

# Implementasi Metode K-Means Clustering Dengan Davies Bouldin Index Pada Analisis Faktor Penyebab Perceraian

Esty Purwaningsih<sup>1,\*</sup>, Ela Nurelasari<sup>1</sup>

<sup>1,2</sup> Program Studi Sistem Informasi; Universitas Bina Sarana Informatika; Jl. Kramat Raya No.98, Senen, Jakarta Pusat; e-mail: [esty.epw@bsi.ac.id](mailto:esty.epw@bsi.ac.id) , [ela.eur@bsi.ac.id](mailto:ela.eur@bsi.ac.id)

\* Korespondensi: e-mail: [esty.epw@bsi.ac.id](mailto:esty.epw@bsi.ac.id)  
No. Telp:0856-9488-2867

Diterima: 29 Mei 2023; Review:07 Juni 2023; Disetujui: 26 Juni 2023

Cara sitasi: Purwaningsih Esty, Nurelasari Ela, 2023. Implementasi Metode K-Means Clustering Dengan Davies Bouldin Index Pada Analisis Faktor Penyebab Perceraian. Information Management for Educators and Professionals. Vol 7 (2): 134-143.

**Abstrak:** Pada dasarnya, perceraian adalah pelepasan hubungan pernikahan antara pasangan. Di negara ini, jumlah kasus perceraian telah mencapai puncaknya dalam enam tahun terakhir. Banyak alasan dapat menyebabkan perceraian, seperti masalah keuangan, meninggalkan pasangan, kekerasan dalam rumah tangga, atau poligami. Dalam penelitian ini, metode clustering K-Means, yang dibantu oleh indeks Davies Bouldin, menunjukkan keunggulan dalam penyelesaian masalah clustering. Software Rapid Studio digunakan untuk mengolah data sekunder. Data diuji dengan nilai  $k=3$ ,  $k=5$ , dan  $k=7$ . Hasil penelitian menunjukkan bahwa kelompok  $k=3$  memperoleh nilai  $-0,419$ , kelompok  $k=5$  memperoleh nilai  $-0,423$ , dan kelompok  $k=7$  memperoleh nilai  $-0,337$ . Dengan demikian, dapat disimpulkan bahwa metode clustering K-Means yang menggunakan indeks Davies Bouldin memiliki nilai  $k=7$ , yang merupakan cluster terbaik dibandingkan dengan nilai  $k=3$  dan  $k=5$ . Cluster berikut dihasilkan dari penelitian yang dilakukan pada metode K-Means dengan nilai  $k=7$  menggunakan Davies Bouldin Indeks: Cluster\_0 terdiri dari "Provinsi Jawa Barat", Cluster\_1 terdiri dari "Kota Tasikmalaya", Cluster\_2 terdiri dari "Cirebon" dan "Indramayu", Cluster\_3 terdiri dari "Tasikmalaya", "Kuningan" dan "Subang", Cluster\_4 terdiri dari "Bogor", "Cianjur", "Sumedang".

**Kata kunci:** pengelompokan, k-means, davies bouldin, perceraian

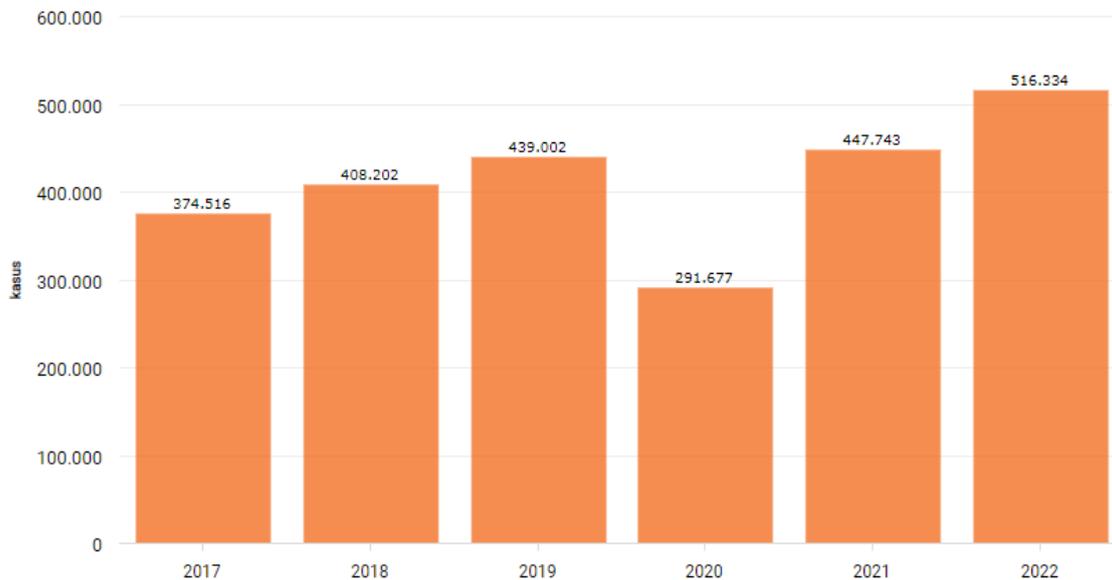
**Abstract:** Basically, divorce is the release of the marital relationship between partners. In this country, the number of divorce cases has reached its peak in the last six years. Many reasons can lead to divorce, such as financial problems, leaving a partner, domestic violence, or polygamy. In this study, the K-Means clustering method, which is assisted by the Davies Bouldin index, shows an advantage in solving clustering problems. Rapid Studio software is used to process secondary data. The data were tested with the values  $k=3$ ,  $k=5$ , and  $k=7$ . The results showed that the  $k=3$  group obtained a value of  $-0,419$ , the  $k=5$  group obtained a value of  $-0,423$ , and the  $k=7$  group obtained a value of  $-0,337$ . Thus, it can be concluded that the K-Means clustering method using the Davies Bouldin index has a value of  $k=7$ , which is the best cluster compared to the values of  $k=3$  and  $k=5$ . The following clusters were generated from research conducted on the K-Means method with a value of  $k=7$  using the Davies Bouldin Index: Cluster\_0 consists of "Provinsi Jawa Barat", Cluster\_1 consists of "Kota Tasikmalaya", Cluster\_2 consists of "Cirebon" and "Indramayu", Cluster\_3 consists of "Tasikmalaya", "Kuningan" and "Subang", Cluster\_4 consists of "Bogor", "Cianjur", "Sumedang".

**Keywords:** clustering, davies bouldin, divorce, k-means

## 1. Pendahuluan

Menikah dan hidup bahagia adalah impian hampir semua orang. Selama menjalani rumah tangga perselisihan pasti terjadi namun semua berharap perselisihan tidak berakhir pada perceraian [1]. Retaknya hubungan ikatan suami istri, yang disebabkan karena tidak ada lagi kesamaan visi, kesamaan prinsip disebut dengan perceraian. Dasar dari perceraian salah satunya ikatan dari pasangan terputus [2]. Perceraian dapat disebabkan banyak faktor, salah satu diantaranya ada karena masalah finansial, perselingkuhan, atau bahkan perbedaan. Perceraian menimbulkan banyak rasa sakit dan kesedihan juga menjadi momok bagi kehidupan anak-anak yang terlibat dalam situasi tersebut.

Laporan Statistik Indonesia menunjukkan bahwa terdapat 516.344 perceraian yang terjadi di Indonesia pada tahun 2022. Jumlah kasus dari perceraian di negara ini mencapai titik tertinggi dalam enam tahun terakhir. Mayoritas perceraian di dalam negeri pada tahun 2022 adalah cerai gugat yang diajukan oleh pihak istri dan sudah diputuskan oleh Pengadilan. Pada tahun sebelumnya, total kasus dari perceraian di negara ini mencapai 388.358, yang merupakan 75,21% dari total kasus. Sebaliknya, cerai talak -permohonan cerai yang diajukan oleh pihak suami dan diputuskan oleh pengadilan adalah penyebab 127.986 kasus, atau 24,78% dari semua perceraian. Jawa Barat memiliki jumlah kasus perceraian tertinggi pada tahun 2022, dengan 113.643 kasus, menurut provinsinya. Jawa Tengah dengan 85.412 kasus dan Jawa Timur dengan 102.065 kasus masing-masing [3].



Sumber: <https://databoks.katadata.co.id/>

Gambar 1. Grafik Jumlah Kasus Perceraian di Indonesia (2017-2022)

Lima provinsi tidak mengalami kasus perceraian pada tahun 2022. Kepulauan Riau, Bali, Kalimantan Utara, Sulawesi Barat, dan Papua Barat termasuk di antaranya. Pada tahun 2022, dilaporkan terjadi 448.126 perceraian di Indonesia berdasarkan faktor penyebabnya. Perselisihan dan pertengkarannya adalah penyebab utama perceraian pada tahun 2022. Jumlahnya 284.169 kasus, atau 63,41% dari faktor penyebab total kasus perceraian di negara asal mereka. Faktor lain yang berkontribusi pada perceraian adalah masalah ekonomi, meninggalkan pasangan, kekerasan dalam rumah tangga, atau poligami.

Peneliti sebelumnya telah banyak melakukan penelitian tentang faktor-faktor yang menyebabkan perceraian. Salah satu contohnya adalah penggunaan Divorce Predictors Scale (DPS), yang didasarkan pada terapi pasangan Gottman, untuk memprediksi perceraian. Algoritma Multilayer Perceptron Neural Network dan C4.5 Decision Tree digunakan untuk menyelidiki keberhasilan DPS. Tingkat keberhasilan tertinggi dengan jaringan saraf RBF adalah 98,23% ketika metode klasifikasi langsung diterapkan pada kumpulan data perceraian. Tingkat akurasi tertinggi sebesar 98,82% dengan JST dicapai setelah menggunakan metode seleksi fitur berbasis korelasi pada kumpulan data yang sama untuk memilih enam fitur yang paling efektif[4].

Metode K-means clustering juga telah digunakan dalam penelitian sebelumnya. Berdasarkan hasil penelitian dan analisis yang dilakukan [1] terhadap Analisis Kasus Perceraian Pada Pengadilan Negeri Bekasi Menggunakan Algoritma K-Means Clustering dengan mengelompokkan kasus perceraian berdasarkan nilai  $y$ , yang menunjukkan lama pernikahan, dan nilai  $x$ , yang menunjukkan perbedaan usia pasangan, dapat disimpulkan bahwa: 1) Hasil uji menggunakan Rapid menunjukkan bahwa hipotesis yang dibuat sebelumnya mengenai klaster area tidak tepat. 2) Berdasarkan data dari tahun 2017 hingga 2020, kasus perceraian di Kota Bekasi cenderung dikaitkan dengan usia pernikahan yang lama. 3) Hasil pengujian menggunakan indeks bouldin Davies, koefisien siluet, dan elbow menunjukkan bahwa jumlah cluster yang ideal adalah dua cluster. Dengan indeks Davies bouldin 0,766, coefficient silhouette 0,5440509737745332, dan coefficient elbow 111.08085465495077. Untuk mencapai hasil terbaik, nilai Silhouette Coefficient dan Davies Bouldin Index sangat tinggi.

Di Indonesia, perceraian telah menjadi hal yang biasa. Angka perceraian mencapai 18,8% dari 1,9 juta peristiwa yang tercatat pada tahun 2017. Studi kasus tersebut bertujuan untuk mengklasifikasikan wilayah Indonesia yang paling sering melakukan perceraian. Ada empat variabel yang digunakan: perselisihan dan pertengkaran terus-menerus, masalah ekonomi, meninggalkan salah satu pihak, dan pelanggaran hak asasi manusia (KDRT). Untuk memastikan metode Clustering yang digunakan akurat, penelitian ini diuji dengan software Rapid. Hasil clustering menunjukkan tingkat penyebab perceraian dari 29 data wilayah Indonesia, dengan 3 wilayah dengan cluster tertinggi dan 26 wilayah dengan cluster terendah. [5].

Analisa kasus perceraian di desa di Kabupaten Kuningan dilakukan dengan metode data mining clustering dengan algoritma KMeans. Data dikelompokkan dengan cara yang sama dengan metode clustering. Diharapkan bahwa jumlah cluster yang dibentuk akan lebih optimum. Ini dicapai dengan menggunakan nilai terkecil dari Davies Bouldin Index. Hasil penelitian menunjukkan bahwa empat kelompok terdiri dari desa atau kelurahan dengan tingkat perceraian tertinggi, tertinggi, sedang, rendah, dan terendah [6].

Di Cimahi, jumlah perceraian terus meningkat setiap tahunnya, rata-rata 800 kasus per bulan. Sebagian besar dari perceraian ini terjadi di rumah tangga, dengan 75 persen dari kasus tersebut merupakan kasus isbat nikah dan waris. Berdasarkan masalah ini, perlu dilakukan prediksi untuk menentukan jumlah perceraian bulanan. Pengolahan data untuk memprediksi perceraian adalah salah satu metode yang digunakan untuk menemukan perceraian. Ini termasuk penggunaan teknik pengolahan data seperti algoritma Naive Bayes dan K-Nearest Neighbor. Algoritma ini dapat memprediksi dengan sangat akurat. Dengan melakukan perbandingan, Anda dapat menemukan algoritma mana yang memiliki akurasi terbaik. Tujuan dari perbandingan algoritma adalah untuk menemukan algoritma yang paling akurat dan paling cepat untuk melakukan prediksi terhadap suatu masalah. Hasil dari perbandingan algoritma Naive Bayes dan K-Nearest Neighbor menunjukkan bahwa algoritma Naive Bayes memiliki akurasi 72,5%, sedangkan algoritma K-Nearest Neighbor memiliki akurasi 57,5% [7].

Naive Bayes juga digunakan dalam penelitian perceraian, dan untuk menemukan estimasi parameter yang diperlukan untuk proses prediksi, teknik ini hanya membutuhkan sedikit data pelatihan [5]. Algoritma C4.5 digunakan untuk pemrosesan data mining, yang mengumpulkan dan mengklasifikasikan data menggunakan representasi struktur pohon. Dalam representasi ini, setiap node, setiap cabang, dan setiap daun menunjukkan nilai atribut masing-masing. Aturan yang saling berhubungan ditemukan melalui penelusuran data sejarah untuk mengkategorikan penyebab perceraian berdasarkan karakteristik sebelumnya [8].

Dalam penelitian ini, tiga kelompok ditentukan: kelompok pertama memiliki produk yang paling laris, kelompok kedua memiliki produk yang laris, dan kelompok ketiga memiliki produk yang kurang laris. Studi ini mengambil lima kategori product, masing-masing dengan sepuluh jenis produk. Hasil perhitungan yang dilakukan dengan Microsoft Excel menunjukkan bahwa kelompok 1 memiliki 18 anggota, kelompok 2 memiliki 6 anggota, dan kelompok 3 memiliki 26 anggota. Cluster 1' memiliki 22 anggota, cluster 2' memiliki 5 anggota, dan cluster 3 memiliki 23 anggota, menurut perhitungan Rapidminer. Hasil evaluasi kelompok dengan DBI cukup baik dengan nilai 0,431. Hasil perhitungan Microsoft Excel mendapatkan nilai 62%, 67%, dan 59% akurasi. Perhitungan Rapidminer menunjukkan akurasi 64% dan recall 81% dan *precision* 88% [9].

Tujuan dari studi kasus ini adalah untuk mengetahui faktor penyebab perceraian di wilayah Jawa Barat, serta bagaimana metode K-Means dengan Davies Bouldin Index berfungsi untuk mengkluster faktor penyebab dari perceraian di sekitar wilayah Jawa Barat.

## 2. Metode Penelitian

Peneliti menggunakan data sekunder dari kemenag, yang dapat diakses melalui <https://jabar.bps.go.id/>. Data yang dikumpulkan meliputi jumlah kasus perceraian di wilayah Jawa Barat serta faktor penyebabnya. Data ini terdiri dari berbagai provinsi di Jawa Barat, seperti Bogor, Sukabumi, Cianjur, Bandung, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Sumedang, Indramayu, Subang, Purwakarta, Karawang, Bekasi, Bandung Barat, Pengandaran, Bogor, Sukabumi, Bandung, Cianjur, Bandung, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Sumedang, Indramayu, Subang, Purwakarta, Zina, mabuk, madat, judi, meninggalkan salah satu, dipenjara, poligami, KDRT, cacat, perselisihan, kawin paksa, murtad, dan ekonomi adalah tiga belas variabel dalam dataset.

Sampel yang diambil untuk kasus penyebab dari perceraian ini meliputi wilayah Provinsi Jawa Barat, Bogor, Cianjur, Garut, Tasikmalaya, Ciamis, Kuningan, Cirebon, Majalengka, Sumedang, Indramayu, Subang, Karawang, Bandung Barat, Kota Bandung, Kota Tasikmalaya. Sedangkan variable yang digunakan pada faktor penyebab perceraian diantaranya mabuk, judi, ditinggal, dipenjara, poligami, kdrt, perselisihan, 'murtad, dan 'ekonomi dimana akan diolah dengan menggunakan metode K-Means melalui aplikasi RapidMiner.

Tabel 1. Sampel Data Penyebab Perceraian Wilayah Jawa Barat

wilayah	mabuk	judi	ditinggal	dipenjara	poligami	kdrt	perselisihan	murtad	ekonomi
Provinsi Jawa Barat	103	86	0	141	151	323	0	306	40603
Bogor	10	7	618	8	23	55	0	48	1251
Cianjur	2	1	41	6	2	3	0	8	1278
Garut	2	1	189	0	2	5	0	38	568
Tasikmalaya	1	1	117	2	0	5	0	24	2680
Ciamis	0	0	493	4	13	24	743	3	3838
Kuningan	0	0	18	2	0	6	104	0	2662
Cirebon	2	6	272	10	4	7	368	4	6437
Majalengka	2	4	121	1	1	0	191	0	3628
Sumedang	1	0	46	17	2	6	0	3	1206
Indramayu	2	10	93	11	7	9	940	7	6937
Subang	6	5	163	14	17	19	0	3	2499
Karawang	28	21	191	10	10	35	0	4	1840
Bandung Barat	0	3	127	4	5	16	0	13	656
Kota Bandung	20	9	401	17	10	63	0	31	1705
Kota Tasikmalaya	1	2	41	4	1	9	749	12	1111

Sumber: <https://opendata.jabarprov.go.id/>

Data mining merupakan cabang ilmu komputer yang masih diperdebatkan apakah termasuk dalam bidang ilmu mana data mining menggabungkan teknik dari berbagai disiplin ilmu [7]. Data mining adalah suatu proses penggalian data yang sangat besar dan tidak diketahui sebelumnya. Menurut fungsinya, data mining dibagi menjadi beberapa metode: asosiasi, pengklusteran (clustering), prediksi, estimasi, dan klasifikasi [2].

Menurut Shimonski, cluster merupakan kumpulan mesin yang bekerja 'sama sebagai satu entitas sumber daya serta layanan kedalam jaringan serta memastikan bahwa sumber daya dapat diakses oleh klien ketika terjadi defect software ataupun hardware pada server [10]

Menurut Aggarwal dan Reddy, pengelompokan atau clustering bermanfaat untuk tugas penambangan data dan pembelajaran mesin. Bisa dianggap sebagai langkah sebelum pemrosesan atau tugas eksplorasi [11]. Data dikelompokkan ke dalam kumpulan berdasarkan algoritma K-Means. Cluster yang sama terdiri dari data dengan atribut yang sama, dan cluster yang berbeda terdiri dari data dengan atribut yang berbeda [12]. Nilai-nilai pada cluster (K) ditetapkan secara acak oleh algoritma kmeans, dan dari nilai tersebut menjadikan pusat cluster atau dapat disebut juga sebagai centroid, mean, atau means. Yang selanjutnya, gunakan rumus Euclidian untuk menghitung jarak antara setiap data dan centroid sampai anda menemukan jarak terdekat antara keduanya. Setiap data harus diklasifikasikan didasarkan pada kedekatan dengan centroid. Sampai tidak mengubah nilai pada centroid, lanjutkan prosedur ini [13].

Komponen utama algoritma K-means adalah: Tentukan jumlah cluster yang dibutuhkan dengan menghitung nilai k., Pembentukan centroid yang dapat dibangkitkan secara acak dengan k,

Gunakan persamaan jarak geometris berikut untuk menghitung jarak data ke- masing-masing centroid:

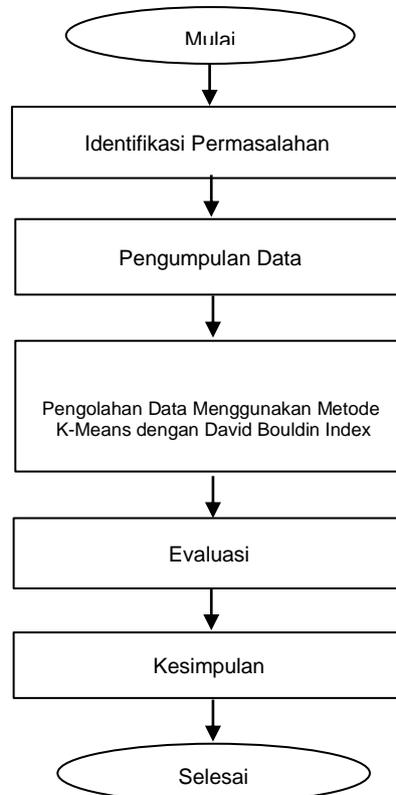
$$d(x_i, \mu_j) = \sqrt{\sum (x_i - \mu_j)^2} \dots\dots\dots (1)$$

Merge setiap data berdasarkan jarak' terdekat antara centroid dan data, Identifikasi posisi centroid (k) baru, Apabila posisi dari centroid baru dengan centroid lama tidak sama, maka kembali ke- langkah 3.

Metode pengelompokan berdasarkan partisi yang dikenal sebagai algoritma K-Means biasanya menggunakan penghitungan jarak untuk mengevaluasi kesamaan. Inti algoritma K-means adalah memilih titik pusat k acak dan membagi data sesuai dengan jarak antara titik pusat k. [14]

Metode K-Means digunakan oleh peneliti untuk menganalisis penyebab perceraian di wilayah Jawa Barat. Dalam metode K-Means, algoritma clustering cukup sederhana digunakan untuk membagi dataset ke dalam berbagai cluster k. Algoritma ini sangat mudah digunakan dan dijalankan, relatif cepat, mudah diubah, dan banyak digunakan. Studi kasus ini bertujuan untuk mengetahui tingkat perceraian, khususnya di wilayah jawa barat, dan untuk mengetahui seberapa baik metode K-Means berfungsi untuk mengkluster tingkat faktor penyebab perceraian di wilayah jawa barat.

Peneliti bereksperimen dengan menggunakan model clustering dengan metode K-Means untuk menganalisis dari suatu faktor penyebab perceraian di wilayah Jawa Barat. Agar menjadi lebih terarah dengan baik, maka penelitian harus memiliki kerangka pemikiran. Berikut adalah metodologi penelitian yang dilakukan peneliti. Tahapan yang dilakukan dalam memecahkan masalah dalam penelitian sebagai berikut:



Sumber: Hasil Penelitian (2023)

Gambar 2. Tahapan Penelitian

1. Identifikasi Permasalahan  
 Identifikasi dari masalah pada studi kasus penelitian ini adalah untuk menganalisis faktor penyebab perceraian di wilayah Jawa Barat dan untuk mengetahui fungsi metode K-Means dengan Davies Bouldin Index (DBI) untuk mengkluster faktor penyebab dari perceraian.
2. Pengumpulan Data  
 Data yang dikumpulkan menggunakan data sekunder. Data diadopsi dari Open Data Jabar melalui <https://opendata.jabarprov.go.id/>. Data yang dikumpulkan meliputi jumlah kasus perceraian di wilayah Jawa Barat serta faktor penyebabnya. Data ini terdiri dari berbagai provinsi di Jawa Barat, seperti -Bogor, -Sukabumi, -Cianjur, -Bandung, -Garut, -Tasikmalaya, -Ciamis, -Kuningan, -Cirebon, -Majalengka, -Sumedang, -Indramayu, -Subang, -Purwakarta, -Karawang, -Bekasi, -Bandung Barat, -Pengandaran, -Bogor, -Sukabumi, -Bandung, -Cianjur, -Bandung, -Garut, -Tasikmalaya, -Ciamis, -Kuningan, -Cirebon, -Majalengka, -Sumedang, -Indramayu, -Subang, -Purwakarta.  
 Faktor penyebab perceraian yang dijadikan sebagai variable dalam dataset pada penelitian ini terdiri dari Zina, mabuk, madat, judi, meninggalkan salah satu, dipenjara, poligami, KDRT, cacat, perselisihan, kawin paksa, murtad, dan ekonomi.
3. Pengolahan Data  
 Penelitian ini menggunakan model clustering dengan metode K-means yang menggunakan algoritma Davies Bouldin Index pada data faktor penyebab perceraian di wilayah Jawa Barat. Pengolahan data menggunakan software Rapid Studio. Hasil Analisa studi kasus pada penelitian ini dapat berguna untuk dijadikan evaluasi dan agar dapat meminimalisir faktor penyebab perceraian.
4. Evaluasi  
 Pengujian dari studi kasus pada penelitian ini menggunakan software Rapid Studio dengan merelasikan dataset yang akan diuji. Informasi yang didapat dan dihasilkan dari proses uji data dapat dilihat kedalam bentuk yang pastinya mudah dimengerti.
5. Kesimpulan  
 Dari hasil uji penelitian selesai maka dapat diberikan kesimpulan terhadap permasalahan yang diujikan agar pihak-pihak yang membutuhkan informasi tersebut dapat menggunakannya.

**3. Hasil dan Pembahasan**

*Rapid* sebagai *software* untuk pengujian data yang digunakan dalam penelitian ini. Adapun tahapan dari proses penelitian dan hasil implementasi yang menggunakan *RapidStudio* adalah sebagai berikut :

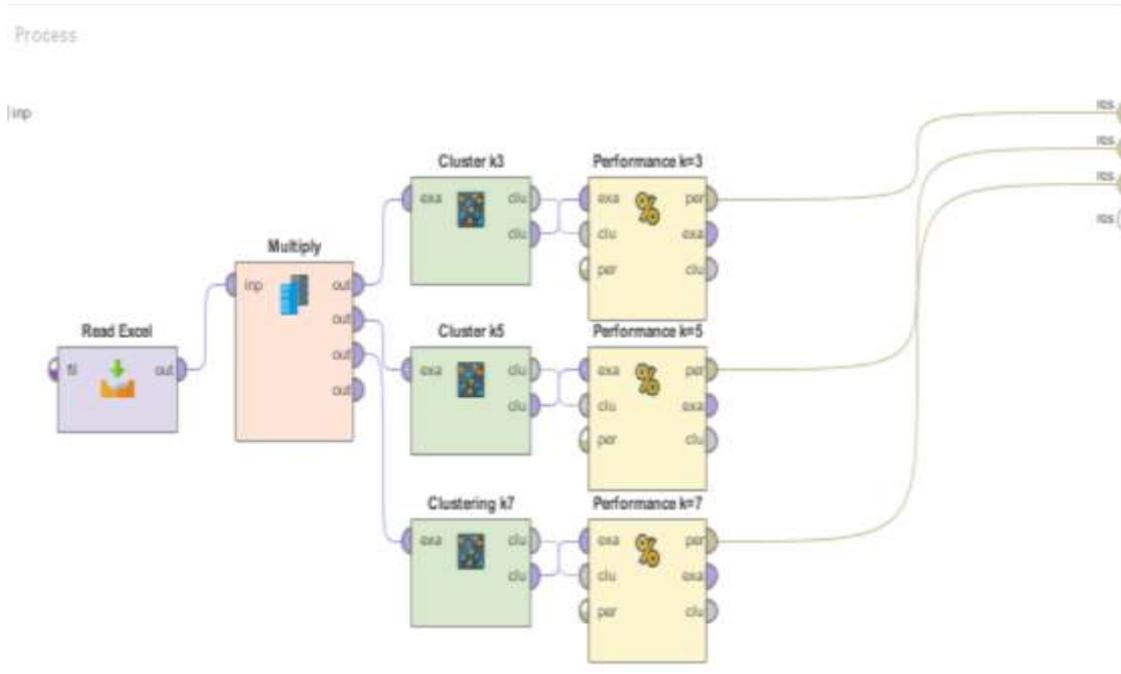
Pemilihan Operator: Operator yang dipakai untuk implementasi pengujian data adalah : *Read Excel*, dengan mengupload data yang berfungsi untuk pembaca data; Dimana data diolah dengan menggunakan Microsoft Excel.

	A	B	C	D	E	F	G	H	I	J
1	wilayah	mabuk	judi	ditinggal	dipenjara	poligami	kdr	perselisi...	murtad	ekonomi
2	Provinsi ...	103	86	0	141	151	323	0	306	40603
3	Bogor	10	7	618	8	23	55	0	48	1251
4	Cianjur	2	1	41	6	2	3	0	8	1278
5	Garut	2	1	189	0	2	5	0	38	568
6	Tasikma...	1	1	117	2	0	5	0	24	2680
7	Ciamis	0	0	493	4	13	24	743	3	3838
8	Kuningan	0	0	18	2	0	6	104	0	2662
9	Cirebon	2	6	272	10	4	7	368	4	6437
10	Majaleng...	2	4	121	1	1	0	191	0	3628
11	Sumeda...	1	0	46	17	2	6	0	3	1206
12	Indramayu	2	10	93	11	7	9	940	7	6937
13	Subang	6	5	163	14	17	19	0	3	2499
14	Karawang	28	21	191	10	10	35	0	4	1840

Sumber: Hasil Penelitian (2023)

Gambar 3. Import Dataset

*Clustering* : merupakan implementasi dari metode *k-means*. Masukkan data yang nantinya akan diolah berupa Microsoft excel dengan menggunakan *Rapid Miner* dan pilih modelling segmentation k-means untuk mengelompokkan data dengan mengisi nilai (k) yang akan digunakan. Pada penelitian ini menguji nilai (k) dengan k=3, k=5 dan k=7.



Sumber: Hasil Penelitian (2023)

Gambar 4. Pemilihan Operator

Hasil Olahan

Pemilihan modelling segmentation k-means dilakukan dari masing-masing nilai k diantaranya k=3, k=5 dan k=7, didapat nilai performance, yaitu:

Tabel 2. Nilai Performance

Criterion	k=5	k=3	k=7
Avg. within centroid distance	-994769,276	-171338,677	-74481,840
Avg. within centroid distance_cluster_0	-2331331,375	-203700,972	-0,000
Avg. within centroid distance_cluster_1	-0,000	-0,000	-0,000
Avg. within centroid distance_cluster_2	-599180,264	-152316,000	-152316,000
Avg. within centroid distance_cluster_3		-194121,600	-12954,444
Avg. within centroid distance_cluster_4		-121986,500	-119611,520
Avg. within centroid distance_cluster_5			-121986,500
Avg. within centroid distance_cluster_6			-3091,750
Davies Bouldin	-0,419	-0,423	-0,337

Sumber: Hasil Penelitian (2023)

Avg. within centroid distance, rata-rata dalam jarak cluster yang dihitung dengan rata-rata jarak antara centroid dan semua contoh cluster. Sedangkan untuk evaluasi cluster digunakan metode Davies-Bouldin Index, dimana semakin kecil nilai Davies Bouldin Index (DBI) yang diperoleh (non negative  $\geq 0$ ), maka semakin baik pula nilai cluster yang didapatkan dari pengelompokan dengan algoritma clustering. Dari hasil pengolahan data terlihat dari tabel performance untuk kriteria davies bouldin index yang terbaik adalah pada cluster k=7 senilai -0,337 dibandingkan dengan cluster k=3 senilai -0,419 dan cluster k=5 senilai -0,423.

Pada hasil cluster data dengan nilai  $k=7$ , dapat ditampilkan centroid table dengan penjabaran setiap cluster, yaitu:

Tabel 3. Centroid Table  $k=7$

attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5	cluster_6
mabuk	103	1	2	2,333	12,200	1	1
judi	86	2	8	2	7,600	2	2
ditinggal	0	41	182,500	99,333	259,400	307	158
dipenjara	141	4	10,500	6	11,600	2,500	2
poligami	151	1	5,500	5,667	9,400	7	3.500
kdr	323	9	8	10	32,400	12	10,500
perselisihan	0	749	654	34,667	0	467	0
Murtad	306	12	5,500	9	18,800	1,500	25,500
ekonomi	40603	1111	6687	2613,667	1456	3733	612

Sumber: Hasil Penelitian (2023)

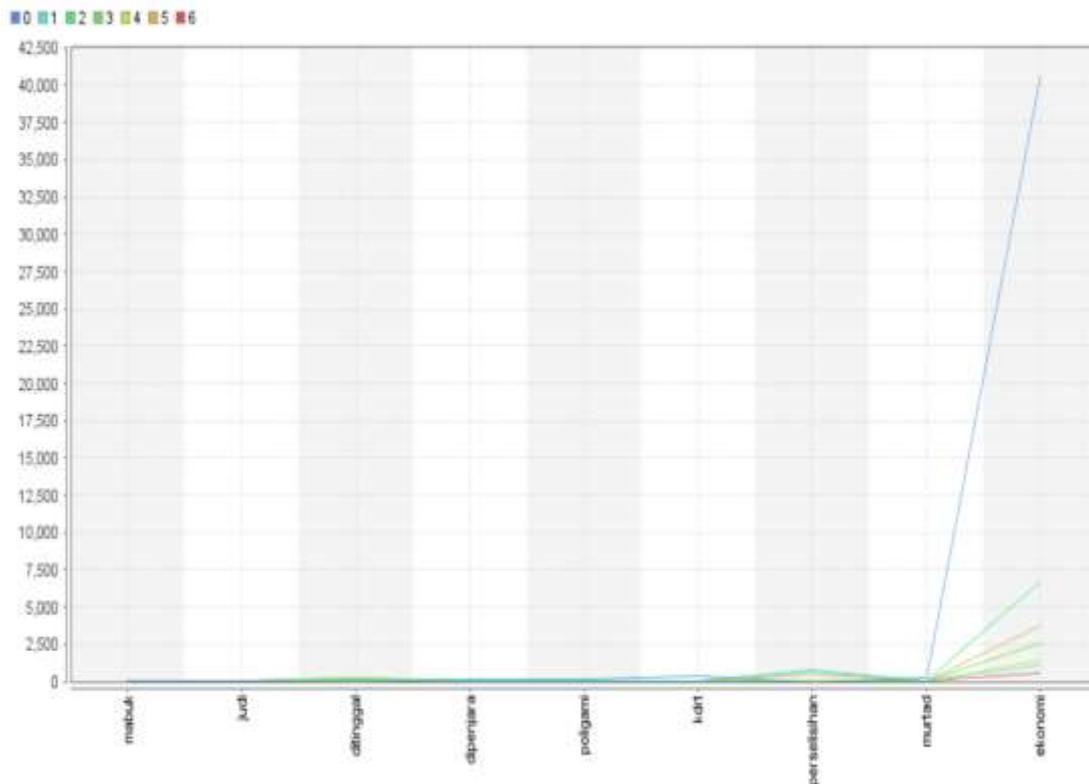
Masing-masing *cluster* diperoleh dari hasil pengujian data yang didapat, diantaranya:

Tabel 4. Item Cluster Model

Cluster Model	
Cluster 0	1 items
Cluster 1	1 items
Cluster 2	2 items
Cluster 3	3 items
Cluster 4	5 items
Cluster 5	2 items
Cluster 6	2 items
Total number of items	16

Sumber: Hasil Penelitian (2023)

Dalam proses uji pengolahan dataset faktor penyebab perceraian, didapat grafik yang menggambarkan penyebaran dari masing-masing *cluster*.



Sumber: Hasil Penelitian (2023)

Gambar 5. Plot Cluster Model  $k=7$

#### 4. Kesimpulan

Untuk melakukan pengelompokan factor penyebab perceraian dapat dilakukan dengan menerapkan metode Clustering K-Means dengan menggunakan davies bouldin index. Data-data yang diperoleh, diolah dengan menggunakan software Rapid Studio untuk menentukan nilai centroid dari nilai k yang terbaik. Dari hasil pengujian data dengan menggunakan davies bouldin index dihasilkan bahwa cluster k=3 mendapat nilai -0,419, sedangkan cluster k=5 mendapat nilai -0,423 dan cluster k=7 mendapat nilai -0,337. Sehingga dapat disimpulkan bahwa metode clustering K-Means dengan menggunakan davies bouldin index nilai k=7 adalah cluster terbaik dibandingkan dengan nilai k=3 dan k=5. Penelitian pada metode K-Means dengan nilai k=7 menggunakan davies bouldin index dihasilkan cluster sebagai berikut: cluster\_0 terdiri dari “Provinsi Jawa Barat”, cluster\_1 terdiri dari “Kota Tasikmalaya”, cluster\_2 terdiri dari “Cirebon” dan “Indramayu”, cluster\_3 terdiri dari “Tasikmalaya”, “Kuningan” dan “Subang”, cluster\_4 terdiri dari “Bogor”, “Cianjur”, “Sumedang”, “Karawang”, dan “Kota Bandung”, cluster\_5 terdiri dari “Ciamis” dan “Majalengka”, cluster\_6 terdiri dari “Garut” dan “Bandung Barat”. Penelitian ini dapat membantu pemerintah menangani kasus perceraian di Indonesia dengan membuat program yang tepat untuk masing-masing provinsi untuk mengurangi tingkat perceraian.

#### Ucapan Terima Kasih

Alhamdulillah, terima kasih kepada Allah Subhana huwata'ala, keluarga dan teman serta instansi yang mendukung untuk menyelesaikan penelitian pada studi kasus ini.

#### Referensi

- [1] U. D. Rahayu *et al.*, “Analisis Kasus Perceraian Pada Pengadilan Negeri Bekasi Menggunakan Algoritma K-Means Clustering,” *J. Univ. Pendidik. Indones.*, vol. 6, no. 1, pp. 165–172, 2022.
- [2] E. windarman; sapri; suryana, “Implementation of the Naïve Bayes Algorithm for Divorce Prediction at the Tais Religious Court Implementasi Algoritma Naïve Bayes Untuk Prediksi Perceraian Pada Pengadilan Agama Tais,” *J. Komputer, Inf. dan Teknol.*, vol. 2, no. 2, pp. 501–510, 2022.
- [3] E. F. Santika, “Kasus Perceraian di Indonesia Melonjak Lagi pada 2022, Tertinggi dalam Enam Tahun Terakhir,” 2023. [Online]. Available: [https://databoks.katadata.co.id/datapublish/2023/03/01/kasus-perceraian-di-indonesia-melonjak-lagi-pada-2022-tertinggi-dalam-enam-tahun-terakhir#:~:text=Menurut laporan Statistik Indonesia%2C jumlah,2021 yang mencapai 447.743 kasus. \[Accessed: 01-Mar-2023\]](https://databoks.katadata.co.id/datapublish/2023/03/01/kasus-perceraian-di-indonesia-melonjak-lagi-pada-2022-tertinggi-dalam-enam-tahun-terakhir#:~:text=Menurut laporan Statistik Indonesia%2C jumlah,2021 yang mencapai 447.743 kasus. [Accessed: 01-Mar-2023]).
- [4] M. K. Yontem, K. Adem, Ilhan, and S. T. Ve Kilicarslan, “Divorce Prediction Using Correlation Based Feature Selection and Artificial Neural Networks,” *International Congress on Politic, Economic and Social*, vol. 9, no. 1. pp. 259–273, 2019.
- [5] N. Nurhayati, F. Azzahra, S. Ramadani, S. D. Hastuti, and E. Irawan, “Analisis Metode Klastering Pada Kasus Penyebab Perceraian Berdasarkan Provinsi Dengan Teknik K-Means,” *KOMIK (Konferensi ...)*, vol. 4, no. 1, pp. 278–284, 2020.
- [6] Y. Sopyan, A. D. Lesmana, and C. Juliane, “Analisis Algoritma K-Means dan Davies Bouldin Index dalam Mencari Cluster Terbaik Kasus Perceraian di Kabupaten Kuningan,” *Build. Informatics, Technol. Sci.*, vol. 4, no. 3, pp. 1464–1470, 2022.
- [7] M. Irfan, W. Uriawan, O. T. Kurahman, M. A. Ramdhani, and I. A. Dahlia, “Comparison of Naive Bayes and K-Nearest Neighbor methods to predict divorce issues,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 434, no. 1, 2018.
- [8] W. J. Andari and E. Bululolo, “Implementasi Algoritma C4.5 Mengetahui Penyebab Perceraian Dalam Pernikahan (Studi Kasus: Pengadilan Agama Medan Kelas I-A),” *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 3, p. 365, 2020.
- [9] F. Amin, D. S. Anggraeni, and Q. Aini, “Penerapan Metode K-Means dalam Penjualan Produk Souq.Com,” *Appl. Inf. Syst. Manag.*, vol. 5, no. 1, pp. 7–14, 2022.
- [10] D. Ayu, M. Wati, D. Puspitasari, and E. Purwaningsih, “Metode Clustering Pada Model Algoritma K-Means Untuk Pemilihan Alat Kontrasepsi,” vol. 3, no. 2, pp. 129–138, 2019.
- [11] E. Purwaningsih, “Analisis Kecelakaan Berlalu Lintas Di Kota Jakarta Dengan Menggunakan Metode K-Means,” *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*,

- vol. 5, no. 1, pp. 139–144, 2019.
- [12] W. A. Eko, "Implementasi data mining dalam pengelompokan data peserta didik di sekolah untuk memprediksi calon penerima beasiswa dengan menggunakan algoritma k- means (studi kasus sman 16 bekasi)," vol. 21, no. 3, 2016.
- [13] R. Tri Vulandari, *Data Mining Teori dan Aplikasi Rapidminer*, 1st ed. surakarta: GAVA MEDIA, 2017.
- [14] W. Wu and M. Peng, "A Data Mining Approach Combining K-Means Clustering with Bagging Neural Network for Short-Term Wind Power Forecasting," *IEEE Internet Things J.*, vol. 4, no. 4, pp. 979–986, 2017.