statistik Jawa Barat dari tahun 2018-2020. Proses pengolahan data dibantu dengan software Rapidminer. Dari 4 cluster yang terbentuk, cluster 2 memiliki rata-rata tertinggi sehingga berdasarkan penelitian ini pemerintah harus memberikan perhatian khusus untuk menurunkan tingkat pengangguran pada daerah yang berada pada kluster tersebut. Hasil evaluasi pada clustering menggunakan metode Davies Bouldin Index menunjukkan Davies Boulding sebesar 0.328.

Penelitian yang lain yaitu[10], penelitian ini menggunakan algoritma *fuzzy kmeans* untuk mengelompokkan data dengan menggunakan atribut Tingkat pengangguran Terbuka (TPT), Tingkat Partisipasi Angkatan Kerja (TPAK), Produk Domestik Regional Bruto (PDRB) dan Indeks Pembangunan Manusia (IPM). Penelitian ini membentuk 2 kelompok data atau cluster yaitu cluster daerah yang memiliki tingkat pengangguran rendah dan tinggi.

Dari kasus dan penjabaran di atas, maka penelitian ini akan mengelompokkan daerah di Kabupaten/Kota yang ada di Jawa Barat berdasarkan atribut Tingkat Pengangguran Terbuka (TPT), Tingkat Partisipasi Angkatan Kerja (TPAK), dan Jumlah Penanaman Modal Asing (PMA). Penelitian ini akan menggunakan bantuan *tools excel* dan bahasa pemrograman Python. Hasil dari penelitian ini diharapkan dapat membantu pemerintah daerah sekitar untuk meratakan ketersediaan lapangan pekerjaan di tiap Kabupaten/Kota di Jawa Barat.

#### A. Pengangguran

Suatu kondisi ketika seseorang yang termasuk kedalam angkatan kerja atau telah masuk ke usia siap kerja namun belum mulai bekerja dapat dikatakan sebagai pengangguran[10]. Seseorang dapat dikategorikan sebagai pengangguran apabila orang tersebut tidak memiliki pekerjaan dan sedang mencari pekerjaan, orang yang tidak memiliki pekerjaan dan dalam tahap persiapan usaha, orang yang tidak punya

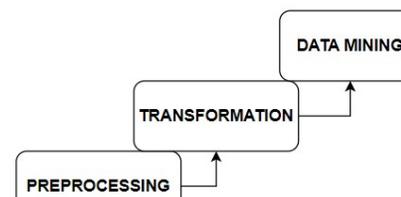
pekerjaan dan merasa putus asa sehingga tidak mencari pekerjaan, dan orang yang belum mulai bekerja padahal sudah memiliki pekerjaan[5].

#### B. Data Mining

Suatu teknik untuk mengolah data besar berdasarkan pola sehingga dapat menjadi pengetahuan yang informatif dan menarik dikenal dengan istilah data mining[11].

#### C. Knowledge Discovery in Database (KDD).

*Knowledge Discovery in Database (KDD)* adalah suatu cara memanfaatkan database untuk mendapatkan informasi yang berbasis pengetahuan (*knowledge base*), kemudian informasi yang diperoleh tersebut digunakan untuk membantu *stakeholder* untuk mengambil keputusan[12]. Pada metode ini terdapat 5 alur yang harus dilalui yaitu selection data, preprocessing yang terdiri dari data cleaning dan intergration, kemudian tahap transformation, data mining dan evaluation[13].



Gambar 1. Rancangan Penelitian KDD

#### D. Clustering.

Salah satu metode atau teknik menambang data (data mining) yang digunakan untuk menganalisa data dengan mengelompokkan data disebut clustering, tujuan dari clustering yaitu mengelompokkan data yang memiliki karakteristik sama ke satu wilayah yang sama[14].

#### E. K-Means

Algoritma *K-Means* adalah suatu algoritma yang digunakan untuk mengelompokkan data, pengelompokkan data pada algoritma ini didasarkan pada pencarian pusat *cluster* secara iterative[15]. Huruf *K* pada nama algoritma ini berarti jumlah cluster yang akan dibuat, dan means berarti nilai sementara yang menjadi *center* atau pusat *cluster* (centroid). Cluster terbentuk dari perhitungan jarak tiap data dari *centroid* menggunakan rumus *Eucliden*[16].

Tahapan *Kmeans* yaitu penentuan jumlah *cluster*, mengalokasikan anggota atau data pada *cluster*, menghitung *centroid* setiap cluster berdasarkan nilai tengah keseluruhan anggota cluster, peruntukkan tiap data ke *centroid* terdekat[3].

## II. METODE PENELITIAN

Alur pada penelitian ini akan berlandaskan kepada metode KDD. Adapun tahapan dan proses yang dilakukan pada penelitian ini yaitu :

### 1. Data Selection

Tahap pertama pada penelitian ini yaitu pengumpulan data. Data yang akan diolah dan dianalisis serta dilakukan modeling pada penelitian ini bersumber pada website Badan Pusat Statistik (BPS) Jawa Barat dan website Open Data Jabar tahun 2019-2021.

### 2. Preprocessing and Cleansing

Tahap selanjutnya yaitu *data preprocessing*. Dalam tahapan ini mencakup beberapa hal yaitu menghilangkan duplikasi data, memperbaiki kesalahan data, melakukan *handling missing value* dan sebagainya. Pada penelitian ini data yang digunakan sudah cukup bersih dan baik untuk selanjutnya masuk ke tahap *transformation*.

### 3. Transformation

Tahap ini merupakan tahap *coding*, proses *coding* yaitu proses kreatif untuk mengelola data sehingga menghasilkan pola informasi yang dibutuhkan melalui database. Pada penelitian ini tahap transformasi dilakukan menggunakan Bahasa pemrograman python

### 4. Data Mining

Tahap berikutnya yaitu data mining. Pada tahap ini teknik, metode, atau algoritma digunakan. Seperti pada penjelasan sebelumnya penelitian ini akan menjalankan algoritma *K-Means* yang merupakan algoritma yang sering digunakan untuk melakukan clustering pada data mining

### 5. Evaluation

Tahap terakhir adalah tahap evaluasi, disini hasil clustering yang telah didapat dari proses *elbow method* akan diperiksa apakah hasilnya bertentangan dengan asumsi atau hipotesis sebelumnya. Pada penelitian ini tahap evaluasi akan dilakukan dengan menghitung nilai *silhouette coefficient*

## III. HASIL DAN PEMBAHASAN

### A. Data Selection

Data yang dipilih pada penelitian ini adalah data yang bersifat *open source*, data didapat dari website resmi Badan Pusat Statistik (BPS) Jawa Barat dan Open Data Jabar tahun 2019 - 2021. Adapun data yang digunakan sebagai atribut yaitu data TPT, TPAK, an PMA dalam satuan USD, tipe data pada setiap data adalah numerik

Tabel 1 merupakan sampel data yang diunduh langsung dari website BPS.

Wilayah	2019	2020	2021
Bogor	9,11	14,29	12,22
Sukabumi	8,05	9,6	9,51
Cianjur	9,81	11,05	9,32
Bandung	5,51	8,58	8,32
Garut	7,35	8,95	8,68
Tasikmalaya	6,31	7,12	6,16
...	...	...	...
Kota Banjar	6,16	6,73	6,09

Tabel 1. Sampel Data TPT di Jawa Barat

### B. Preprocessing

Pada tahap preprocessing nilai pada tiap data dihitung rata rata nya, sehingga menghasilkan nilai rata rata pada tiap atribut yang nantinya akan digunakan pada pengelompokan daerah. Tahap perhitungagn nilai rerata tiap data dibantu menggunakan *Microsoft Excel*.

Wilayah	2019	2020	2021	rata rata
Bogor	9,11	14,29	12,22	11,87
Sukabumi	8,05	9,6	9,51	9,05
Cianjur	9,81	11,05	9,32	10,06
Bandung	5,51	8,58	8,32	7,47
Garut	7,35	8,95	8,68	8,33
Tasikmalaya	6,31	7,12	6,16	6,53
Ciamis	5,16	5,66	5,06	5,29
....	...	...	...	...
Kota Banjar	6,16	6,73	6,09	6,33

Tabel 2. Sampel Data Rerata Tingkat PengangguranTerbuka (TPT)

Setelah nilai rerata telah didapatkan, nilai rerata dari data TPT, TPAK dan PMA. Ketiga nilai rerata tersebut digabungkan kedalam satu file baru. Seperti pada tabel 3 dengan nama data final. Data final ini lah yang nanti akan digunakan untuk mengelompokan daerah di jawa barat

Wilayah	TPT	TPAK	PMA (USD)
Bogor	11,87	63,54	278112133
Sukabumi	9,05	63,05	14163867
Cianjur	10,06	67,75	8766733
Bandung	7,47	64,21	163047967
Garut	8,33	62,16	31711600
Tasikmalaya	6,53	68,06	0
...	...	...	...
Kota Banjar	6,33	66,34	0

Tabel 3. Data Final Rerata

Tahap selanjutnya yaitu memastikan data dalam keadaan bersih sehingga data yang akan masuk ke tahap pemodelan dapat menghasilkan keluaran yang sesuai. Gambar 2 menuntukan bahwa data yang dimiliki sudah tidak terdapat nilai yang kosong atau null

```
memory usage: 1.
[67] 1 df.isnull().s
Wilayah 0
```

Gambar 2. Proses Preprocessing

C. Transformation

Pada tahapan ini data numerik yang nilainya beragam ditransformasi agar menjadi selaras, proses ini dilakukan untuk menghindari penyimpangan yang besar pada data, misalnya pada data PMA yang memiliki nilai besar sedangkan atribut lain memiliki besaran nilai yang lebih kecil, agar algoritma dapat berjalan dengan baik diperlukan proses normalisasi data, tahap normalisasi dilakukan menggunakan *StandarScaler* yang merupakan *class* yang ada pada library *sklearn*. Gambar 3 menunjukkan hasil normalisasi menggunakan *StandarScaler*.

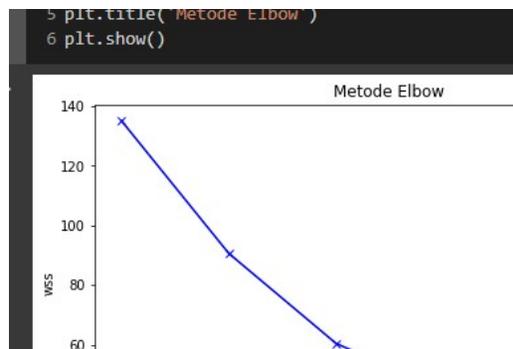
```
kode + Ieks
TPT    TPAK    PMA
0    1.3730  -0.4591
1   -0.0114  -0.6059
2    0.4828   0.8015
```

Gambar 3. Hasil Normalisasi *StandarScaler*

Wilayah	TPT	TPAK	PMA (USD)	cluster
Bogor	11,87	63,54	278112133	1
Sukabumi	9,05	63,05	14163867	1
Cianjur	10,06	67,75	8766733	0
Bandung	7,47	64,21	163047967	0
Garut	8,33	62,16	31711600	1
Tasikmalaya	6,53	68,06	0	0
Ciamis	5,29	69,54	3867	0
Kuningan	10,86	62,33	1032200	1
Cirebon	10,75	64,50	415834500	1
Majalengka	5,31	68,37	26114167	0
Sumedang	8,92	67,12	2751033	0
Indramayu	8,62	69,03	39808533	0
Subang	9,31	67,38	176371633	0
Purwakarta	10,50	61,97	251988967	1
Karawang	11,01	64,22	1228731267	2
Bekasi	10,21	64,71	1922073000	2
Bandung Barat	10,71	60,88	218788767	1
Pangandaran	4,28	75,54	2233	0
Kota Bogor	11,21	62,09	8462267	1
Kota Sukabumi	10,48	58,36	343500	1
Kota Bandung	10,28	65,31	187450733	1
Kota Cirebon	10,18	63,11	576567	1
Kota Bekasi	9,95	65,05	205751133	1

D. Data Mining

Teknik data mining diterapkan pada tahap ini, penelitian ini menggunakan metode teknik *data mining clustering* menggunakan algoritma *K-Means* untuk memproses data. Hal pertama yang dilakukan ketika melakukan clustering data adalah menentukan jumlah banyaknya cluster (*k*). Untuk menentukan jumlah cluster dapat dilakukan dengan melihat grafik seperti gambar 4, penentuan nilai *k* ini dikenal dengan metode *elbow*.



Gambar 4. Grafik Metode Elbow

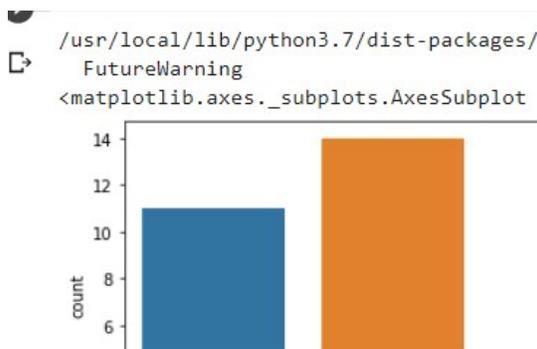
Dari grafik pada gambar 4 terlihat bahwa siku terbentuk pada cluster 3. Sehingga nilai *k=3* atau jumlah cluster 3 pada data dianggap optimal.

Setelah menentukan *k* optimal menggunakan metode *elbow*, tahap selanjutnya adalah mengolah data dengan menjalankan model clustering menggunakan algoritma *k-mean*. Adapun hasil dari clustering kabupaten/kota berdasarkan TPT, TPAK, dan PMA pada dilihat dari tabel 4 dibawah ini.

Kota Depok	8,58	63,85	60711767	1
Kota Cimahi	11,49	62,57	54787833	1
Kota Tasikmalaya	7,48	65,85	95033	0
Kota Banjar	6,33	66,34	0	0

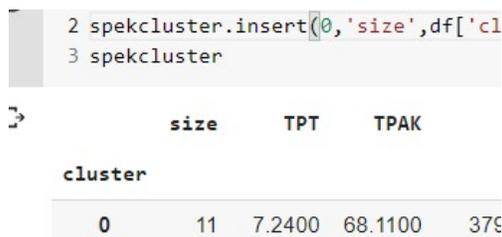
Tabel 4. Hasil Clustering

Melalui tabel 4 terlihat dengan jelas bahwa terdapat 11 kabupaten/kota yang masuk ke dalam cluster 0. Anggota pada cluster ini diantaranya, Cianjur, Bandung, Ciamis, Tasikmalaya, Majalengka, Sumedang, Indramayu, Subang, Pangandaran, Kota Tasikmalaya dan Kota Banjar. Pada cluster 1 terdapat 14 Kota/Kabupaten yang masuk kedalamnya, adapun diantaranya adalah Sukabumi, Bogor, Sukabumi, Garut, Kuningan, Purwakarta, Bandung Barat, Kota Bogor, Kota Bandung, Kota Sukabumi, Kota Cirebon, Kota Bekasi, Kota Depok dan Kota Cimahi. Sedangkan daerah yang masuk ke dalam cluster 2 adalah Bekasi dan Karawang.



Gambar 5. Visualisasi jumlah anggota tiap cluster

Gambar 5 menunjukkan visualisasi dari jumlah kabupaten/kota atau anggota pada tiap cluster, dapat disimpulkan bahwa cluster 1 memiliki anggota terbanyak.



Gambar 6. karakteristik Setiap Cluster

Pada gambar 6 dapat terlihat bagaimana karakteristik pada setiap cluster yang terbentuk, agar dapat diambil kesimpulan untuk memetakan daerah guna pemetaan lapangan pekerjaan baru di Provinsi Jawa Barat..

Cluster 0 dapat dikategorikan sebagai kelompok daerah dengan prioritas rendah, karena cluster 0 merupakan daerah dengan tingkat pengangguran yang rendah, tingkat partisipasi angkatan kerja yang tinggi, dengan jumlah penanam modal asing atau luar negeri yang cukup rendah.

Cluster 1 dapat dikategorikan sebagai kelompok daerah dengan prioritas tinggi, karena apabila ditinjau dari karakteristiknya, cluster 1 merupakan daerah dengan tingkat pengangguran yang lumayan tinggi, tingkat partisipasi angkatan kerja yang rendah, dan jumlah penanam modal luar negeri yang tidak terlalu tinggi atau sedang. Pemerintah dapat fokus untuk mencari investor di daerah kelompok ini, agar dapat memperluas lapangan pekerjaan baru serta dapat mengurangi jumlah pengangguran.

Cluster 2 dapat dikatakan sebagai kelompok daerah prioritas sedang, karena berdasarkan karakteristiknya cluster ini merupakan daerah dengan tingkat pengangguran yang tinggi, tingkat partisipasi kerja yang tidak terlalu rendah dan tinggi, dan jumlah penanam modal asing yang tinggi. Cluster ini menunjukkan bahwa daerah karawang dan bekasi memiliki dana investor yang banyak tapi masih memiliki tingkat pengangguran yang tinggi. Dari sini pemerintah mungkin dapat membuat kebijakan bagi para investor atau membuka peluang usaha yang lebih luas.

E. Evaluation

Setelah melakukan teknik data mining dan mendapatkan hasil clustering, selanjutnya adalah tahap evaluation, tahap ini dilakukan untuk melihat keberhasilan model yang telah diterapkan. Silhouette coefficient merupakan salah satu teknik pada evaluasi performa k-means.

```

3 print("Silhouette score untuk k = "+str
4     +str(metrics.silhouette_score(mc

```

Gambar 7. Uji Silhouette Coefficient

Silhouette score yang terlihat pada gambar 7 menunjukkan hasil evaluasi dari model clustering yang telah dilakukan dengan nilai k=3 atau jumlah cluster 3 adalah jumlah cluster optimal karena menunjukkan nilai 0.461 dan paling mendekati angka 1 dibanding cluster lain.

IV. KESIMPULAN

Dari proses data mining yang telah dilakukan menggunakan data pengangguran dan indikatornya di daerah jawa barat, dapat ditarik dan diambil kesimpulan bahwa:

1. Pengelompokan Kabupaten/Kota untuk pemetaan lapangan pekerjaan baru dan mengurangi pengangguran berdasarkan

- atribut TPT, TPAK dan PMA dapat dilakukan dengan menggunakan teknik penambangan data atau data mining clustering dengan menggunakan algoritma *K-Means*
2. Berdasarkan jumlah sampel data sebanyak 27 Kabupaten/Kota di Jawa Barat, dihasilkan 3 cluster
  3. Hasil evaluasi menggunakan *silhouette coefficient* menunjukkan bahwa pembentukan kelompok atau cluster sebanyak 3 ( $k=3$ ) adalah jumlah cluster yang optimal.
  4. Hasil clustering dapat membantu pemerintah daerah setempat untuk memperluas lapangan pekerjaan baru berdasarkan prioritas nya.

#### DAFTAR PUSTAKA

- [1] D. Viky Mouren, "Pengaruh Tingkat Pendidikan, Pertumbuhan Ekonomidan Jumlah Penduduk Terhadap Pengangguran Di Kabupaten Toraja Utara," vol. 22, no. 1, pp. 131–143, 2022.
- [2] A. Akramunnisa and F. Fajriani, "Hierarchical Clustering Analysis dalam Pengelompokan Tingkat Pengangguran di Sulawesi Selatan," *Pros. Semant.*, vol. 2, no. 1, pp. 270–275, 2019, [Online]. Available: <https://journal.uncp.ac.id/index.php/semantik/article/view/1525/1336%0Ahttps://journal.uncp.ac.id/index.php/semantik/article/view/1525>
- [3] D. N. Alfiansyah, V. R. S. Nastiti, and N. Hayatin, "Penerapan Metode K-Means pada Data Penduduk Miskin Per Kecamatan Kabupaten Blitar," *REPOSITOR*, vol. 4, no. 1, 2022.
- [4] N. N. Fransiska, D. S. Anggraeni, and U. Enri, "Pengelompokan Data Kemiskinan Provinsi Jawa Barat Menggunakan Algoritma K-Means dengan Silhouette Coefficient," *Temat. (Jurnal Teknol. Inf. Komunikasi)*, vol. 9, no. 1, pp. 29–35, 2022, [Online]. Available: <https://doi.org/10.38204/tematik.v9i1.921>
- [5] B. P. S. J. Barat, "Badan Pusat Statistik Jawa Barat."
- [6] H. Widayanti, "Analisis Faktor-faktor yang Mempengaruhi Tingkat Pengangguran Terbuka (TPT) di Kabupaten/Kota Provinsi Jawa Barat Tahun 2012-2019," UNIVERSITAS ISLAM INDONESIA, 2021.
- [7] A. A. Putrie, R. Sanjaya, P. Studi, S. Informasi, U. Adhirajasa, and R. Sanjaya, "Pengelompokan Kabupaten / Kota Berdasarkan Indikator Tingkat Pengangguran Menggunakan Algoritma K-Means Clustering ( Studi Kasus : Provinsi Jawa Barat )," vol. 2, no. 2, pp. 111–121, 2021.
- [8] L. Maulida, "Penerapan Datamining Dalam Mengelompokkan Kunjungan Wisatawan Ke Objek Wisata Unggulan Di Prov. Dki Jakarta Dengan K-Means," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 2, no. 3, p. 167, 2018, doi: 10.14421/jiska.2018.23-06.
- [9] M. G. Sadewo, A. P. Windarto, and D. Hartama, "Penerapan Datamining Pada Populasi Daging Ayam Ras Pedaging Di Indonesia Berdasarkan Provinsi Menggunakan K-Means Clustering," *InfoTekJar (Jurnal Nas. Inform. dan Teknol. Jaringan)*, vol. 2, no. 1, pp. 60–67, 2017, doi: 10.30743/infotekjar.v2i1.164.
- [10] S. Sarbaini, W. Saputri, Nazaruddin, and F. Muttakin, "Cluster Analysis Menggunakan Algoritma Fuzzy K-Means Untuk Tingkat Pengangguran Di Provinsi Riau," *J. Teknol. dan Manaj. Ind. Terap.*, vol. 1, no. II, pp. 78–84, 2022, [Online]. Available: <http://jurnal-tmit.com/index.php/home/article/view/30>
- [11] A. Y. Pratama, Y. Umidah, and A. Voutama, "Analisis Sentimen Media Sosial Twitter Dengan Algoritma K-Nearest Neighbor Dan Seleksi Fitur Chi-Square ( Kasus Omnibus Law Cipta Kerja )," vol. 5, no. September, pp. 897–910, 2021.
- [12] E. Saraswati, Y. Umidah, and A. Voutama, "Penerapan Algoritma Artificial Neural Network untuk," *Gener. J.*, vol. 5, no. 2, pp. 109–118, 2021.
- [13] V. Ramadhan and A. Voutama, "Clustering Menggunakan Algoritma K-Means Pada Penyakit ISPA di Puskesmas Kabupaten Karawang," *J. Pendidik. dan Konseling*, vol. 4, 2022.
- [14] M. Wahyudi and S. Pujiastuti, "Penerapan Data Mining Dalam Mengelompokkan Data Pengangguran Terbuka Menurut Provinsi Menggunakan," *Pros. Semin. Nas. Ris. dan Inform. Sci.*, vol. 2, pp. 432–440, 2020.
- [15] E. Satria, H. S. Tambunan, I. S. Saragih, I. S. Damanik, Fany, and T. E. Sitanggang, "Penerapan Clustering dalam Mengelompokkan Jumlah Kunjungan Wisatawan Mancanegara Dengan Metode K-Means," *Pros. Semin. Nas. Ris. Inf. Sci.*, no. September, pp. 462–471, 2019.
- [16] Y. R. Sari, A. Sudewa, D. A. Lestari, and T. I. Jaya, "Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer," *CESS (Journal Comput. Eng. Syst. Sci.)*, vol. 5, no. 2, p. 192, 2020, doi: 10.24114/cess.v5i2.18519.