JDSP
Jurnal Data Science Penusa
E-ISSN: 1234-5678

Clustering Dalam Penentuan Kelayakan Penerima Bantuan Program Keluarga Harapan (PKH) di Kantor Desa Gunung Manumpak B Menggunakan K-Means

Iche Mayang Sari Br Perangin-angin[1)*], Megaria Purba, S.SI.,M.Kom[2)]
[1,2]Teknologi Informasi, STMIK Pelita Nusantara Medan
[1,2]sariichemayang@gmail.com, megariapurba18@gmail.com[2]

Received: 30 September 2025 Accepted: 25 Juni 2026 Published: 10 July 2026
Pp. 255-267

***Penulis Koresponden**

Kata Kunci: data mining, PKH, K-means, clustering

JDSP: Jurnal Data Science Penusa is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0).

Abstrak : Program Keluarga Harapan (PKH) merupakan bantuan sosial bersyarat yang ditujukan untuk meningkatkan kesejahteraan masyarakat miskin melalui dukungan pendidikan, kesehatan, dan ekonomi rumah tangga. Di Desa Gunung Manumpak B, proses penentuan kelayakan penerima bantuan masih dilakukan secara manual dan cenderung subjektif, sehingga menimbulkan potensi ketidaktepatan sasaran dan kecemburuan sosial. Penelitian ini bertujuan untuk menerapkan algoritma K-Means dalam pengelompokan data keluarga calon penerima PKH agar proses seleksi lebih objektif, efisien, dan transparan. Tahapan penelitian dilakukan mulai dari pengumpulan data, pra-pemrosesan, perhitungan jarak menggunakan *Euclidean Distance*, pembentukan klaster, hingga evaluasi hasil clustering. Sistem berbasis web dibangun menggunakan PHP dan MySQL untuk memfasilitasi proses input data, perhitungan otomatis, serta penyajian hasil analisis dalam bentuk tabel dan grafik. Data penelitian terdiri atas 50 keluarga dengan atribut sosial ekonomi seperti pekerjaan, penghasilan, jumlah tanggungan, kepemilikan rumah, dan sumber listrik. Hasil penelitian menunjukkan bahwa algoritma K-Means mampu membagi data ke dalam dua klaster utama, yaitu *layak* menerima bantuan dengan jumlah 28 keluarga (56%) dan *tidak layak* menerima bantuan dengan jumlah 22 keluarga (44%). Distribusi ini menunjukkan bahwa sebagian besar keluarga yang dianalisis memenuhi kriteria kelayakan sesuai parameter yang digunakan. Dengan demikian, penerapan K-Means terbukti efektif dalam membantu perangkat desa menentukan penerima bantuan PKH secara lebih terstruktur, adil, dan tepat sasaran, sekaligus mendukung transparansi dalam pengambilan keputusan berbasis data.

PENDAHULUAN

Perkembangan teknologi informasi berperan penting dalam meningkatkan efisiensi dan efektivitas pengelolaan data di berbagai bidang. Digitalisasi mendorong transformasi administrasi, pelayanan publik, dan pengambilan keputusan berbasis data. Dalam pemerintahan, teknologi ini mendukung transparansi, akuntabilitas, serta ketepatan sasaran, khususnya pada program bantuan sosial, melalui sistem pendukung keputusan yang mampu mengolah dan mengelompokkan data masyarakat secara objektif.

Di Kantor Desa Gunung Manumpak B, proses penentuan kelayakan penerima bantuan PKH masih dilakukan secara manual dan sangat bergantung pada pertimbangan subjektif perangkat desa. Kondisi ini sering menimbulkan permasalahan seperti ketidaktepatan sasaran, tumpang tindih penerima, hingga munculnya kecemburuan sosial antarwarga. Kurangnya pemanfaatan teknologi juga menyebabkan rendahnya validitas data serta efisiensi penyaluran bantuan. Oleh karena itu, dibutuhkan sebuah pendekatan berbasis teknologi yang mampu menganalisis dan mengelompokkan data masyarakat secara objektif sehingga calon penerima bantuan dapat ditentukan secara lebih adil, tepat sasaran, dan transparan.

Sejumlah penelitian telah menunjukkan efektivitas clustering dalam menentukan kelayakan penerima bantuan. (Oop Sofiyah, R., and Danar Dana 2023) membuktikan bahwa metode K-Means dapat mengelompokkan data penerima bantuan sosial di Kecamatan Sidoarjo sehingga memudahkan identifikasi tingkat kebutuhan ekonomi. Penelitian (Dwiguna and Bahtiar 2024) juga menggunakan K-Means untuk pengelompokan warga miskin di Kelurahan Tamalanrea, Makassar, dan menemukan bahwa algoritma ini mampu memberikan klasifikasi akurat sebagai dasar kebijakan penyaluran. Hasil-hasil tersebut menegaskan bahwa teknik data mining, khususnya clustering, layak diimplementasikan dalam konteks serupa.

Penelitian ini menggunakan algoritma K-Means karena kemampuannya mengelompokkan data ke dalam beberapa klaster berdasarkan kemiripan. Metode ini bekerja dengan menentukan centroid sebagai pusat klaster, lalu mengelompokkan data berdasarkan kedekatannya. Kelebihannya terletak pada perhitungan yang sederhana namun efektif, serta sesuai untuk data berjumlah besar dengan variabel numerik kompleks. Tujuan penelitian ini

adalah memanfaatkan K-Means untuk mengelompokkan data warga agar penentuan kelayakan penerima bantuan lebih objektif, efisien, dan transparan. Dengan sistem berbasis data mining ini, diharapkan kualitas pengambilan keputusan dalam penyaluran bantuan sosial desa meningkat sekaligus meminimalisasi potensi kesalahan distribusi.

Pendekatan clustering dengan algoritma K-Means memungkinkan Pemerintah Desa Gunung Manumpak B dalam menentukan kelayakan penerima PKH secara lebih objektif. Integrasi metode ini ke dalam sistem berbasis data mining mendukung pengelompokan masyarakat berdasarkan kondisi sosial ekonomi, sehingga proses seleksi penerima bantuan menjadi lebih efisien, transparan, dan tepat sasaran. Penerapan teknologi ini membantu perangkat desa mengambil keputusan yang akurat serta mengurangi potensi kesalahan distribusi bantuan.

TINJAUAN LITRATUR

1. Program Keluarga Harapan (PKH)

Program Keluarga Harapan (PKH) adalah bantuan sosial bersyarat yang diluncurkan sejak 2007 untuk keluarga miskin dengan tujuan meningkatkan kualitas pendidikan, kesehatan, dan kesejahteraan. PKH terbukti efektif mengurangi kemiskinan kronis serta mendorong peningkatan Indeks Pembangunan Manusia (IPM). Peserta diwajibkan memanfaatkan layanan kesehatan dan pendidikan, sementara bantuan tunai disesuaikan dengan kategori, seperti ibu hamil, anak usia dini, pelajar, penyandang disabilitas, dan lansia (Suhendri et al., 2022). Sebagai program perlindungan sosial, PKH berperan penting dalam memperkecil ketimpangan dan mendukung penanggulangan kemiskinan nasional (Setyawardani, Paat, and Lesawengen 2020).

2. Tujuan PKH

Tujuan utama Program Keluarga Harapan (PKH) adalah menanggulangi kemiskinan melalui peningkatan kualitas sumber daya manusia, khususnya bagi masyarakat sangat miskin. Dalam jangka pendek, PKH meringankan beban pengeluaran rumah tangga lewat bantuan tunai bersyarat, sedangkan dalam jangka panjang mendorong akses pendidikan, kesehatan, dan gizi guna memutus rantai kemiskinan antar generasi. Secara khusus, PKH ditujukan untuk meningkatkan kualitas layanan pendidikan dan kesehatan, taraf pendidikan, serta status gizi peserta (Setyawardani et al. 2020).

3. Data Mining

Data mining adalah proses penemuan pengetahuan tersembunyi dalam basis data dengan memanfaatkan teknik statistik, matematika, kecerdasan buatan, dan machine learning. Tujuannya adalah menemukan informasi relevan sebagai dasar pengambilan keputusan yang lebih efektif. Proses ini dikenal juga sebagai Knowledge Discovery in Database (KDD) yang meliputi tahapan pemilihan data, pembersihan, transformasi, analisis dengan algoritma, serta interpretasi hasil. Data mining telah banyak dimanfaatkan di berbagai bidang, seperti bisnis dan pemerintahan, untuk memahami pola data, meningkatkan efisiensi, serta mendukung keputusan strategis (Cahyo 2023).

Dalam data mining, tugas analisis dapat dikelompokkan menjadi tiga, yaitu deskripsi untuk menggambarkan pola atau tren, estimasi untuk menentukan nilai numerik berdasarkan variabel lain, dan prediksi untuk memperkirakan kejadian di masa depan (Cahyo, 2023). Proses data mining umumnya melalui tahapan pemilihan data, pembersihan, transformasi, analisis dengan algoritma, serta interpretasi hasil (Zyen et al. 2023). Salah satu teknik penting adalah clustering, yaitu pengelompokan data ke dalam beberapa klaster berdasarkan kemiripan. Clustering membantu memahami data dalam jumlah besar secara sistematis, dengan hasil yang dipengaruhi jumlah klaster yang ditentukan. Analisis ini bertujuan mengidentifikasi pola tersembunyi dan menyajikan struktur alami data secara lebih akurat (Sudarsono & Lestari, 2021 ; (Hutagalung 2022).

4. Algoritma Apriori

K-Means merupakan algoritma clustering yang bersifat iteratif dengan menentukan jumlah klaster (k) sebagai pusat awal atau centroid. Algoritma ini mudah diimplementasikan, relatif cepat, dan efisien dalam mengolah data berskala besar. Namun, kelemahannya terletak pada pemilihan centroid awal yang bergantung pada inisialisasi data (Yulia and Silalahi 2021). Pada dasarnya, K-Means mengelompokkan data tanpa label (unsupervised) dengan menghitung jarak tiap data ke centroid menggunakan Euclidean Distance. Adapun prosedur perhitungan Algoritma K- Means *Clustering* analysis diuraikan sebagai berikut (Rohmah 2021). Menurut (Tabianan, Velu, and Ravi 2022), algoritma K-Means dapat digunakan untuk segmentasi pelanggan berdasarkan perilaku pembelian sehingga membantu mempertahankan pelanggan, meningkatkan keuntungan, serta mendukung strategi pemasaran.

- Pilih secara acak k buah data sebagai pusat cluster:
- Jarak antara data dan pusat cluster dihitung menggunakan Euclidian Distance. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai berikut:

  $D_{euclidean}\ (x,y) = \sqrt{\sum(xi - yi)^2}$ ·····(1)

  Keterangan:

  d (x,y) : jarak antara data pada titik x

E-ISSN: 1234-5678

dan y x : titik data objek
y : titik data centroid
i : jumlah atribut data

- Data ditempatkan dalam cluster yang terdekat, dihitung dari tengah cluster
- Pusat cluster baru akan ditentukan bila semua data telah ditetapkan dalam cluster terdekat. Rumus menghitung titik pusat cluster baru:

$$v = \sum_{n=1}^{n} xi; i = 1,2,3 \dots \dots \dots \dots \dots n \quad (2)$$

Dimana:
v adalah centroid pada cluster
Xi adalah objek ke-i
N adalah banyaknya objek/jumlah objek yang menjadi anggota cluster

- Proses penentuan pusat cluster dan penempatan data dalam cluster diulangi sampai nilai centroid tidak berubah lagi.

5. Database

Database adalah himpunan data yang tersusun dalam bentuk atribut dan record untuk menghasilkan informasi yang dibutuhkan pengguna. Pengelolaan database dilakukan dengan Database Management System (DBMS), yang memudahkan penyimpanan, pencarian, pembaruan, serta penghapusan data secara terstruktur. Dengan demikian, database berfungsi sebagai media penyimpanan yang sistematis untuk mengolah data menjadi informasi yang bermanfaat (Kunci 2022).

6. Web

Website adalah kumpulan halaman yang menyajikan informasi dalam bentuk teks, gambar, animasi, suara, atau kombinasi, yang dapat bersifat statis maupun dinamis. Halaman-halaman tersebut saling terhubung melalui hyperlink sebagai penghubung dan hiperteks sebagai media navigasinya, sehingga membentuk satu kesatuan struktur informasi (Solahudin 2021).

7. AppServ

AppServ adalah paket perangkat lunak yang memudahkan instalasi komponen server web, seperti Apache, PHP, phpMyAdmin, dan MySQL. Sementara itu, XAMPP berfungsi sebagai server lokal dengan komponen utama berupa Apache, MySQL, PHP, dan Perl, yang memungkinkan pengembangan dan pengujian aplikasi web secara lokal sebelum diterapkan ke server sebenarnya .(Anjani & Astuty, 2023).

8. PHP

PHP (Hypertext Preprocessor) merupakan bahasa pemrograman open source yang diperkenalkan oleh Rasmus Lerdorf pada tahun 1995 dan banyak dimanfaatkan dalam pembuatan aplikasi web dinamis. Sebagai bahasa scripting di sisi server (*server-side*), PHP memproses data pada server terlebih dahulu sebelum hasilnya ditampilkan di browser pengguna. Bahasa ini dapat disisipkan dalam HTML, mudah dipelajari, serta memiliki sintaks mirip dengan C, Java, dan Perl (Solahudin 2021) ; (Maya Sufiani, 2021 : 24).

9. MySQL

MySQL adalah sistem manajemen basis data relasional (RDBMS) open source yang dilisensikan di bawah GNU General Public License (GPL). Database ini menggunakan struktur tabel, kolom, dan baris untuk menyimpan serta mengelola data, dan banyak digunakan pada pengembangan web karena kemampuannya mengolah data dalam jumlah besar dengan cepat. MySQL juga mendukung pembuatan fitur dinamis seperti form login, manajemen kata sandi, dan formulir interaktif lainnya (Solahudin 2021).

10. Unified Modeling Language (UML)

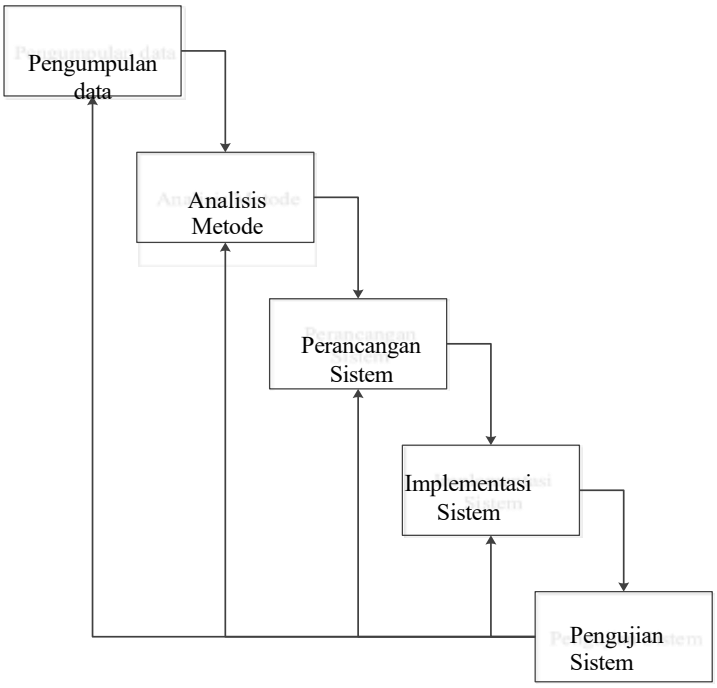
Unified Modeling Language (UML) merupakan bahasa pemodelan visual yang berfungsi untuk menggambarkan sistem perangkat lunak berbasis objek dengan cara yang terstandarisasi. UML berfungsi sebagai media komunikasi dalam pengembangan sistem, membantu pengembang menyusun rancangan secara sistematis (Alfina and Harahap 2019). Beberapa jenis diagram yang digunakan meliputi:

- *Use Case Diagram* untuk menggambarkan fungsi sistem dan aktor yang terlibat,
- *Class Diagram* untuk menjelaskan hubungan antar kelas beserta atribut dan operasinya,
- *Activity Diagram* untuk memodelkan alur kerja atau aktivitas dalam sistem,
- *Sequence Diagram* untuk menunjukkan interaksi antar objek berdasarkan urutan waktu (Janis et al. 2020).

METODE PENELITIAN

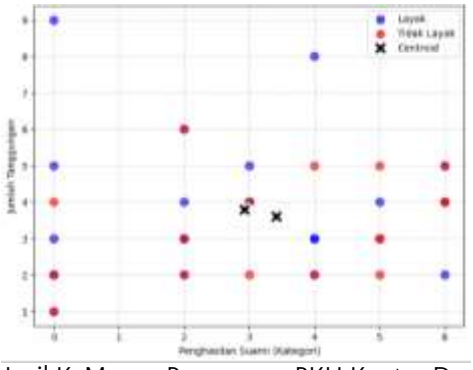
Penelitian ini menggunakan algoritma K-Means untuk mengelompokkan data masyarakat dalam menentukan kelayakan penerima bantuan PKH di Desa Gunung Manumpak B. Tahapan penelitian dilakukan

melalui pengumpulan data, pra-pemrosesan, perhitungan jarak menggunakan Euclidean Distance, pembentukan klaster, hingga evaluasi hasil clustering (Yuliana, Fathoni, and Kurniawati 2022):



Gambar 1 Karangka Diagram Alur

1. **Pengumpulan Data**
   Data diperoleh melalui studi pustaka, observasi, dan wawancara. Data utama meliputi kondisi ekonomi, sosial, serta kriteria penerima PKH.
2. **Analisis Metode K-Means**
   Data masyarakat diolah dengan algoritma K-Means untuk dikelompokkan ke dalam beberapa klaster berdasarkan tingkat kelayakan penerima bantuan.
3. **Perancangan Sistem**
   Sistem dirancang menggunakan pemodelan UML (use case, activity, class, dan sequence diagram) serta desain basis data dan antarmuka.
4. **Implementasi Sistem**
   Implementasi dilakukan dengan bahasa pemrograman PHP dan basis data MySQL untuk menghasilkan sistem pengelompokan penerima PKH.
5. **Pengujian Sistem**
   Pengujian dilakukan dengan metode *black box testing* untuk memastikan semua fungsi berjalan sesuai kebutuhan pengguna.



Gambar 2 *Clustering* Hasil K-Means Penerapan PKH Kantor Desa Gunung Manupak B

HASIL PENELITIAN

1. Data Penjualan

Data yang digunakan dalam penelitian ini adalah data masyarakat calon penerima Program Keluarga Harapan (PKH) di Desa Gunung Manumpak B. Data penelitian terdiri 50 data masyaraka, namun hanya 25 data masyarakat yang ditampilkan dalam jurnal ini,

Table 1
Hasil Transformasi Data

No	Nama	V1	V2	V3	V4	V5	V6	V7
1	Kevin Sinaga	5	5	5	5	5	1	2
2	Mhd Irham	6	8	6	7	2	1	2
3	Winda Maria Ulfa	6	8	6	6	3	1	2
4	Rudi Alamsyah	6	6	5	5	2	2	2
5	Muslimah	6	5	7	6	4	2	2
6	Dini Aflina	5	8	6	7	3	2	2
7	Muhibuddin	5	8	5	7	3	2	2
8	Syarifah Annisa	5	8	5	7	3	1	2
9	Sufratul Aini	4	8	6	7	4	1	2
10	Megi Herman	6	6	5	6	5	4	2
11	Fatimah Zuhra	6	8	6	7	8	1	2
12	Naswatun Zikra	4	4	6	6	6	1	2
13	Putri Raudhah	5	5	6	7	4	1	2
14	Khairil Ihsan	5	5	6	6	2	2	2
15	Yuhardi	6	8	6	7	4	1	2
16	Lisa Nazmi	8	6	7	6	2	1	2
17	Rhafiri Rahman	3	8	5	7	6	1	2
18	Muhammad Zuhdi Arifin	5	8	6	7	3	1	2
19	Wahyu Sudirman	5	8	7	5	5	1	2
20	Lailatul Fajar	6	5	6	6	5	1	2
21	Novita Ristiani	6	1	5	6	4	1	2
22	Muhammad Nur Hidayat	4	5	6	7	2	1	2
23	Rizki Oktara	4	5	5	7	3	1	2
24	Gita Anggreini Pratiwi	8	8	7	7	1	4	2
25	Neni Ralulita Br. Siregar	5	8	6	7	2	1	2

Kemudian untuk menentukan nilai centroid 2 (Tidak Layak) adalah sebagai berikut :
Jarak data masyarakat 1 Kevin Sinaga dengan nilai rata-rata:

$d(ai,bj) = \sqrt{(5-5{,}4)^2 + (5-6{,}64)^2 + (5-5{,}92)^2 + (5-6{,}56)^2 + (5-3{,}7)^2 + (1-1{,}42)^2 + (2-2)^2} = 2{,}82772$

Jarak data masyarakat 2 Mhd Irham dengan nilai rata-rata:

$d(ai,bj) = \sqrt{(6-5{,}4)^2 + (8-6{,}64)^2 + (6-5{,}92)^2 + (7-6{,}56)^2 + (2-3{,}7)^2 + (1-1{,}42)^2 + (2-2)^2}$

= 2,340085 dan seterusnya
Maka dengan demikian diperoleh nilai Centroid 2 Tidak Layak dengan nilai tertinggi 5,796206 yaitu data dari Novita Ristiani. Selanjutnya adalah mencari nilai Centroid 1 :

Untuk mencari nilai Centroid 1 :

- Jarak data masyarakat 1 Kevin Sinaga dengan nilai rata-rata:
  $d(ai,bj) = \sqrt{(5-6)^2 + (5-1)^2 + (5-5)^2 + (5-6)^2 + (5-4)^2 + (1-1)^2 + (2-2)^2} = 4{,}3588$ dan seterusnya. Maka dengan demikian dapat diketahui nilai dari Centroid awal dapat dilihat pada Tabel 2.

Tabel 2
Centroid Awal

Centroid	Pekerjaan Suami	Pekerjaan Istri	Penghasilan Suami	Penghasilan Istri	Jumlah Tanggungan	Kepemilikan Rumah	Sumber Listrik	Cluster
1	8	8	7	7	1	4	2	C1 (Layak)
2	6	1	5	6	4	1	2	C2 (Tidak Layak)

Jarak data masyarakat 1 Kevin Sinaga dengan *centroid* 1, 2 :

$d(ai,bj) = \sqrt{(5-8)^2 + (5-8)^2 + (5-7)^2 + (5-7)^2 + (5-1)^2 + (1-4)^2 + (2-2)^2} = 7{,}14$

$d(ai,bj) = \sqrt{(5-6)^2 + (5-1)^2 + (5-5)^2 + (5-6)^2 + (5-4)^2 + (1-1)^2 + (2-2)^2} = 4{,}36$ dan seterusnya.

Tabel 3
Hasil Perhitungan Jarak Data Dengan *Centroid* Awal Iterasi 1

No	Nama	*Cluster* 1	*Cluster* 2	Jarak Terdekat	*Cluster*
1	Kevin Sinaga	7,14	4,36	4,36	2
2	Mhd Irham	3,87	7,42	3,87	1
3	Winda Maria Ulfa	4,36	7,14	4,36	1
4	Rudi Alamsyah	4,58	5,57	4,58	1
5	Muslimah	5,20	4,58	4,58	2
6	Dini Aflina	4,24	7,35	4,24	1
7	Muhibuddin	4,58	7,28	4,58	1
8	Syarifah Annisa	5,10	7,21	5,10	1
9	Sufratul Aini	5,92	7,42	5,92	1
10	Megi Herman	5,39	5,92	5,39	1
11	Fatimah Zuhra	7,94	8,19	7,94	1
12	Naswatun Zikra	8,25	4,24	4,24	2
13	Putri Raudhah	6,08	4,36	4,36	2
14	Khairil Ihsan	5,00	4,80	4,80	2
15	Yuhardi	4,80	7,14	4,80	1
16	Lisa Nazmi	3,87	6,08	3,87	1
17	Rhafiri Rahman	7,94	7,94	7,94	2
18	Muhammad Zuhdi Arifin	4,80	7,28	4,80	1
19	Wahyu Sudirman	6,16	7,48	6,16	1
20	Lailatul Fajar	6,32	4,24	4,24	2
21	Novita Ristiani	8,72	0,00	0,00	2
22	Muhammad Nur Hidayat	6,00	5,10	5,10	2
23	Rizki Oktara	6,48	4,69	4,69	2
24	Gita Anggreini Pratiwi	0,00	8,72	0,00	1
25	Neni Ralulita Br. Siregar	4,47	7,48	4,47	1

Setelah semua data masuk ke cluster terdekat, hitung pusat cluster baru dari rata-rata anggotanya untuk iterasi berikutnya:

Tabel 4
Sampel Centroid

Centroid	Pekerjaan Suami	Pekerjaan Istri	Penghasilan Suami	Penghasilan Istri	Jumlah Tanggungan	Kepemilikan Rumah	Sumber Listrik	Cluster
1	5,58	7,61	5,97	6,68	3,58	1,42	2,00	C1 (Layak)
2	5,11	5,05	5,84	6,37	3,89	1,42	2,00	C2 (Tidak Layak)

Jarak data masyarakat 1 Kevin Sinaga dengan *centroid* 1, 2 :

$d(ai,bj)= \sqrt{(5-5{,}58)^2+(5-7{,}61)^2+(5-5{,}97)^2+(5-6{,}68)^2+(5-3{,}58)^2+(1-1{,}42)^2+(2-2)^2}= 3{,}62$ dan seterusnya.

Tabel 5
Hasil Perhitungan Jarak Data Dengan *Centroid* Awal Iterasi 2

No	Nama	*Cluster* 1	*Cluster* 2	Jarak Terdekat	*Cluster*
1	Kevin Sinaga	3,62	2,00	2,00	2
2	Mhd IRHAM	1,76	3,70	1,76	1
3	Winda Maria Ulfa	1,14	3,26	1,14	1
4	Rudi Alamsyah	3,06	2,86	2,86	2
5	Muslimah	3,01	1,62	1,62	2
6	Dini Aflina	1,13	3,20	1,13	1
7	Muhibuddin	1,48	3,31	1,48	1
8	Syarifah Annisa	1,43	3,28	1,43	1
9	Sufratul Aini	1,76	3,24	1,76	1
10	Megi Herman	3,58	3,23	3,23	2
11	Fatimah Zuhra	4,49	5,19	4,49	1
12	Naswatun Zikra	4,69	2,66	2,66	2
13	Putri Raudhah	2,76	0,79	0,79	2

No	Nama	*Cluster* 1	*Cluster* 2	Jarak Terdekat	*Cluster*
14	Khairil Ihsan	3,23	2,02	2,02	2
15	Yuhardi	0,88	3,18	0,88	1
16	Lisa Nazmi	3,56	3,81	3,56	1
17	Rhafiri Rahman	3,73	4,34	3,73	1
18	Muhammad Zuhdi Arifin	1,05	3,18	1,05	1
19	Wahyu Sudirman	2,56	3,65	2,56	1
20	Lailatul Fajar	3,11	1,54	1,54	2
21	Novita Ristiani	6,76	4,27	4,27	2
22	Muhammad Nur Hidayat	3,48	2,33	2,33	2
23	Rizki Oktara	3,30	1,82	1,82	2
24	Gita Anggreini Pratiwi	4,53	5,82	4,53	1
25	Neni Ralulita Br. Siregar	1,81	3,59	1,81	1

Setelah semua data ditempatkan pada cluster terdekat, pusat cluster dihitung kembali dari rata-rata anggotanya sehingga diperoleh centroid baru untuk iterasi berikutnya. Nilai rata-rata tersebut ditunjukkan pada tabel berikut:

Tabel 6
Sampel Centroid

Centroid	Pekerjaan Suami	Pekerjaan Istri	Penghasilan Suami	Penghasilan Istri	Jumlah Tanggungan	Kepemilikan Rumah	Sumber Listrik	Cluster
1	5,24	7,93	5,97	6,83	3,93	1,31	2,00	C1 (Layak)
2	5,62	4,86	5,86	6,19	3,38	1,57	2,00	C2 (Tidak Layak)

Proses pembaruan centroid dan perhitungan jarak dilanjutkan hingga **iterasi keenam**. Pada iterasi ke-5 dan ke-6 tidak terjadi perubahan, sehingga proses pengelompokan dihentikan. Hasil akhir menunjukkan dua cluster, yaitu *layak* dan *tidak layak* menerima bantuan.

Tabel 7
Data Layak dan Tidak Layak

Cluster 1 (Layak)		*Cluster* 2 (Tidak Layak)	
No	**Nama**	**No**	**Nama**
]	Mhd Irham	1	Kevin Sinaga
2	Winda Maria Ulfa	2	Rudi Alamsyah
3	Dini Aflina	3	Muslimah
4	Muhibuddin	4	Megi Herman
5	Syarifah Annisa	5	Naswatun Zikra
6	Sufratul Aini	6	Putri Raudhah
7	Fatimah Zuhra	7	Khairil Ihsan
8	Yuhardi	8	Lisa Nazmi
9	Rhafiri Rahman	9	Lailatul Fajar
10	Muhammad Zuhdi Arifin	10	Novita Ristiani
11	Wahyu Sudirman	11	Muhammad Nur Hidayat
12	Gita Anggreini Pratiwi	12	Rizki Oktara
13	Neni Ralulita Br. Siregar	13	Juli Sarifah Hanum
14	Anzan Hario	14	Faratul Nanda
15	Mutia Rahmi	15	Ryan Miranda
16	Chairul Bariyah	16	Mika Indawati
17	Muhammad Akmal	17	Yulia
18	Resa Issanti	18	Rahmat Maulidan
19	Muhammad Rizal Bakri	19	Muhammad Yanis
20	Salma Sahara	20	Frillyanti Chairunnisa
21	Ari Maulana	21	Nova Rosmalita
22	Maulina	22	Hanif Syah Reza Funna
23	Muhammad Rizky		
24	Recha Ariska		
25	Zara Zettyra. Rd		
26	Risna Handayyani Is		
27	Elia Rahmi		

28	Lia Risma		

Hasil ini menggambarkan kelayakan penerimaan bantuan PKH. Aplikasi yang dikembangkan menyediakan antarmuka yang intuitif untuk input data dan penentuan kelayakan penerimaan PKH, dengan menu utama sebagai pusat navigasi menuju berbagai fitur. Sistem aplikasi ini dilengkapi dengan beberapa menu utama yang dirancang untuk mendukung proses penentuan minat beli menggunakan metode K-Means. Adapun menu-menu utama yang tersedia antara lain sebagai berikut:

1. Menu *Login*
   Halaman login merupakan antarmuka awal aplikasi yang berfungsi sebagai *form input username* dan *password* bagi admin untuk mengakses sistem.



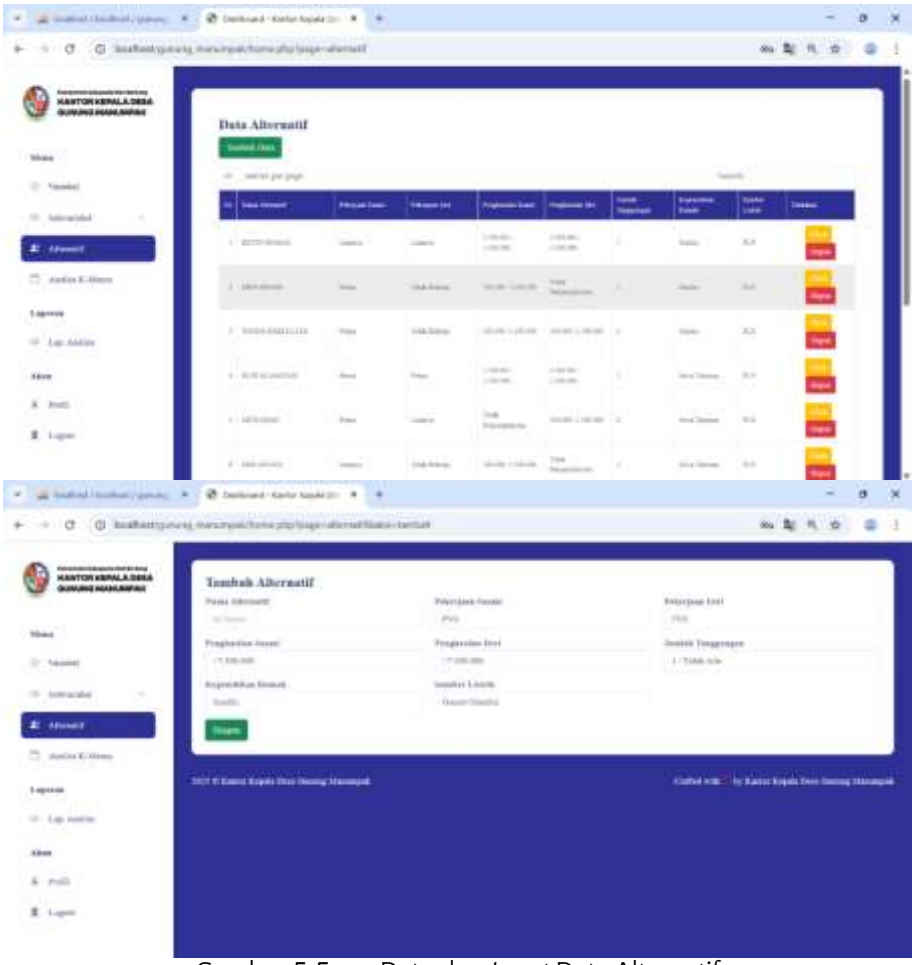
Gambar 2 menu *Login*

2. Menu Profil
   Form menu profil menampilkan informasi admin yang sedang login, termasuk identitas dan hak akses. Fitur ini mendukung pengelolaan akun dan keamanan system.



Gambar 4 Menu Profil

3. Form Data dan Input Data Alternatif
   Form Data Alternatif digunakan untuk menampilkan data mentah penerima bantuan PKH, sedangkan Form Input Data Alternatif berfungsi untuk menambahkan data masyarakat calon penerima PKH. Kedua form ini memuat atribut penting seperti pendapatan, jumlah tanggungan, pekerjaan, dan kondisi rumah, yang menjadi input utama sebelum diproses dengan algoritma K-Means.



Gambar 5 *From* Data dan *Input* Data Alternatif

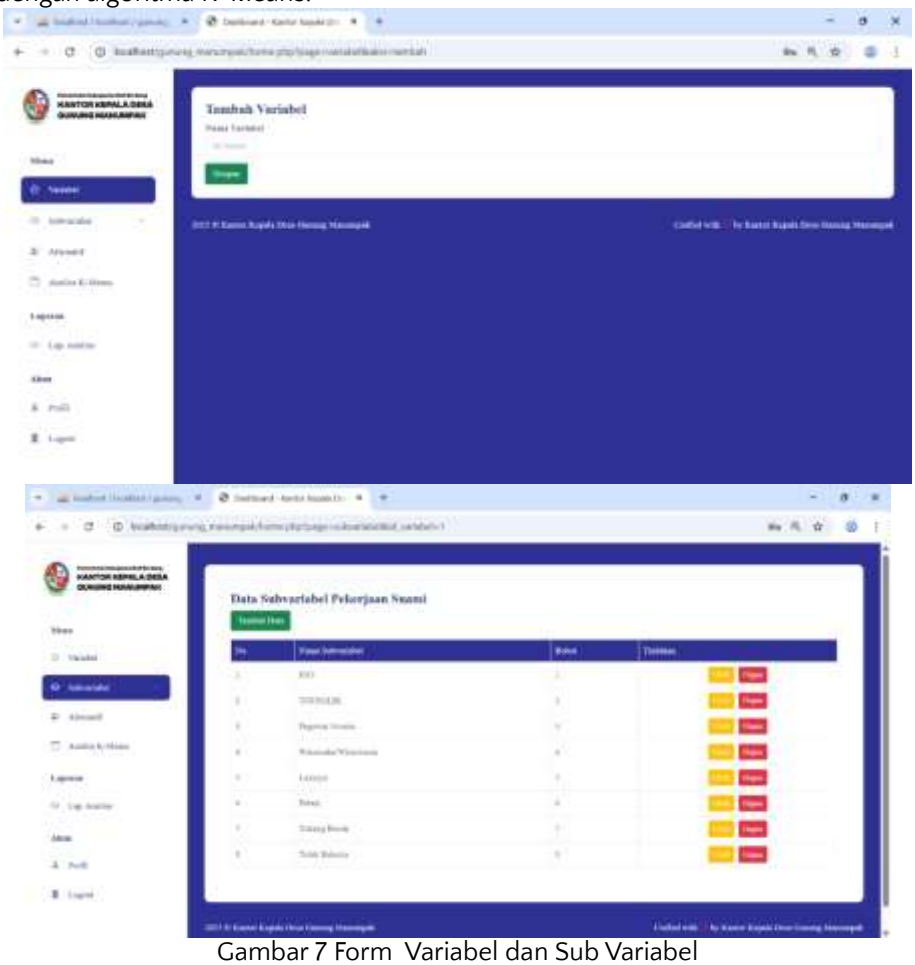
4. Form Data Variabel
Form data variabel digunakan untuk mengelola variabel yang dipakai dalam analisis clustering, seperti pekerjaan, penghasilan, jumlah tanggungan, kepemilikan rumah, dan sumber listrik.



Gambar 6 Form Data Variabel

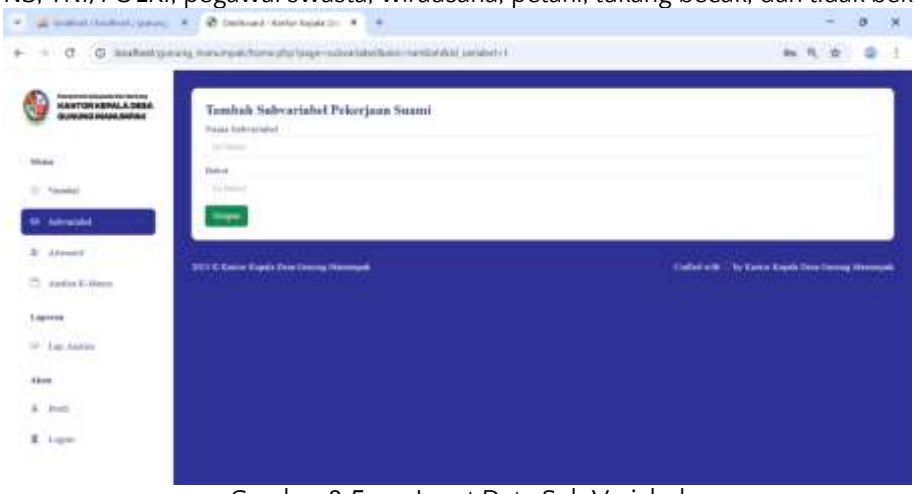
5. Form Variabel dan Sub Variabel
Form Input Variabel digunakan untuk memasukkan dan mengelola variabel yang dipakai dalam analisis clustering, sedangkan Form Data Sub Variabel berfungsi untuk menampilkan serta mengatur sub variabel

yang mendukung variabel utama. Kedua form ini saling melengkapi dalam mempersiapkan data sebelum diproses dengan algoritma K-Means.



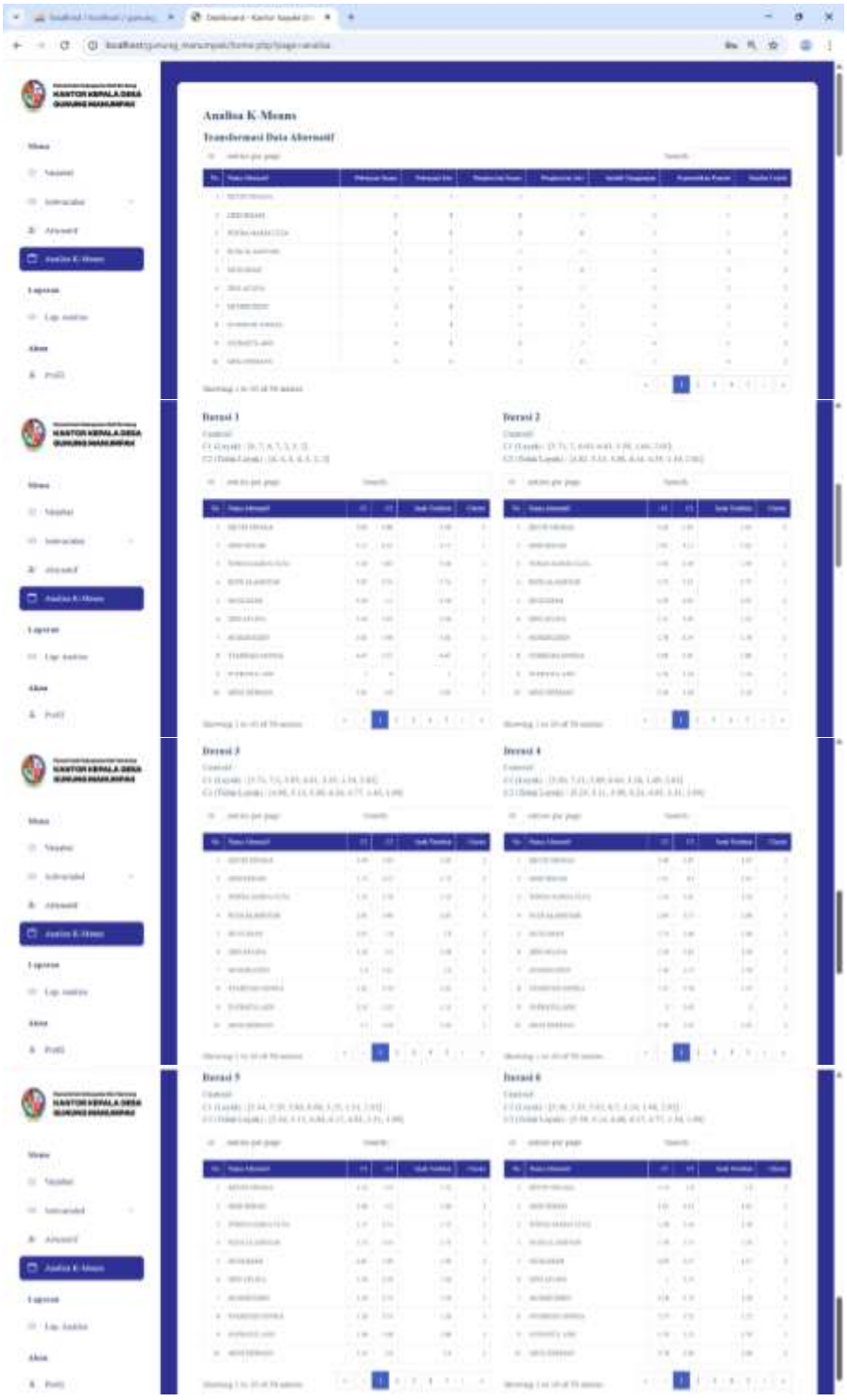
Gambar 7 Form Variabel dan Sub Variabel

6. *Form* Input Data Sub Variabel
Form ini digunakan untuk memasukkan sub variabel dari setiap variabel, misalnya pekerjaan suami dengan kategori PNS, TNI/POLRI, pegawai swasta, wirausaha, petani, tukang becak, dan tidak bekerja.



Gambar 8 *Form* Input Data Sub Variabel

7. *Form* Analisis Metode
Tampilan ini merupakan tampilan *form* Analisa metode yang berfungsi untuk melakukan proses Analisa metode.


Analisa K-Means
Transformasi Data Alternatif
Iterasi 1
Iterasi 2
Iterasi 3
Iterasi 4
Iterasi 5
Iterasi 6



Gambar 9 From Analisis Metode

8. *Form* Laporan Analisa
Form ini berfungsi menampilkan laporan hasil analisis, di mana saat admin memilih opsi laporan analisis, sistem akan menampilkan data laporan tersebut.



Gambar 10 From Laporan Analisa

KESIMPULAN

Penelitian ini menunjukkan bahwa algoritma K-Means berhasil diterapkan untuk mengelompokkan data masyarakat Desa Gunung Manumpak B berdasarkan variabel sosial ekonomi seperti pekerjaan, penghasilan, jumlah tanggungan, kepemilikan rumah, dan sumber listrik. Hasil clustering menghasilkan dua kategori, yaitu layak dan tidak layak menerima bantuan. Sistem berbasis web yang dibangun dengan PHP dan MySQL terbukti mampu mengolah data secara terstruktur, menampilkan hasil pengelompokan, serta memvisualisasikannya secara otomatis. Namun, penelitian ini masih memiliki keterbatasan pada jumlah data yang digunakan serta penerapan hanya satu algoritma. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar serta membandingkan dengan metode lain agar hasil pengelompokan lebih akurat dan komprehensif.

REFERENCES

Alfina, Ommi, and Fitriana Harahap. 2019. "Pemodelan Uml Sistem Pendukung Keputusan Dalam Menetukan Siswa Siswa Tunagrahita." *Methomika : Jurnal Manajemen Informatika & Komputerisasi Akuntansi* 3(2):143–50.

Cahyo, Arisa Dwi. 2023. "METODE NAIVE BAYES UNTUK KLASIFIKASI MASA STUDI SARJANA | Jurnal Teknologi Pintar." 3(4).

Dwiguna, Eddiwan, and Agus Bahtiar. 2024. "Penerapan Data Mining Untuk Menentukan Penerima Bantuan Blt Menggunakan Metode Clustering K-Means Pada Desa Pamulihan." *JATI (Jurnal Mahasiswa Teknik Informatika)* 8(2):1382–88. doi: 10.36040/jati.v8i2.9029.

Hutagalung, Juniar. 2022. "Pemetaan Siswa Kelas Unggulan Menggunakan Algoritma K-Means Clustering." *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)* 9(1):606–20. doi: 10.35957/jatisi.v9i1.1516.

Janis, Janiver W., Dringhuzen J. Mamahit, Brave A. Sugiarso, Arthur M. Rumagit, Teknik Elektro, Universitas Sam, and Ratulangi Manado. 2020. "Rancang Bangun Aplikasi Online Sistem Pemesanan Jasa Tukang Bangunan Berbasis Lokasi." *Jurnal Teknik Informatika* 15(1):1–12. doi: 10.35793/jti.15.1.2020.29023.

Kunci, Kata. 2022. "Inventaris Merupakan Kegiatan Mencatat Barang Dengan Benar Agar Memudahkan Kegiatan Pengawasan Jika Sewaktu-Waktu Dibutuhkan Arsip . Pengolahan Data Inventaris Perlu Dilakukan Agar Tertib Dokumen Dan Tertib Administrasi . Di Kantor Desa Salulemo Masih Men." 12.

Oop Sofiyah, Siti, Nining R., and Raditya Danar Dana. 2023. "Analisis Efektivitas Pelayanan Publik Menggunakan K-Means Clustering Di Kecamatan Sukagumiwang." *JATI (Jurnal Mahasiswa Teknik Informatika)* 7(2):1291–96. doi: 10.36040/jati.v7i2.6536.

Rohmah, Ai. 2021. "Analisis Penentuan Hambatan Pembelajaran Daring Dengan Metode Algoritma K-Means Clustering (Studi Kasus : Smk Yaspim Gegerbitung)." *Jurnal Rekayasa Teknologi Nusa Putra* 4(2):30–35. doi: 10.52005/rekayasa.v4i2.122.

Setyawardani, Diyah Tri Rezki, Cornelius J. Paat, and Lisbeth Lesawengen. 2020. "Dampak Bantuan PKH Terhadap Masyarakat Miskin Di Kelurahan Bumi Nyiur Kecamatan Wanea Kota Manado." *Jurnal Kebiijakan Publik* 13(2):1–14.

Solahudin, Muhamad. 2021. "Rancang Bangun Sistem Informasi Akademik Sekolah (SIAS) Berbasis Website." *Journal of Computer and Information Technology* 4(2):107–13.

Sudarsono, Bernadus Gunawan, and Sri Poedji Lestari. 2021. "Clustering Penerima Beasiswa Yayasan Untuk Mahasiswa Menggunakan Metode K-Means." *Jurnal Media Informatika Budidarma* 5(1):258. doi: 10.30865/mib.v5i1.2670.

Tabianan, Kayalvily, Shubashini Velu, and Vinayakumar Ravi. 2022. "K-Means Clustering Approach for Intelligent Customer Segmentation Using Customer Purchase Behavior Data." *Sustainability (Switzerland)* 14(12):1–15. doi: 10.3390/su14127243.

Yulia, and Mesri Silalahi. 2021. "Penerapan Data Mining Clustering Dalam Mengelompokan Buku Dengan Metode K-Means." *Indonesian Journal of Computer Science* 10(1). doi: 10.33022/ijcs.v10i1.3008.

Yuliana, Devi Tri, M. Ivan Ariful Fathoni, and Naning Kurniawati. 2022. "Penentuan Penerima Kartu Indonesia Pintar KIP Kuliah Dengan Menggunakan Metode K-Means Clustering." *Journal Focus Action of Research Mathematic (Factor M)* 5(1):127–41. doi: 10.30762/f_m.v5i1.570.

Zyen, Akhmad Khanif, Gentur Wahyu Nyipto Wibowo, Nur Aeni Widiastuti, and Buang Budi Wahono. 2023. "Klasterisasi Penerima Bantuan Fasilitas Sekolah Di MI Datuk Singaraja Menggunakan Metode K-Means." *Jurnal Teknik Informatika* 2(1):25–38.