

Penerapan Algoritma K-Means Pada Penyebaran Covid-19 Di Provinsi Jambi

Reza Pahlevi, Muhammad Wahyu Prayogi, Errissya Rasywir*, Yovi Pratama

Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Dinamika Bangsa, Jambi, Indonesia
Email: ¹paaahlevi@gmail.com, ²wahyuprayogi@gmail.com, ³*errissya.jurnal@gmail.com, ⁴yovi.pratama@gmail.com
Email Penulis Korespondensi: errissya.jurnal@gmail.com

Submitted: 25/10/2022; Accepted: 29/11/2022; Published: 30/11/2022

Abstrak—Tujuan dalam penelitian menggunakan metode K-Means Cluster adalah untuk mengetahui tingkat persebaran kasus Covid-19 kategori tinggi, sedang, dan rendah pada masing-masing daerah di Provinsi Jambi. Ada beberapa aspek yang bisa diukur seperti jumlah penduduk, kepadatan penduduk, kasus positif terinfeksi Covid-19, pasien yang sembuh, dan pasien yang meninggal dunia. Metode pengumpulan data yang digunakan adalah metode dokumentasi berupa data sekunder yang diperoleh dari website Pemerintah Provinsi Jambi. Data yang digunakan yaitu positif, sembuh, dan meninggal dan dianalisis menggunakan aplikasi WEKA. Dari hasil penelitian dengan metode K-Means menggunakan 3 cluster. Cluster 0 merupakan Cluster dengan kategori penyebaran tingkat tinggi yakni ada pada Kota Jambi. Cluster 1 merupakan Cluster dengan kategori penyebaran tingkat sedang yakni terdiri dari Batanghari, Merangin, Muaro Jambi, Tanjab Barat. Cluster 2 merupakan Cluster dengan kategori penyebaran tingkat rendah yakni terdiri dari Bungo, Kerinci, Sarolangun, Sungai Penuh, Tanjab Timur, Tebo.

Kata Kunci: Algoritma; K-Means; WEKA; Covid-19; Provinsi Jambi

Abstract—The purpose of this study using the K-Means Cluster method is to determine the level of distribution of Covid-19 cases in the high, medium, and low categories in each region in Jambi Province. There are several aspects that can be measured such as population, population density, positive cases infected with Covid-19, recovered patients, and patients who died. The data collection method used is the documentation method in the form of secondary data obtained from the Jambi Provincial Government website. The data used were positive, recovered, and died and were analyzed using the WEKA application. From the results of research with the K-Means method using 3 clusters. Cluster 0 is a cluster with a high level of distribution category, which is in the city of Jambi. Cluster 1 is a cluster with a medium level distribution category consisting of Batanghari, Merangin, Muaro Jambi, Tanjab Timur. Cluster 2 is a cluster with a low-level distribution category consisting of Bungo, Kerinci, Sarolangun, Sungai Penuh, Tanjab Barat, Tebo.

Keywords: Algorithm; K-Means; WEKA; Covid-19; Jambi Province

1. PENDAHULUAN

COVID-19 adalah penyakit yang menular, dan ditandai oleh gejala pada bagian pernapasan akut (coronavirus 2 (severe acute respiratory syndrome coronavirus 2 atau SARS-CoV-2). Virus ini merupakan keluarga besar Coronavirus yang dapat menyerang hewan [1]–[3]. Penularan virus ini terjadi jika adanya kontak antar sesama manusia. Penyakit yang disebabkan oleh virus ini adalah penyakit infeksi saluran pernafasan, seperti flu, MERS (Middle East Respiratory Syndrome), dan SARS (Severe Acute Respiratory Syndrome) [4].

Di Indonesia sendiri kasus COVID – 19 semakin hari semakin meningkat semenjak pertama kali tercatat kasus orang positif pada 20 maret 2020 sampai 16 juli 2022 sudah tercatat ada 6.131.413 kasus korban Positif, sedangkan untuk pasien yang sembuh mencapai 5.947.980 kasus, dan untuk korban yang meninggal sudah terkonfirmasi ada 156.839 jiwa sampai saat ini [5]. Hal ini tentunya membuat keresahan masyarakat yang akan berdampak pada segala aspek baik pada bidang pendidikan, ekonomi maupun sosial, oleh karenanya negara harus menerapkan sistem lockdown dan pandemi secara global termasuk di negara Indonesia. Karena cepatnya penyebaran virus corona ini pemerintah menerapkan beberapa kebijakan untuk mencegah dan memperlambat penyebarannya. Beberapa kebijakan ini di antaranya: penerapan isolasi mandiri bagi orang yang sudah terkena COVID-19 selama minimal 14 hari, penerapan protokol 3M (menjaga jarak dari kerumunan orang, selalu mencuci tangan dan memakai masker) [5]. Selain itu ada juga pembuatan kelompok relawan tanggap COVID-19 sebagai upaya untuk menghambat penyebarannya. Kelompok relawan ini berperan untuk memberikan penyuluhan, dan pelatihan kepada warga akan bahaya dan bagai mana cara menanggapi virus corona ini. Maka dari itu pihak relawan harus di bekali dengan pengetahuan tentang COVID-19 dan bagaimana cara menanggulangnya agar mereka dapat meneruskan pengetahuan tersebut kepada masyarakat [4].

Penelitian menggunakan metode K–Means pernah dilakukan Diah Novita Sari dan rekannya untuk mengetahui tingkat keparahan Virus Corona di 34 provinsi yang ada di Indonesia. Dengan mengacu pada tiga variabel berikut: Case Fatality Rate (CFR), Total kasus dan Kepadatan penduduk. Kemudian Dilakukan *Clustering Analysis* menggunakan metode K–Means. Dari hasil penelitian tersebut disimpulkan bahwa provinsi dengan tingkat keparahan tertinggi yaitu DKI Jakarta dan Jawa Timur [6].

Penelitian menggunakan metode K-Means juga pernah dilakukan oleh Wiyli Yustanti dan tim untuk pengelompokan data persebaran Virus Corona di jawa timur. Hasil dari penelitiannya menunjukkan bahwa jumlah *Cluster* optimum nya 5 *Cluster*. Dilakukan sebuah uji mean vektor menggunakan statistic Wilks Lambda yang digunakan untuk membuktikan kelima *Cluster* yang dibentuk berbeda secara signifikan. Hasil dari penelitian tersebut ialah kelima *Cluster* terbukti memiliki perbedaan yang signifikan dengan tingkat kepercayaan 95% [7].

Adapun alasan mengapa penulis menggunakan metode K-Means ini dikarenakan metode KMeans merupakan metode yang cukup populer, mempunyai keakuratan data yang cukup baik. Metode ini juga dapat digunakan dalam kuantitas yang besar dan praktis tanpa harus memperhatikan urutan objek [8].

Jumlah kasus Covid-19 yang terus meningkat di setiap wilayah Indonesia. Salah satunya di Provinsi Jambi tercatat 11 Juli 2022 jumlah konfirmasi 38193. 9 (0,02) kasus aktif ,881 (2,30%) kasus kematian dan 37143 (97,68%) kasus sembuh, data ini diperoleh dari situs satgas covid Provinsi Jambi [4].

Nama	Positif	Sembuh	Meninggal
batanghari	3,165	3,048	114
bungo	2,077	2,046	31
kerinci	590	564	22
kota jambi	9,738	9,482	254
merangin	1,679	1,585	94
muaro jambi	4,131	4,060	68
sarolangun	1,036	1,027	9
sungai penuh	1,408	1,388	20
tanjab barat	2,419	2,353	63
tanjab timur	1,406	1,370	36
tebo	2,137	2,067	31

Gambar 1. Data Covid-19 Provinsi Jambi Per Tanggal 11 Juli 2022 [5]

Penyebaran kasus COVID-19 yang merata di Provinsi Jambi, merupakan penyebaran yang cukup cepat dan berdampak negatif pada seluruh bidang. Luasnya wilayah Provinsi Jambi memungkinkan diperlukannya pengelompokan bagian bagian berdasarkan kota dan kabupaten di Provinsi Jambi. Pengelompokan ini akan menghasilkan titik-titik pusat penyebaran kasus COVID-19.

Clustering merupakan suatu proses pengelompokan record suatu observasi, atau mengelompokkan kelas yang memiliki kesamaan objek. Adapun Perbedaan *Clustering* dengan klasifikasi adalah tidak adanya variabel target dalam melakukan suatu pengelompokan pada proses *Clustering*. *Clustering* sering dilakukan sebagai cara awal dalam proses data mining pada saat melakukan sebuah metode analisis [4]. *Clustering* merupakan sebuah metode pada data mining yang digunakan untuk pengelompokan data yang memiliki karakteristik mirip ataupun sama menjadi sebuah *Cluster*. Analisis *Cluster* merupakan teknik multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya [4].

K-Means merupakan salah satu algoritma *Clustering* yang masuk dalam kelompok Unsupervised learning yang digunakan untuk mengelompokkan data kedalam beberapa kelompok dengan sistem partisi. Algoritma ini menerima masukan berupa data tanpa label kelas. Pada algoritma K-Means, komputer menerima data-data yang tidak diketahui kelasnya terlebih dahulu lalu mengelompokkannya. Input yang diterima ialah data dan jumlah kelompok (*Cluster*) yang diinginkan [9]. Algoritma ini akan memasukan data kedalam beberapa kelompok tersebut.

2. METODOLOGI PENELITIAN

2.1 Clustering

Clustering adalah metode pengelompokan data yang sering digunakan sebagai salah satu metode machine learning [10]–[12]. *Clustering* adalah proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan *Cluster* [1], [12], [13]. Oleh karena itu, metode *Clustering* ini sangat berguna untuk menemukan kelompok yang tidak dikenal dalam data [9].

2.2 K-Means

K-Means merupakan algoritma *Clustering* yang berulang-ulang [9]. K-means merupakan salah satu bentuk pengelompokan yang paling sederhana. Prosesnya yang sederhana dan mudah untuk mengelompokkan data yang diberikan melalui sejumlah *Cluster*. Penentuan *centroid* dilakukan dengan cara mengambil data pertama sebagai *centroid* pertama, data kedua sebagai *centroid* kedua, dan seterusnya hingga jumlah *centroid* yang diperlukan. Langkah berikutnya adalah dengan menghitung jarak dari titik yang akan di *Cluster* ke setiap *centroid* yang ada dan dikelompokkan sesuai dengan jarak terdekat kepada *centroid*-nya. Bila semua titik Paradigma sudah masuk kedalam pengelompokan maka langkah pertama selesai.

K-Means merupakan teknik *Clustering* yang diperoleh dari sebuah dataset dengan cara menghitung jarak dari setiap titik ke pusat *Cluster* secara iterative. Algoritma K-Means memiliki beberapa aturan yaitu [14]:

1. Total *Cluster* yang dibutuhkan

2. Jenis atribut adalah numeric.

Proses untuk *Clustering* menggunakan algoritma k-means adalah sebagai berikut :

1. Tentukan nilai k dalam total *Cluster* yang ingin dibentuk.
2. Pilih nilai *centroid Cluster (centroid)* berawal dari k.
3. Hitung jarak setiap data input dan setiap *centroid* menggunakan rumus jarak Euclidean untuk menemukan jarak terdekat antara setiap titik data dan *centroid*.
4. Mengklasifikasikan/mengelompokkan setiap item data berlandaskan jarak (jarak minimum) ke *centroid*.
5. Perbarui mediannya. Nilai pusat baru.
6. Mengulangi langkah ke 3 sampai 5 anggota setiap *Cluster* tetap tidak berubah. Berikut ini merupakan rumus terhadap perhitungan algoritma K-Means *Clustering* :

$$de = \sqrt{(x_i - s_i)^2 + (y_i - t_i)^2} \tag{1}$$

2.3 Tools Eksperimen

Pada penelitian ini, tools yang gunakan untuk melakukan eksperimen adalah tools WEKA. Aplikasi WEKA merupakan sebuah perangkat lunak yang menerapkan berbagai algoritma *machine learning* untuk melakukan beberapa proses yang berkaitan dengan sistem temu kembali informasi atau *data mining* [12], [13], [15]. Beberapa fitur unggulan yang dimiliki oleh WEKA yaitu:

1. *Classification*

Di dalam WEKA terdapat banyak algoritma yang mendukung untuk proses klasifikasi sebuah objek serta pengguna dimudahkan dalam melakukan implementasi secara langsung. *User* dapat melakukan *load dataset*, melakukan pemilihan algoritma untuk klasifikasi, kemudian diberikan beberapa representasi data yang mewakili hasil akurasi, tingkat kesalahan dari proses klasifikasi.

2. *Regression*

Regression merupakan sebuah proses yang dapat melakukan suatu prediksi terhadap berbagai pola yang sudah terbentuk sebelumnya yang dijadikan sebagai model data. Tujuan dari *regression* adalah menciptakan suatu variabel baru yang mewakili suatu representasi perkembangan data pada masa yang akan datang. WEKA mendukung proses *regression* dan hal tersebut dipermudah dengan *user interface/user experience* yang sederhana.

3. *Clustering*

Clustering merupakan salah satu cabang konsep dari *unsupervised method* dari *machine learning* yang bertujuan untuk melakukan pengelompokan data dan juga menjelaskan hubungan/relasi yang ada di antara data tersebut dan memaksimalkan kesamaan antar satu kelas/*Cluster* tetapi meminimumkan kesamaan antar kelas/*Cluster*. *Clustering* digunakan untuk analisa suatu data dan diharapkan menghasilkan suatu representasi data yang mewakili suatu pola yang terbentuk akibat relasi yang ada antar data.

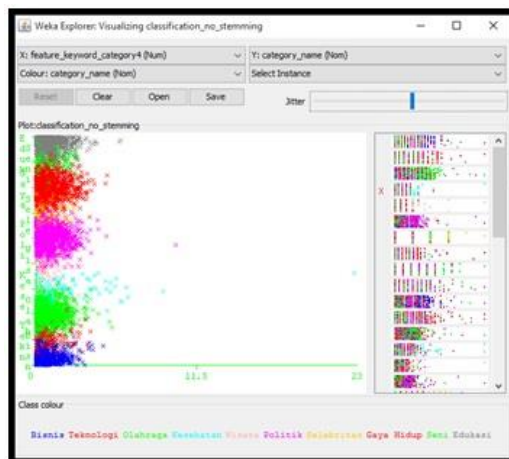
Di dalam WEKA tersedia beberapa pendekatan algoritma untuk menangani permasalahan *Clustering* dan pada fitur ini juga terdapat bagian kesimpulan dari proses *Clustering* data yang memberikan secara garis besar perhitungan dan hasil yang diberikan dalam implementasi algoritma *Clustering* [16]–[18].

1. *Association Rules*

Association Rules merupakan metode yang digunakan untuk menemukan berbagai relasi antara banyaknya variabel yang terdapat di dalam sebuah basis data dengan jumlah yang besar.

2. *Visualization*

WEKA memiliki fitur untuk memberikan sebuah representasi data hasil sebuah proses *data mining* dalam bentuk gambar atau *chart* yang juga dapat dilakukan pemilihan berbagai parameter yang mendukung dalam membentuk representasi data yang ada dalam aplikasi WEKA.



Gambar 2. Fitur Visualisasi WEKA [19], [20]

3. DataProcessing

WEKA menyediakan fitur dalam hal data preprocessing yaitu stemming dan stopwords removal. Proses stemming dan stopwords removal yang ada di dalam perangkat lunak WEKA berbasis Bahasa Inggris, sehingga untuk implementasi bahasa diluar bahasa Inggris diharuskan untuk melakukan proses preprocessing data di luar aplikasi WEKA. Beberapa algoritma stemming yang telah disediakan oleh WEKA adalah Iterated Lovins Stemmer, Lovins Stemmer dan Snowball Stemmer. Data yang digunakan pada Weka adalah dengan format ekstensi .arff. Anda bisa membuka file dengan ekstensi ini dengan berbagai macam text editor, misalnya Notepad. Contoh file berformat .arff adalah sebagai berikut:

```

@relation iris
@attribute sepallength real
@attribute sepalwidth real
@attribute petalength real
@attribute petalwidth real
@attribute class {Iris-setosa, Iris-versicolor, Iris-virginica}
@data
5.1,3.5,1.4,0.2,Iris-setosa
4.9,3.0,1.4,0.2,Iris-setosa
...
7.0,3.2,4.7,1.4,Iris-versicolor
6.4,3.2,4.5,1.5,Iris-versicolor
...
6.3,3.3,6.0,2.5,Iris-virginica
5.8,2.7,5.1,1.9,Iris-virginica
    
```

Gambar 3. Cuplikan Data .arff

2.4 Teknik Pengumpulan Data

Data yang digunakan untuk penelitian ini adalah data jumlah penyebaran virus COVID-19 di Provinsi Jambi yang diperoleh dari website Corona - Pemerintah Provinsi Jambi (<http://corona.jambiprov.go.id/v2/index.php>). Data penyebaran virus COVID-19 yang digunakan pada penelitian ini adalah data penyebaran pada tanggal 11 Juli 2022. Untuk rangkuman data jumlah penyebaran virus COVID-19 dapat dilihat pada tabel 1.

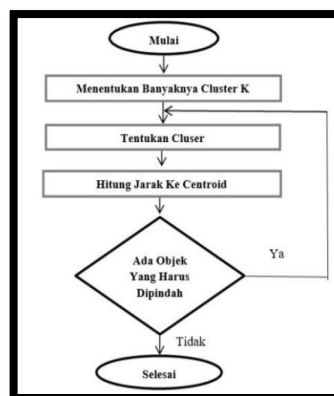
Tabel 1. Data Penyebaran Covid-19 Provinsi Jambi Per Tanggal 11 Juli 2022

Nama	Positif	Sembuh	Meninggal
Batanghari	3165	3048	114
Bungo	2077	2046	31
Kerinci	590	564	22
Kota Jambi	9738	9482	254
Menrangin	1679	1585	94
Muaro Jambi	4131	4060	68
Sarolangun	1036	1027	9
Sungai Penuh	1408	1388	20
Tanjab Barat	2419	2353	63
Tanjab Timur	1406	1370	36
Tebo	2137	1067	31

3. HASIL DAN PEMBAHASAN

3.1 Implementasi Algoritma K-means Menggunakan Aplikasi WEKA

Ketika data siap di proses, algoritma K-Means memproses data pada Gambar 2 dibawah ini.



Gambar 4. Tahapan Algoritma K-Means

Data yang digunakan ditampilkan dalam Tabel 2 dibawah ini:

Tabel 2. Data Penyebaran Covid-19 Provinsi Jambi Per Tanggal 11 Juli 2022

Data Nomor	Nama	Positif	Sembuh	Meninggal
1	Batanghari	3165	3048	114
2	Bungo	2077	2046	31
3	Kerinci	590	564	22
4	Kota Jambi	9738	9482	254
5	Merangin	1679	1585	94
6	Muaro Jambi	4131	4060	68
7	Sarolangun	1036	1027	9
8	Sungai Penuh	1408	1388	20
9	Tanjab Barat	2419	2353	63
10	Tanjab Timur	1406	1370	36
11	Tebo	2137	1067	31

3.2 Hasil Perhitungan K-Means Clustering Menggunakan Aplikasi WEKA

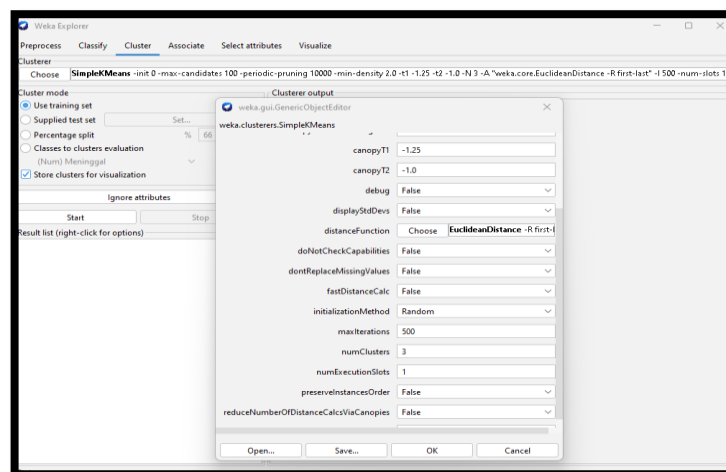
Berikut ini kami tampilkan hasil perhitungan k-means *Clustering* dengan menggunakan aplikasi WEKA untuk Penerapan Algoritma K-Means Pada Penyebaran Covid-19 Di Provinsi Jambi:

```
@relation DataCovid
@attribute Nama string
@attribute Positif numeric
@attribute Sembuh numeric
@attribute Meninggal numeric

@data
Batanghari,3165,3048,114
Bungo,2077,2046,31
Kerinci,590,564,22
Kota_Jambi,9738,9482,254
Menrangin,1679,1585,94
Muaro_Jambi,4131,4060,68
Sarolangun,1036,1027,9
Sungai_Penuh,1408,1388,20
Tanjab_Barat,2419,2353,63
Tanjab_Timur,1406,1370,36
Tebo,2137,1067,31
```

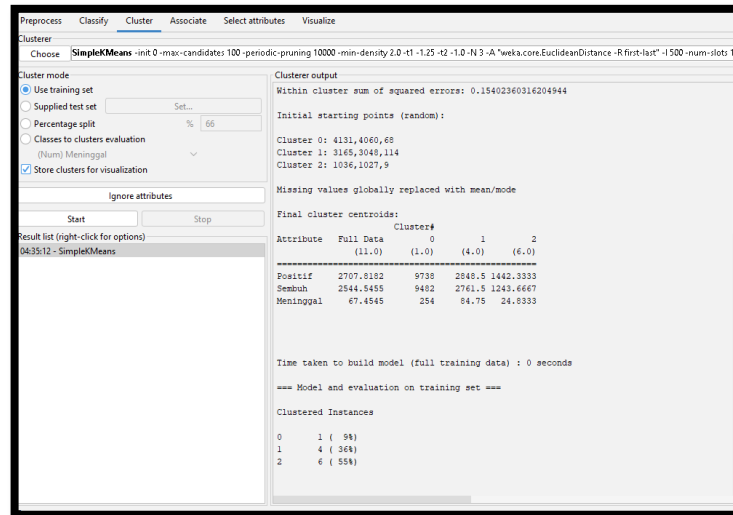
Gambar 5. Cuplikan File .arff Data Covid Provinsi Jambi

Selanjutnya menentukan parameter dalam pengolahan kelompok data. Pada penelitian ini menggunakan K=3.



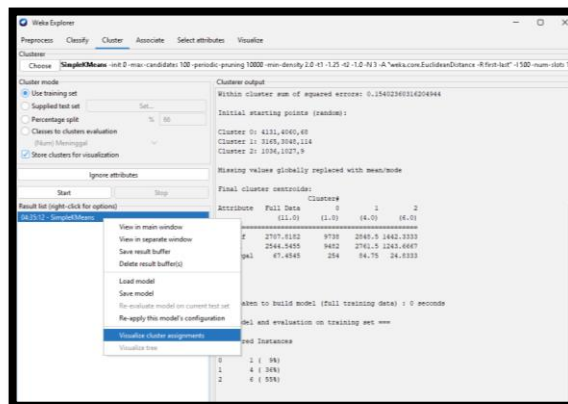
Gambar 6. Jumlah Cluster Yang Digunakan Penerapan Algoritma K-Means Pada Penyebaran Covid-19 Di Provinsi Jambi

Hasil running tools menghasilkan *output* seperti pada tampilan 'Clusterer Output' seperti gambar dibawah ini.



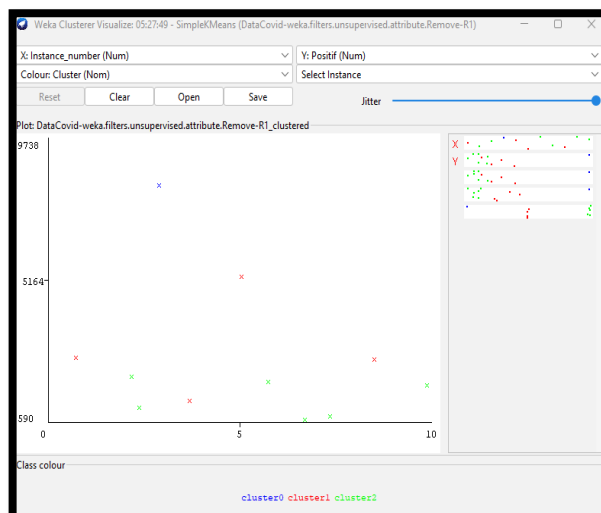
Gambar 7. Hasil Cluster Dan Hasil Keluaran Penerapan Algoritma K-Means Pada Penyebaran Covid-19 Di Provinsi Jambi

Tampilan ‘Visualize Cluster assignments’ dari data Penerapan Algoritma K-Means Pada Penyebaran Covid-19 Di Provinsi Jambi dapat dilihat di bawah ini.

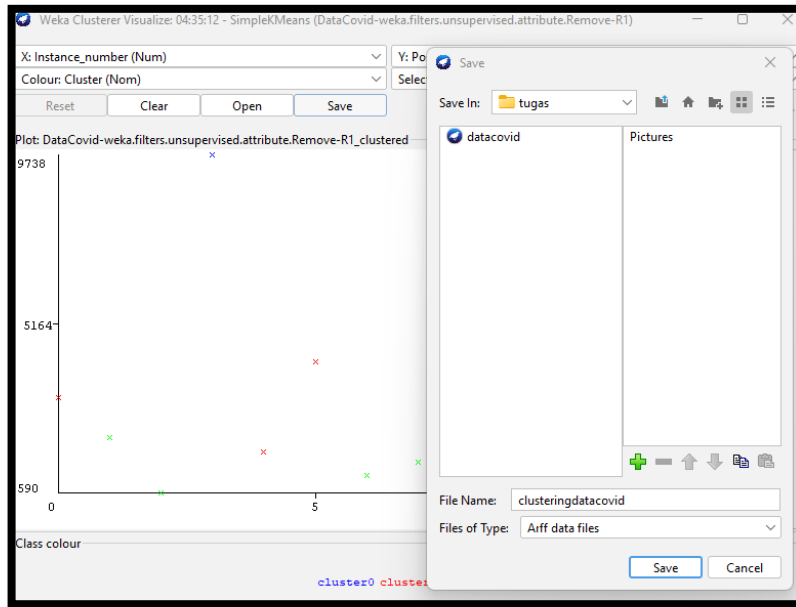


Gambar 8. Visualisasi Dari Hasil Keluaran Penerapan Algoritma K-Means Pada Penyebaran Covid-19 Di Provinsi Jambi

Berikut tampilan hasil Cluster yang diperoleh pada Penerapan Algoritma K-Means Pada Penyebaran Covid-19 Di Provinsi Jambi.



Gambar 9. Visualisasi Hasil Penerapan Algoritma K-Means Pada Penyebaran Covid-19 Di Provinsi Jambi



Gambar 10. Menyimpan Visualisasi Hasil Penerapan Algoritma K-Means Pada Penyebaran Covid-19 Di Provinsi Jambi

The screenshot shows the ARFF-Viewer interface with the following data table:

No.	1: Instance_number Numeric	2: Positif Numeric	3: Sembuh Numeric	4: Meninggal Numeric	5: Cluster Nominal
1	0.0	3165.0	3048.0	114.0	cluster1
2	1.0	2077.0	2046.0	31.0	cluster2
3	2.0	590.0	564.0	22.0	cluster2
4	3.0	9738.0	9482.0	254.0	cluster0
5	4.0	1679.0	1585.0	94.0	cluster1
6	5.0	4131.0	4060.0	68.0	cluster1
7	6.0	1036.0	1027.0	9.0	cluster2
8	7.0	1408.0	1388.0	20.0	cluster2
9	8.0	2419.0	2353.0	63.0	cluster1
10	9.0	1406.0	1370.0	36.0	cluster2
11	10.0	2137.0	1067.0	31.0	cluster2

Gambar 11. Hasil Cluster Data Penerapan Algoritma K-Means Pada Penyebaran Covid-19 Di Provinsi Jambi

3.3 Hasil Perhitungan

Berikut hasil *Clustering* Penerapan Algoritma K-Means Pada Penyebaran Covid-19 Di Provinsi Jambi menggunakan aplikasi WEKA dijadikan dalam bentuk tabel:

Tabel 3. Hasil Perhitungan Menggunakan Aplikasi WEKA

Data Nomor	Nama	Positif	Sembuh	Meninggal	Cluster
1	Batanghari	3165	3048	114	Cluster2
2	Bungo	2077	2046	31	Cluster3
3	Kerinci	590	564	22	Cluster3
4	Kota Jambi	9738	9482	254	Cluster1
5	Merangin	1679	1585	94	Cluster2
6	Muaro Jambi	4131	4060	68	Cluster2
7	Sarolangun	1036	1027	9	Cluster3
8	Sungai Penuh	1408	1388	20	Cluster3
9	Tanjab Barat	2419	2353	63	Cluster2
10	Tanjab Timur	1406	1370	36	Cluster3
11	Tebo	2137	1067	31	Cluster3

Berdasarkan Tabel 3 didapatkan bahwa perhitungan ini terdiri dari 3 *Cluster* yaitu :

1. *Cluster1* terdiri dari Kota Jambi
2. *Cluster2* terdiri dari Batanghari, Merangin, Muaro Jambi, Tanjab Barat
3. *Cluster3* terdiri dari Bungo, Kerinci, Sarolangun, Sungai Penuh, Tanjab Timur, Tebo

Jadi dapat disimpulkan bahwa tingkat penyebaran covid-19 tertinggi ada pada Kota Jambi, dan penyebaran tingkat sedang ada pada Batanghari, Merangin, Muaro Jambi, dan Tanjab Barat, serta tingkat penyebaran terendah ada pada Bungo, Kerinci, Sarolangun, Sungai Penuh, Tanjab Timur, dan Tebo

4. KESIMPULAN

Dalam proses algoritma K-Means, *centroid* pertama ditentukan secara acak. Oleh karena itu, setiap data memiliki peluang yang sama untuk terpilih sebagai *centroid* pertama. Algoritma K-Means suboptimal terdiri dari penentuan *centroid* awal. Perubahan *centroid* awal mempengaruhi nilai keseluruhan untuk hasil *Clustering* yang optimal dan mengubah validasi. Berdasarkan hasil perhitungan yang diperoleh menunjukkan bahwa aplikasi WEKA dapat digunakan dalam menentukan *Clustering*. Hasil perhitungan *Clustering* penyebaran covid-19 di Provinsi Jambi dengan aplikasi WEKA menggunakan 3 atribut yakni, jumlah kasus positif, kasus sembuh, dan kasus meninggal dan dikelompokkan menjadi 3 kelompok *Clustering* yaitu, penyebaran tingkat tinggi, tingkat sedang dan tingkat rendah. *Cluster 0* merupakan *Cluster* dengan kategori penyebaran tingkat tinggi yakni ada pada Kota Jambi. *Cluster 1* merupakan *Cluster* dengan kategori penyebaran tingkat sedang yakni terdiri dari Batanghari, Merangin, Muaro Jambi, Tanjab Barat. *Cluster 2* merupakan *Cluster* dengan kategori penyebaran tingkat rendah yakni terdiri dari Bungo, Kerinci, Sarolangun, Sungai Penuh, Tanjab Timur, Tebo.

REFERENCES

- [1] Hendrawan, A. Haris, E. Rasywir, and Y. Pratama, "Sistem Pakar Diagnosis Penyakit Tanaman Karet dengan Metode Fuzzy Mamdani Berbasis Web," *J. Media Inform. Budidarma*, vol. 4, no. 4, pp. 1225–1234, 2020, doi: 10.30865/mib.v4i4.2521.
- [2] E. Rasywir, R. Sinaga, and Y. Pratama, "Analisis dan Implementasi Diagnosis Penyakit Sawit dengan Metode Convolutional Neural Network (CNN)," *J. Paradig. UBSI*, vol. 22, no. 2, pp. 117–123, 2020.
- [3] F. Fachrudin, E. Rasywir, Hendrawan, Y. Pratama, D. Kisbianty, and M. R. Borroek, "Real Time Detection on Face Side Image with Ear Biometric Imaging Using Integral Image and Haar- Like Feature," *2018 Int. Conf. Electr. Eng. Comput. Sci.*, pp. 165–170, 2018.
- [4] A. R. Setiawan, "Lembar Kegiatan Literasi Sainifik untuk Pembelajaran Jarak Jauh Topik Penyakit Coronavirus 2019 (COVID-19)," *Edukatif J. Ilmu Pendidik.*, vol. 2, no. 1, pp. 28–37, 2020, doi: 10.31004/edukatif.v2i1.80.
- [5] N. Dwitri, J. A. Tampubolon, S. Prayoga, F. Ilmi Zer, and D. Hartama, "Penerapan Algoritma K-Means Dalam Menentukan Tingkat Penyebaran Pandemi Covid-19 Di Indonesia," *Jti (Jurnal Teknol. Informasi)*, vol. 4, no. 1, pp. 101–105, 2020.
- [6] H. Prastiwi, J. Pricilia, and E. Raswir, "Implementasi Data Mining Untuk Menentuksn Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM)," *J. Inform. Dan Rekayasa Komput.*, vol. 1, no. April, pp. 141–148, 2022.
- [7] Dodo Zaenal Abidin, S. Nurmaini, R. F. Malik, Jasmir, E. Rasywir, and Y. Pratama, "A Model of Preprocessing For Social Media Data Extraction," in *2019 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, 2019, vol. 1, no. 1, pp. 1346–1358, doi: 10.1109/JPROC.2015.2447016.
- [8] K. Celikmih, O. Inan, and H. Uguz, "Failure Prediction of Aircraft Equipment Using Machine Learning with a Hybrid Data Preparation Method," *Sci. Program.*, vol. 2020, pp. 1–10, 2020, doi: 10.1155/2020/8616039.
- [9] H. Priyatman, F. Sajid, and D. Haldivany, "Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan," *JEPIN*, vol. 5, no. 1, pp. 62–66, 2019.
- [10] E. Rasywir, Y. Pratama, H. Hendrawan, and M. Istoningtyas, "Removal of Modulo as Hashing Modification Process in Essay Scoring System Using Rabin-Karp," *2018 Int. Conf. Electr. Eng. Comput. Sci.*, pp. 159–164, 2018.
- [11] Y. Pratama and E. Rasywir, "Automatic Cost Estimation Analysis on Datawarehouse Project with Modified Analogy Based Method," in *Proceedings of 2018 International Conference on Electrical Engineering and Computer Science, ICECOS 2018*, 2019, pp. 171–176, doi: 10.1109/ICECOS.2018.8605195.
- [12] H. Hendrawan, A. Haris, E. Rasywir, and Y. Pratama, "Diagnosis Penyakit Tanaman Karet dengan Metode Fuzzy Mamdani," *J. Paradig. UBSI*, vol. 22, no. 2, pp. 132–138, 2020.
- [13] M. R. Borroek, E. Rasywir, Y. Pratama, Fachrudin, and M. Istoningtyas, "Analysis on Knowledge Layer Application for Knowledge Based System," in *Proceedings of 2018 International Conference on Electrical Engineering and Computer Science, ICECOS 2018*, 2019, pp. 177–182, doi: 10.1109/ICECOS.2018.8605262.
- [14] L. Buch and A. Andrzejak, "Learning-Based Recursive Aggregation of Abstract Syntax Trees for Code Clone Detection," *SANER 2019 - Proc. 2019 IEEE 26th Int. Conf. Softw. Anal. Evol. Reengineering*, pp. 95–104, 2019, doi: 10.1109/SANER.2019.8668039.
- [15] Fachrudin, Saparudin, E. Rasywir, and Y. Pratama, "Network and layer experiment using convolutional neural network for content based image retrieval work," *Telkommika (Telecommunication Comput. Electron. Control.*, vol. 20, no. 1, pp. 118–128, 2022, doi: 10.12928/TELKOMNIKA.v20i1.19759.
- [16] B. Demir and L. Bruzzone, "A Novel Active Learning Method in Relevance Feedback for Content-Based Remote Sensing Image Retrieval," *Geosci. Remote Sensing, IEEE Trans.*, vol. 53, no. 5, pp. 2323–2334, 2015, doi: 10.1109/TGRS.2014.2358804.

- [17] W. Xu, S. Wu, M. J. Er, C. Zheng, and Y. Qiu, “New non-negative sparse feature learning approach for content-based image retrieval,” *IET Image Process.*, vol. 11, no. 9, 2017, doi: 10.1049/iet-ipr.2016.0726.
- [18] P. Alkhairi and A. P. Windarto, “Penerapan K-Means Cluster pada Daerah Potensi Pertanian Karet Produktif di Sumatera Utara,” *Semin. Nas. Teknol. Komput. Sains*, pp. 762–767, 2019.
- [19] F. Fachruddin and Y. Pratama, “Eksperimen Seleksi Fitur Pada Parameter Proyek Untuk Software Effort Estimation dengan K-Nearest Neighbor,” *J. Inform. J. Pengemb. IT*, vol. 2, no. 2, pp. 53–62, 2017, [Online]. Available: <http://ejournal.poltektegal.ac.id/index.php/informatika/article/view/510>.
- [20] V. N. Sari, L. Y. Astri, and E. Rasywir, “Analisis Dan Penerapan Algoritma Naive Bayes Untuk Evaluasi,” *J. Ilm. Mhs. Tek. Inform.*, vol. 2, no. 1, pp. 53–68, 2020.