



Clustering Data Penduduk Miskin Menggunakan Metode K-Means Pada Desa Teluk Agung Kecamatan Indramayu Kabupaten Indramayu

Mohammad Farhan Rizaldy¹, Tati Suprpti², Gifthera Dwilestari³

^{1,2,3}STMIK IKMI Cirebon

ARTICLE INFO

Article history:

Received July 12, 2025

Revised August 02, 2025

Accepted October 30, 2025

Available online December 30, 2025

Kata Kunci :

Desa Teluk Agung, Penduduk Miskin, Clustering, K-Means, Knowledge Discovery in Database (KDD), Davies Bouldin (DBI), PK3E

Keywords:

Teluk Agung Village, Poor Population, Clustering, K-Means, Knowledge Discovery in Database (KDD), Davies Bouldin (DBI), PK3E



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright ©2025 by Author. Published by LPPM Universitas Islam Syekh-Yusuf

ABSTRAK

Metode K-Means merupakan salah satu metode Data Mining yang banyak digunakan dalam penelitian pengelompokan. Berdasarkan hasil penelitian yang telah dilakukan untuk membangun sebuah model proses Analisa Pengelompokan Data Kemiskinan Menggunakan Pendekatan Metode K-Means Di Desa Teluk Agung Kecamatan Indramayu Kabupaten Indramayu, dapat menggunakan tools RapidMiner dengan membuat operator dan mengolah parameter yang digunakan untuk pengelompokan kategori kelas P3KE. Operator yang digunakan dalam penelitian ini yaitu operator Read Excel, Set Role, Select Attributes, Replace Missing Values, Nominal to Numerical, Multiply, Clustering (K-Means) dan Performance. Operator yang digunakan berjumlah 8 operator dengan menerapkan tahapan Knowledge Discovery in Database (KDD). Penelitian kali ini akan mengaplikasikan Indeks Davies Bouldin (DBI) sebagai salah satu cara optimasi jumlah cluster untuk mengelompokkan data, dari perobaan nilai cluster terbaik maka yang mendekati angka 0 adalah K9 dengan nilai DBI -2.257, dari sini bisa kita simpulkan bahwasannya kurang lebih 43 item dari cluster 2 - 10 itu masuk ke dalam kategori P3KE, dan selain 43 item tadi bisa diartikan masih belum termasuk kedalam kategori P3KE.

ABSTRACT

The K-Means method is one of the Data Mining methods that is widely used in clustering research. Based on the results of research that has been conducted to build a process model for Poverty Data Clustering Analysis Using the K-Means Method Approach in Teluk Agung Village, Indramayu District, Indramayu Regency, can use RapidMiner tools by creating operators and processing parameters used for clustering the P3KE class category. The operators used in this study are Read Excel, Set Role, Select Attributes, Replace Missing Values, Nominal to Numerical, Multiply, Clustering (K-Means) and Performance operators. The operators used are 8 operators by applying the stages of Knowledge Discovery in Database (KDD). This research will apply the Davies Bouldin Index (DBI) as a way of optimising the number of clusters to group data, from the best cluster value experiment, the closest to 0 is K9 with a DBI value of -2.257, from this we can conclude that approximately 43 items from clusters 2 - 10 are included in the P3KE category, and other than the 43 items can be interpreted as still not included in the P3KE category.

1. INTRODUCTION

Salah satu bentuk inovasi dalam kependudukan adalah mengembangkan aplikasi yang dapat dimanfaatkan baik dalam menganalisis tingkat kemakmuran dalam suatu daerah. Sebagai upaya meningkatkan mutu kesejahteraan melalui pengelompokan penduduk dalam kelas sesuai dengan kriteria penduduk. Pemecahan permasalahan pengelompokan data penduduk dengan data yang semakin banyak menjadi kurang efisien dan dibutuhkan pengelompokan data penduduk yang juga memiliki prestasi untuk masuk kelas unggulan telah ditetapkan. Pengelompokan data penduduk pada kelompok yang tepat akan mampu mengimprovisasi hasil pembelajaran menjadi lebih baik. Metode ini telah banyak digunakan dalam berbagai penelitian, seperti yang pernah dilakukan oleh, Kesulitan dalam menentukan wilayah mana yang mengalami tingkat kemiskinan yang paling tinggi dan normal serta wilayah dengan tingkat kemiskinan rendah menjadi sebuah alasan untuk melakukan pengclusteran ini. Metode cluster yang digunakan adalah clustering k-means sebagai salah satu permodelan deskriptif. (Febianto & Palasara, 2019). Proses pengelompokan melibatkan pembagian item data menjadi kelompok-kelompok yang lebih kecil, dengan masing-masing kelompok memiliki kesamaan yang signifikan. menjelaskan clustering

*Corresponding author.

E-mail addresses: mohfarhan1133@gmail.com

sebagai proses membagi data yang tidak berlabel menjadi kelompok data dengan kesamaan. (Bahauddin et al., 2021)

Kemiskinan merupakan masalah yang serius, terutama di negara berkembang seperti Indonesia. Dalam jangka panjang, kemiskinan di suatu wilayah akan menghambat pembangunan nasional. Kemiskinan adalah suatu kondisi kehidupan dimana seseorang atau rumah tangga mengalami kekurangan sehingga tidak mampu memenuhi kebutuhan hidup minimum atau memadai. (Febianto & Palasara, 2019)

Pada penelitian sebelumnya dalam jurnal nya yang berjudul "*K-MEANS Clustering Analysis Pada Persebaran Tingkat Pengangguran Kabupaten/Kota Di Sulawesi Selatan*" Metode pengelompokan yang sering digunakan dalam penelitian adalah clustering analysis. Clustering analysis yang sering dijumpai dalam penelitian adalah clustering analysis menggunakan algoritma K-means. Pengelompokan dengan menggunakan k-means diterapkan di berbagai bidang, diantaranya adalah kesehatan, pertanian, dan pendidikan. (Akramunnisa & Fajriani, 2020)

Indikator yang lain yang lain adalah IPM, yaitu suatu indeks yang diperoleh berdasarkan tingkat pendidikan, tingkat kesehatan, dan perekonomian penduduk. IPM dalam hal ini mewakili tingkat pendidikan, tingkat kesehatan pada penelitian sebelumnya. Hasil yang diperoleh dari pengelompokan TPT, IPM dan UMK dikategorikan menjadi 2 cluster, yaitu cluster 1 merupakan kelompok dengan tingkat pengangguran rendah dan cluster 2 merupakan tingkat pengangguran tinggi. Hasil analisis yang diperoleh dapat dijadikan sebagai gambaran bagi pemerintah daerah/kota bahkan tingkat provinsi dalam menangani masalah pengangguran dengan melihat aspek yang berpengaruh terhadap perubahan tingkat pengangguran, yaitu UMK dan IPM. (Akramunnisa & Fajriani, 2020)

Berdasarkan pada permasalahan yang telah dipaparkan maka judul yang diajukan adalah "*Clustering Data Penduduk Miskin Menggunakan Metode K-MEANS Pada Desa Teluk Agung Kecamatan Indramayu Kabupaten Indramayu*". Penentuan judul tersebut akan memberikan solusi atas permasalahan yang dihadapi pegawai di Balai desa Teluk Agung, dengan melakukan analisa pengelompokan mampu memberikan hasil analisis yang lebih baik dan detail, sehingga dapat membantu meningkatkan efisiensi dan efektivitas dalam mengelola data penduduk miskin.

Metode *K-Means* merupakan salah satu metode *Data Mining* yang banyak digunakan dalam penelitian pengelompokan. Namun metode *K-Means* memiliki beberapa kekurangan, salah satunya yaitu dalam penentuan jumlah cluster. Penelitian kali ini akan mengaplikasikan Indeks *Davies Bouldin (DBI)* sebagai salah satu cara optimasi jumlah cluster untuk mengelompokkan propinsi berdasar potensi desa dengan banyaknya jenis industri yang dimiliki wilayahnya. Data yang digunakan adalah data banyaknya desa atau kelurahan menurut keberadaan dan jenis industri kecil dan mikro (desa). (Waruwu et al., 2023)

Dapat menjadi sumber informasi bagi pembaca bahwasannya data yang tercantum dalam kategori desil 1,2, dan 3 bisa terkelompokan dalam bentuk *Cluster* yang merupakan hasil dari metode *K-Means*

2. METHODS

Metode penelitian ini menggunakan metode eksperimen. Metode eksperimen digunakan untuk proses pengujian terhadap variabel yang terdapat pada dataset berdasarkan hasil observasi yang sudah dilakukan pada desa Teluk Agung Kecamatan Indramayu Kabupaten Indramayu guna mendapatkan informasi dari dataset yang digunakan sebagai objek dalam penelitian.

Untuk melaksanakan penelitian ini menggunakan metode *K-Means*. *K-Means* merupakan Algoritma yang efektif untuk menganalisis data dalam jumlah besar (Novitasari et al., 2023). Tahapan dalam metode penelitian ini berupa:

1. Observasi

Merupakan pengamatan mengenai suatu objek tertentu secara cermat secara langsung di lokasi penelitian tersebut berada.

2. Data Selection

Merupakan proses yang bertujuan untuk memilih data yang akan digunakan dalam proses *Clustering*.

Adapun turunan dari Data Selection antara lain:

1. Operator Set Role
2. Operator Select Attributes
3. Data Preprocessing
4. Data Transformation
5. Data Mining

Merupakan Pada tahapan data mining data dibagi menjadi dua tahapan yaitu:

- a. *Clustering (K-Means)*

b. Evaluasi *David Bouldin Index* (DBI).

3. Hasil Clustering.

3. RESULT AND DISCUSSION

Dari metode penelitian ini menghasilkan dua hasil yang tujuannya berbeda yakni, yang pertama untuk menunjukkan nilai K terbaik pada K9 dan nilai dari David Bouldin Index adalah -2.257 merupakan angka yang mendekati nilai 0. Berdasarkan hasil penelitian yang telah dilakukan hasil dari pengelompokan penduduk miskin melalui pekerjaan dan pendidikan hasil uji *cluster* 2 sampai dengan 10, untuk nilai K terbaik yang mendekati angka 0 yakni pada *cluster* 9 dan mendapatkan nilai DBI sebesar -2.257, kelompok pada pengelompokan data penduduk miskin berdasarkan variable Penghasilan perbulan menggunakan algoritma *K-Means*. Berikut ini adalah tabel hasil perolehan nilai dari setiap *cluster* menggunakan pendekatan *K-Means* :

Tabel Cetroid dan DBI

dan hasil yang kedua merupakan pengelompokan jumlah cluster untuk mengetahui kategori desil yang merupakan program P3KE di desa Teluk agung Kabupaten Indramayu yang terdapat 3 kelas Desil, yaitu Desil 1 untuk kategori miskin extreme yang dimana penghasilannya dibawah Rp.351.000.00,-, dan untuk kategori Desil 2 yaitu yang memiliki penghasilan dibawah Rp.550.000.00,-, dan untuk kategori Desil 3 yaitu memiliki penghasilan dibawah Rp.1.000.000.00,-, dan jika diterapkan dalam pengelompokan data maka bisa kita lihat tabel dibawah.

	Desil 1 (item)	Desil 2 (item)	Desil 3 (item)	Tidak termasuk kategori desil (item)
Cluster 0	2	3	34	56
Cluster 1	2	0	14	29
Cluster 2	0	1	3	59
Cluster 3	0	0	0	4
Cluster 4	2	4	9	5
Cluster 5	3	1	4	4
Cluster 6	0	0	0	9
Cluster 7	0	0	0	4
Cluster 8	1	0	4	32

Cluster	Centroid Distance	Davies Bouldin Index
K2	-2.928	-2.853
K3	-2.757	-3.046
K4	-2.581	-2.543
K5	-2.527	-2.600
K6	-2.378	-2.508
K7	-2.399	-2.457
K8	-2.319	-2.564
K9	-2.331	-2.257
K10	-2.257	-2.463

Tabel Pembagian kelompok P3KE

4. CONCLUSION

Berdasarkan hasil analisis dan pembahasan dari penelitian yang telah dilakukan untuk Analisa Pengelompokkan Data Kemiskinan Menggunakan Pendekatan Metode K-Means Di Desa Teluk Agung Kecamatan Indramayu Kabupaten Indramayu maka dapat ditarik kesimpulan sebagai berikut.

1. Berdasarkan hasil penelitian yang telah dilakukan untuk membangun sebuah model proses Analisa Pengelompokkan Data Kemiskinan Menggunakan Pendekatan Metode K-Means Di Desa Teluk Agung Kecamatan Indramayu Kabupaten Indramayu, dapat menggunakan tools RapidMiner dengan membuat operator dan mengolah parameter yang digunakan untuk pengelompokkan P3KE. Operator yang digunakan dalam penelitian ini yaitu operator Read Excel, Set Role, Select Attributes, Replace Missing Values, Nominal to Numerical, Multiply, Clustering (K-Means) dan Performance. Operator yang digunakan berjumlah 8 operator dengan menerapkan tahapan Knowledge Discovery in Database (KDD).

2. Berdasarkan hasil penelitian clustering data penduduk miskin menggunakan metode K-Means pada Desa Teluk Agung Kecamatan Indramayu Kabupaten Indramayu, dengan sejumlah data yang dipakai kurang lebih 290 record dan 10 atribut diantaranya atribut No, NIK, Nama, Jenis Usia, Status, Hub Keluarga, Pendidikan, Pekerjaan, dan Penghasilan Perbulan dari proses pengclustoran dengan cluster 2 sampai 10 pada penelitian dengan software rapid miner menggunakan algoritma K-Means di dapatkan K9 dengan jumlah anggota Cluster 0: 95 items, Cluster 1: 45 items, Cluster 2: 63 items, Cluster 3: 4 items, Cluster 4: 20 items, Cluster 5: 12 items, Cluster 6: 9 items, Cluster 7: 4 items, Cluster 8: 37 items, Cluster 9: 95 items, dari perobaan nilai cluster terbaik maka yang mendekati angka 0 adalah K9 dengan nilai DBI - 2.257, dari sini bisa kita simpulkan bahwasannya kurang lebih 43 item dari cluster 2-10 itu masuk ke dalam kategori P3KE, dan selain 43 item tadi bisa diartikan masih belum termasuk kedalam kategori P3KE.

5. REFERENCES

- AKRAMUNNISA, A., & FAJRIANI, F. (2020). K-MEANS CLUSTERING ANALYSIS PADA PERSEBARAN TINGKAT PENGANGGURAN KABUPATEN/KOTA DI SULAWESI SELATAN. *JURNAL VARIAN*, 3(2), 103–112. [HTTPS://DOI.ORG/10.30812/VARIAN.v3i2.652](https://doi.org/10.30812/VARIAN.v3i2.652)
- Bahauddin, A., Fatmawati, A., & Permata Sari, F. (2021). Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means. *Jurnal Manajemen Informatika Dan Sistem Informasi*, 4(1), 1–8. <https://doi.org/10.36595/misi.v4i1.216>
- Febianto, N. I., & Palasara, N. (2019). Analisa Clustering K-Means Pada Data Informasi Kemiskinan Di Jawa Barat Tahun 2018. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 8(2), 130–140. <https://doi.org/10.32736/sisfokom.v8i2.653>
- Filki, Y. (2022). Algoritma K-Means Clustering dalam Memprediksi Penerima Bantuan Langsung Tunai (BLT) Dana Desa. *Jurnal Informatika Ekonomi Bisnis*, 4, 166–171. <https://doi.org/10.37034/infeb.v4i4.166>
- Handoko, S., Fauziah, F., & Handayani, E. T. E. (2020). Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering. *Jurnal Ilmiah Teknologi Dan Rekayasa*, 25(1), 76–88. <https://doi.org/10.35760/tr.2020.v25i1.2677>
- Kasim, R. J., Bahri, S., & Amir, S. (2021). Implementasi Metode K-Means Untuk Clustering Data Penduduk

- Miskin Dengan Systematic Random Sampling. *Prosiding SISFOTEK*, 5(1), 95–101.
- Novitasari, N., Nuris, N. D., & Herdiana, R. (2023). Penerapan Algoritma K-Means untuk Clustering Data Jumlah Penduduk Miskin Berdasarkan Kota/Kabupaten di Jawabarar menggunakan Rapidminer. *Jurnal Informatika Terpadu*, 9(1), 68–73. <https://doi.org/10.54914/jit.v9i1.660>
- Nugroho Arif Sudibyo, Ardymulya Iswardani, Kartika Sari, & Siti Suprihatiningsih. (2020). Penerapan Data Mining Pada Jumlah Penduduk Miskin Di Indonesia. *Jurnal Lebesgue : Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, 1(3), 199–207. <https://doi.org/10.46306/lb.v1i3.42>
- Sari, Y. R., Sudewa, A., Lestari, D. A., & Jaya, T. I. (2020). Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer. *CESS (Journal of Computer Engineering, System and Science)*, 5(2), 192. <https://doi.org/10.24114/cess.v5i2.18519>
- Sunia, D., Kurniabudi, & Alam Jusia, P. (2019). Penerapan Data Mining Untuk Clustering Data Penduduk Miskin Menggunakan Algoritma K-Means. *Jurnal Ilmiah*
- Akramunnisa, A., & Fajriani, F. (2020). K-Means Clustering Analysis pada PersebaranTingkat Pengangguran Kabupaten/Kota di Sulawesi Selatan. *Jurnal Varian*, 3(2), 103–112. <https://doi.org/10.30812/varian.v3i2.652>
- Bahauddin, A., Fatmawati, A., & Permata Sari, F. (2021). Analisis Clustering Provinsi Di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Algoritma K-Means. *Jurnal Manajemen Informatika Dan Sistem Informasi*, 4(1), 1–8. <https://doi.org/10.36595/misi.v4i1.216>
- Febianto, N. I., & Palasara, N. (2019). Analisa Clustering K-Means Pada Data Informasi Kemiskinan Di Jawa Barat Tahun 2018. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 8(2), 130–140. <https://doi.org/10.32736/sisfokom.v8i2.653>
- Filki, Y. (2022). Algoritma K-Means Clustering dalam Memprediksi Penerima Bantuan Langsung Tunai (BLT) Dana Desa. *Jurnal Informatika Ekonomi Bisnis*, 4, 166–171. <https://doi.org/10.37034/infv4i4.166>
- Handoko, S., Fauziah, F., & Handayani, E. T. E. (2020). Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering. *Jurnal Ilmiah Teknologi Dan Rekayasa*, 25(1), 76–88. <https://doi.org/10.35760/tr.2020.v25i1.2677>
- Kasim, R. J., Bahri, S., & Amir, S. (2021). Implementasi Metode K-Means Untuk Clustering Data Penduduk Miskin Dengan Systematic Random Sampling. *Prosiding SISFOTEK*, 5(1), 95–101.
- Novitasari, N., Nuris, N. D., & Herdiana, R. (2023). Penerapan Algoritma K-Means untuk Clustering Data Jumlah Penduduk Miskin Berdasarkan Kota/Kabupaten di Jawabarar menggunakan Rapidminer. *Jurnal Informatika Terpadu*, 9(1), 68–73. <https://doi.org/10.54914/jit.v9i1.660>
- Nugroho Arif Sudibyo, Ardymulya Iswardani, Kartika Sari, & Siti Suprihatiningsih. (2020). Penerapan Data Mining Pada Jumlah Penduduk Miskin Di Indonesia. *Jurnal Lebesgue : Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, 1(3), 199–207. <https://doi.org/10.46306/lb.v1i3.42>
- Sari, Y. R., Sudewa, A., Lestari, D. A., & Jaya, T. I. (2020). Penerapan Algoritma K-Means Untuk Clustering Data Kemiskinan Provinsi Banten Menggunakan Rapidminer. *CESS (Journal of Computer Engineering, System and Science)*, 5(2), 192. <https://doi.org/10.24114/cess.v5i2.18519>
- Sunia, D., Kurniabudi, & Alam Jusia, P. (2019). Penerapan Data Mining Untuk Clustering Data Penduduk Miskin Menggunakan Algoritma K-Means. *Jurnal Ilmiah Mahasiswa Teknik Informatika*, 1(2), 121–134.
- Waruwu, A., Yetri, M., & Setiawan, F. (2023). Implementasi Data Mining Dalam Mengelompokkan Data penduduk Kurang Mampu Menggunakan Metode K-Means Clustering. *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, 2(6), 945. <https://doi.org/10.53513/jursi.v2i6.8965>
- Yulia, & Silalahi, M. (2021). Penerapan Data Mining Clustering Dalam Mengelompokkan Buku Dengan Metode K-Means. *Indonesian Journal of Computer Science*, 10(1). <https://doi.org/10.33022/ijcs.v10i1.3008>
- Mahasiswa Teknik Informatika, 1(2), 121–134.
- Waruwu, A., Yetri, M., & Setiawan, F. (2023). Implementasi Data Mining Dalam Mengelompokkan Data penduduk Kurang Mampu Menggunakan Metode K-Means Clustering. *Jurnal Sistem Informasi Triguna Dharma (JURSI TGD)*, 2(6), 945. <https://doi.org/10.53513/jursi.v2i6.8965>
- Yulia, & Silalahi, M. (2021). Penerapan Data Mining Clustering Dalam Mengelompokkan Buku Dengan Metode K-Means. *Indonesian Journal of Computer Science*, 10(1). <https://doi.org/10.33022/ijcs.v10i1.3008>