



## Menentukan Kelompok Prioritas Penerima Bantuan Program Keluarga Harapan (PKH) dengan Penerapan Data Mining Menggunakan Metode Clustering K-Means di Desa Dagang Kelambir

Rahmat Syahroni<sup>1\*</sup>, Didi Febrina<sup>2</sup>

<sup>1,2</sup> Universitas Negeri Medan, Indonesia

\*Penulis Korespondensi: [rahmatsyahroni0207@gmail.com](mailto:rahmatsyahroni0207@gmail.com)

**Abstract.** *The Keluarga Harapan Program (PKH) is a government conditional social assistance program aimed at improving the welfare of low-income families. However, its implementation often faces issues in accurately targeting beneficiaries. Therefore, an objective method is needed to classify people based on their level of need. This study applies the K-Means clustering method to group PKH beneficiary families in Dagang Kelambir Village into clusters with similar characteristics. The variables used include age, number of dependents, occupation, income, and housing ownership status, with a total of 201 residents analyzed. Data processing was conducted using the R programming language, including normalization and the Elbow method to determine the optimal number of clusters. The results show three clusters: 124 families in the non-priority group, 39 in the secondary priority group, and 38 in the primary priority group. This method is effective in identifying priority groups objectively and supports more accurate and efficient distribution of PKH assistance.*

**Keywords:** *Data Mining; Data Normalization; Family Hope Program; K-Means Clustering; Social Assistance.*

**Abstrak.** Program Keluarga Harapan (PKH) merupakan bantuan sosial bersyarat dari pemerintah yang ditujukan untuk meningkatkan kesejahteraan masyarakat kurang mampu. Namun, dalam pelaksanaannya masih terdapat kendala, terutama dalam menentukan penerima bantuan agar tepat sasaran. Oleh sebab itu, dibutuhkan suatu metode yang objektif untuk mengelompokkan masyarakat berdasarkan tingkat kebutuhannya. Penelitian ini bertujuan menerapkan metode *clustering K-Means* untuk mengelompokkan keluarga penerima PKH di Desa Dagang Kelambir ke dalam kelompok dengan karakteristik yang serupa. Metode yang digunakan adalah K-Means dengan variabel meliputi usia, jumlah tanggungan, pekerjaan, penghasilan, serta status kepemilikan rumah. Data yang diolah berjumlah 201 warga dan dianalisis menggunakan bahasa pemrograman R melalui proses normalisasi serta penentuan jumlah klaster dengan metode Elbow. Hasil penelitian menunjukkan bahwa data terbagi menjadi tiga klaster, yaitu klaster non-prioritas sebanyak 124 keluarga, klaster prioritas menengah sebanyak 39 keluarga, dan klaster prioritas utama sebanyak 38 keluarga. Penerapan metode ini terbukti mampu mengidentifikasi kelompok prioritas secara objektif sehingga mendukung penyaluran bantuan PKH yang lebih adil, tepat sasaran, dan efisien di tingkat desa.

**Kata kunci:** Bantuan Sosial; Data Mining; Keluarga Harapan Program; K-Means Clustering; Normalisasi Data.

### 1. LATAR BELAKANG

Program Keluarga Harapan (PKH) merupakan salah satu program bantuan sosial dari pemerintah yang bertujuan untuk mengurangi angka kemiskinan serta meningkatkan kualitas hidup masyarakat kurang mampu melalui pemberian bantuan bersyarat (Supriana et al., 2023). Program ini menasar keluarga miskin yang memenuhi kriteria tertentu, seperti memiliki ibu hamil, anak usia sekolah, lanjut usia, maupun penyandang disabilitas. Namun, dalam praktiknya, penentuan penerima bantuan PKH seringkali masih menghadapi berbagai permasalahan, seperti ketidaktepatan sasaran, kurangnya objektivitas dalam penilaian, serta keterbatasan dalam pengolahan data masyarakat secara akurat dan efisien (Tri Yunarni et al., 2019).

Di Desa Dagang Kelambir, proses penentuan penerima bantuan PKH umumnya masih dilakukan secara manual berdasarkan observasi dan pertimbangan subjektif aparatur desa. Hal ini berpotensi menimbulkan ketidaksesuaian antara kondisi ekonomi masyarakat dengan bantuan yang diterima. Selain itu, banyaknya data masyarakat dengan berbagai karakteristik seperti usia, jumlah tanggungan, pekerjaan, penghasilan, serta status kepemilikan rumah membuat proses seleksi menjadi semakin kompleks dan rentan terhadap kesalahan (Sudiby et al., 2020).

Seiring dengan perkembangan teknologi, pemanfaatan data mining menjadi solusi yang efektif dalam membantu proses pengambilan keputusan berbasis data. Data mining memungkinkan pengolahan data dalam jumlah besar untuk menemukan pola atau informasi tersembunyi yang dapat digunakan sebagai dasar penentuan kebijakan (Surahman & Hayati, 2023). Salah satu metode yang dapat digunakan adalah *clustering*, yaitu teknik pengelompokan data berdasarkan kemiripan karakteristik tertentu tanpa menggunakan label kelas sebelumnya (Sudirman & Sulaeman, 2020).

Metode K-Means merupakan salah satu algoritma *clustering* yang banyak digunakan karena memiliki keunggulan dalam kesederhanaan proses dan efisiensi dalam pengolahan data (Sudiby et al., 2020). Metode ini bekerja dengan cara mengelompokkan data ke dalam beberapa cluster berdasarkan jarak terdekat terhadap titik pusat (*centroid*). Dengan menggunakan variabel seperti usia, jumlah tanggungan, pekerjaan, penghasilan, dan status kepemilikan rumah, metode K-Means dapat membantu mengelompokkan masyarakat ke dalam kategori tertentu yang mencerminkan tingkat kelayakan sebagai penerima bantuan (Mustika et al., 2021).

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menentukan kelompok prioritas penerima bantuan Program Keluarga Harapan (PKH) di Desa Dagang Kelambir dengan menerapkan teknik data mining menggunakan metode *clustering K-Means*. Diharapkan hasil penelitian ini dapat memberikan rekomendasi yang lebih objektif, akurat, dan tepat sasaran dalam penyaluran bantuan sosial, serta menjadi bahan pertimbangan bagi pemerintah desa dalam meningkatkan efektivitas program PKH (Yulia & Nugroho, 2022). Seiring dengan perkembangan teknologi dan metode analisis data, penggunaan algoritma seperti K-Means *clustering* dapat meningkatkan akurasi dan objektivitas dalam seleksi penerima bantuan, sekaligus mengurangi kemungkinan kesalahan manusia (Kurniawan & Suprati, 2023).

Metode K-Means *clustering* memungkinkan pembagian data ke dalam kelompok-kelompok yang lebih homogen berdasarkan karakteristik tertentu, seperti yang telah diterapkan

dalam penelitian oleh Fadilah et al. (2022) yang menggunakan teknik serupa untuk mengelompokkan kelas siswa berdasarkan pemahaman mereka dalam bimbingan belajar. Penerapan metode ini dalam konteks penerima bantuan sosial diharapkan dapat mempermudah penentuan kelompok prioritas dan memastikan bantuan diberikan sesuai dengan kebutuhan (Putri et al., 2022). Penelitian oleh Amaliyah & Rianti (2023) juga mengungkapkan pentingnya penerapan data mining untuk menentukan kelompok prioritas penerima bantuan PKH di desa dengan metode K-Means yang terbukti efektif dalam meningkatkan ketepatan penyaluran bantuan.

## **2. KAJIAN TEORITIS**

### **Data Mining dan *Clustering***

Data mining merupakan suatu proses untuk menggali pola, hubungan, serta informasi penting dari sekumpulan data berukuran besar dengan memanfaatkan berbagai metode seperti statistik, kecerdasan buatan, dan machine learning. Proses ini dilakukan melalui analisis data secara mendalam guna menemukan pola tersembunyi, kecenderungan, maupun anomali yang sulit dikenali melalui teknik analisis konvensional. Tujuan utama dari data mining adalah untuk membantu proses pengambilan keputusan yang lebih optimal dengan menghadirkan informasi bernilai dari hasil pengolahan data tersebut. Dengan adanya data mining, suatu organisasi dapat menghasilkan keputusan yang lebih akurat, berbasis data, dan terarah, sehingga mampu meningkatkan efektivitas operasional serta memperkuat daya saingnya (Fadilah et al., 2022).

*Clustering* adalah teknik dalam data mining yang bertujuan untuk mengelompokkan objek-objek ke dalam grup-grup (cluster) berdasarkan kemiripan atau kedekatan karakteristik mereka. Teknik ini mengidentifikasi kelompok-kelompok dalam data sehingga objek-objek dalam satu kelompok memiliki kesamaan yang tinggi satu sama lain, sementara objek-objek di kelompok yang berbeda memiliki perbedaan yang signifikan. *Clustering* sangat berguna dalam berbagai aplikasi, termasuk analisis pemasaran untuk segmentasi pelanggan, pengenalan pola dalam gambar atau teks, dan analisis sosial-ekonomi untuk mengidentifikasi kelompok populasi dengan karakteristik tertentu (Sudirman & Sulaeman, 2020).

### **Metode *Clustering K-Means***

K-Means merupakan algoritma clustering yang banyak digunakan karena sederhana dan efisien dalam proses pengelompokan data. Metode ini membagi data ke dalam sejumlah  $k$  cluster berdasarkan kedekatan jarak terhadap *centroid* sebagai pusat cluste.

Berikut adalah langkah-langkah dan rumus dalam metode *K-Means*:

- a. Inisialisasi: Pilih  $k$  *centroid* awal secara acak.
- b. Assignment Step: Alokasikan setiap data  $x_i$  ke *cluster* dengan *centroid* terdekat menggunakan rumus jarak Euclidean:

$$d(x_i, c_j) = \sqrt{\sum_{m=1}^n (x_{im} - c_{jm})^2}$$

- c. Update Step: Hitung ulang posisi *centroid* dengan mengambil rata-rata dari semua data dalam *cluster* tersebut:

$$c_j = \frac{1}{|C_j|} \sum_{x_i \in C_j} x_i$$

dimana  $c_j$  adalah *centroid* baru dari *cluster*  $j$ ,  $C_j$  adalah himpunan data dalam *cluster*  $j$ , dan  $|C_j|$  adalah jumlah data dalam *cluster*  $j$ .

Iterasi: Ulangi langkah 2 dan 3 hingga posisi *centroid* stabil atau perubahan posisi *centroid* sangat kecil.

### 3. METODE PENELITIAN

Penelitian ini menggunakan desain deskriptif kuantitatif yang dilaksanakan di Kantor Desa Dagang Kelambir selama kurang lebih dua bulan, dengan populasi seluruh kepala keluarga Dusun II sebanyak 201 keluarga dan teknik pengambilan sampel berupa total sampling. Data yang digunakan merupakan data sekunder dari basis data desa serta data primer yang dikumpulkan melalui wawancara dan observasi, mencakup variabel usia, jumlah tanggungan, pekerjaan, penghasilan, dan status kepemilikan rumah. Selanjutnya, data diolah dengan melakukan normalisasi pada variabel numerik serta transformasi data kategorikal, kemudian dianalisis menggunakan metode *Clustering K-Means* dengan bantuan bahasa pemrograman R. Hasil analisis digunakan untuk mengidentifikasi dan mengelompokkan keluarga berdasarkan karakteristiknya sehingga dapat memberikan rekomendasi penentuan prioritas penerima bantuan Program Keluarga Harapan (PKH) secara lebih tepat sasaran.

#### 4. HASIL DAN PEMBAHASAN

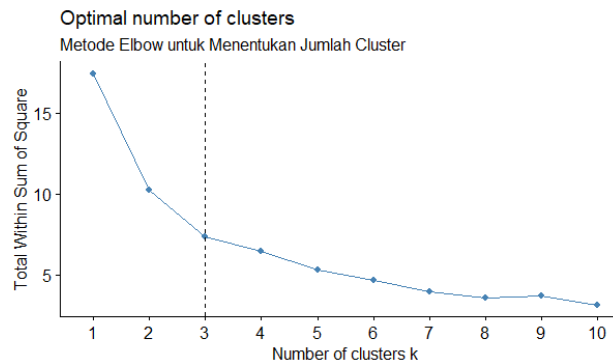
##### Statistik Deskriptif

**Tabel 1.** Statistik Deskriptif.

Statistik	Usia	Tanggungan	Penghasilan (Rp)
Minimum	21	0	0
Maksimum	91	5	10.000.000
Rata-rata (Mean)	48,56	1,67	2.710.448
Median	48	2	2.900.000
Simpangan Baku (SD)	12,62	1,02	1.154.487

Berdasarkan Tabel 4.6, dapat dilihat bahwa: Usia warga Dusun II berkisar antara 21 hingga 91 tahun, dengan rata-rata usia sekitar 48,56 tahun dan median 48 tahun. Ini menunjukkan bahwa sebagian besar responden berada pada usia produktif hingga lanjut usia. Jumlah tanggungan keluarga berkisar antara 0 hingga 5 orang, dengan rata-rata sekitar 1,67 tanggungan. Hal ini mencerminkan beban ekonomi yang beragam antar keluarga. Penghasilan bulanan menunjukkan rentang yang sangat lebar, dari Rp 0 hingga Rp 10.000.000, dengan rata-rata Rp 2.710.448 dan median Rp 2.900.000. Simpangan baku yang tinggi (Rp 1.154.487) menunjukkan adanya ketimpangan penghasilan antar warga.

##### Penentuan Jumlah Kluster



**Gambar 1.** Grafik Metode Elbow.

**Tabel 2.** Tabel Nilai SSE.

Jumlah Kluster (k)	SSE	Selisih SSE
1	372.8681	—
2	273.2156	-99.6525
3	182.527	-90.6886
4	104.4199	-78.1071
5	84.25909	-20.1608
6	59.38243	-24.8767
7	39.22164	-20.1608
8	31.60315	-7.61849
9	30.24278	-1.36037
10	23.9965	-6.24628

Berdasarkan tabel di atas, Tabel menunjukkan nilai Sum of Squared Errors (SSE) untuk berbagai jumlah kluster dari 1 hingga 10, serta selisih penurunan SSE antar jumlah kluster. SSE adalah total kesalahan kuadrat dalam setiap kluster, yang digunakan untuk menilai seberapa baik data diklusterisasi. Semakin rendah nilai SSE, semakin baik model kluster membagi data. Dari tabel tersebut terlihat bahwa nilai SSE mengalami penurunan drastis dari kluster ke-1 hingga ke-3. SSE menurun sebesar 99.65 saat berpindah dari 1 ke 2 kluster. Kemudian turun lagi sebesar 90.69 dari 2 ke 3 kluster. Selanjutnya, penurunan menjadi lebih kecil dan bertahap, seperti 78.11 dari kluster ke-3 ke ke-4, dan semakin kecil pada kluster-kluster berikutnya. Penurunan yang signifikan hanya terjadi hingga kluster ke-3. Setelah itu, penambahan jumlah kluster tidak memberikan pengurangan SSE yang cukup besar untuk membenarkan kompleksitas model yang meningkat. Hal ini juga didukung oleh plot elbow, yang menunjukkan adanya "siku" (*elbow point*) pada kluster ke-3. Dengan demikian, berdasarkan nilai SSE dan visualisasi elbow plot, jumlah kluster optimal untuk analisis kluster dalam penelitian ini adalah sebanyak 3 kluster. Ini menjadi dasar dalam proses klusterisasi gabungan antara hasil PCA untuk variabel numerik dan MCA untuk variabel kategorikal.

### Pengelompokkan menggunakan K-Means

- a. Proses pertama dilakukan dengan memasukkan variabel-variabel data yang akan digunakan dalam pengelompokan. Data tersebut sebelumnya telah melalui tahap normalisasi agar setiap variabel memiliki skala yang sebanding. Penentuan *centroid* awal dilakukan secara acak dengan jumlah tiga titik pusat yang mewakili masing-masing *cluster* pada setiap variabel. Nilai *centroid* awal yang digunakan:

**Tabel 3.** *Centroid* Awal.

<i>Centroid</i>	No	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$
<i>Centroid 1</i>	4	-0.00034	-0.01417	-0.01141	-0.07522	0.352278
<i>Centroid 2</i>	47	0.013494	0.023486	0.035711	0.341202	-1.16288
<i>Centroid 3</i>	150	-0.14583	0.255414	0.008067	-1.18825	0.4534

- b. Menghitung jarak setiap data terhadap *centroid* dari masing-masing *cluster*. Perhitungan jarak pada penelitian ini dilakukan dengan menggunakan *euclidean distance*. *Euclidean distance* lebih efektif dalam mengidentifikasi pola data dibanding *manhattan distance* dan *minkowski distance*. rumus *euclidean distance* :

$$D(i,j) = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2}.$$

- c. Dalam penelitian ini, metode yang digunakan adalah *Euclidean Distance* karena dinilai lebih mampu menggambarkan kedekatan antar data dibandingkan metode lain seperti

Manhattan Distance maupun *Minkowski Distance*. Perhitungan jarak dilakukan dengan mengukur akar dari jumlah kuadrat selisih setiap variabel antar data dan *centroid*.

- d. Mengelompokkan data ke dalam *cluster* yang paling dekat dengan *centroid*. Setiap data akan dibandingkan jaraknya terhadap seluruh *centroid*, kemudian dimasukkan ke dalam *cluster* dengan nilai jarak terkecil. Hasil koordinat *centroid* yang baru disajikan pada Tabel di bawah ini.

**Tabel 4.** *Centroid* pada Iterasi 1.

<i>Centroid</i>	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$
<i>Centroid 1</i>	-0.161	0.0343	-0.0540	-0.621	0.464
<i>Centroid 2</i>	0.0135	0.0235	0.0357	0.341	-1.16
<i>Centroid 3</i>	0.0711	-0.0271	0.00983	0.144	0.303

Berdasarkan tabel di atas, masih ada anggota *cluster* yang berubah, sehingga nilai *centroid* perlu diperbaharui lagi. Nilai *centroid* terbaru telah dihitung sebagai berikut.

**Tabel 5.** *Centroid* pada Iterasi 2.

<i>Centroid</i>	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$
<i>Centroid 1</i>	-0.0829	0.00542	-0.0444	-0.524	0.309
<i>Centroid 2</i>	0.0135	0.0235	0.0357	0.341	-1.16
<i>Centroid 3</i>	0.116	-0.0264	0.0402	0.535	0.424

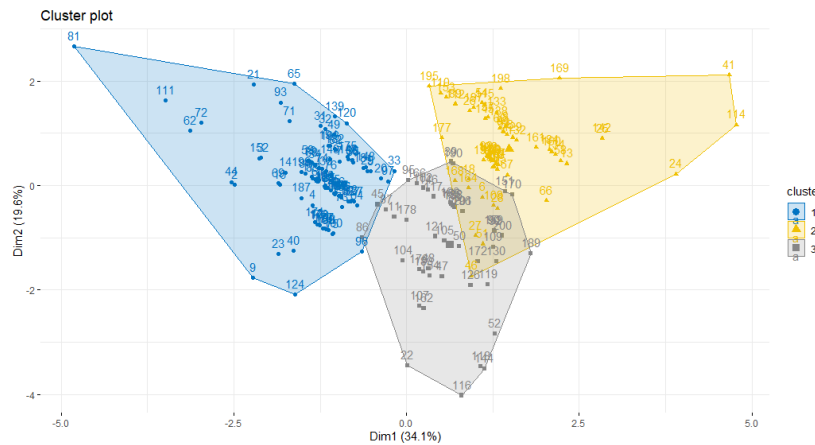
Berdasarkan tabel di atas, hasil pengelompokan menunjukkan masih adanya perpindahan anggota klaster, sehingga diperlukan pembaruan *centroid* selanjutnya.

**Tabel 6.** *Centroid* pada Iterasi 3.

<i>Centroid</i>	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$
<i>Centroid 1</i>	-0.0187	0.00663	-0.0253	-0.446	0.275
<i>Centroid 2</i>	0.000169	0.0154	0.0373	0.296	-1.19
<i>Centroid 3</i>	0.0669	-0.0442	0.0415	1.21	0.592

Tabel di atas menunjukkan nilai *centroid* terbaru. Hasil pengelompokan menunjukkan bahwa anggota klaster sudah tidak berubah lagi, sehingga proses klasterisasi telah selesai.

## Pengelompokan Menggunakan *K-Means Clustering*



**Gambar 2.** Scatter Plot.

Berdasarkan hasil pengelompokan *K means clustering*, dapat diidentifikasi kelompok prioritas penerima bantuan Program Keluarga Harapan (PKH). Tabel berikut menyajikan keterangan masing-masing klaster berdasarkan karakteristik utama dan status prioritasnya.

**Tabel 7.** Kesimpulan Prioritas Penerima Bantuan.

<i>Cluster</i>	Keterangan	Prioritas Bantuan	Penerima
1	Penghasilan tertinggi, kondisi ekonomi lebih stabil	Bukan Prioritas	
2	Penghasilan menengah, sebagian belum punya rumah sendiri, mayoritas buruh harian lepas.	Prioritas Sekunder	
3	penghasilan terendah, tidak ada yang memiliki rumah sendiri, pekerjaan beragam tapi dominan IRT dan karyawan, dan tanggungan lebih banyak.	Prioritas Utama	

## 5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian, dapat disimpulkan bahwa penerapan metode K-Means dengan menggunakan variabel sosial ekonomi seperti penghasilan, jumlah tanggungan, status pekerjaan, usia, dan kepemilikan rumah mampu mengelompokkan 201 warga Dusun II Desa Dagang Kelambir ke dalam tiga klaster secara objektif dalam menentukan kelayakan penerima bantuan PKH. Hasil pengelompokan menunjukkan bahwa *cluster* 1 terdiri dari 124 orang dengan kondisi ekonomi relatif stabil sehingga tidak menjadi prioritas, *cluster* 2 sebanyak 39

orang berada pada kondisi menengah sebagai prioritas sekunder, dan *cluster* 3 sebanyak 38 orang merupakan kelompok paling rentan yang menjadi prioritas utama penerima bantuan.

## DAFTAR REFERENSI

- Amaliyah, S., & Rianti, S. (2023). Penerapan data mining untuk menentukan kelompok prioritas penerima bantuan PKH menggunakan metode *clustering K-Means* pada Desa Kuala Dendang. *Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM)*, 3(1), 453-458. <https://doi.org/10.33998/jakakom.2023.3.1.802>
- Amaliyah, S., & Rianti, S. (2023). Penerapan data mining untuk menentukan kelompok prioritas penerima bantuan PKH menggunakan metode *clustering K-Means* pada Desa Kuala Dendang. *Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM)*, 3(1), 453-458. <https://doi.org/10.33998/jakakom.2023.3.1.802>
- Aprilia, S., Elmayati, E., & Sobri, A. (2023). Data mining penerapan metode *K-Means clustering* untuk mengelompokkan penerima bantuan Kartu Indonesia Pintar (Studi kasus di Universitas Musi Rawas). *ESCAF*, 1095-1100.
- Auliasari, K., & Kertaningtyas, M. (2019). Penerapan algoritma K-Means untuk segmentasi konsumen menggunakan R. *Jurnal Teknologi Dan Manajemen Informatika*, 5(2). <https://doi.org/10.26905/jtmi.v5i2.3644>
- Fadilah, A. G., Wahyuni, F. S., & Achmadi, S. (2022). Implementasi metode *K-Means clustering* untuk pengelompokan kelas berdasarkan pemahaman siswa pada bimbingan belajar Matematika Saschio Banyuwangi. *Jurnal Mahasiswa Teknik Informatika*, 6(2). <https://doi.org/10.36040/jati.v6i2.5450>
- Fadilah, A. G., Wahyuni, F. S., & Achmadi, S. (2022). Implementasi metode K-Means clustering untuk pengelompokan kelas berdasarkan pemahaman siswa pada bimbingan belajar Matematika Saschio Banyuwangi. *Jurnal Mahasiswa Teknik Informatika*, 6(2). <https://doi.org/10.36040/jati.v6i2.5450>
- Fitriyah, H., Safitri, E. M., Muna, N., Khasanah, M., Aprilia, D. A., & Nurdiansyah, D. (2023). Implementasi algoritma clustering dengan modifikasi metode Elbow untuk mendukung strategi pemerataan bantuan sosial di Kabupaten Bojonegoro. *Jurnal Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, 4(3), 1598-1607. <https://doi.org/10.46306/lb.v4i3.453>
- Kurniawan, R., & Suprapti, T. (2023). Penerapan algoritma K-Means *clustering* untuk identifikasi kelayakan penerima bantuan Program Keluarga Harapan (PKH) di Desa Tambaksari Ciamis. *Jurnal Mahasiswa Teknik Informatika*, 7(6). <https://doi.org/10.36040/jati.v7i6.8197>
- Kurniawan, R., & Suprapti, T. (2023). Penerapan algoritma K-Means *clustering* untuk identifikasi kelayakan penerima bantuan Program Keluarga Harapan (PKH) di Desa Tambaksari Ciamis. *Jurnal Mahasiswa Teknik Informatika*, 7(6). <https://doi.org/10.36040/jati.v7i6.8197>
- Mustika, P., Dewi, C., Emban, P. S., Wijayanti, G. A., Aulia, N., & Nooraeni, R. (2021). Comparison of DBSCAN and K-Means *clustering* for grouping the village status in Central Java 2020. *Jurnal Matematika, Statistika & Komputasi*, 17(3), 394-404. <https://doi.org/10.20956/j.v17i3.11704>

- Putri, F. I., Damayanti, R., & Kismiantini, K. (2022, May). Penerapan algoritma K-Means untuk mengelompokkan Kecamatan di Kabupaten Gunungkidul berdasarkan Program Keluarga Harapan. *In Prosiding Seminar Nasional Matematika dan Statistika (Vol. 2)*.
- Rahmawati, R., Fitri, T., Rahma, A. N., Marzuki, C. C., & Zukrianto, Z. (2023). Pengelompokan data penerima bantuan Program Keluarga Harapan (PKH) di Kecamatan Tuah Madani Pekanbaru menggunakan fuzzy C-Means. *Jurnal Sains Matematika dan Statistika*, 9(1), 1-10. <https://doi.org/10.24014/jsms.v9i1.19820>
- Sudibyo, N. A., Iswardani, A., Sari, K., & Suprihatiningsih, S. (2020). Penerapan data mining pada jumlah penduduk miskin di Indonesia. *Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, 1(3), 199-207. <https://doi.org/10.46306/lb.v1i3>
- Sudibyo, N. A., Iswardani, A., Sari, K., & Suprihatiningsih, S. (2020). Penerapan data mining pada jumlah penduduk miskin di Indonesia. *Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, 1(3), 199-207. <https://doi.org/10.46306/lb.v1i3>
- Sudirman, & Sulaeman, I. (2020). *Clustering* data mining algoritma K-Means dalam pengolahan nilai PAT pada mata pelajaran Bahasa Inggris, Bahasa Indonesia dan Matematika dengan pembelajaran virtual learning (Studi kasus siswa kelas XI SMA AL Maghfirah Telajung Bekasi). *SIGMA: Jurnal Teknologi Pelitas Bangsa*, 11(1), 17-20.
- Supriana, I. W., Raharja, M. A., & Bimantara, I. M. S. (2023). Pengembangan sistem prediksi bantuan Program Keluarga Harapan (PKH) berbasis machine learning. *Sintech Journal*, 6(1), 1-11. <https://doi.org/10.31598/sintechjournal.v6i1.1297>
- Surahman, A., & Hayati, U. (2023). Implementasi algoritma Naïve Bayes untuk prediksi penerima bantuan sosial. *Jurnal Mahasiswa Teknik Informatika*, 7(1). <https://doi.org/10.36040/jati.v7i1.6302>
- Tri Yunarni, B. R., Mintasrihardi, M., & Setiawati, Y. (2019). Efektivitas Program Keluarga Harapan (PKH) dalam mengurangi angka kemiskinan (Studi: Desa Daha Kecamatan Hu'u Kabupaten Dompu). *JlAP (Jurnal Ilmu Administrasi Publik)*, 7(2), 193. <https://doi.org/10.31764/jiap.v7i2.1276>
- Yulia, A. F., & Nugroho, H. W. (2022). Implementasi algoritma K-Means classifier sebagai pendukung keputusan penerima dana bantuan siswa miskin (Studi kasus: SMKN Sukoharjo). *Seminar Nasional Hasil Penelitian Dan Pengabdian Masyarakat*, 48-57.
- Yulia, A. F., & Nugroho, H. W. (2022). Implementasi algoritma K-Means classifier sebagai pendukung keputusan penerima dana bantuan siswa miskin (Studi kasus: SMKN Sukoharjo). *Seminar Nasional Hasil Penelitian Dan Pengabdian Masyarakat*, 48-57.