

## PENERAPAN ALGORITMA K-MEANS UNTUK *CLUSTERING* DATA KEMISKINAN PROVINSI BANTEN MENGGUNAKAN RAPIDMINER

Yunita Ratna Sari<sup>1</sup>, Arby Sudewa<sup>2</sup>, Diah Ayu Lestari<sup>3</sup>, Tri Ika Jaya<sup>4</sup>

<sup>1,2,3,4</sup>Fakultas Teknologi Informasi Universitas Budi Luhur  
Jl. Ciledug Raya, Petukangan Utara, Jakarta Selatan, 12260. DKI Jakarta,  
Indonesia.

<sup>1</sup>11712502341@student.budiluhur.ac.id, <sup>2</sup>1712501244@student.budiluhur.ac.id,  
<sup>3</sup>1712500576@student.budiluhur.ac.id, <sup>4</sup>tri.ikajaya@budiluhur.ac.id

---

**Abstrak**— Angka kemiskinan di Provinsi Banten tergolong rendah di tingkat nasional. Hal ini dibuktikan dengan persentase penduduk miskin Banten pada September 2019 sebesar 4,94% berada di bawah rata-rata nasional pada periode yang sama sebesar 9,22%. Penelitian ini memanfaatkan teknik data mining dengan menggunakan metode K-Means *Clustering*. Data penelitian ini diambil dari situs Badan Pusat Statistik (BPS) dari tahun 2015-2019 yang terdiri dari 8 Kabupaten/Kota dengan 3 variabel. Variabel yang digunakan adalah jumlah penduduk miskin (ribu jiwa), rata-rata lama pendidikan sekolah (tahun), dan pengeluaran per kapita yang disesuaikan (ribu rupiah/tahun). Semua data tersebut kemudian diolah dengan Rapidminer dan dilakukan 3 *cluster*, yaitu: tingkat *cluster* sedang (C0), tingkat *cluster* tinggi (C1), dan tingkat *cluster* rendah (C2). Hasil dari perhitungan rapidminer menunjukkan Kabupaten Tangerang, Kota Cilegon, dan Kota Serang masuk sebagai anggota *cluster* 0, Kabupaten Pandeglang, Kabupaten Lebak, dan Kabupaten Serang berada pada *cluster* 1, Kota Tangerang, dan Kota Tangerang Selatan berada di *cluster* 2.

**Kata Kunci**— Data Mining, Kemiskinan, Clustering, K-means

**Abstract**— The poverty rate in Banten Province is classified as low at the national level. This is evidenced by the percentage of poor people in Banten in September 2019 of 4.94% which is below the national average in the same period of 9.22%. This study utilizes data mining techniques using the K-Means Clustering method. This research data was taken from the Central Statistics Agency (BPS) website from 2015-2019 which consisted of 8 districts / cities with 3 variables. The variables used are the number of poor people (thousand inhabitants), the average length of school education (years), and adjusted per capita expenditure (thousand rupiah / year). All data is then processed by Rapidminer and performed 3 clusters, namely: medium cluster level (C0), high cluster level (C1), and low cluster level (C2). The results of the rapidminer calculation show that Tangerang Regency, Cilegon City, and Serang City are included as members of cluster 0, Pandeglang Regency, Lebak Regency, and Serang Regency are in cluster 1, Tangerang City and South Tangerang City are in cluster 2.

**Keywords**— Data Mining, Poverty, Clustering, K-means

---

### I. PENDAHULUAN

Provinsi Banten memiliki berbagai masalah yang dihadapi salah satunya adalah kemiskinan. Dalam 5 tahun terakhir (2015-2019), kemiskinan di Provinsi Banten rata-rata mengalami penurunan. Menurut data yang disajikan oleh Badan Pusat Statistik (BPS) dari hasil Survey Sosial Ekonomi Nasional (SUSENAS) persentase penduduk miskin di Banten pada September 2019 mencapai 4,94%. Kondisi tersebut berada di bawah rata-rata nasional pada periode yang sama sebesar 9,22% [1]. Meskipun cenderung mengalami penurunan, disparitas atau ketimpangan sosial (terutama dalam pembangunan infrastruktur)

antar wilayah Banten Utara dan Selatan dinilai cukup tinggi perbedaannya. Disparitas dapat terjadi akibat berbagai indikator pendukung yang mempengaruhi perbedaan tingkat kemiskinan di setiap kabupaten atau kota. Kemiskinan ini juga dapat diakibatkan karena kurangnya perkembangan pendidikan di setiap daerah yang ada di provinsi Banten. Hal ini bisa dipicu karena kurang perhatiannya pemerintah pada tingkat pendidikan yang memadai, sehingga banyak sekali masyarakat yang tidak menempuh pendidikan yang layak dan hal ini menjadikan tingkat jumlah rata-rata yang lanjut sekolah setiap daerah berbeda-beda. Lalu pendapatan perkapita juga sangat berpengaruh pada pembangunan yang ada disebuah daerah agar maju.

Namun di provinsi banten ini masih adanya tidak meratanya pendapatan setiap daerahnya.

Untuk mengatasinya, dibutuhkan penanganan yang berbeda antar wilayah sesuai pada kondisi masing-masing kabupaten/kota. Salah satu cara yang bisa dilakukan adalah dengan mengelompokkan karakteristik suatu wilayah berdasarkan indikator kemiskinan[1]. Pengelompokkan karakteristik wilayah dapat dilihat dari seberapa lama tingkat pendidikan penduduknya, atau seberapa banyak jumlah penduduk miskin dalam suatu wilayah, dan lain sebagainya sesuai dengan indikator kemiskinan telah ditentukan publikasi data kemiskinan Kota/Kabupaten yang disajikan oleh BPS.

Melihat permasalahan di atas, penulis tertarik melakukan penelitian dengan memanfaatkan Data Mining untuk mengelompokkan data masyarakat miskin yang ada di Provinsi Banten. Data mining atau penambangan data merupakan *machine learning* dalam ilmu sains yang dapat mengekstrak pengetahuan atau menemukan pola dari suatu data[5]. Data mining memiliki segudang manfaat untuk mengelola data mentah menjadi *knowledge* atau kumpulan informasi yang dapat dijadikan sebagai penunjang keputusan secara efektif[2].

## II. TINJAUAN PUSTAKA

### A. Data Mining

*Data Mining* merupakan bukan sesuatu hal yang baru untuk mengolah data menjadi sebuah *knowledge*. Yang menjadi salah satu faktor yang menjadi kesulitan untuk menggambarkan *Data Mining* adalah faktanya bahwa Data Mining menurun banyak aspek dari bidang-bidang ilmu yang sudah ada terlebih dahulu. Diawali dari beberapa disiplin ilmu, Data Mining bertujuan untuk mengubah pandangan data tradisional sehingga bisa menangani:

1. Jumlah data yang sangat besar
2. Dimensi data yang tinggi
3. Data yang heterogen dan berbeda sifat

Data mining memiliki segudang manfaat untuk mengelola data mentah menjadi *knowledge* atau kumpulan informasi yang dapat dijadikan sebagai penunjang keputusan secara efektif[12].

Pengelompokan Data Mining dibagi menjadi beberapa kelompok [13] yaitu :

#### 1. Deskripsi

Deskripsi merupakan cara untuk mengidentifikasi pola dan kecenderungan yang terdapat dalam data yang dimiliki.

#### 2. Estimasi

Estimasi memperkirakan sebuah nilai yang belum diketahui seperti memperkirakan penghasilan seseorang tanpa mengetahui orang tersebut. Estimasi ini seperti klasifikasi namun bedanya targetnya belum diketahui

#### 3. Prediksi

Prediksi memperkirakan sebuah nilai yang belum diketahui dan juga memperkirakan nilai untuk masa mendatang.

#### 4. Klasifikasi

Dalam klasifikasi sudah terdapat target variable kategori, misal penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu tinggi, sedang, dan rendah.

#### 5. Pengklasteran

Merupakan pengelompokan record, pengamatan, atau memperhatikan dan membentuk kelas objek-objek yang memiliki kesamaan dan kemiripan.

#### 6. Asosiasi

Asosiasi berguna untuk menemukan atribut yang muncul dalam satu waktu. Seperti yang sering terjadi di dunia perbelanjaan khususnya Retail.

### B. K-Means Clustering

K-Means merupakan salah satu metode yang digunakan pada *clustering*. *Clustering* mempunyai peranan penting dalam aplikasi data mining [6]. Biasanya *clustering* dimanfaatkan untuk text mining serta dalam analisis web. Metode *clustering* bekerja dengan cara yang cukup sederhana yaitu dengan membagi data ke dalam grup-grup yang memiliki kemiripan antar objeknya[4]. Dalam penelitian ini, penggunaan klasterisasi dapat mengidentifikasi suatu wilayah dengan menemukan pola-pola karakteristik hingga menemukan keterkaitan yang menarik antara atribut-tribut data[10].

Berdasarkan penelitian terdahulu yang telah dilakukan oleh [7],[8],[9], algoritma k-means dinilai mempunyai tingkat akurasi yang cukup tinggi sehingga mampu dipakai guna menghasilkan *knowledge* kelompok karakteristik wilayah pada penelitian ini. K-Means termasuk algoritma *clustering* dengan proses berulang-ulang. Huruf K diartikan sebagai jumlah cluster yang hendak dibuat. Selanjutnya nilai K ditetapkan secara acak. Sedangkan means adalah nilai sementara yang menjadi pusat dari klaster atau disebut juga dengan *centroid*. Setiap data yang ada dihitung jaraknya terhadap masing-masing *centroid* dengan memakai rumus Euclidean hingga dihasilkan jarak terdekat dari setiap data dengan *centroid*[10].

### C. RapidMiner

*RapidMiner* merupakan perangkat lunak yang dapat diakses oleh siapa saja dan bersifat terbuka (*open source*). *RapidMiner* ini dijadikan sebuah solusi untuk menganalisa terhadap *data processing*. Pada *RapidMiner* ini digunakan berbagai teknik seperti teknik deskriptif dan prediksi. *RapidMiner* ini menggunakan bahasa Java untuk pengoperasian nya[11]

Berkat adanya kecanggihan teknologi algoritma komputasi dan analisis data berbasis komputer,

datamining dapat diolah menggunakan *software* Rapidminer. Perhitungan algoritma K-Means akan dilakukan dengan menggunakan rapidminer. Hasil dari penelitiannya ini diharapkan dapat membantu pemerintah dalam mengenal tingkat kemiskinan di setiap wilayah Kabupaten / Kota pada Provinsi Banten. Sehingga, pemerintah bisa lebih tepat dan tanggap dalam mengambil kebijakan untuk menanggulangi kesenjangan antar wilayah terutama dalam hal pemerataan pendidikan.

### III. METODELOGI PENELITIAN

#### A. Tahap Pengumpulan Data

Data yang digunakan selama penelitian ini diambil berdasarkan dokumen-dokumen keterangan rata-rata lama sekolah masyarakat kabupaten/kota provinsi Banten, pengeluaran per kapita yang disesuaikan menurut kabupaten/kota provinsi Banten, dan jumlah penduduk miskin menurut kabupaten/kota provinsi Banten yang dihasilkan oleh Publikasi Statistik Indonesia dan diolah oleh Badan Pusat Statistik Nasional. Data yang digunakan tersebut adalah data dari tahun 2015 sampai 2019 (periode Maret).

#### B. Tahap Pengolahan Data

Data yang telah didapat akan dilakukan proses

pengolahan terlebih dahulu untuk dapat dilakukan proses clustering. Dalam tahap ini, akan dilakukan proses perhitungan rata-rata dari setiap data yang digunakan. Sehingga menghasilkan rata-rata berdasarkan 3 kriteria penilaian yakni:

- 1) Rata-rata lama sekolah masyarakat kabupaten/kota yang ada di provinsi Banten.
- 2) Rata-rata pengeluaran per kapita yang disesuaikan menurut kabupaten/kota provinsi Banten.
- 3) Rata-rata jumlah penduduk miskin menurut kabupaten/kota yang ada di provinsi Banten.

#### C. Tahap Analisis

Pada tahapan ini dilakukan analisis data yang telah dikumpulkan dengan *tools* RapidMiner. RapidMiner merupakan aplikasi berbasis dekstop yang bersifat terbuka (*open source*). Data yang diperoleh kemudian diolah dengan menggunakan perhitungan bobot dari tiap indeks. Pada langkah sebelumnya, telah ditentukan data akan dikelompokkan ke dalam 3 *cluster* yakni cluster tingkat pendistribusian tinggi dan cluster tingkat pendistribusian rendah. Pada tahapan inilah akan dianalisis hasilnya.

### IV. ANALISA DAN PEMBAHASAN

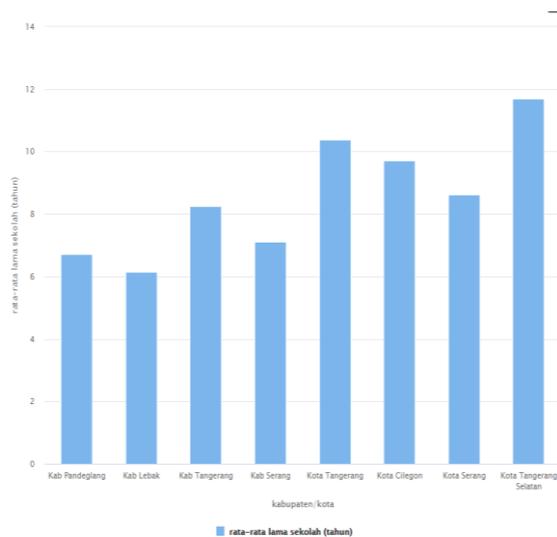
#### A. Tahap Input Data

TABEL I DATA BPS

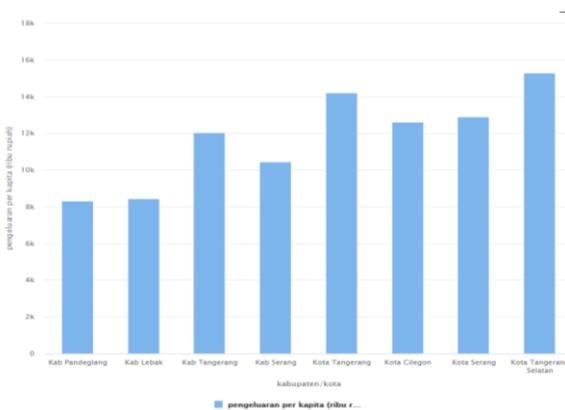
Kabupaten/Kota	Rata-Rata Lama Sekolah Menurut Kabupaten/Kota (Tahun)					Pengeluaran per Kapita yang Disesuaikan Menurut Kabupaten/Kota (ribu rupiah/tahun)					Jumlah Penduduk Miskin Menurut Kabupaten/Kota (Ribu Jiwa)				
	2015	2016	2017	2018	2019	2015	2016	2017	2018	2019	2015	2016	2017	2018	2019
Kab Pandeglang	6.6	6.62	6.63	6.72	6.96	7730	8138	8358	8613	8719	124.42	115.9	117.31	116.16	114.09
Kab Lebak	5.86	6.19	6.2	6.21	6.31	8111	8308	8372	8634	8850	126.42	111.21	111.08	108.81	107.93
Kab Tangerang	8.22	8.23	8.24	8.27	8.28	11727	11863	11914	12179	12476	191.12	182.52	191.62	190.05	193.97
Kab Serang	6.9	6.98	7.17	7.18	7.33	10004	10317	10466	10693	10802	74.85	67.92	69.1	64.46	61.54
Kota Tangerang	10.2	10.28	10.29	10.51	10.65	13766	13911	14104	14443	14860	102.56	102.88	105.34	103.49	98.37
Kota Cilegon	9.67	9.68	9.69	9.73	9.74	12127	12326	12562	12900	13230	16.96	14.9	14.89	13.96	13.2
Kota Serang	8.59	8.6	8.61	8.62	8.67	12289	12660	12914	13261	13418	40.19	36.4	36.97	36.21	36.21
Kota Tangerang Selatan	11.57	11.58	11.77	11.78	11.8	14588	14972	15291	15672	15988	25.89	26.38	28.73	28.21	29.16

TABEL II. NILAI RATA-RATA TAHUN 2015-2019

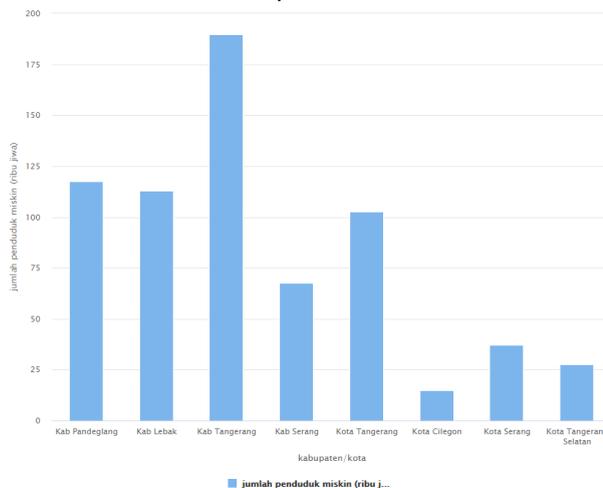
Kabupaten/Kota	Rata-Rata Lama Sekolah Menurut Kabupaten/Kota Dari 2015 sampai 2019 (Tahun)	Rata-Rata Pengeluaran per Kapita yang Disesuaikan Menurut Kabupaten/Kota Dari 2015 Sampai 2019 (ribu rupiah/tahun)	Rata-Rata Jumlah Penduduk Miskin Menurut Kabupaten/Kota Dari 2015 Sampai 2019 (Ribu Jiwa)
Kab Pandeglang	6.706	8311.6	117.576
Kab Lebak	6.154	8455	113.09
Kab Tangerang	8.248	12031.8	189.856
Kab Serang	7.112	10456.4	67.574
Kota Tangerang	10.386	14216.8	102.528
Kota Cilegon	9.702	12629	14.782
Kota Serang	8.618	12908.4	37.196
Kota Tangerang Selatan	11.7	15302.2	27.674



Gbr 1. Visualisasi Data rata-rata lama sekolah masyarakat menurut kabupaten/kota provinsi Banten dari 2015 sampai 2019



Gbr 2. Visualisasi Data rata-rata pengeluaran per kapita yang disesuaikan menurut kabupaten/kota provinsi Banten dari 2015 sampai 2019



Gbr 3. Visualisasi data rata-rata jumlah penduduk miskin menurut kabupaten/kota provinsi Banten dari 2015 sampai 2019

### B. Centroid Data

Sebelum melakukan proses *clustering* dengan menggunakan algoritma K-Means, maka yang perlu dilakukan adalah menentukan titik tengah atau *centroid data*. Dalam proses menentukan *centroid data* maka RapidMiner Studio akan otomatis menentukan nilai-nilai dari *centroid data* tersebut.

### C. Clustering data dengan menggunakan RapidMiner Studio

Hasil dari *clustering* data terkait kabupaten/kota yang memiliki jumlah penduduk miskin terbanyak dengan rata-rata lama sekolah, dan pengeluaran per kapita di provinsi Banten menggunakan RapidMiner Studio :

1. Visualisasi Data dengan menggunakan Data View. Dari data yang ada maka dapat ditampilkan data awal :

- Kabupaten Pandeglang masuk kedalam cluster 1 dengan rata-rata lama sekolah 6,706 (tahun), rata-rata pengeluaran per kapita 8311,600 (ribu rupiah/tahun), dan rata-rata jumlah penduduk miskin 117,576 (ribu jiwa).
- Kabupaten Lebak masuk kedalam cluster 1 dengan rata-rata lama sekolah 6,154 (tahun), rata-rata pengeluaran per kapita 8455 (ribu rupiah/tahun), dan rata-rata jumlah penduduk miskin 113,090 (ribu jiwa).
- Kabupaten Tangerang masuk kedalam cluster 0 dengan rata-rata lama sekolah 8,248 (tahun), rata-rata pengeluaran per kapita 12031,800 (ribu rupiah/tahun), dan rata-rata jumlah penduduk miskin 189,856 (ribu jiwa).
- Kabupaten Serang masuk kedalam cluster 1 dengan rata-rata lama sekolah 7,112 (tahun), rata-rata pengeluaran per kapita 10456,400 (ribu rupiah/tahun), dan rata-rata jumlah penduduk miskin 67,574 (ribu jiwa).
- Kota Tangerang masuk kedalam cluster 2 dengan rata-rata lama sekolah 10,386 (tahun), rata-rata pengeluaran per kapita 14216,800 (ribu rupiah/tahun), dan rata-rata jumlah penduduk miskin 102,528 (ribu jiwa).
- Kota Cilegon masuk kedalam cluster 0 dengan rata-rata lama sekolah 9,702 (tahun), rata-rata pengeluaran per kapita 12629 (ribu rupiah/tahun), dan rata-rata jumlah penduduk miskin 14,782 (ribu jiwa).
- Kota Serang masuk kedalam cluster 0 dengan rata-rata lama sekolah 8,618 (tahun), rata-rata pengeluaran per kapita 12908,400 (ribu rupiah/tahun), dan rata-rata jumlah penduduk miskin 37,196 (ribu jiwa).
- Kota Tangerang Selatan masuk kedalam cluster 2 dengan rata-rata lama sekolah 11,700 (tahun), rata-rata pengeluaran per kapita 15302,200

(ribu rupiah/tahun), dan rata-rata jumlah penduduk miskin 27,674 (ribu jiwa).

Row No.	kabupaten/kota	cluster	rata-rata lama sekolah (tahun)	pengeluaran per kapita (ribu rupiah)	jumlah penduduk miskin (ribu jiwa)
1	Kab Pandeglang	cluster_1	9.706	8311.600	117.576
2	Kab Lebak	cluster_1	6.154	8455	113.000
3	Kab Tangerang	cluster_0	8.248	12071.600	189.856
4	Kab Serang	cluster_1	7.112	18456.400	67.574
5	Kota Tangerang	cluster_2	10.386	14216.600	192.528
6	Kota Cilegon	cluster_0	9.702	12629	14.792
7	Kota Serang	cluster_0	8.618	12988.400	37.186
8	Kota Tangerang Selatan	cluster_2	11.706	15392.200	27.674

Gbr 4. Visualisasi data awal menggunakan Data View

2. Menampilkan data cluster dengan menggunakan Cluster Model. Dari data awal maka terdapat 3 cluster antara lain :

- *Cluster 0* : yang terdiri dari 3 items.
- *Cluster 1* : yang terdiri dari 3 items.
- *Cluster 2* : yang terdiri dari 2 items.



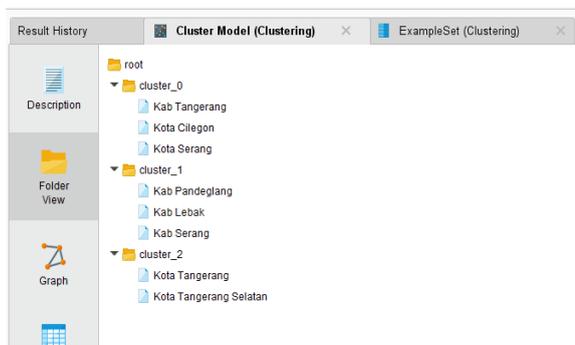
### Cluster Model

Cluster 0: 3 items  
 Cluster 1: 3 items  
 Cluster 2: 2 items  
 Total number of items: 8

Gbr 5. Tampilan data cluster menggunakan Cluster Model

3. Mengelompokkan data berdasarkan clusternya menggunakan Folder View. Dari data awal maka dapat dikelompokkan menjadi 3 cluster antara lain :

- *Cluster 0* : terdiri dari Kabupaten Tangerang, Kota Cilegon, dan Kota Serang.
- *Cluster 1* : terdiri dari Kabupaten Pandeglang, Kabupaten Lebak, dan Kabupaten Serang.
- *Cluster 2* : terdiri dari Kota Tangerang, dan Kota Tangerang Selatan.



Gbr 6. Tampilan kelompok data berdasarkan *cluster* menggunakan Folder View

4. Menampilkan *centroid* data yang digunakan dengan menggunakan *Centroid Table*. Berikut *centroid* data yang digunakan :

Attribute	cluster_0	cluster_1	cluster_2
rata-rata lama sekolah (tahun)	8.656	6.657	11.043
pengeluaran per kapita (ribu rupiah)	12523.867	8074.333	14759.530
jumlah penduduk miskin (ribu jiwa)	68.811	66.413	66.101

Gbr 7. Menampilkan *centroid* data dengan menggunakan *Centroid Table*

Berdasarkan gambar 7 diatas, maka dapat disimpulkan bahwa *centroid* data yang digunakan pada RapidMiner Studio adalah sebagai berikut :

1. *Cluster 0* : Kabupaten/kota yang masyarakatnya memiliki rata-rata waktu sekolah sedang, pengeluaran per kapita sedang, dan jumlah penduduk miskin sedang.
2. *Cluster 1* : Kabupaten/kota yang masyarakatnya memiliki rata-rata waktu sekolah sebentar, pengeluaran per kapita sedikit, jumlah penduduk miskin banyak.
3. *Cluster 2* : Kabupaten/kota yang masyarakatnya memiliki rata-rata waktu sekolah terlama, pengeluaran per kapita besar, jumlah penduduk miskin sedikit.

Hasil akhir dari proses klusterisasi dengan bantuan Rapid Miner Studio ini adalah berupa sebuah knowledge atau pengetahuan terkait kabupaten atau kota mana di Provinsi Tangerang yang memiliki jumlah penduduk miskin terbanyak, dengan didukung data-data lain seperti data rata-rata lama sekolah setiap masyarakatnya, dan pengeluaran per kapita dari setiap masyarakatnya. Dari data awal yang hanya berupa data terkait rata-rata lama sekolah, pengeluaran per kapita, dan jumlah penduduk miskin. Kemudian kita olah dengan cara menghitung rata-rata setiap data tersebut dari tahun 2015 sampai 2019. Dilanjutkan dengan proses klusterisasi sehingga menghasilkan hasil akhir berupa sebuah *knowledge* atau pengetahuan.

TABEL III  
HASIL AKHIR PENGOLAHAN TERHADAP DATA AWAL

Kabupaten atau Kota	Masyarakatnya memiliki rata-rata waktu sekolah sedang, pengeluaran per kapita sedang, dan jumlah penduduk miskin sedang	Masyarakatnya memiliki rata-rata waktu sekolah sebentar, pengeluaran per kapita sedikit, jumlah penduduk miskin banyak	Masyarakatnya memiliki rata-rata waktu sekolah terlama, pengeluaran per kapita besar, jumlah penduduk miskin sedikit
Kabupaten Pandeglang		1	
Kabupaten Lebak		1	
Kabupaten Tangerang	1		
Kabupaten Serang		1	
Tangerang			1
Kota Cilegon	1		
Kota Serang	1		
Kota Tangerang Selatan			1

Dengan bantuan software RapidMiner Studio yang digunakan selama penelitian ini, maka didapat sebuah *knowledge* sebagai berikut :

1. Kabupaten/kota yang masyarakatnya memiliki rata-rata waktu sekolah sedang, pengeluaran per kapita sedang, dan jumlah penduduk miskin sedang adalah : Kabupaten Tangerang, Kota Cilegon, dan Kota Serang.
2. Kabupaten/kota yang masyarakatnya memiliki rata-rata waktu sekolah sebentar, pengeluaran per kapita sedikit, jumlah penduduk miskin banyak adalah : Kabupaten Pandeglang, Kabupaten Lebak, dan Kabupaten Serang.
3. Kabupaten/kota yang masyarakatnya memiliki rata-rata waktu sekolah terlama, pengeluaran per kapita besar, jumlah penduduk miskin sedikit adalah : Kota Tangerang, dan Kota Tangerang Selatan.

## V. KESIMPULAN

Pada penelitian ini diperoleh hasil 3 cluster. Cluster(0) merupakan cluster dengan tingkat kemiskinan sedang karena memiliki jumlah penduduk miskin yang sedang dengan pengeluaran per kapita sedang dan rata-rata lama waktu sekolah (pendidikan) juga dalam tingkat sedang. Cluster(1) merupakan cluster dengan tingkat kemiskinan tertinggi karena memiliki jumlah penduduk miskin terbanyak dengan tingkat pendidikan yang rendah serta pengeluaran per kapitanya yang paling sedikit dibandingkan dengan cluster lain. Selanjutnya, cluster(2) merupakan cluster dengan tingkat kemiskinan terendah karena memiliki jumlah penduduk miskin yang paling sedikit dengan tingkat pendidikan tinggi serta pengeluaran per kapitanya paling besar dibandingkan dengan 2 cluster lainnya. Dalam hal ini, pemerintah dapat memprioritaskan Kabupaten Pandeglang, Kabupaten Lebak serta Kabupaten Serang dalam memberikan bantuan terutama dalam biaya beasiswa pendidikan maupun

dana sosial serta perbaikan infrastruktur lainnya demi kesejahteraan hidup penduduk Banten.

## REFERENSI

- [1] (<https://www.rmolbanten.com/read/2020/01/16/14437/Penduduk-Miskin-Banten-Turun-13-Ribu-Jiwa->, 19 Agustus 2020).
- [2] N. I. Febianto dan N. Palasara, "Analisa Clustering K-Means Pada Data Informasi Kemiskinan Di Jawa Barat Tahun 2018," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 8, no. 2, p. 130, 2019, doi: 10.32736/sisfokom.v8i2.653.
- [3] J. A. Talingdan, "Data mining using clustering algorithm as tool for poverty analysis," *ACM Int. Conf. Proceeding Ser.*, vol. Part F147956, pp. 56–59, 2019, doi: 10.1145/3316615.3316672.
- [4] Heni Sulastri, A. I. G. (2017). Penerapan Data Mining Dalam Pengelompokan Penderita Thalassaemia, 02, 299–305.
- [5] Tan, P.N., Steinbach, M., Kumar, V. (2006) Introduction to Data Mining. Boston:Pearson Education
- [6] Han, J., Kamber, M., Pei, J.: Data Mining Concept and Techniques, 3rd ed. Morgan Kaufmann-Elsevier, Amsterdam (2012)
- [7] A. P. Windarto, "Implementation of Data Mining on Rice Imports by Major Country of Origin Using Algorithm Using K-Means Clustering Method," *Int. J. Artif. Intell. Res.*, vol. 1, no. 2, pp. 26–33, 2017.
- [8] J. Informatika, W. Mega, and P. Duhita, "CLUSTERING MENGGUNAKAN METODE K-MEANS UNTUK," vol. 15, no. 2, 2015.
- [9] A. K. Wardhani, "Implementasi Algoritma K-Means untuk Pengelompokan Penyakit Pasien pada Puskesmas Kajen Pekalongan," *J. Transform.*, vol. 14, no. 1, pp. 30–37, 2016.
- [10] Z. Aras and Sarjono, "Analisis Data Mining Untuk Menentukan Kelompok Prioritas Penerima Bantuan Bedah Rumah Menggunakan Metode Clustering K-Means( Studi Kasus: Kantor Kecamatan Bahar Utara)," *J. Manaj. Sist. Inf.*, vol. 1, no. 2, pp. 159–170, 2016.
- [11] L. Elvitaria dan M. Havenda, "Memprediksi Tingkat Peminat Ekstrakurikuler pada Siswa SMK Analisis Kesehatan Abdurrah Menggunakan Algoritma C4.5 (Studi Kasus: SMK Analisis Kesehatan Abdurrah)," *RABIT(Jurnal Teknologi dan Sistem Informasi Univrab)*, vol. 2, no. 2, pp. 220-233, 2019

- [12] Sijabat, Alimancon. 2015, "Penerapan Data Mining Untuk Pengolahan Data Siswa Dengan Menggunakan Metode Decision Tree," *Jurnal Ilmiah Teknologi dan Informasi ASIA* Vol. 8 No 2, Agustus 2014.
- [13] Andriani, Anik. 2013, "Aplikasi Data Mining Market Basket Analysis Penjualan Suku Cadang Sepeda Motor Menggunakan Metode Association Rules Pada PT.Sejahtera Motor Gemilang.," *Jurnal Ilmiah Teknik Informatika Universitas Nusantara PGRI, Kediri*.