

## CLUSTERING PENDUDUK MISKIN MENGGUNAKAN ALGORITMA K-MEANS PADA WILAYAH JAWA BARAT

**Dewanti Fuji Astri, Martanto**

Manajemen Informatika, STMIK IKMI Cirebon

Jalan Perjuangan No. 10B Karyamulya Kec. Kesambi Kota Cirebon, Jawa Barat 45131

<sup>1</sup>*Dewantifa07@gmail.com*, <sup>2</sup>*martantomusijo@gmail.com*

### ABSTRAK

Abstrak tersebut membahas kemiskinan sebagai status sosial yang melibatkan aspek ekonomi, pendapatan, pekerjaan, pendidikan, kepemilikan harta, dan tempat tinggal. Pemerintah Indonesia menjalankan berbagai program pengentasan kemiskinan, seperti Program Keluarga Harapan (PKH) dan Bantuan Langsung Tunai (BLT). PKH memberikan bantuan keuangan bersyarat kepada Rumah Tangga Sangat Miskin (RTSM), termasuk kelompok seperti ibu hamil, anak-anak, lansia, dan penyandang disabilitas. Angka kelahiran yang tinggi dapat mempengaruhi pertumbuhan penduduk dan tingkat pengangguran di suatu wilayah. Program bantuan pemerintah kadang-kadang tidak tepat sasaran dan dapat menimbulkan kecemburuan sosial. Penelitian ini mengusulkan penggunaan metode K-Means Clustering untuk mengatasi masalah pengelompokan keluarga miskin berdasarkan berbagai kategori. Hasilnya, terdapat dua cluster dengan tingkat kemiskinan tinggi dan rendah. penerapan Algoritma *K-Means Clustering* didapatkan pengelompokan menjadi 2 *Cluster* yaitu *Cluster 0* dengan tingkat kemiskinan tinggi sebanyak 90 data dan *Cluster 1* dengan tingkat kemiskinan rendah sebanyak 45 data.

**Kata kunci :** Penduduk miskin, *K-Means*, Jawa Barat

### 1. PENDAHULUAN

Selama ini pemerintah telah berupaya untuk mengatasi masalah kemiskinan. Salah satunya adalah dukungan terhadap rumah tangga miskin, antara lain Program Keluarga Harapan (PKH), Bantuan Langsung Tunai (BLT), penerbitan kartu jaminan kesehatan, dan lain-lain. PKH memberikan bantuan keuangan bersyarat kepada Rumah Tangga Sangat Miskin termasuk Ibu Hamil, Anak Kecil, Anak Sekolah Dasar (SD), Sekolah Menengah Pertama (SMP), Sekolah Menengah atas Umum (SMA), lansia, dan penyandang disabilitas. Besaran bantuan yang diterima PKH ditentukan berdasarkan tingkat kemiskinan keluarga. Parameter kemiskinan dipertimbangkan berdasarkan jumlah kategori RTSM yang dicakup serta kewajiban peserta untuk membuat komitmen penting di bidang kesehatan dan pendidikan.[1]

Namun permasalahan mendasar yang penting dan dapat menghambat keberhasilan program ini adalah sasaran yang buruk. Salah satu penyebabnya adalah belum adanya informasi mengenai situasi kemiskinan di setiap kabupaten/kota. Karena data distribusi terstruktur terkait masyarakat miskin masih belum efisien dan akurat, maka jika ada program

dukungan pemerintah tidak tepat sasaran dan terkadang dapat menimbulkan kecemburuan di kalangan masyarakat, misalnya ada keluarga yang mempunyai akses terhadap distribusi terstruktur. Peduli terhadap masyarakat miskin dan menerima program dukungan.

Penelitian terdahulu ditinjau dari jurnal yang ditulis oleh[2] menggunakan judul “Implementasi Data Mining menggunakan algoritma K-Means Clustering dalam Pengelompokan keluarga Miskin Desa Tapung Jaya” berisi data keluarga miskin di wilayah tersebut dengan menggunakan pengelompokan K-Means dengan metode elbow yang berguna dalam memberikan rekomendasi jumlah klaster terbaik agar pengelompokan lebih cepat. Penelitian selanjutnya menunjukkan bahwa tingkat kemiskinan masih relatif tinggi di beberapa wilayah di Indonesia. Oleh karena itu, perlu digunakan cara ini untuk menurunkan angka kemiskinan di suatu daerah. Apalagi di tahun 2020 ini kita mengalami pandemi Covid-19, resesi ekonomi nasional yang menyebabkan pengangguran dimana-mana sehingga membuat perekonomian tidak stabil.

Tabel 1. Perkembangan Tingkat Kemiskinan di Jawa Barat

	2017	2018	2019	2020	2021	2022	2024
Maret	4,17 juta orang	3,62 juta orang	3,40 juta orang	3,92 juta orang	4,20 juta orang	4,07 juta orang	3,89 juta orang
September	3,77 juta orang	3,54 juta orang	3,88 juta orang	4,19 juta orang	4,00 juta orang	4,05 juta orang	

## 2. TINJAUAN PUSTAKA

### 2.1. Landasan Teori

#### 2.1.1. Algoritma

Menurut Kamus Besar Bahasa Indonesia (KBBI), algoritma juga dapat dijabarkan sebagai suatu prosedur terstruktur untuk menyelesaikan permasalahan matematis dengan langkah-langkah yang terbatas. Selain itu, algoritma juga dapat diinterpretasikan sebagai rangkaian keputusan logis untuk menyelesaikan suatu masalah. [3].

#### 2.1.2. Pengelompokan Data (Clustering)

*Clustering* adalah metode pengelompokan data yang diterapkan untuk mengenali kelompok (*cluster*) yang terbentuk melalui penggabungan elemen-elemen yang memiliki karakteristik serupa [4]. *Clustering* merupakan teknik pengelompokan data yang umumnya diterapkan dalam data mining atau penambangan data. Clustering juga sering disebut sebagai segmentasi data. Keuntungan dari penggunaan clustering adalah sebagai metode segmentasi data yang sangat efektif untuk memprediksi dan menganalisis berbagai permasalahan bisnis khusus, termasuk pemetaan zona wilayah, segmentasi pasar, dan strategi pemasaran. [5]

#### 2.1.3. K-Means

Algoritma *K-means* adalah metode klasterisasi yang mengelompokkan data berdasarkan pusat klaster terdekat atau *centroid*. Tujuan dari algoritma *K-means* adalah mengelompokkan data dengan maksimalkan kesamaan di antara klaster. Hal ini dilakukan dengan mengukur kemiripan klaster berdasarkan jarak terpendek antara data dan centroid. [6]

#### 2.1.4. Pengertian Kemiskinan

Kemiskinan merujuk pada situasi dimana seseorang atau keluarga tidak dapat memenuhi kebutuhan dasar sehari-hari karena pendapatan yang rendah, yang menghambat kenyamanan dan kemakmuran. Ketika tingkat kemiskinan suatu Negara berhasil atau berada pada tingkat rendah, Negara tersebut dianggap sejahtera. Jika kemiskinan tidak dapat diatasi, kesejahteraan tidak dapat dicapai oleh upaya pemerintah atau lembaga public. Dengan kata lain, kegagalan program kesejahteraan terjadi ketika ada komunitas yang tidak dapat memenuhi standar kebutuhan minimumnya dan hidup di bawah tingkat kesejahteraan yang ditetapkan. [7].

#### 2.1.5. Program Keluarga Harapan (PKH)

Program Keluarga Harapan (PKH) merupakan inisiatif perlindungan sosial yang memberikan bantuan uang tunai kepada Rumah Tangga Sangat Miskin (RTSM). Penerima manfaat PKH melibatkan ibu hamil, nifas, atau menyusui, anak balita atau anak usia 5 hingga 7 tahun, anak usia SD atau SMP, dan anak usia 15 hingga 18 tahun yang belum menyelesaikan pendidikan dasar. PKH berfungsi sebagai upaya pemberantasan kemiskinan dengan memberikan

perlindungan sosial di bidang pendidikan dan kesehatan kepada RTSM..[8].

#### 2.1.6. RapidMiner

RapidMiner merupakan suatu platform perangkat lunak yang sangat kuat untuk pembelajaran mesin dan ilmu data, menyediakan berbagai alat untuk pemodelan, evaluasi, implementasi, dan persiapan data. Meskipun dirancang untuk mudah digunakan, RapidMiner memungkinkan individu tanpa pengalaman pemrograman untuk dengan cepat membangun dan menguji berbagai model. Dengan antarmuka drag-and-drop, pengguna dapat membuat alur kerja untuk memproses dan menganalisis data. Keunggulan perangkat lunak ini terletak pada dukungan terhadap berbagai sumber data, seperti file datar, basis data, dan platform big data seperti Hadoop dan Spark. RapidMiner juga menyertakan berbagai operator bawaan yang mencakup semua langkah dalam proses pemrosesan data, seperti pemodelan, pembersihan data dan pemilihan fitur.[9]

## 3. METODE PENELITIAN

### 3.1. Sumber Data

Data yang digunakan dalam penelitian ini berasal dari website resmi Daerah Jawa Barat, [data.jabarprov.go.id](http://data.jabarprov.go.id) yang saya pilih pada tanggal 1 November 2023 untuk mengumpulkan informasi tentang penduduk miskin di Daerah Jawa Barat. Penelitian ini menggunakan data sekunder. Dataset ini berisi indeks kedalaman kemiskinan untuk masing-masing Kabupaten atau Kota pada Provinsi Jawa Barat dari tahun 2018 sampai 2022. Data ini berisi id, nama Kabupaten dan Kota, tahun, kode Kabupaten dan Kota, serta kedalaman kemiskinan di tahun tersebut.

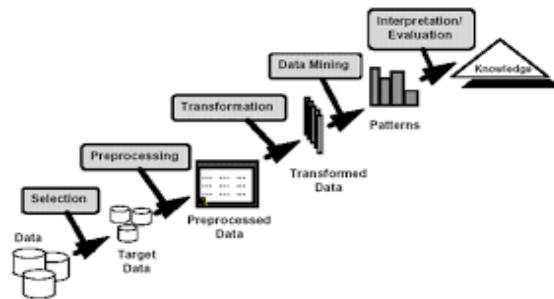
### 3.2. Teknik Pengumpulan Data

Dalam penelitian ini, data sekunder dimanfaatkan karena berasal dari sumber lain. Secara lebih spesifik, penelitian ini memperoleh informasi melalui analisis literatur, yang mencakup pengumpulan data dari berbagai sumber seperti buku, jurnal, artikel, dan internet. [10]. Terdapat 135 data dari tahun 2018 hingga 2022 dari Badan Pusat Statistik Provinsi Jawa Barat. Data ini diambil dari situs web resmi open data Jawa Barat.

### 3.3. Tahapan Perancangan

Tahapan perancangan yang dilakukan dalam penelitian yang diterapkan dalam penelitian ini menggunakan metode perancangan KDD. Data mining, yang juga dikenal sebagai *Knowledge Discovery in Database* (KDD), merupakan suatu proses ekstraksi infomasi yang potensial, implisit, dan belum diketahui dari sejumlah data. Hasil dari proses data mining, yang melibatkan identifikasi trend dan pola data, selanjutnya diolah dengan akurat untuk menghasilkan informasi yang dapat dipahami dengan mudah [11]. Proses *Knowledge Discovery in Database* secara garis besar terdiri dari 5 tahapan yaitu data

*selection, pre-processing/cleaning, transformation, data mining dan interpretation/evaluation.* Tahapan dari KDD dapat dilihat pada gambar berikut :



Gambar 1. Tahapan KDD

Berikut ini merupakan tahapan dari proses *Knowledge Discovery in Database (KDD)* :

#### 1) Data Selection

Langkah pemilihan data, yang juga dikenal sebagai “seleksi”, merujuk pada tahap dimana data dipilih dari kumpulan data operasional sebelum dilakukan proses pengumpulan informasi KDD. Data yang diambil untuk kegiatan data mining berasal dari data operasional dan dikompilasi dalam satu dokumen. [12]

#### 2) Data Preprocessing

*Preprocessing data* merupakan tahap awal dalam proses data mining. Pada tahap ini dilakukan tindakan penghapusan data duplikat, pemeriksaan kesalahan data, dan koreksi kesalahan seperti kesalahan penulisan[13]

#### 3) Data Transformation

Tahap mengonversi data kedalam suatu format khusus agar dapat dimanfaatkan dalam proses data mining.

#### 4) Data Mining

*Data mining* merupakan suatu proses pengidentifikasi pola atau informasi yang menarik dari sejumlah data tertentu dengan menerapkan teknik khusus, bertujuan untuk menghasilkan informasi yang memiliki nilai atau manfaat.. [14]

#### 5) Interpretation/evaluation

Tahap ini mengurangi keperluan untuk memverifikasi apakah informasi atau data yang dipresentasikan sejalan dengan fakta atau hipotesis sebelumnya.[12].

## 4. HASIL DAN PEMBAHASAN

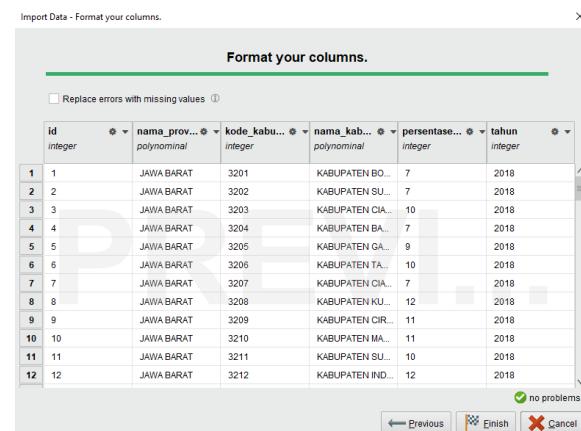
### 4.1. Hasil Penerapan Algoritma K-Means Clustering di Jawa Barat

Dalam bagian hasil penelitian ini, tahapan KDD (*Know Discovery in Database*) digunakan dengan alat *RapidMiner* versi 10.1. Data ini diambil dari [opendatajabar.com](http://opendatajabar.com). Algoritma Clustering K-means digunakan untuk uji coba ini.

#### 4.1.1. Data Selection

Pada tahap awal, hanya beberapa atribut yang dapat diakses, yaitu id, nama\_provinsi, kode\_kabupaten\_kota,nama\_kabupaten\_kota,persenta se\_pendudu, dan tahun, sehingga total dataset memiliki 135 daftar. Langkah pertama dalam proses pemilihan data adalah mengimport data penduduk miskin dalam format excel (xlsx) ke dalam aplikasi rapidminer. Untuk dapat melakukannya dengan menggunakan import data atau menggunakan konfigurasi wizard di bagian operator Read Excel, yang dapat membaca data yang telah diimport.

Berikut merupakan gambar dari Data Selection:



Gambar 2. Import data pada operator Read Excel

Langkah selanjutnya menggunakan operator set role yaitu proses menentukan atribut mana saja yang akan digunakan. Berikut merupakan bentuk operator set role:



Gambar 3. Operator Set Role

Hasil data yang telah dimasukan menggunakan operator set role terdapat pada gambar 4

Row No.	id	nama_provinsi	kode_kabupaten_kota	nama_kabupaten_kota	persentase_se_pendudu	tahun	Set Role
1	1	JAWA BARAT	3201	KABUPATEN BO...	7	2018	Set Role
2	2	JAWA BARAT	3202	KABUPATEN SU...	7	2018	Set Role
3	3	JAWA BARAT	3203	KABUPATEN CIA...	10	2018	Set Role
4	4	JAWA BARAT	3204	KABUPATEN BA...	7	2018	Set Role
5	5	JAWA BARAT	3205	KABUPATEN GA...	9	2018	Set Role
6	6	JAWA BARAT	3206	KABUPATEN TA...	10	2018	Set Role
7	7	JAWA BARAT	3207	KABUPATEN CI...	7	2018	Set Role
8	8	JAWA BARAT	3208	KABUPATEN KU...	12	2018	Set Role
9	9	JAWA BARAT	3209	KABUPATEN CIR...	11	2018	Set Role
10	10	JAWA BARAT	3210	KABUPATEN MA...	11	2018	Set Role
11	11	JAWA BARAT	3211	KABUPATEN SU...	10	2018	Set Role
12	12	JAWA BARAT	3212	KABUPATEN IND...	12	2018	Set Role
13	13	JAWA BARAT	3213	KABUPATEN ...	9	2018	Set Role
14	14	JAWA BARAT	3214	KABUPATEN ...	8	2018	Set Role
15	15	JAWA BARAT	3215	KABUPATEN ...	8	2018	Set Role

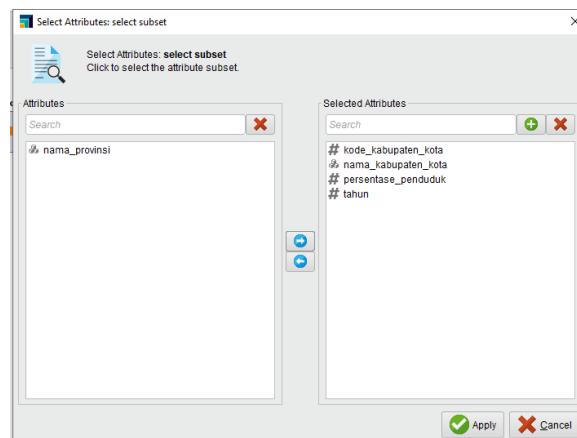
Gambar 4. Data Pada Operator Set Role

#### 4.1.2. Data Cleaning / Preprocessing

Pada proses tahapan Preprocessing yaitu untuk menghapus atribut data yang tidak diperlukan dan tidak digunakan agar saat diproses tidak menimbulkan kesalahan atau masalah pada data tersebut.

Gambar 5. Hasil Data Preprocessing

Pada gambar 5 proses data cleaning atau data Preprocessing didapatkan hasil bahwa tidak ada kesalahan pada data tersebut. Kemudian langkah selanjutnya adalah menggunakan operator *Select Attributes* yaitu operator yang berfungsi untuk memilih *subset* atribut yang akan digunakan dari *ExampleSet* dan menghapus atribut lainnya yang tidak digunakan.



Gambar 6. Hasil Data Select Attributes

#### 4.1.3. Data Transformation

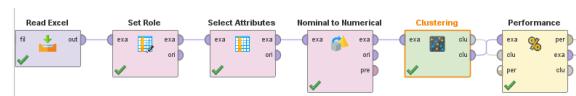
Gambar 7. Nominal to numerical

Dalam proses tahapan Data Transformation ini mengubah tipe atribut data yang semula *non-numeric* diubah menjadi tipe data *numeric*. Berikut hasil

penampilan data Tranformation menggunakan operator *Nominal to Numerical*.

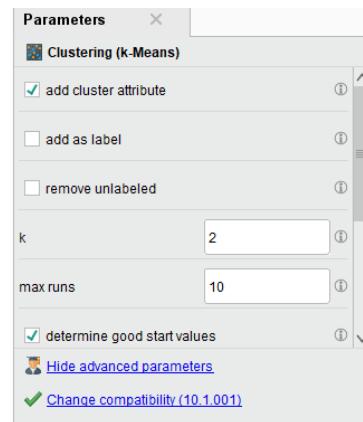
#### 4.1.4. Data Mining

Pada tahapan Data mining menggunakan proses *Clustering* menggunakan algoritma K-Means dengan operator yang digunakan yaitu *Clustering K-Means* dan untuk melakukan pengujian dan mengevaluasi hasil dari *Clustering* menggunakan operator *Cluster Disance Perfomance* dengan metode yang digunakan yaitu *David Bouldin Index (DBI)* untuk menentukan jumlah cluster terbaik. Berikut merupakan gambar model Algoritma K-means:



Gambar 8. Model Data Mining

Pada saat proses melakukan *clustering* K-means yaitu untuk menentukan K yang paling optimal ketika menggunakan operator performance untuk menghitung DBI dan parameter yang digunakan dimulai dari K = 2 sampai K = 10. Pada proses *clustering* terdapat parameter yang harus disesuaikan terlebih dahulu. Tampilan parameter *clustering* pada gambar berikut ini :



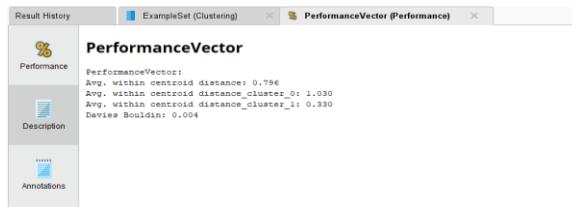
Gambar 9. Operator K-means Clustering

K adalah jumlah cluster yang digunakan. Jumlah percobaan cluster yang ada dicontoh yaitu dari 2-10 cluster. Kemudian, *max runs* adalah jumlah maksimum run K-means dan jumlah yang ada dicontoh yaitu 10.

#### 4.1.5. Interpretation/ Evaluation

Operator *Perfomance Cluster Distance* digunakan dalam penelitian ini. Tahap evaluasi ini dilakukan untuk mengevaluasi keberhasilan model yang telah digunakan. Sepuluh percobaan dilakukan dengan operator *Perfomance Cluster Distance* ini untuk menentukan nilai K yang ideal dengan parameter max/run. Nilai DBI (*Davies Bouldin Index*) dari setiap cluster dibatasi untuk menentukan hasil yang paling baik berdasarkan jarak terdekat yang telah

dikelompokkan. Berdasarkan hasil implementasi algoritma K-means, diperoleh hasil performance :



Gambar 10. Hasil Perfomance Vector

Pengujian dilakukan dengan menggunakan DBI terhadap 2 *cluster* dengan keseluruhan data sebanyak 135 record. Dengan hasil *Avg. Centroid distance* bernilai 0.796, hasil rata-rata *cluster\_0* bernilai 1.030, rata-rata dalam *cluster\_1* bernilai 0,330, kemudian nilai pengukuran jarak antar titik dalam *cluster* didapat nilai DBI 0,004.

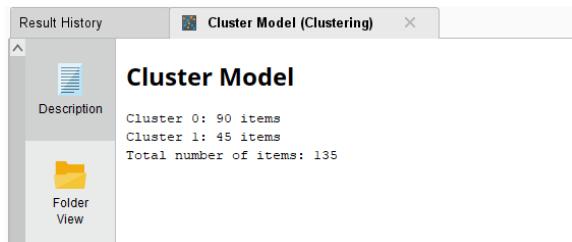
#### 4.2. Analisis Hasil Pengelompokan Algoritma K-means di Jawa Barat

Berikut merupakan hasil penerapan Algoritma K-means yang sudah dilakukan : Pada hasil Evaluasi setelah dilakukan proses clustering dengan menggunakan cluster 2 sampai dengan 10 rekapitulasi dapat dilihat nilai DBI pada table sebagai berikut :

Tabel 2. Hasil Menentukan Nilai K

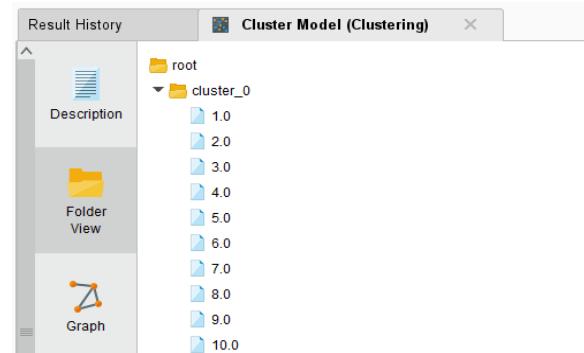
Nilai K	AVG	DBI
2	0.796	0.004
3	0.330	0.016
4	0.242	0.021
5	0.183	0.028
6	0.153	0.031
7	0.138	0.034
8	0.125	0.037
9	0.120	0.038
10	0.140	0.040

Dijelaskan pada table 2 merupakan proses menentukan Cluster Optimun dan nilai DBI dari mengelompokan tingkat kemiskinan di Jawa Barat menggunakan Algoritma K-mens. Perlu dilakukan uji coba untuk menentukan nilai K dimulai nilai Cluster 2-10 dan didapat yang mendekati angka 0 adalah ada di K 2 dengan nilai DBI 0.004. Jumlah *Cluster* sebanyak 2 (dua) diperoleh dengan total data item sebanyak 135 items. *Cluster* 0 berjumlah 90 record, sedangkan *Cluster* 1 berjumlah 45 record. Berikut merupakan hasil gambar *Cluster* model :



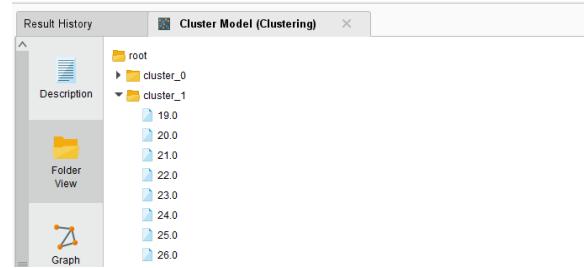
Gambar 11. Hasil Cluster Model

Sedangkan untuk melihat pengelompokan *Cluster* 0 dan *Cluster* 1 dalam bentuk Folder yang terdapat pada bagian Folder View. Berikut gambar 11 dari Folder View pada *Cluster* 0 :



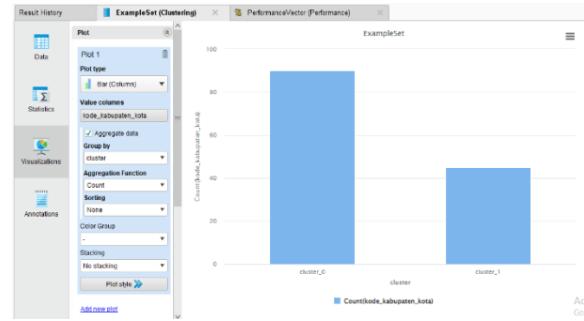
Gambar 12. Folder View Cluster 0

Berikut gambar 13 Folder View pada *Cluster* 1 :



Gambar 13. Folder View Cluster 1

Berikut merupakan tampilan Visualizations gambar grafik histogram hasil Cluster dari keseluruhan data :



Gambar 14. Tampilan Histogram Cluster

Berdasarkan histogram pada Gambar 14 terdapat 2 Cluster, yaitu Cluster 0 dengan jumlah Kota atau Kabupaten dengan tingkat kemiskinan tinggi sebanyak 90 data, dan Cluster 1 dengan jumlah Kota atau Kabupaten dengan tingkat kemiskinan rendah sebanyak 45 data. *Cluster* 0 dengan kelompok Kota atau Kabupaten dengan tingkat kemiskinan tinggi sebanyak 90 data, dengan gambar *Cluster* sebagai berikut :

1	id	nama provinsi	kode kabupaten_kota	nama kabupaten_kota	percentase_penduduk	tahun
2	1	JAWA BARAT	3201	KABUPATEN BOGOR	7,14	2018
3	2	JAWA BARAT	3202	KABUPATEN SUKABUMI	6,76	2018
4	3	JAWA BARAT	3203	KABUPATEN CIANJUR	9,81	2018
5	4	JAWA BARAT	3204	KABUPATEN BANDUNG	6,65	2018
6	5	JAWA BARAT	3205	KABUPATEN GARUT	9,27	2018
7	6	JAWA BARAT	3206	KABUPATEN TASIKMALAYA	9,85	2018
8	7	JAWA BARAT	3207	KABUPATEN CIAMIS	7,22	2018
9	8	JAWA BARAT	3208	KABUPATEN KUNINGAN	12,22	2018
10	9	JAWA BARAT	3209	KABUPATEN CIREBON	10,7	2018
11	10	JAWA BARAT	3210	KABUPATEN MAJALENGKA	10,79	2018
12	11	JAWA BARAT	3211	KABUPATEN SUMEDANG	9,76	2018
13	12	JAWA BARAT	3212	KABUPATEN INDRAMAYU	11,89	2018
14	13	JAWA BARAT	3213	KABUPATEN SUBANG	8,67	2018
15	14	JAWA BARAT	3214	KABUPATEN PURWAKARTA	7,99	2018
16	15	JAWA BARAT	3215	KABUPATEN KARAWANG	8,06	2018
17	16	JAWA BARAT	3216	KABUPATEN BEKASI	4,37	2018
18	17	JAWA BARAT	3217	KABUPATEN BANDUNG BARAT	10,06	2018
19	18	JAWA BARAT	3218	KABUPATEN PANGANDARAN	8,12	2018
20	28	JAWA BARAT	3201	KABUPATEN BOGOR	6,66	2019
21	29	JAWA BARAT	3202	KABUPATEN SUKABUMI	6,22	2019
22	30	JAWA BARAT	3203	KABUPATEN CIANJUR	9,15	2019
23	31	JAWA BARAT	3204	KABUPATEN RANDUONG	5,94	2019

Gambar 15. Cluster 0

Cluster 1 dengan kelompok Kota atau Kabupaten dengan tingkat kemiskinan rendah dengan jumlah sebanyak 45 data, dengan gambar *Cluster* sebagai berikut :

1	id	nama provinsi	kode kabupaten_kota	nama kabupaten_kota	percentase_penduduk	tahun
2	19	JAWA BARAT	3271	KOTA BOGOR	5,93	2018
3	20	JAWA BARAT	3272	KOTA SUKABUMI	7,12	2018
4	21	JAWA BARAT	3273	KOTA BANDUNG	3,57	2018
5	22	JAWA BARAT	3274	KOTA CIREBON	8,88	2018
6	23	JAWA BARAT	3275	KOTA BEKASI	4,11	2018
7	24	JAWA BARAT	3276	KOTA DEPOK	2,14	2018
8	25	JAWA BARAT	3277	KOTA CIAMIH	4,94	2018
9	26	JAWA BARAT	3278	KOTA TASIKMALAYA	12,71	2018
10	27	JAWA BARAT	3279	KOTA BANJAR	5,7	2018
11	46	JAWA BARAT	3271	KOTA BOGOR	5,77	2019
12	47	JAWA BARAT	3272	KOTA SUKABUMI	6,67	2019
13	48	JAWA BARAT	3273	KOTA BANDUNG	3,38	2019
14	49	JAWA BARAT	3274	KOTA CIREBON	8,41	2019
15	50	JAWA BARAT	3275	KOTA BEKASI	3,81	2019
16	51	JAWA BARAT	3276	KOTA DEPOK	2,07	2019
17	52	JAWA BARAT	3277	KOTA CIAMIH	4,39	2019
18	53	JAWA BARAT	3278	KOTA TASIKMALAYA	11,6	2019
19	54	JAWA BARAT	3279	KOTA BANJAR	5,5	2019
20	73	JAWA BARAT	3271	KOTA BOGOR	6,68	2020
21	74	JAWA BARAT	3272	KOTA SUKABUMI	7,7	2020
22	75	JAWA BARAT	3273	KOTA BANDUNG	3,99	2020
23	76	JAWA BARAT	3274	KOTA CIREBON	9,52	2020

Gambar 16. Cluster 1

## 5. KESIMPULAN DAN SARAN

Setelah menerapkan Data Mining metode K-Means Clustering dalam melakukan pengelompokan data Kemiskinan di Provinsi Jawa Barat dengan jumlah data sebanyak 135 data, pengelompokan data dilakukan dengan menggunakan tools RapidMiner. Dapat diuraikan kesan sebagai berikut : Dari hasil perhitungan menggunakan metode K-Means Clustering, didapatkan informasi yaitu berupa tingkat kemiskinan yang terbagi menjadi 2 Cluster yaitu Cluster 0 dengan tingkat kemiskinan tinggi sebanyak 90 data dan Cluster 1 dengan tingkat kemiskinan rendah sebanyak 45 data. Setelah menentukan pengelompokan menggunakan K-Means Clustering ini, diharapkan dengan adanya informasi ini pemerintah Jawa Barat dengan cepat menentukan kebijakan yang tepat untuk menyalurkan beberapa bantuan kepada masyarakat yang membutuhkan, dan didapatkan pada Cluster 0 dengan tingkat kemiskinan tinggi lebih membutuhkan bantuan dikarenakan terdapat banyak masyarakat yang masih kekurangan di wilayah tersebut.

Penelitian ini masih banyak kekurangan untuk peningkatan yang dapat digunakan di masa mendatang. Berikut ini merupakan saran untuk pengembangan penelitian lebih lanjut, yaitu : Untuk membandingkan hasil akhir dari proses peningkatan cluster, penelitian ini dapat menggunakan algoritma tambahan seperti K-Medoids atau metode Fpgrowth. Untuk Pemerintah Jawa Barat harus memberikan

perhatian lebih besar pada masyarakat dan meninjau kembali ke lapangan untuk membuat kebijakan untuk memerangi kemiskinan di wilayah tersebut. Ini termasuk membangun infrastruktur masyarakat umum seperti jalan umum dan peningkatan program pendidikan, serta mengembangkan program ekonomi seperti bantuan kepada UMKM.

## DAFTAR PUSTAKA

- [1] D. RAHAYU UTAMI, “JURUSAN TEKNIK INFORMATIKA FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM MALANG 2018 APLIKASI MONITORING KELUARGA MISKIN MENGGUNAKAN METODE K-MEANS CLUSTERING BERBASIS MOBILE GIS (STUDI KASUS : PKH KEC KEDUNGKANDANG KOTA MALANG),” 2018.
- [2] D. Sebagai *et al.*, “IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING DALAM PENGELOMPOKAN KELUARGA MISKIN DESA TAPUNG JAYA TUGAS AKHIR,” 2022.
- [3] Y. Suharya and H. Widia, “Implementasi Digital Signature Menggunakan Algoritma Kriptografi RSA Untuk Pengamanan Data Di Smk Wirakarya 1 Ciparay,” *Computing*, vol. 07, no. 01, pp. 20–29, 2020.
- [4] A. S. Jinggo Pratama, Abdul Khamid, and Yesy Diah Rosita, “Pencarian Rute Optimal Wisata Mojokerto Dalam Kasus Traveling Salesman Problem Menggunakan Algoritma Genetika,” *Jurnal Informatika Teknologi dan Sains (Jinteks)*, vol. 5, no. 2, pp. 283–288, 2023, doi: 10.51401/jinteks.v5i2.2447.
- [5] hanif sri Yulianto, “Pengertian Clustering, Metode, dan Syarat-syaratnya,” *bola.com*, 2022.
- [6] Y. Prihati, Suwarno, and A. Dharmawan, “Implementasi Algoritma K-Means Untuk Pemetaan Prestasi Akademik Siswa Disekolah Dasar Terang Bagi Bangsa Pati,” *Kinabalu*, vol. 11, no. 2, pp. 50–57, 2019.
- [7] H. F. D. Laoh, Elvira Rosa. Kalangi, Josep Bintang. Siwu, “Pengaruh Produk Domestic Regional Bruto Dan Indeks Pembangunan Manusia,” *Jurnal Berkala Ilmiah Efisiensi*, vol. 23, no. 1, pp. 85–96, 2023.
- [8] S. A. S. Risna Resnawaty, “Program Keluarga Harapan ( Pkh ) : Antara Perlindungan Sosial,” *Prosiding Ks: Riset & Pkm*, vol. 4, no. 1, pp. 1–140, 2017.
- [9] M. R. Nahjan, N. Heryana, and A. Voutama, “IMPLEMENTASI RAPIDMINER DENGAN METODE CLUSTERING K-MEANS UNTUK ANALISA PENJUALAN PADA TOKO OJ CELL,” 2023.
- [10] M. A. Sholihhudin, “Implementasi Data Mining Menggunakan Algoritma K-Means Clustering

- Dalam Pengelompokan Keluarga Miskin Desa Tapung Jaya,” 2022.
- [11] M. Marsono, “Penerapan Data Mining Pengaturan Pola Tata Letak Barang Pada Berkah Swalayan Untuk Strategi Penjualan Menggunakan Algoritma Apriori,” *InfoTekJar (Jurnal Nasional Informatika dan Teknologi Jaringan)*, vol. 3, no. 2, pp. 170–175, 2019, doi: 10.30743/infotekjar.v3i2.908.
- [12] N. Septiani, S. Anwar, and R. Herdiana, “Penerapan Algoritma K-Means Clustering Untuk Harga Rumah di Jakarta Selatan,” *Trending: Jurnal Ekonomi, Akuntansi dan Manajemen*, vol. 1, no. 2, 2023.
- [13] E. Priyanti, “Penerapan Algoritma Naïve Bayes Untuk Klasifikasi Bakteri Gram- Negatif,” *Jurnal Teknik Komputer*, vol. III, no. 2, pp. 68–76, 2017.
- [14] Yuli Mardi, “Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5 Data mining merupakan bagian dari tahapan proses Knowledge Discovery in Database ( KDD ) . Jurnal Edik Informatika,” *Jurnal Edik Informatika*, vol. 2, no. 2, pp. 213–219, 2019.