

# Analisis K-Medoids Dalam Pengelompokan Rasio Murid dengan Guru, Murid dengan Rombel, dan Rasio Rombel dengan Kelas Jenjang Pendidikan SD dan SMP Menurut Provinsi

Dewinta Marthadinata Sinaga<sup>1\*</sup>, Agus Perdana Windarto<sup>1</sup>, Dedy Hartama<sup>2</sup>

<sup>1</sup> Program Studi Sistem Informasi, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

<sup>2</sup> Program Studi Teknik Informatika, STIKOM Tunas Bangsa, Pematangsiantar, Indonesia

Email: <sup>1\*</sup>dewintamarthadinatasinaga@gmail.com, <sup>2</sup>agus.perdana@amiktunasbangsa.ac.id,

<sup>3</sup>dedyhartama@amiktunasbangsa.ac.id

**Abstrak**—Rombongan belajar merupakan tempat pertemuan antara peserta didik dan guru dalam suatu kelas pada satuan pendidikan. Sehingga, sebuah rombongan belajar dapat dikatakan sah apabila memiliki minimal 20 peserta didik dan adanya guru yang mengajar. Jumlah peserta didik dan jumlah rombongan belajar berbeda-beda pada tiap sekolah yang disesuaikan dengan jumlah peserta didik yang terdaftar disekolah tersebut. Beragamnya kondisi dan keadaan sekolah di Indonesia yang memungkinkan aturan mengenai jumlah peserta didik dan jumlah rombongan belajar tidak dapat diterapkan secara menyeluruh. Hal ini dapat menyebabkan rendahnya kualitas pendidikan di Indonesia. Sumber data yang digunakan dari situs Badan Pusat Statistik Indonesia yaitu data rasio murid dengan guru, murid dengan rombongan belajar, dan rasio rombongan belajar dengan kelas untuk tingkat sd dan tingkat smp pada tahun 2016-2017. Tujuan dari penelitian ini untuk membuat sebuah model pengelompokan menggunakan algoritma k-medoids. Metode k-medoid memiliki kesamaan dengan metode k-means yaitu sama-sama termasuk metode partitioning. Metode partitioning merupakan metode pengelompokan data kedalam sejumlah cluster tanpa adanya struktur hirarki antara satu dengan yang lainnya. Dari hasil penelitian diatas dapat disimpulkan bahwa pengelompokan rasio murid dengan guru, murid dengan rombongan belajar, dan rasio rombongan belajar dengan kelas jenjang pendidikan sd dan smp menurut provinsi dapat menerapkan metode k-medoid kmeans. Data yang diolah menjadi 2 cluster yaitu cluster tinggi dan cluster rendah. Dari hasil perhitungan dapat menyimpulkan dari 34 provinsi di Indonesia bahwa cluster tinggi (C1) diperoleh 5 provinsi yaitu Kep. Bangka Belitung, DKI Jakarta, Jawa Barat, Banten, Papua Barat dan 29 provinsi sebagai cluster rendah (C2) yaitu Aceh, Sumatera Utara, Riau, Jambi, Indonesia Selatan, Bengkulu, Lampung, Kep. Riau, Jawa Tengah, DI Yogyakarta, Jawa Timur, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, dan Papua. Proses implementasi menggunakan aplikasi RapidMiner 5.3 digunakan untuk membantu menemukan nilai yang akurat. Diharapkan penelitian ini menjadi masukan kepada pemerintah untuk mengambil kebijakan dalam meningkatkan mutu pendidikan di Indonesia pada masa yang akan datang.

**Kata Kunci:** Datamining, Clustering; K-Medoid; Rombongan Belajar; RapidMiner 5.3

**Abstract**—Study groups are a meeting place between students and teachers in a class in the education unit. So that, a study group can be said to be valid if it has at least 20 students and there are teachers who teach. The number of students and the number of study groups varies in each school that is adjusted to the number of students registered at the school. The various conditions and conditions of schools in Indonesia that allow rules regarding the number of learners and the number of study groups cannot be applied in their entirety. This can cause the low quality of education in Indonesia. Sources of data used from the Indonesian Central Statistics Agency website include data on student to teacher ratio, student with class, and class to class ratio for elementary and junior high school levels in 2016-2017. The purpose of this study is to create a grouping model using the k-medoids algorithm. The k-medoid method has similarities with the k-means method, which includes the partitioning method. Partitioning method is a method of grouping data into a number of clusters without a hierarchical structure between one another. From the results of the study above it can be concluded that the grouping of student to teacher ratios, students with rombongan belajar, and rombongan belajar ratios with elementary and junior high school education classes by province can apply the k-medoidkmeans method. The data is processed into 2 clusters, namely high and low clusters. From the results of calculations it can be concluded from 34 provinces in Indonesia that a high cluster (C1) obtained 5 provinces namely Kep. Bangka Belitung, DKI Jakarta, West Java, Banten, West Papua and 29 provinces as low clusters (C2) namely Aceh, North Sumatra, Riau, Jambi, South Indonesia, Bengkulu, Lampung, Kep. Riau, Central Java, DI Yogyakarta, East Java, Bali, West Nusa Tenggara, East Nusa Tenggara, West Kalimantan, Central Kalimantan, South Kalimantan, East Kalimantan, North Kalimantan, North Sulawesi, Central Sulawesi, South Sulawesi, Southeast Sulawesi, Gorontalo, West Sulawesi, Maluku, North Maluku, and Papua. The implementation process using the RapidMiner 5.3 application is used to help find accurate values. It is hoped that this research will provide input to the government to take policies in improving the quality of education in Indonesia in the future.

**Keywords:** Datamining; Clustering; K-Medoid; Study Group; RapidMiner 5.3

## 1. PENDAHULUAN

Rombongan belajar merupakan tempat pertemuan antara peserta didik dan guru dalam suatu kelas pada satuan pendidikan. Sehingga, sebuah rombongan belajar dapat dikatakan sah apabila memiliki minimal 20 peserta didik dan adanya guru yang mengajar. Jumlah peserta didik dan jumlah rombongan belajar berbeda-beda pada tiap sekolah yang disesuaikan dengan jumlah peserta didik yang terdaftar disekolah tersebut. Jumlah peserta didik dalam satu rombongan belajar untuk jenjang pendidikan SD dalam satu kelas berjumlah paling sedikit 20 peserta didik dan paling banyak 28 peserta didik dengan jumlah rombongan belajar paling sedikit 6 dan paling banyak

24 rombongan belajar dan masing-masing tingkat paling banyak 4 rombongan belajar sedangkan untuk jenjang pendidikan SMP berjumlah paling sedikit 20 peserta didik dan paling banyak 32 peserta didik dengan jumlah rombongan belajar paling sedikit 3 dan paling banyak 33 rombongan belajar dan masing-masing tingkat paling banyak 11 rombongan belajar [1]. Jumlah rombongan belajar menjadi salah satu indikator penting untuk menentukan jumlah jam mengajar yang harus dipenuhi oleh guru. Beragamnya kondisi dan keadaan sekolah di Indonesia yang memungkinkan aturan mengenai jumlah peserta didik dan jumlah rombongan belajar tidak dapat diterapkan secara menyeluruh. Hal ini dapat menyebabkan rendahnya kualitas pendidikan di Indonesia. Untuk itu pemerintah perlu meningkatkan kualitas pemerataan tenaga pendidik dan memaksimalkan dana pendidikan untuk kebutuhan penambahan sarana dan prasarana sekolah dan tenaga pendidik.

Berdasarkan hal tersebut, peneliti bermaksud untuk melakukan pengelompokan rasio murid dengan guru, murid dengan rombel, dan rasio rombel dengan kelas jenjang pendidikan SD dan SMP menurut provinsi dengan memanfaatkan ilmu komputer. Banyak cabang ilmu komputer yang dapat menyelesaikan permasalahan tersebut. Salah satunya adalah *Artificial Intelligence (AI)*. Diantaranya adalah jaringan saraf tiruan[2]–[4], sistem pendukung keputusan[5]–[8], datamining[9]–[11], sistem pakar[12][13], logika fuzzy[14]–[16] dan lain-lain. Dari banyaknya cabang AI, peneliti menggunakan teknik datamining dengan metode *K-Medoids*. Algoritma *K-Medoids* adalah metode *clustering* yang digunakan untuk memecah data set menjadi kelompok-kelompok. Kelebihan dari algoritma ini mampu mengatasi kelemahan dari metode *K-Means* yang sensitive terhadap *outlier*. Banyak penelitian terdahulu yang dilakukan dengan menggunakan metode *K-Means* dalam menyelesaikan permasalahan, seperti yang dilakukan oleh[17], pada penelitian tersebut dapat disimpulkan bahwa teknik datamining dapat diterapkan pada kasus pengelompokan indeks harga konsumen menurut kota. Sehingga diperoleh penilaian untuk indeks harga konsumen *cluster* tinggi (C1) sebanyak 14 kota yakni Banda Aceh, Sibolga, Pematangsiantar, Medan, Padangsidempuan, Palembang, Jakarta, Jember, Surabaya, Pontianak, Palangkaraya, Palu, Makassar, Kendari dengan centroid data 2321,2486, untuk *cluster* sedang (C2) sebanyak 29 kota yakni Lhokseumawe, Padang, Pekanbaru, Batam, Jambi, Bengkulu, Bandar Lampung, Bandung, Cirebon, Tasikmalaya, Serang, Purwokerto, Surakarta, Semarang, Tegal, Yogyakarta, Kediri, Malang, Denpasar, Mataram, Kupang, Sampit, Banjarmasin, Balikpapan, Samarinda, Manado, Ambon, Ternate, Jayapura dengan centroid data 1974, 7593 dan untuk *cluster* rendah (C3) sebanyak 23 kota yakni Dumai, Tanjung Pinang, Pangkal Pinang, Bogor, Sukabumi, Bekasi, Depok, Tangerang, Cilegon, Sumenep, Probolinggo, Madiun, Bima, Maumere, Singkawang, Tarakan, Gorontalo, Watampone, Pare-pare, Mamuju, Palopo, Manokwari, Sorong dengan centroid data 661, 0513. Tujuan dari penelitian adalah untuk menganalisa apakah teknik ilmu komputer dapat diterapkan dalam dalam pengelompokan rasio murid dengan guru, murid dengan rombel, dan rasio rombel dengan kelas jenjang pendidikan SD dan SMP menurut provinsi. Data yang digunakan dalam penelitian ini berdasarkan dokumen-dokumen yang dihasilkan oleh Badan Pusat Statistik Nasional melalui situs <https://www.bps.go.id>. Hasil dari penelitian ini dapat menjadi pertimbangan bagi pemerintah untuk mengembangkan dan meningkat mutu pendidikan di Indonesia.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahap Pengumpulan Data

Pengumpulan data dalam penelitian ini diambil berdasarkan data sekunder yang diperoleh dari Badan Pusat Statistik Nasional (BPS) dengan subjek data 2016-2017 yang terdiri dari 34 provinsi.

### 2.2 Tahap Pengolahan Data

Pengolahan data awal yang merupakan tahap untuk mempersiapkan data yang diperoleh sebelumnya dan akan digunakan pada tahap selanjutnya. Data yang sudah dikumpulkan akan diolah menjadi 2 *cluster* yaitu *cluster* tinggi dan *cluster* rendah. Kemudian data akan diuji dengan menggunakan aplikasi *RapidMiner* 5.3 untuk mencari keakuratan data.

### 2.3 Tahap Clustering

*Clustering* merupakan klasifikasi tanpa pengawasan dan merupakan proses partisisekumpulan objek data dari satu set menjadi beberapa kelas. Hal ini dapat dilakukan dengan menerapkan berbagai persamaan dan langkah-langkah mengenai jarak algoritma, yaitu dengan Euclidean Distance. Analisis kluster ialah metode yang dipakai untuk membagi rangkaian data menjadi beberapa grup berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya[18].

### 2.4 Data Mining

Data Mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran computer (machine learning) untuk menganalisis dan mengekstraksi pengetahuan (knowledge) secara otomatis[19]. Data mining mulai ada sejak 1990-an sebagai cara yang efektif untuk mengambil pola dan informasi yang sebelumnya tidak diketahui dari suatu data set. Teknik data mining digunakan untuk menemukan hubungan anatar data untuk

melakukan pengklasifikasian yang digunakan dalam klasifikasi nilai-nilai dari beberapa variabel, membagi data yang diketahui menjadi kelompok-kelompok yang mempunyai kesamaan karakteristik (*clustering*)[20].

### 2.5 Algoritma K-Medoid

Metode *K-medoid* dikembangkan oleh Leonard Kaufman dan Peter J. Rousseeuw pada tahun 1987. Algoritma *k-medoid* sering disebut juga algoritma *Partitioning Around Medoid (PAM)*. Metode *k-medoid* memiliki kesamaan dengan metode *k-means* yaitu sama-sama termasuk metode *partitioning*. Metode *partitioning* merupakan metode pengelompokkan data kedalam sejumlah *cluster* tanpa adanya struktur hirarki antara satu dengan yang lainnya. Metode *k-medoid* memiliki keunggulan dibandingkan dengan metode *k-means*. *K-medoid* memiliki kinerja yang lebih optimal jika jumlah data yang digunakan berjumlah sedikit. Algoritma ini menggunakan objek pada kumpulan objek yang mewakili sebuah *cluster*. Objek yang terpilih untuk mewakili sebuah *cluster* disebut *k-medoid*[21]. Beberapa langkah-langkah dalam perhitungan algoritma *k-medoid* adalah:

- Inisialisasi pusat *cluster* sebanyak  $k$  (jumlah *cluster*)
- Alokasikan setiap data (objek) ke *cluster* terdekat menggunakan persamaan ukuran jarak *Euclidian Distance* dengan persamaan sebagai berikut:  

$$d(x, y) = \sum_{i=1}^n (x_i - y_i)^2$$
- Pilih secara acak objek pada masing-masing *cluster* sebagai kandidat *medoid* baru
- Hitung jarak setiap objek yang berada pada masing-masing *cluster* dengan kandidat *medoid* baru
- Hitung total simpangan ( $S$ ) dengan menghitung nilai total *distance* baru – total *distance* lama. Jika  $S < 0$  maka objek dengan data *cluster* untuk membentuk sekumpulan  $k$  objek sebagai *medoid*.
- Ulangi langkah 3 sampai 5 hingga tidak terjadi perubahan *medoids*, sehingga didapatkan *cluster* beserta anggota *cluster* masing-masing.

## 3. HASIL DAN PEMBAHASAN

Pada pengelompokkan ini, data yang sudah dikumpulkan akan dihitung terlebih dahulu berdasarkan data rasio murid dengan guru, murid dengan rombel, dan rasio rombel dengan kelas jenjang pendidikan SD dan SMP menurut provinsi yang ada di Indonesia yang ditunjukkan pada Tabel 1.

**Tabel 1.** Data Rasio Murid Dengan Guru, Murid Dengan Rombel, Dan Rasio Rombel Dengan Kelas Jenjang Pendidikan SD Dan SMP Menurut Provinsi Tahun 2016-2017

| No | Provinsi             | SD                 |                      |                      | SMP                |                      |                      |
|----|----------------------|--------------------|----------------------|----------------------|--------------------|----------------------|----------------------|
|    |                      | Rasio Murid - Guru | Rasio Murid - Rombel | Rasio Rombel - Kelas | Rasio Murid - Guru | Rasio Murid - Rombel | Rasio Rombel - Kelas |
| 1  | Aceh                 | 10                 | 20                   | 0,99                 | 9                  | 24                   | 0,90                 |
| 2  | Sumatra Utara        | 17                 | 24                   | 1,04                 | 17                 | 31                   | 0,97                 |
| 3  | Sumatra Barat        | 15                 | 21                   | 1,01                 | 12                 | 27                   | 0,95                 |
| 4  | Riau                 | 16                 | 24                   | 1,10                 | 15                 | 28                   | 1,00                 |
| 5  | Jambi                | 14                 | 21                   | 1,05                 | 13                 | 26                   | 0,96                 |
| 6  | Sumatera Selatan     | 16                 | 24                   | 1,14                 | 15                 | 31                   | 1,02                 |
| 7  | Bengkulu             | 15                 | 21                   | 1,02                 | 13                 | 26                   | 0,96                 |
| 8  | Lampung              | 15                 | 23                   | 1,07                 | 15                 | 29                   | 0,97                 |
| 9  | Kep. Bangka Belitung | 19                 | 25                   | 0,96                 | 20                 | 31                   | 0,95                 |
| 10 | Kep. Riau            | 18                 | 25                   | 1,18                 | 18                 | 30                   | 1,01                 |
| 11 | Lainnya              | 23                 | 29                   | 1,12                 | 21                 | 31                   | 1,03                 |

Data tersebut kemudian diakumulasikan berdasarkan 2 kriteria yaitu tingkat SD dan tingkat SMP seperti pada tabel 2 berikut.

**Tabel 2.** Nilai rata-rata setelah data alternatif diakumulasikan

| No | Provinsi         | SD      | SMP     |
|----|------------------|---------|---------|
| 1  | Aceh             | 10,3300 | 11,3000 |
| 2  | Sumatra Utara    | 14,0133 | 16,3233 |
| 3  | Sumatra Barat    | 12,3367 | 13,3167 |
| 4  | Riau             | 13,7000 | 14,6667 |
| 5  | Jambi            | 12,0167 | 13,3200 |
| 6  | Sumatera Selatan | 13,7133 | 15,6733 |
| 7  | Bengkulu         | 12,3400 | 13,3200 |
| 8  | Lampung          | 13,0233 | 14,9900 |

| No | Provinsi             | SD      | SMP     |
|----|----------------------|---------|---------|
| 9  | Kep. Bangka Belitung | 14,9867 | 17,3167 |
| 10 | Kep. Riau            | 14,7267 | 16,3367 |
| 11 | Lainnya              | 17,7067 | 17,6767 |

### 3.1 Centroid Data

Dalam penerapan algoritma K-Medoid nilai centroid dapat ditentukan secara acak dari data yang didapat dengan ketentuan bahwa pengelompokan yang diinginkan adalah 2 yakni *cluster* tingkat tinggi (C1) dan *cluster* tingkat rendah (C2), maka nilai titik tengah atau centroid juga terdapat dua titik. Penentuan titik *cluster* ini dilakukan dengan mengambil nilai secara acak. Data awal akan diolah menggunakan rumus Euclidian Distance dengan centroid awal yang telah ditentukan. Berikut ini adalah centroid awal yang digunakan.

**Tabel 2.** Data Centroid Awal

|    |                |         |         |
|----|----------------|---------|---------|
| C1 | Kep Riau       | 14,7267 | 16,3367 |
| C2 | Sulawesi Utara | 10,9967 | 13,3267 |

adapun hasilnya yang ditunjukkan pada Tabel 3.

**Tabel 3.** Data setelah diolah

| No | Provinsi             | SD      | SMP     | C1     | C2     | Jarak Terdekat | Klaster |
|----|----------------------|---------|---------|--------|--------|----------------|---------|
| 1  | Aceh                 | 10,3300 | 11,3000 | 9,7508 | 4,0631 | 4,0631         | C2      |
| 2  | Sumatra Utara        | 14,0133 | 16,3233 | 3,9335 | 2,4094 | 2,4094         | C2      |
| 3  | Sumatra Barat        | 12,3367 | 13,3167 | 6,9172 | 1,2351 | 1,2351         | C2      |
| 4  | Riau                 | 13,7000 | 14,6667 | 5,0114 | 0,7439 | 0,7439         | C2      |
| 5  | Jambi                | 12,0167 | 13,3200 | 7,1664 | 1,5101 | 1,5101         | C2      |
| 6  | Sumatera Selatan     | 13,7133 | 15,6733 | 4,4677 | 1,7063 | 1,7063         | C2      |
| 7  | Bengkulu             | 12,3400 | 13,3200 | 6,9125 | 1,2305 | 1,2305         | C2      |
| 8  | Lampung              | 13,0233 | 14,9900 | 5,3993 | 1,0436 | 1,0436         | C2      |
| 9  | Kep. Bangka Belitung | 14,9867 | 17,3167 | 2,7438 | 3,6897 | 2,7438         | C1      |
| 10 | Kep. Riau            | 14,7267 | 16,3367 | 3,2675 | 2,7025 | 2,7025         | C2      |
| 11 | Lainnya              | 17,7067 | 17,6767 | 0,0000 | 5,6884 | 0,0000         | C1      |

### 3.2 Clustering

Dengan menggunakan centroid tersebut maka dapat di *cluster* data yang telah didapat menjadi 2 *cluster*. Proses *cluster* dengan mengambil jarak terdekat dari setiap data yang diolah. Dari data awal didapatkan pengelompokan pada iterasi 1 untuk 2 *cluster* tersebut. Pada iterasi awal pengelompokan yaitu 2 *cluster* yang ditunjukkan pada Tabel 4.

**Tabel 4.** Hasil *Cluster* Iterasi Pertama

| No | Klaster 1            | Klaster 2     |
|----|----------------------|---------------|
| 1  | Sumatra Utara        | Aceh          |
| 2  | Riau                 | Sumatra Barat |
| 3  | Sumatera Selatan     | Jambi         |
| 4  | Lampung              | Bengkulu      |
| 5  | Kep. Bangka Belitung | DI Yogyakarta |
| 6  | Kep. Riau            | Jawa Timur    |
| 7  | DKI Jakarta          |               |
| 8  | Jawa Barat           |               |
| 9  | Jawa Tengah          |               |
| 10 | Banten               |               |
| 11 | Lainnya              |               |

Proses *clustering* berhenti jika Simpangan > 0, tetapi jika Simpangan < 0 maka tukar objek dengan data *cluster* untuk membentuk sekumpulan k objek baru sebagai medoid. Proses perhitungan simpangan dengan cara menjumlahkan nilai total distance baru – total distance lama yang berisikan jarak setiap objek yang berada pada masing-masing *cluster* dengan kandidat medoid baru. Dengan jumlah simpangan yang ditunjukkan pada Tabel 5.

**Tabel 5.** Jumlah *Distance* Iterasi Pertama

| No | Klaster 1 | Klaster 2 |
|----|-----------|-----------|
| 1  | 0,7135    | 2,1335    |
| 2  | 1,9604    | 1,34      |
| 3  | 1,2112    | 1,02      |

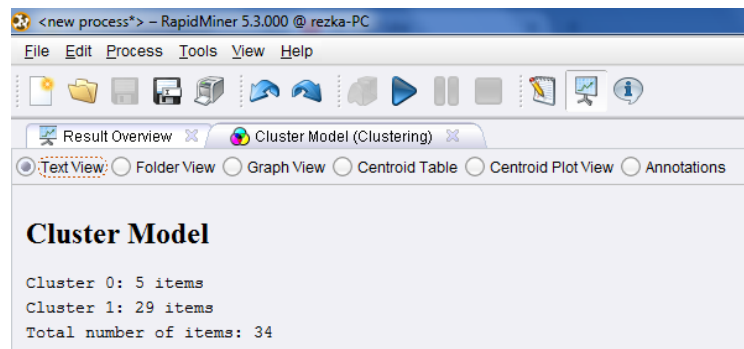
| No | Klaster 1 | Klaster 2   |
|----|-----------|-------------|
| 4  | 2,1714    | 1,3433      |
| 5  | 1,0139    | 2,1229      |
| 6  | 0         | 2,2479      |
| 7  | 3,2674    | 1,5042      |
| 8  | 3,0789    | 2,2406      |
| 9  | 1,4343    | 1,2011      |
| 10 | 3,0612    | 0,0138      |
| 11 | Lainnya   | 1,2527      |
|    | 6,0565    | 15,1673     |
|    |           | 51,51474871 |

Kemudian hasil *cluster* sampai pada iterasi kedua yang ditunjukkan pada tabel 6 berikut:

**Tabel 6.** Jumlah *Distance* Iterasi Kedua

| No | Klaster 1 | Klaster 2 |
|----|-----------|-----------|
| 1  | 2,7438    | 4,0631    |
| 2  | 0,0000    | 2,4094    |
| 3  | 1,4096    | 1,2351    |
| 4  | 0,9263    | 0,7439    |
| 5  | 1,7018    | 1,5101    |
| 6  | Lainnya   | 1,7063    |
|    | 6,7815    | 53,7431   |
|    |           | 60,5246   |

Dari hasil penjumlahan data diatas distance baru – distance lama menghasilkan simpangan  $< 0$ , maka proses pengelompokkan berhenti pada ietrasi kedua dengan hasil simpangan 9,0098. Berdasarkan dari penjelasan tersebut dapat disimpulkan bahwa data yang ada bersifat valid. Hal ini dibuktikan dengan hasil akhir perhitungan manual dan dengan menggunakan aplikasi RapidMiner 5.3 dapat menampilkan hasil yang sama. Berikut adalah tampilan *cluster* model berupa teks yang ada dalam aplikasi RapidMiner 5.3 dari *cluster* yang terbentuk:



**Gambar 1.** *Cluster Model*

## 4. KESIMPULAN

Dari hasil penelitian di atas dapat disimpulkan bahwa pengelompokkan rasio murid dengan guru, murid dengan rombongan belajar, dan rasio rombongan belajar dengan kelas jenjang pendidikan SD dan SMP menurut provinsi dapat menerapkan metode k-medoid/kmeans. Data yang diolah menjadi 2 *cluster* yaitu *cluster* tinggi dan *cluster* rendah. Dari hasil perhitungan dapat menyimpulkan dari 34 provinsi di Indonesia bahwa *cluster* tinggi (C1) diperoleh 5 provinsi yaitu Kep. Bangka Belitung, DKI Jakarta, Jawa Barat, Banten, Papua Barat dan 29 provinsi sebagai *cluster* rendah (C2) yaitu Aceh, Sumatera Utara, Riau, Jambi, Indonesia Selatan, Bengkulu, Lampung, Kep. Riau, Jawa Tengah, DI Yogyakarta, Jawa Timur, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Selatan, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, dan Papua.

## REFERENCES

- [1] Kemendikbud, "Peraturan Menteri Pendidikan Dan Kebudayaan Republik Indonesia Nomor 17 Tahun 2017:

- Penerimaan peserta didik baru pada taman kanak-kanak, sekolah dasar, sekolah menengah pertama, sekolah menengah atas, sekolah menengah kejuruan, atau bentuk lain,” *Kementerian Pendidikan dan Kebudayaan*, pp. 1–20, 2017.
- [2] E. Hartato, D. Sitorus, and A. Wanto, “Analisis jaringan saraf tiruan untuk prediksi luas panen biofarmaka di indonesia,” *semanTIK*, vol. 4, no. 1, pp. 49–56, 2018.
  - [3] B. K. Sihotang, A. Wanto, T. Informatika, P. Studi, and T. Informatika, “Analisis Jaringan Syaraf Tiruan Dalam Memprediksi Jumlah Tamu Pada Hotel Non Bintang,” vol. 17, no. 4, pp. 333–346, 2018.
  - [4] A. A. Fardhani, D. I. N. Simanjuntak, and A. Wanto, “Prediksi Harga Eceran Beras Di Pasar Tradisional Di 33 Kota Di Indonesia Menggunakan Algoritma Backpropagation,” *Jurnal Infomedia*, vol. 3, no. 1, 2018.
  - [5] K. Terbaik and B. Konsumen, “Jurnal Sains dan Informatika,” *Jurnal Sains dan Informatika*, vol. 4, no. 1, pp. 129–135, 2018.
  - [6] L. Marlinda, “SISTEM PENDUKUNG KEPUTUSAN PEMILIHAN TEMPAT WISATA YOGYAKARTA MENGGUNAKAN METODE ELIMINATION ET CHOIX TRADUISAN LA REALIT  (ELECTRE),” *Jurnal.Umj.Ac.Id/Index.Php/Semnastek*, no. November, pp. 1–7, 2016.
  - [7] F. Syahputra, M. Mesran, I. Lubis, and A. P. Windarto, “Sistem Pendukung Keputusan Pemilihan Guru Berprestasi Kota Medan Menerapkan Metode Preferences Selection Index (Studi Kasus : Dinas Pendidikan Kota Medan),” *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, vol. 2, no. 1, pp. 147–155, 2018.
  - [8] D. R. Sari, A. P. Windarto, D. Hartama, and S. Solikhun, “Sistem Pendukung Keputusan untuk Rekomendasi Kelulusan Sidang Skripsi Menggunakan Metode AHP-TOPSIS,” *Jurnal Teknologi dan Sistem Komputer*, vol. 6, no. 1, p. 1, 2018.
  - [9] H. Sulastrı and A. I. Gufroni, “Penerapan Data Mining Dalam Pengelompokan Penderita Thalassaemia,” *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 3, no. 2, pp. 299–305, 2017.
  - [10] M. G. Sadewo, A. Eriza, A. P. Windarto, and D. Hartama, “Algoritma K-Means Dalam Mengelompokkan Desa / Kelurahan Menurut Keberadaan Keluarga Pengguna Listrik dan Sumber Penerangan Jalan Utama Berdasarkan Provinsi,” pp. 754–761, 2019.
  - [11] A. P. Windarto, “Penerapan Datamining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering Method,” *Techno.Com*, vol. 16, no. 4, pp. 348–357, 2017.
  - [12] A. H. Aji, M. T. Furqon, and A. W. Widodo, “Sistem Pakar Diagnosa Penyakit Ibu Hamil Menggunakan Metode Certainty Factor ( CF ),” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 2, no. 5, pp. 2127–2134, 2018.
  - [13] M. Turnip, “Sistem Pakar Diagnosa Penyakit THT Menggunakan Metode Backward Chaining,” *Riau Journal Of Computer Science*, vol. 1, no. 1, pp. 1–8, 2015.
  - [14] M. Sugeno and T. Takagi, “[15] Fuzzy identification of systems and its applications to modeling and control,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 15, no. 1, pp. 116–132, 1985.
  - [15] W. Munandar *et al.*, “Metode Fuzzy Multiple Attribute Decision Making ( Fmadm ),” no. 03, pp. 45–58.
  - [16] I. Tri, R. Yanto, M. Akmar, and T. Herawan, “Engineering Applications of Artificial Intelligence A modified Fuzzy k -Partition based on indiscernibility relation for categorical data clustering,” *Engineering Applications of Artificial Intelligence*, vol. 53, pp. 41–52, 2016.
  - [17] D. M. Sinaga, A. P. Windarto, and D. Hartama, “Pengelompokkan Indeks Harga Konsumen Menurut Kota Dengan Datamining Clustering,” pp. 328–336, 2019.
  - [18] M. G. Sadewo, A. P. Windarto, and D. Hartama, “PENERAPAN DATAMINING PADA POPULASI DAGING AYAM RAS PEDAGING DI INDONESIA BERDASARKAN PROVINSI MENGGUNAKAN K-MEANS,” *Nasional Informatika dan Teknologi Jaringan*, vol. 2, no. 1, pp. 60–67, 2016.
  - [19] E. D. Sikumbang, “Penerapan Data Mining Penjualan Sepatu Menggunakan Metode Algoritma Apriori,” *Jurnal Teknik Komputer AMIK BSI (JTK)*, vol. Vol 4, No., no. September, pp. 1–4, 2018.
  - [20] Z. Mustofa and I. S. Suasana, “Algoritma Clustering K-Medoids pada E-Government Bidang Information and Communication Teknologi Dalam Penentuan Status Edgi,” *Jurnal Teknologi Informasi dan Komunikasi*, vol. 9, no. 1, pp. 1–10, 2018.
  - [21] S. R. Ningsih, I. S. Damanik, A. P. Windarto, H. S. Tambunan, J. Jalaluddin, and A. Wanto, “Analisis K-Medoids Dalam Pengelompokkan Penduduk Buta Huruf Menurut Provinsi,” *Prosiding Seminar Nasional Riset Information Science (SENARIS)*, vol. 1, no. September, p. 721, 2019.