

Penerapan Algoritma K-Medoids Dalam Menentukan Cluster Kabupaten/Kota Berdasarkan Migrasi Penduduk Jawa Barat

Bahrul Ulum

STMIK IKMI Cirebon

Edi Tohidi

STMIK IKMI Cirebon

Nisa Dienwati Nuris

STMIK IKMI Cirebon

Alamat: Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45135

Korespondensi penulis: ulumb353@gmail.com

Abstract. *One of the three main factors influencing population dynamics is migration, along with births and deaths. Changes in population structure will definitely be influenced by migration. If in-migration is more than out-migration, the population will increase, but if out-migration is less than in-migration, the population will decrease. Therefore, it is necessary to know the grouping of regions based on population migration levels. To carry out this grouping, it is necessary to use Data mining methods. In this research, the Data mining used is Clustering using the K-Medoids algorithm. This method divides each district into predetermined groups. The K-Medoids method was chosen because it uses physical data that is not abstract and clear, which is suitable for the problem of grouping population migration data. By grouping migration levels based on districts/cities in West Java, it will be known which districts/cities in West Java have high levels of incoming migration, medium migration and high outmigration. Then recommendations can be given to the local government according to the migration level category.*

Keywords: *Migration, Data mining, Clustering, K-Medoids*

Abstrak. Satu dari tiga faktor utama yang mempengaruhi dinamika kependudukan adalah migrasi, bersama dengan kelahiran dan kematian. Perubahan stuktur penduduk pasti akan dipengaruhi oleh migrasi. Jika migrasi masuk lebih banyak daripada migrasi keluar, jumlah penduduk akan meningkat, tetapi jika migrasi keluar lebih sedikit daripada migrasi masuk, jumlah penduduk akan menurun. Oleh karena itu, perlu mengetahui pengelompokan wilayah-wilayah berdasarkan tingkat migrasi penduduk. Untuk melakukan pengelompokan tersebut maka perlu dilakukan dengan memakai metode Data mining. Pada penelitian ini Data mining yang digunakan yaitu Clustering dengan menggunakan algoritma K-Medoids. Metode ini membagi tiap kabupaten ke dalam kelompok yang telah ditetapkan. Metode K-Medoids dipilih karena menggunakan data fisik yang tidak abstrak dan jelas, yang cocok untuk masalah pengelompokan data migrasi penduduk. Dengan mengelompokkan tingkat migrasi berdasarkan Kabupaten/Kota di Jawa Barat, maka akan di ketahui Kabupaten/Kota mana saja di Jawa Barat yang memiliki Tingkat migrasi masuk tinggi, migrasi sedang, dan migrasi keluar tinggi. Kemudian dapat diberi rekomendasi untuk pemerintah daerah setempat sesuai kategori Tingkat migrasinya.

Kata kunci: Migrasi, Data mining, Clustering, K-Medoids

LATAR BELAKANG

Satu dari tiga faktor utama yang mempengaruhi dinamika kependudukan adalah migrasi, bersama dengan kelahiran dan kematian. Perpindahan penduduk yang bertujuan untuk menetap dari satu tempat ke tempat lain disebut migrasi. Perubahan struktur penduduk pasti akan dipengaruhi oleh migrasi. Jika migrasi masuk lebih banyak daripada migrasi keluar, jumlah penduduk akan meningkat, tetapi jika migrasi keluar lebih sedikit daripada migrasi masuk, jumlah penduduk akan menurun. Dalam kasus seperti itu, berbagai kemungkinan buruk dapat terjadi jika tidak dikendalikan segera.

KAJIAN TEORITIS

A. Data mining

Data Mining adalah bidang ilmu yang mempelajari cara mengekstrak pengetahuan atau menemukan pola dalam data. Data mining merupakan kata lain yang digunakan untuk menggambarkan proses penggalian nilai dari database yang sangat besar, yang melibatkan penggunaan teknik otomatisasi untuk mempelajari dan mengidentifikasi hubungan yang kompleks dalam kumpulan data yang sangat besar, seperti data tabulasi pada teknologi manajemen basis data relasional atau domain data spasial, berbasis teks, dan multimedia (Amelia dkk., 2023).

B. Clustering

Clustering adalah teknik pengelompokan data dalam Data mining. Tujuan dari Clustering adalah untuk mengelompokkan Cluster objek, titik, dokumen, dan pola. Perbandingan antara kumpulan objek yang memiliki persamaan dengan kelompok lain. Karena hasil validasi yang lebih baik daripada algoritma k-means, maka dalam penelitian ini algoritma K-Medoids digunakan dalam pengelompokan data migrasi masuk dan keluar Jawa Barat tahun 2022 (Widiawati dkk., 2022).

C. K-Medoids

Merupakan Varian dari K-Means yang lebih kuat terhadap kebisingan dan outlier. K-Medoids merupakan metode pengelompokan partisi yang menggunakan titik aktual dalam kluster untuk mewakilinya. Objek yang mewakili sebuah kluster disebut dengan Medoids. Medoids merupakan objek yang letaknya terpusat di dalam suatu kluster dengan jarak minimum ke titik lain sehingga robust terhadap outlier. Kluster dibangun dengan

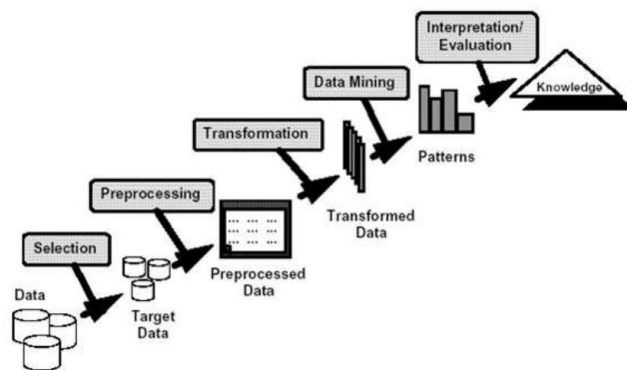
menghitung kedekatan yang dimiliki antara Medoids dengan objek non Medoids (Sari & Ediwijoyo, 2021).

D. MIGRASI

Satu dari tiga faktor utama yang mempengaruhi dinamika kependudukan adalah migrasi, bersama dengan kelahiran dan kematian. Perpindahan penduduk yang bertujuan untuk menetap dari satu tempat ke tempat lain disebut migrasi. Perubahan struktur penduduk pasti akan dipengaruhi oleh migrasi. Jika migrasi masuk lebih banyak daripada migrasi keluar, jumlah penduduk akan meningkat, tetapi jika migrasi keluar lebih sedikit daripada migrasi masuk, jumlah penduduk akan menurun. Dalam kasus seperti itu, berbagai kemungkinan buruk dapat terjadi jika tidak dikendalikan segera (Afidah & Masrukan, 2023).

METODE PENELITIAN

Pada penelitian ini metode yang digunakan adalah metode *Knowledge Discovery Process* (KDD) sebagai proses dimana teknik Data mining digunakan guna mengidentifikasi informasi serta pola yang relevan dalam sekumpulan data, proses Data mining ini melibatkan penggunaan algoritma untuk menemukan pola yang dapat mengidentifikasi dalam suatu proses database.



Gambar 1 Tahapan KDD

1. Data Selection

Data Selection adalah proses pemilihan (sampling) data dari suatu dataset yang harus dilakukan sebelum data mining di KDD. Data pemrosesan pesan disimpan dalam file terpisah dari database aplikasi. pemilihan atribut yang sesuai dengan apa yang dibutuhkan seperti jumlah pen, tahun, dan atribut lainnya yang relevan.

2. Data Preprocessing

Data preprocessing adalah teknik penambangan data. Sebelum tahap konstruksi. Pertama, data mentah diproses. Preprocessing data atau data preprocessing dilakukan dengan menghilangkan pesan error.

3. Transformation

Transformasi data digunakan dalam pemrosesan data untuk mengubah data menjadi format yang sesuai. Banyak teknik transformasi data termasuk normalisasi, pemilihan atribut, dan diskritisasi.

4. Data Mining

Sebuah proses di mana berbagai teknik digunakan untuk mengekstraksi pola potensial dari data yang berguna.

5. Interpretation/Evaluation

Langkah ini melibatkan pemeriksaan apakah model atau data yang dihasilkan konsisten dengan informasi atau asumsi sebelumnya.

HASIL DAN PEMBAHASAN

A. Hasil

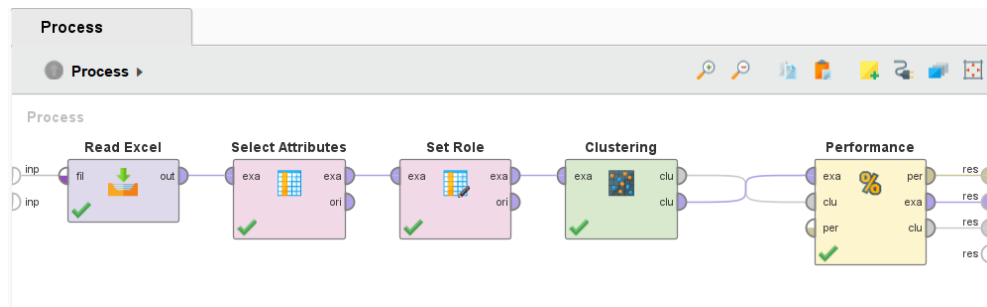
Pada bagian ini akan membahas tahapan-tahapan hasil yang diperoleh untuk pengklasteran Kabupaten/Kota berdasarkan migrasi masuk dan keluar penduduk di Jawa Barat, dengan menggunakan metode *Cluster K-Medoids*.

a) Dataset

Dataset yang digunakan dalam penelitian ini adalah data Migrasi Masuk dan Keluar di Jawa Barat pada tahun 2022. Data ini diperoleh dengan mengunjungi situs web resmi Open Data Jabar dan mengidentifikasi sumber-sumber data sekunder yang relevan dengan topik atau penelitian.

Tabel 4. 1 Data Migrasi Masuk dan Keluar

No.	Kode Kabupaten/Kota	Nama Kabupaten/Kota	Migrasi Masuk	Migrasi Keluar
1	3201	Kabupaten Bogor	28934	25692
2	3202	Kabupaten Sukabumi	13615	16138
3	3203	Kabupaten Cianjur	12449	16245
4	3204	Kabupaten Bandung	34389	31333
5	3205	Kabupaten Garut	15750	18906
6	3206	Kabupaten Tasikmalaya	12980	17325
7	3207	Kabupaten Ciamis	9948	10240
8	3208	Kabupaten Kuningan	5835	6292
9	3209	Kabupaten Cirebon	10416	11805
10	3210	Kabupaten Majalengka	7411	7283
26	3278	Kota Tasikmalaya	9773	7673
27	3279	Kota Banjar	2243	2427

b) *Data Selection*

Gambar 4. 1 Design Process RapidMiner

Data Selection merupakan tahapan memilih data yang akan diolah untuk penelitian. Proses ini dilakukan untuk memilih atribut- atribut dari data yang akan digunakan dalam proses Data mining . Tahap ini dilakukan proses Select Attributes, atribut yang digunakan dalam pengolahan data terdapat 3 atribut yaitu: Kabupaten/Kota, migrasi masuk, dan migrasi keluar.

c) *Data Preprocessing*

Di *RapidMiner*, preprocessing digunakan untuk menghapus data yang tidak konsisten dan rusak menjadi data yang nantinya dapat digunakan dalam pemrosesan data. Data cleaning merupakan tindakan menghilangkan atribut yang tidak akan digunakan dan menghapus kolom yang berisi nilai null yang tidak relevan dalam perhitungan selanjutnya. Dengan demikian, data yang akan diolah menjadi data yang sepenuhnya relevan. Ini mungkin melibatkan pemrosesan awal seperti pengisian data yang hilang atau penghapusan entri yang tidak valid

d) *Transformation*

Transformasi data adalah proses untuk merubah *type* atribut data akan diubah menjadi bentuk yang sesuai untuk proses Data mining . Karena *type* atribut data dalam penelitian ini sudah sesuai untuk proses Data mining dan data sudah dalam berbentuk null atau tetap berbentuk numerik maka tahap *Transformation* dilakukan ketahap selanjutnya .

e) *Proses Clustering*

Tahap pertama yang perlu dilakukan adalah mengkonfigurasi Operator apa saja yang akan digunakan dalam *RapidMiner*. operator yang digunakan untuk pengelompokan Kabupaten/Kota berdasarkan migrasi masuk dan keluar di Jawa Barat

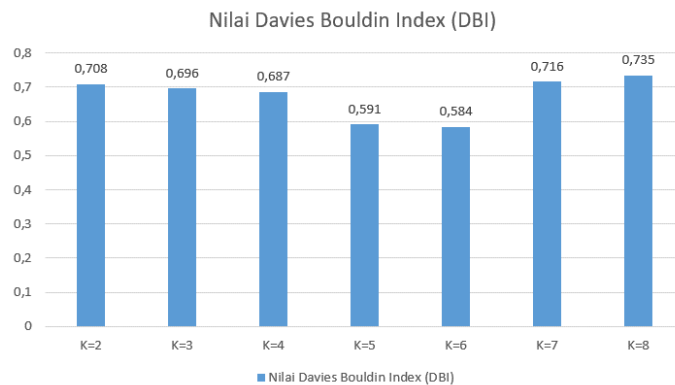
adalah operator *Read Excel*, *Select Attributes*, *Set Role*, operator *Clustering (K-Medoids)* dan *performance*.

Tahap kedua yaitu menjalankan semua operator sesuai fungsinya. Operator pertama yang dijalankan yaitu *Read Excel* yang berfungsi untuk membaca data utama *file excel*, selanjutnya Operator *Select Attributes* yang berfungsi untuk memilih subset dari *ExampleSet* dan menghapus atribut lainnya, kemudian operator *Set Role* yang berfungsi untuk menentukan peran atau jenis data dari suatu atribut dalam *dataset*, operator selanjutnya yaitu *Clustering (K-Medoids)* yang berfungsi untuk mengurutkan data menjadi kelompok (K) yang telah ditentukan, operator terakhir yaitu *Performance* yang bekerja untuk mengevaluasi kinerja model, memberikan daftar tujuan kinerja sesuai dengan tugas yang diberikan. Setelah semua operator siap dijalankan, dilakukanlah proses percobaan untuk menemukan *Cluster (K)* terbaik. Pada percobaan ini *cluster (K)* dimulai dari K=2 sampai K=8. Untuk K=2 menghasilkan *Cluster0* berjumlah 11 items dan *Cluster1* berjumlah 16 items. Untuk K=3 didapatkan *Cluster0* berjumlah 14 items, *Cluster1* berjumlah 10 items dan *Cluster2* berjumlah 3 items. Untuk K=4 didapatkan hasil *Cluster0* berjumlah 7 items, *Cluster1* berjumlah 14 items, *Cluster2* berjumlah 3 items dan *Cluster3* berjumlah 3 items. Selanjutnya untuk K=5 didapatkan hasil yaitu *Cluster0* berjumlah 7 items, *Cluster1* berjumlah 1 items, *Cluster2* berjumlah 13 items, *cluster3* berjumlah 3 items dan *Cluster4* berjumlah 3 items. Kemudian K=6 menghasilkan *Cluster0* berjumlah 3 items, *Cluster1* berjumlah 3 items, *Cluster2* berjumlah 4 items, *Cluster3* berjumlah 1 items, *Cluster4* berjumlah 13 items dan *Cluster5* berjumlah 3 items. Selanjutnya untuk percobaan menggunakan K=7 menghasilkan *Cluster0* berjumlah 6 items, *Cluster1* berjumlah 4 items, *Cluster2* berjumlah 3 items, *Cluster3* berjumlah 3 items, *Cluster4* berjumlah 1 items, *Cluster5* berjumlah 7 items dan *Cluster6* berjumlah 3 items. Percobaan terakhir yaitu menggunakan K=8 dengan hasil *Cluster0* berjumlah 3 items, *Cluster1* berjumlah 6 items, *Cluster2* berjumlah 4 items.

f) Evaluasi DBI

Davies Bouldin Indeks(DBI) adalah metode internal yang digunakan dalam evaluasi klaster. Keandalan internal DBI mengukur efektivitas pengelompokan dengan mengukur kuantitas dan kualitas data yang dikumpulkan. Setelah mempersiapkan data, lanjutkan ke langkah berikutnya, pengolahan data. Data diolah menggunakan algoritma *K-Medoids*. Pada penelitian ini jumlah *cluster* ditetapkan dari *Cluster 2* sampai 8.

Terakhir, jumlah *cluster* yang digunakan dalam penelitian ini ditentukan berdasarkan hasil evaluasi. Analisis data yang digunakan dalam penelitian ini menggunakan *RapidMiner* sebagai perangkat lunak untuk melakukan proses analisis. Hasil DBI menggunakan metode *Clustering* dan algoritma *K-Medoids* pada aplikasi *RapidMiner* dapat dilihat pada Gambar berikut :



Gambar 4. 2 Hasil Nilai DBI

B. Pembahasan

a) Penerapkan Algoritma *K-Medoids* Untuk Menentukan Klaster Kabupaten/Kota Di Jawa Barat Berdasarkan Pola Migrasi Penduduk

Pada penelitian ini dilakukan penerapan algoritma *K-Medoids* dalam mengelompokkan Kabupaten/Kota berdasarkan migrasi masuk dan keluar di Jawa Barat. Metode analisa yang digunakan adalah KDD (*Knowledge Discovery Database*) dan *tools* yang digunakan adalah *RapidMiner*. Tahapan analisa terdiri dari beberapa tahap yaitu:

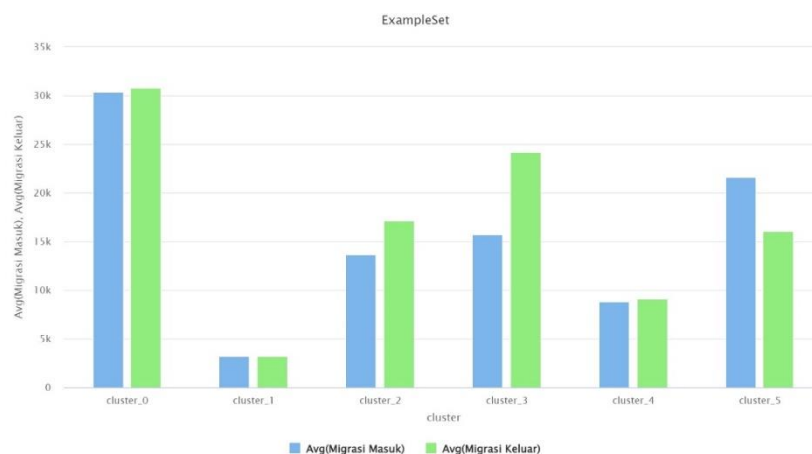
Data Selection , dengan melakukan pemilihan atribut yang sesuai dengan apa yang dibutuhkan seperti nama Kabupaten/Kota, jumlah migrasi masuk dan jumlah migrasi keluar. Selanjutnya tahap, *Preprocessing*, yaitu proses cleaning data yang hilang atau kosong, dan terdapat tiga atribut data yang dihilangkan yaitu nomor dan kode Kabupaten/Kota dikarenakan data tersebut berisi nilai null. Kemudian di tahap *Data Transformation* tidak perlu dilakukan dalam penelitian ini dikarenakan data yang digunakan untuk pengolahan data sudah berbentuk format yang sesuai untuk di proses ke Data mining . Pada tahap Data mining diterapkan algoritma *K-Medoids* untuk mengelompokkan Kabupaten/Kota berdasarkan migrasi masuk dan keluar di Jawa Barat. Selanjutnya dengan melakukan evaluasi dengan *Davies Bouldin Index (DBI)* untuk menentukan jumlah *cluster* yang optimal setelah proses pengelompokan selesai.

b) Evaluasi Algoritma K-Medoids Untuk Menentukan Klaster Kabupaten/Kota Di Jawa Barat Berdasarkan Pola Migrasi Penduduk

Hasil dari proses pengelompokan kabupaten/kota dengan menggunakan algoritma K-Medoids dengan menentukan nilai DBI dari masing masing klaster, dan berdasarkan hasil rekapitulasi *Davies Bouldin Index* (DBI) yang ditunjukkan dengan hasil dari nilai $k=2$ sampai $k=8$, Maka nilai k yang akan dipilih dan menjadi jumlah *cluster* adalah pada percobaan kelima dengan nilai $k=6$ karena memiliki nilai DBI paling kecil yaitu 0.548, karena hasil pengelompokan yang lebih baik diperoleh dengan nilai DBI yang lebih rendah (Hidayat & Fitriana, 2022). Sehingga pada penelitian ini *cluster* yang digunakan sebanyak 6 *cluster*.

c) Rekomendasi Terkait Migrasi Penduduk Di Jawa Barat Berdasarkan Hasil Clustering

Dari hasil pengujian pengelompokan kabupaten/kota berdasarkan jumlah migrasi masuk dan keluar di Jawa Barat menggunakan algoritma k-Medoids, dapat diketahui daerah yang masuk kedalam kategori Migrasi masuk tinggi, Migrasi Sedang dan Migrasi keluar tinggi.

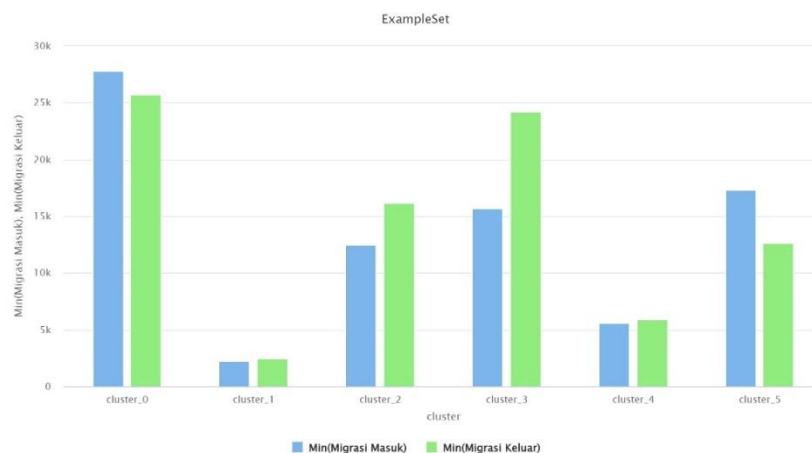


Gambar 4. 3 Migrasi berdasarkan Rata-rata (Average)

Berdasarkan gambar diatas dapat diketahui *Cluster* mana yang termasuk kedalam kategori Migrasi Masuk Tinggi, Migrasi sedang dan Migrasi Keluar Tinggi. *Cluster0* (Kab. Bandung, Kab.Bogor dan Kota Bogor) memiliki Avg(Migrasi Masuk) : 30,362 dan Avg(Migrasi Keluar) : 30,769 maka termasuk kedalam kategori Migrasi Sedang karena meskipun nilai Avg migrasi keluar lebih tinggi dari migrasi masuk tetapi selisihnya tidak terlalu jauh, *Cluster* ini merupakan *cluster* dengan angka Migrasi paling tinggi diantara *Cluster* lain. Kemudian *Cluster1* (Kab.Pangandaran, Kota Banjar dan Kota Cirebon) memiliki Avg(Migrasi Masuk) : 3,172 dan Avg(Migrasi Keluar) : 3,426 maka termasuk kedalam kategori Migrasi Sedang seperti *Cluster0* karena meskipun nilai Avg migrasi keluar lebih tinggi dari

migrasi masuk tetapi selisihnya tidak terlalu jauh dan *cluster* ini merupakan *cluster* terendah diantara *cluster* lain. Selanjutnya *Cluster2* (Kab.Cianjur, Kab. Garut, Kab. Sukabumi dan Kab. Tasikmalaya) memiliki Avg(Migrasi Masuk) : 13,698 dan Avg(Migrasi Keluar) : 17,135 maka termasuk kedalam kategori Migrasi Keluar Tinggi karena selisih nya yaitu 3,437. Kemudian *Cluster3* (Kota Bekasi) memiliki Avg(Migrasi Masuk) : 15,712 dan Avg(Migrasi Keluar) : 24,189 maka termasuk kedalam kategori Migrasi Keluar paling Tinggi karena selisih nya yaitu 8,477.

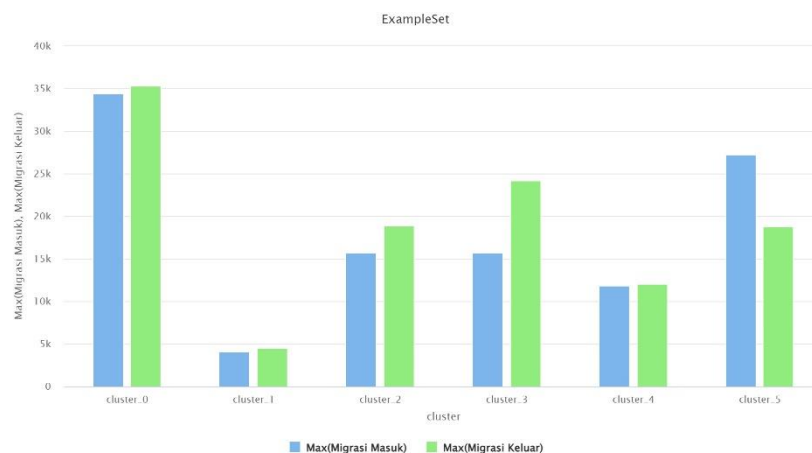
Cluster4 (Kab.Ciamis, Kab.Cirebon, Kab. Indramayu, Kab.Kuningan, Kab. Majalengka, Kab. Purwakarta, Kab. Subang, Kab. Sumedang, Kota Bogor, Kota Cimahi, Kota Depok, Kota Sukabumi, Kota Tasikmalaya) memiliki Avg(Migrasi Masuk) : 8,847 dan Avg(Migrasi Keluar) : 9,130 maka termasuk kedalam kategori Migrasi Sedang karena meskipun nilai Avg migrasi keluar lebih tinggi dari migrasi masuk tetapi selisihnya tidak terlalu jauh. Terakhir yaitu *Cluster5* (Kab. Bandung Barat, Kabupaten Bekasi dan Kab. Karawang) memiliki Avg(Migrasi Masuk) : 21,583 dan Avg(Migrasi Keluar) : 16,034 maka termasuk kedalam kategori Migrasi Masuk Tinggi yaitu selisih 5,549 dengan migrasi keluar, *cluster* ini merupakan satu-satunya *cluster* dengan migrasi masuk lebih tinggi daripada migrasi keluar.



Gambar 4. 4 Migrasi berdasarkan Minimum

Berdasarkan gambar diatas dapat diketahui *Cluster* mana yang termasuk kedalam kategori Migrasi Masuk Tinggi, Migrasi sedang dan Migrasi Keluar Tinggi. *Cluster0* (Kab. Bandung, Kab.Bogor dan Kota Bogor) memiliki Avg(Migrasi Masuk) : 27,763 dan Avg(Migrasi Keluar) : 25,692 maka termasuk kedalam kategori Migrasi Masuk Tinggi karena selisih nya yaitu 2,071, *cluster* ini merupakan *cluster* dengan jumlah migrasi paling banyak diantara *cluster* lainnya. *Cluster1* (Kab.Pangandaran, Kota Banjar dan Kota Cirebon) memiliki Avg(Migrasi Masuk) : 2,243 dan Avg(Migrasi Keluar) : 2,427 maka termasuk kedalam

kategori Migrasi Sedang karena meskipun nilai Avg migrasi keluar lebih tinggi dari migrasi masuk tetapi selisihnya tidak terlalu jauh, *cluster* ini juga merupakan *cluster* dengan migrasi paling rendah. Selanjutnya *Cluster2* (Kab.Cianjur, Kab. Garut, Kab. Sukabumi dan Kab. Tasikmalaya) memiliki Avg(Migrasi Masuk) : 12,449 dan Avg(Migrasi Keluar) : 16,138 maka termasuk kedalam kategori Migrasi Keluar paling Tinggi karena selisih nya yaitu 3,689. *Cluster3* (Kota Bekasi) memiliki Avg(Migrasi Masuk) : 15,712 dan Avg(Migrasi Keluar) : 24,189 maka termasuk kedalam kategori Migrasi Keluar paling Tinggi karena selisih nya yaitu 8,477. Kemudian *Cluster4* (Kab.Ciamis, Kab.Cirebon, Kab. Indramayu, Kab.Kuningan, Kab. Majalengka, Kab. Purwakarta, Kab. Subang, Kab. Sumedang, Kota Bogor, Kota Cimahi, Kota Depok, Kota Sukabumi, Kota Tasikmalaya) memiliki Avg(Migrasi Masuk) : 5,609 dan Avg(Migrasi Keluar) : 5,885 maka termasuk kedalam kategori Migrasi Sedang karena meskipun nilai Avg migrasi keluar lebih tinggi dari migrasi masuk tetapi selisihnya tidak terlalu jauh. Terakhir yaitu *Cluster5* (Kab. Bandung Barat, Kabupaten Bekasi dan Kab. Karawang) memiliki Avg(Migrasi Masuk) : 17,312 dan Avg(Migrasi Keluar) : 12,651 maka termasuk kedalam kategori Migrasi Masuk Tinggi yaitu selisih 4,661 dengan migrasi keluar,



Gambar 4. 5 Migrasi berdasarkan Maximum

Berdasarkan gambar diatas dapat diketahui *Cluster* mana yang termasuk kedalam kategori Migrasi Masuk Tinggi, Migrasi sedang dan Migrasi Keluar Tinggi. *Cluster0* (Kab. Bandung, Kab.Bogor dan Kota Bogor) memiliki Avg(Migrasi Masuk) : 34,389 dan Avg(Migrasi Keluar) : 35,282 maka termasuk kedalam kategori Migrasi Sedang karena meskipun nilai Avg migrasi keluar lebih tinggi dari migrasi masuk tetapi selisihnya tidak terlalu jauh, *Cluster* ini merupakan *cluster* dengan angka Migrasi paling tinggi diantara *Cluster* lain. *Cluster1* (Kab.Pangandaran, Kota Banjar dan Kota Cirebon) memiliki Avg(Migrasi Masuk) :

4,069 dan Avg(Migrasi Keluar) : 4,454 maka termasuk kedalam kategori Migrasi Sedang karena meskipun nilai Avg migrasi keluar lebih tinggi dari migrasi masuk tetapi selisihnya tidak terlalu jauh, *cluster* ini juga merupakan *cluster* dengan migrasi paling rendah. *Cluster2* (Kab.Cianjur, Kab. Garut, Kab. Sukabumi dan Kab. Tasikmalaya) memiliki Avg(Migrasi Masuk) : 15,750 dan Avg(Migrasi Keluar) : 18,906 maka termasuk kedalam kategori Migrasi Keluar paling Tinggi karena selisih nya yaitu 3,156. *Cluster3* (Kota Bekasi) memiliki Avg(Migrasi Masuk) : 15,712 dan Avg(Migrasi Keluar) : 24,189 maka termasuk kedalam kategori Migrasi Keluar paling Tinggi karena selisih nya yaitu 8,477. *Cluster4* (Kab.Ciamis, Kab.Cirebon, Kab. Indramayu, Kab.Kuningan, Kab. Majalengka, Kab. Purwakarta, Kab. Subang, Kab. Sumedang, Kota Bogor, Kota Cimahi, Kota Depok, Kota Sukabumi, Kota Tasikmalaya) memiliki Avg(Migrasi Masuk) : 11,827 dan Avg(Migrasi Keluar) : 12,028 maka termasuk kedalam kategori Migrasi Sedang karena meskipun nilai Avg migrasi keluar lebih tinggi dari migrasi masuk tetapi selisihnya tidak terlalu jauh. Terakhir yaitu *Cluster5* (Kab. Bandung Barat, Kabupaten Bekasi dan Kab. Karawang) memiliki Avg(Migrasi Masuk) : 27,252 dan Avg(Migrasi Keluar) : 18,815 maka termasuk kedalam kategori Migrasi Masuk Tinggi yaitu selisih 8,437 dengan migrasi keluar.

KESIMPULAN DAN SARAN

Berdasarkan hasil pembahasan yang telah diuraikan di atas, dapat ditarik kesimpulan bahwa Metode *Clustering* dengan algoritma *K-Medoids* dapat digunakan untuk mengelompokkan data migrasi penduduk dengan cara mengklasifikasikan Kabupaten/Kota pada *cluster* yang telah ditetapkan. Berdasarkan hasil *Clustering* dengan algoritma *K-Medoids* diketahui *Cluster* terbaiknya yaitu $K=6$ dengan nilai Davies Bouldin atau DBI sebesar: 0.584 yang menghasilkan *Cluster0* berjumlah 2 kabupaten dan 1 Kota, *Cluster1* berjumlah 2 Kota dan 1 Kabupaten, *Cluster2* berjumlah 4 kabupaten, *Cluster3* berjumlah 1 Kota, *Cluster4* berjumlah 8 Kabupaten dan 6 Kota, kemudian *Cluster5* berjumlah 3 Kabupaten. Saran yang disampaikan oleh peneliti dalam penelitian ini adalah pada peneliti selanjutnya untuk menggunakan aplikasi dan algoritma yang berbeda dalam Teknik Data mining mengelompokkan Kabupaten/Kota berdasarkan Jumlah Migrasi Masuk dan Keluar Jawa Barat.

DAFTAR REFERENSI

- Afidah, N. N., & Masrukan. (2023). Penerapan metode *Clustering* dengan algoritma *k-means* untuk pengelompokan data migrasi penduduk tiap kecamatan di kabupaten rembang. *PRISMA, Prosiding Seminar Nasional Matematika*, 6, 729–738.
- Agustian, D. R., & Dermawan, B. A. (2022). Analisis *Clustering* demam berdarah dengue dengan algoritma *K-Medoids* (studi kasus kabupaten karawang). *JIKO (Jurnal Informatika Dan Komputer)*, 6(1), 2477–3964.
- Amalia, D. R., Narasati, R., & Faqih, A. (2019). Perbandingan Hasil Klasifikasi Rasa Minuman Thai Tea yang Paling Digemari Menggunakan *K-means* dan *K-Medoids*. *Prosiding Seminar Nasional Unimus*, 2, 401–407.
- Amelia, F., Iskandar, I., Kurnia Gusti, S., & Haerani, E. (2023). *Clustering* keluarga miskin desa bina baru dengan metode *k-Medoids*. *KREA-TIF: JURNAL TEKNIK INFORMATIKA*, 11(1), 1–13. <https://doi.org/10.32832/krea-tif.v11i1.14104>
- Dinata, R. K., Furqan, & Retno, S. (2023). Pengelompokan daerah padat penduduk untuk penentuan kawasan perumahan di kota lhokseumawe menggunakan *K-Medoids Clustering*. *Jurnal Elektronika Dan Teknologi Informasi*, 4(1), 2721–9380.
- Hidayat, M. K., & Fitriana, R. (2022). Implementasi *k-means* dan *K-Medoids* dalam pengelompokan wilayah potensial produksi daging ayam. *Jurnal Teknologi Industri Pertanian*, 32(3), 239–247. <https://doi.org/10.24961/j.tek.ind.pert.2022.32.3.239>
- Iskak, F. R., Utami, I. T., & Wuryandari, T. (2023). Implementasi algoritma *K-Medoids* dan *k-error* untuk pengelompokan kabupaten/kota di provinsi jawa tengah berdasarkan jumlah produksi peternakan tahun 2020. *Jurnal Gaussian*, 11(3), 366–376. <https://doi.org/10.14710/j.gauss.11.3.366-376>
- Mujiburrahmad, Hamid, A. H., & Nufus, T. (2021). Pola Migrasi Dan Faktor-Faktor Yang Berhubungan Dengan Migrasi Penduduk Di Kecamatan Padang Tiji. *Jurnal Ilmu Sosial Dan Humaniora*, 10(3), 419–429. <https://doi.org/10.23887/jish-undiksha.v10i3.34757>
- Pulungan, N., Suhada, S., & Suhendro, D. (2019). Penerapan Algoritma *K-Medoids* Untuk Mengelompokkan Penduduk 15 Tahun Keatas Menurut Lapangan Pekerjaan Utama. *KOMIK (Konferensi Nasional Teknologi Informasi Dan Komputer)*, 3(1), 329–334. <https://doi.org/10.30865/komik.v3i1.1609>
- Sari, F. D. R., & Ediwijojo, S. P. (2021). Pemetaan Tingkat Kemiskinan di Provinsi Jawa Tengah Berdasarkan Kabupaten/Kota dengan Metode *K-Medoids*. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(4), 1528–1537. <https://doi.org/10.30865/mib.v5i4.3278>
- Widiawati, N., Nurina Sari, B., & Nur Padilah, T. (2022). *Clustering* data penduduk miskin dampak covid-19 menggunakan algoritma *k-Medoids*. *Journal of Applied Informatics and Computing (JAIC)*, 6(1), 55. <http://jurnal.polibatam.ac.id/index.php/JAIC>