JDSP
Jurnal Data Science Penusa

Penerapan Algoritma K-Means Clustering Pengelompokkan Masyarakat Kurang Mampu Pada Desa Sidoarjo Ii Ramunia

Jhosi Sihaloho [1)*], Yuda Perwira[2)]
[1,2)] STMIK Pelita Nusantara, Teknik Informatika,
[1)] sihalohojhosi76@gmail.com@gmail.com
[2)] yudaperwira25@gmail.com

Received: 14 September 2024 Accepted: 20 Februari 2025 Published: 22 Februari 2025
Pg. 347-353

OPEN ACCESS

***Penulis Koresponden**

Kata Kunci: K-Means Clustering, Pengelompokan, Masyarakat Kurang Mampu.

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

Abstrak : Penerapan algoritma K-Means Clustering dalam pengelompokan masyarakat kurang mampu di Desa Sidoarjo II Ramunia merupakan langkah strategis untuk memahami dan mengidentifikasi kebutuhan masyarakat yang membutuhkan perhatian lebih. Penelitian ini bertujuan untuk mengelompokkan individu berdasarkan karakteristik sosial ekonomi, seperti pendapatan, pendidikan, dan jumlah anggota keluarga. Dengan menggunakan metode K-Means, data yang dikumpulkan dari survei masyarakat diolah untuk menghasilkan kluster yang representatif. Proses ini dimulai dengan pengumpulan data primer melalui wawancara dan kuesioner, yang kemudian dianalisis menggunakan perangkat lunak statistik. Hasil dari pengelompokan ini menunjukkan adanya beberapa kluster yang mencerminkan kondisi masyarakat kurang mampu dengan karakteristik yang berbeda. Misalnya, satu kluster mungkin terdiri dari keluarga dengan pendapatan rendah tetapi memiliki tingkat pendidikan yang lebih tinggi, sementara kluster lain mungkin mencakup keluarga dengan pendapatan rendah dan pendidikan yang lebih rendah. Temuan ini memberikan wawasan yang berharga bagi pemerintah desa dalam merancang program bantuan sosial yang lebih tepat sasaran dan efektif. Dengan memahami kebutuhan spesifik dari setiap kluster, intervensi yang dilakukan dapat lebih terarah dan berdampak positif bagi peningkatan kesejahteraan masyarakat. Penelitian ini diharapkan dapat menjadi acuan bagi penelitian lebih lanjut dalam bidang pengelompokan sosial dan pengembangan masyarakat.

PENDAHULUAN

Kemajuan teknologi informasi sudah semakin berkembang pesat disegala bidang kehidupan. Banyak sekali data yang dapat dihasilkan oleh teknologi informasi yang canggih, mulai dari bidang industri, ekonomi, ilmu dan teknologi maupun berbagai aspek kehidupan lainnya. Perkembangan teknologi mendorong perkembangan manusia dalam melakukan aktivitas, Salah satu manfaat dari perkembangan teknologi ini dapat mempermudah memperoleh informasi yang cepat. Perkembangan teknologi tidak akan dapat dihasilkan tanpa ada campur tangan sumber manusia untuk mengelola dan merawatnya dengan baik. Untuk mendapatkan dan menghasilkan informasi, komputer dan teknologinya adalah alat bantu yang paling tepat, salah satunya dalam memperoleh informasi pengelompokkan masyarakat kurang mampu pada Desa Sidoarjo II Ramunia.

Kantor Desa Sidoarjo Ramunia II merupakan pusat administratif yang bertanggung jawab untuk melayani kebutuhan warga, mengelola administrasi desa, dan menyediakan layanan kepada masyarakat sekitar. Masyarakat kurang mampu dapat diklasifikasikan menjadi dua kelompok, yakni mereka yang memiliki pendapatan rendah dan yang rentan terhadap kemiskinan. Hal ini melibatkan identifikasi dan klasifikasi individu atau kelompok yang menghadapi keterbatasan ekonomi sehingga memerlukan dukungan tambahan. Kriteria untuk menentukan tingkat kesejahteraan biasanya mencakup pendapatan, aset, akses terhadap

pendidikan, kesehatan, dan layanan dasar lainnya. Begitu pula pada Desa Sidoarjo II Ramunia, melakukan pengelompokkan masyarakat kurang mampu pada wilayahnya, agar dapat merencanakan serta menyalurkan bantuan kepada masyarakat tersebut.

Namun, Pihak Desa Sidoarjo II Ramunia menghadapi kendala signifikan akibat jumlah penduduk yang besar dan luasnya wilayah desa yang harus didata. Situasi ini dapat menyebabkan kesulitan dan potensi terjadinya kesalahan dalam penentuan atau pengelompokkan masyarakat kurang mampu. Dengan banyaknya data yang harus diolah, perlu dilakukan upaya ekstra untuk memastikan keakuratan dan keberlanjutan proses identifikasi serta pemilihan kriteria yang tepat agar bantuan dan kebijakan yang diterapkan dapat mencapai sasaran dengan lebih efektif. Oleh karena itu diperlukan suatu sistem yang dapat mengelompokkan data, digunakan teknik data mining untuk mengatasi permasalahan diatas.

Data mining adalah serangkaian proses untuk menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data (Tarigan et al., 2022). Data mining pada sistem klasifikasi atau pengelompokkan ada banyak metode yang dapat digunakan salah satunya adalah metode K-Means Clustering. Metode K-Means Clustering adalah metode data yang mempartisi data ke dalam bentuk satu atau lebih cluster sehingga data yang mempunyai karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama yang memiliki karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain (Andrianti & Firmansyah, 2020). Algoritma K-Means dipilih karena efisien, sederhana, dan dapat menangani data besar dengan struktur kelompok yang jelas. Keefisienannya dalam pemrosesan dan pemahaman visual memudahkan identifikasi pola-pola dalam data masyarakat kurang mampu. Meskipun memiliki keterbatasan terhadap outlier, K-Means tetap memberikan wawasan yang berharga tentang distribusi masyarakat dalam kelompok-kelompok tertentu.

Berdasarkan penelitian terdahulu yang berjudul "Penerapan Clustering Data Kurang Mampu Di Desa Situmekar Menggunakan Algoritma K-Means" disimpulkan bahwa Pengolahan data yang dilakukan menghasilkan kelompok penerima bantuan, sehingga pihak desa bisa lebih mudah dalam menentukan penerima bantuan dengan teknik Data Mining (Andrianti & Firmansyah, 2020). Penelitian terdahulu yang berjudul "Pengelompokan Penerima Bantuan Sosial Masyarakat dengan Metode K-Means" dapat disimpulkan bahwa menentukan layak tidak layaknya penerima data bantuan sosial adalah kebaharuan yang belum pernah dilakukan oleh peneliti sebelumnya (Rady Putra & Anggrawan, 2021). Penelitian terdahulu yang berjudul "Penerapan Algoritma K-Means untuk Pengelompokan Produksi Telur Ayam Ras Petelur di Indonesia" disimpulkan bahwa Data jumlah produksi telur ayam ras petelur di Indonesia dikelompokan dengan membuat design pemodelan yang dibentuk menjadi 3 klaster. Didapatkan hasil pada klaster 0 terdiri dari 29 provinsi, klaster 1 sebanyak 4 provinsi dan klaster 2 sebanyak 1 provinsi. Maka dapat disimpulkan bahwa klaster tertinggi berada pada klaster 1, klaster 2 menjadi klaster sedang dan klaster 0 menjadi klaster terendah (Giantika Utami, 2022).

Berdasarkan permasalahan tersebut maka diangkat sebuah penelitian yang berjudul "Penerapan Algoritma K-Means Clustering Pengelompokkan Masyarakat Kurang Mampu Pada Desa Sidoarjo II Ramunia"

TINJAUAN LITERATUR

Data Mining

Data Mining merupakan proses mencari informasi yang menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Data Mining merupakan beberapa ilmu yang menyatukan teknik dari pembelajaran mesin, pengenalan pola, statistik, database dan visualisasi untuk penanganan permasalahan pengambilan informasi dari database yang besar (Andrianti & Firmansyah, 2020).

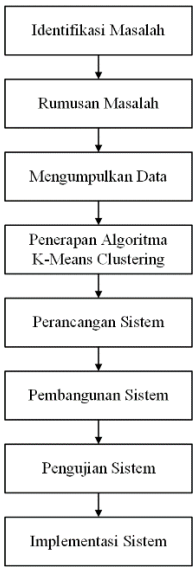
Clustering

Clustering atau analisis pengelompokan merupakan metode pengelompokan berdasarkan kemiripan (kedekatan). Dalam Clustering terdapat proses membagi data dari suatu himpunan tertentu kedalam beberapa kelompok yang memiliki kemiripan karakteristik data dengan karakteristik data pengelompok lain (Ali Hasymi et al., 2021).

K-Means

K-Means merupakan salah satu metode penganalisaan data, dimana data dalam satu kelompok mempunyai karakteristik yang sama akan dikelompokkan menjadi satu cluster (Ali Hasymi et al., 2021).

METODE PENELITIAN



Gambar 1 Kerangka Kerja Penelitian

Dari gambar 1 dapat dijelaskan seperti berikut:

1. Identifikasi Masalah
   Pada tahap ini dilakukan mengidentifikasikan masalah mengenai permasalahan-permasalahan pokok pada penelitian yang dilakukan. Masalah yang timbul adalah akibat jumlah penduduk yang besar dan luasnya wilayah desa yang harus didata. Situasi ini dapat menyebabkan kesulitan dan potensi terjadinya kesalahan dalam penentuan atau pengelompokkan masyarakat kurang mampu. Dengan banyaknya data yang harus diolah, perlu dilakukan upaya ekstra untuk memastikan keakuratan dan keberlanjutan proses identifikasi serta pemilihan kriteria yang tepat agar bantuan dan kebijakan yang diterapkan dapat mencapai sasaran dengan lebih efektif.
2. Rumusan Masalah
   Pada tahap ini ditentukan merumuskan masalah yang akan diselesaikan yaitu, bagaimana menerapkan *K-Means Clustering* dan membangun sistem dalam mengelompokkan masyarakat kurang mampu pada Desa Sidoarjo II Ramunia.
3. Mengumpulkan Data
   Pengumpulan data didapatkan dari sumber data yang dimiliki, diantaranya yaitu:
   a. Observasi
      Kegiatan ini dilakukan di tempat penelitian yaitu pada Desa Sidoarjo II Ramunia. Dari observasi yang peneliti lakukan mencari masalah yang timbul dan mencari solusi agar dapat membantu memudahkan pihak Desa dalam mengelompokkan masyarakat kurang mampu.
   b. Wawancara, dilakukan dengan Bapak Kamler Simbolon, SE., selaku Kepala Desa Sidoarjo II Ramunia.
   c. Studi Literatur
      Pengumpulan data dilakukan dengan cara mengumpulkan jurnal atau buku yang terkait dengan judul penelitian.
4. Penerapan Algoritma *K-Means Clustering*
   Pada tahap ini dilakukan penerapan metode *K-Means Clustering* untuk mengelompokkan masyarakat kurang mampu pada Desa Sidoarjo II Ramunia.
5. Perancangan Sistem
   Dalam tahap ini sistem dirancang menggunakan perancangan UML (*Unifed Modeling Language*).
6. Pembangunan Sistem
   Sistem dibangun berbasis *Website* dengan menggunakan bahasa pemrograman PHP dan MySQL sebagai *Database Management System* (DBMS).
7. Pengujian Sistem
   Pada tahap ini dilakukan pengujian sistem untuk menguji Tingkat akuras penerapan *K-Means* dalam

mengelompokkan masyarakat kurang mampu.

8. Implementasi Sistem
   Pada tahap ini dilakukan implementasi sistem yang sudah selesai dirancang dan dibangun dan setelah diperbaiki kesalahan-kesalahan pada sistem.

PEMBAHASAN

Pada tahap ini dilakukan analisa data serta penerapan metode *K-Means Clustering*, berikut adalah tahapannya:

1. Menentukan Data Masyarakat
   Data masyarakat yang akan dijadikan sampel perhitungan sebanyak 10 orang. Data masyarakat tersebut dapat dilihat pada table 1.

Tabel 1 Data Sampel Masyarakat

Nama	Kondisi Rumah	Penghasilan (Rp)	Pekerjaan	Tanggungan
Naesta Naiborhu	Rusak parah	2.000.000	Buruh Bangunan	3
Narina Sinaga	Layak	1.500.000	Buruh Tani	4
Nelly Lubis	Tidak layak	1.200.000	Penjaga Toko	2
Nally Yunita Marbun	Tidak layak	1.800.000	Pembersih	5
Nerjhon Jacoblev Simarmata	Layak	2.180.000	Buruh Bangunan	1
Nerty Sagala	Tidak Layak	1.530.000	Buruh Tani	3
Nicky Geovany Simanjuntak	Layak	2.750.000	Penjaga Toko	3
Nova Lina Br Lubis	Rusak Parah	1.890.000	Pembersih	1
Novel Risdo Panjaitan	Tidak Layak	2.450.000	Buruh Bangunan	2
Nurcahaya Karokaro	Layak	1.260.000	Buruh Tani	4

2. Menentukan Data Klaster
   Tentukan data klaster yang akan digunakan, pada objek kali ini digunakan 3 klaster, dimana klaster tersebut dapat dilihat pada tabel 2.

Tabel 2 Data Klaster

Nama Klaster	Kode Klaster
Rentan Kurang mampu	C1
Kurang Mampu	C2
Kurang Mampu Ekstrim	C3

3. Persiapan Data: Ubah kondisi rumah menjadi nilai numerik.
   a. Rusak parah = 0
   b. Tidak layak = 1
   c. Layak = 2

Tabel 3 Persiapan Data

Nama	Kondisi Rumah	Penghasilan (Rp)	Tanggungan
Naesta Naiborhu	0	2.000.000	3

Nama	Kondisi Rumah	Penghasilan (Rp)	Tanggungan
Narina Sinaga	2	1.500.000	4
Nelly Lubis	1	1.200.000	2
Nally Yunita Marbun	1	1.800.000	5
Nerjhon Jacoblev Simarmata	2	2.180.000	1
Nerty Sagala	1	1.530.000	3
Nicky Geovany Simanjuntak	2	2.750.000	3
Nova Lina Br Lubis	0	1.890.000	1
Novel Risdo Panjaitan	1	2.450.000	2
Nurcahaya Karokaro	2	1.260.000	4

4. Normalisasi Data

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Penghasilan:
Min = 1.200.000
Max = 2.750.000
Tanggungan:
Min = 1
Max = 5
Setelah normalisasi, data akan terlihat seperti ini:

Tabel 4 Normalisasi

Nama	Kondisi Rumah	Penghasilan (Normalisasi)	Tanggungan (Normalisasi)
Naesta Naiborhu	0	0.327	0.5
Narina Sinaga	2	0.109	0.75
Nelly Lubis	1	0.000	0.25
Nally Yunita Marbun	1	0.246	1.0
Nerjhon Jacoblev Simarmata	2	0.353	0.0
Nerty Sagala	1	0.143	0.5
Nicky Geovany Simanjuntak	2	0.500	0.5
Nova Lina Br Lubis	0	0.246	0.0
Novel Risdo Panjaitan	1	0.327	0.25
Nurcahaya Karokaro	2	0.000	0.75

5. Inisialisasi Pusat Klaster
Tentukan Pusat Klaster
a. Pusat Kluster 1 (C1): (0, 0.5, 0.5)
b. Pusat Kluster 2 (C2): (2, 0.5, 0.5)
c. Pusat Kluster 3 (C3): (1, 0.5, 0.5)

6. Hitung Jarak dan Tetapkan Klaster
Hitung jarak setiap data ke setiap pusat kluster dan tetapkan ke kluster terdekat. Misalnya, untuk Naesta Naiborhu:

$$Jarak\ Ke\ Centroid\ 1 = \sqrt{(0-0)^2 + (0.327-0.5)^2 + (0.5-0.5)^2}$$
$$Jarak\ Ke\ Centroid\ 1 = 0.173$$
$$Jarak\ Ke\ Centroid\ 2 = \sqrt{(0-2)^2 + (0.327-0.5)^2 + (0.5-0.5)^2}$$
$$Jarak\ Ke\ Centroid\ 2 = 2.000$$

$Jarak\ Ke\ Centroid\ 3 = \sqrt{(0-1)^2 + (0.327-0.5)^2 + (0.5-0.5)^2}$
$Jarak\ Ke\ Centroid\ 3 = 1.000$

Lakukan perhitungan yang sama terhadap 9 alternatif lainnya, dari perhitungan ini didapatkan hasil seperti tabel berikut:

Tabel 5 Hasil Klasterisasi

Nama	Klaster 1	Klaster 2	Klaster 3	Klaster
Naesta Naiborhu	0.173	2.000	1.000	C1 (Rentan Kurang Mampu)
Narina Sinaga	2.000	0.447	1.118	C2 (Kurang Mampu)
Nelly Lubis	1.118	1.118	0.353	C3 (Kurang Mampu Ekstrim)
Nally Yunita Marbun	1.118	0.707	0.500	C3 (Kurang Mampu Ekstrim)
Nerjhon Jacoblev Simarmata	2.000	0.500	1.118	C2 (Kurang Mampu)
Nerty Sagala	0.857	1.000	0.357	C3 (Kurang Mampu Ekstrim)
Nicky Geovany Simanjuntak	2.000	0.000	1.000	C2 (Kurang Mampu)
Nova Lina Br Lubis	0.500	2.000	1.118	C1 (Rentan Kurang Mampu)
Novel Risdo Panjaitan	1.118	1.118	0.500	C3 (Kurang Mampu Ekstrim)
Nurcahaya Karokaro	1.000	0.500	1.118	C2 (Kurang Mampu)

Berdasarkan hasil clustering, dapat disimpulkan bahwa terdapat tiga kelompok utama yang mengelompokkan individu berdasarkan tingkat kemampuan ekonomi mereka. Klaster 1, yang terdiri dari Naesta Naiborhu dan Nova Lina Br Lubis, menunjukkan karakteristik sebagai individu dalam kategori "Rentan Kurang Mampu." Klaster 2, yang mencakup Narina Sinaga, Nerjhon Jacoblev Simarmata, Nicky Geovany Simanjuntak, dan Nurcahaya Karokaro, mewakili kelompok "Kurang Mampu" dengan variasi dalam tingkat keterampilan ekonomi mereka. Sementara itu, Klaster 3, yang meliputi Nelly Lubis, Nally Yunita Marbun, Nerty Sagala, dan Novel Risdo Panjaitan, menunjukkan profil "Kurang Mampu Ekstrim," mencerminkan individu dengan kondisi ekonomi yang lebih kritis. Analisis ini membantu dalam memahami kebutuhan dan dukungan yang diperlukan untuk masing-masing kelompok berdasarkan tingkat kerentanan dan kekurangan kemampuan ekonomi mereka.

KESIMPULAN

Berdasarkan uraian permasalahan dan pembahasan pada bab sebelumnya penulis ingin menyampaikan beberapa kesimpulan dan saran yang dapat digunakan sebagai bahan masukan dan pertimbangan selama melakukan perancangan, pembuatan, dan evaluasi dapat diambil kesimpulan sebagai berikut:

1. Penerapan metode K-Means Clustering dalam mengelompokkan masyarakat miskin, maka dapat diambil kesimpulan bahwa Naesta Naiborhu dan Nally Yunita Marbun masuk kategori Kurang Mampu Ekstrim dengan kondisi rumah rusak parah, penghasilan rendah, banyak tanggungan, dan kesehatan buruk. Narina Sinaga, Nerjhon Jacoblev Simarmata, Nicky Geovany Simanjuntak, dan Nurcahaya Karokaro tergolong Rentan Kurang Mampu dengan kondisi rumah layak, penghasilan moderat, beberapa tanggungan, dan kesehatan baik. Nelly Lubis, Nerty Sagala, Nova Lina Br Lubis, dan Novel Risdo Panjaitan masuk kategori Kurang Mampu dengan kondisi rumah tidak layak, penghasilan rendah, beberapa tanggungan, dan kesehatan bervariasi.

2. Perancangan dan pembuatan sistem pakar berbasis website dengan menggunakan bahasa pemograman PHP dan MySQL sebagai database nya.

REFERENCES

Ali Hasymi, M., Faisol, A., & Ariwibisono, F. (2021). Sistem Informasi Geografis Pemetaan Warga Kurang Mampu Di Kelurahan Karang Besuki Menggunakan Metode K-Means Clustering. JATI (Jurnal Mahasiswa Teknik Informatika), 5(1), 284–290. https://doi.org/10.36040/jati.v5i1.3269

Andrianti, F., & Firmansyah, R. (2020). Penerapan Clustering Data Kurang Mampu Di Desa Situmekar Menggunakan Algoritma K-Means. EProsiding Teknik Informatika (PROTEKTIF), 1(1), 88. http://eprosiding.ars.ac.id/index.php/pti

Fu'adi, A., Prianggono, A., Komunitas, A., Pacitan, N., Id, A. A., & Id, A. A. (2022). Analisa dan Perancangan Sistem Informasi Akademik Akademi Komunitas Negeri Pacitan Menggunakan Diagram UML dan EER. Jurnal Ilmiah Teknologi Informasi Asia, 16(1), 45–54. https://jurnal.stmikasia.ac.id/index.php/jitika/article/view/650

Giantika Utami, D. (2022). Penerapan Algoritma K-Means untuk Pengelompokan Produksi Telur Ayam Ras Petelur di Indonesia. 3(1), 72–79. www.bps.go.id.

Handayanna, F. (2023). Penerapan Algoritma K-Means Untuk Klasterisasi Penduduk Miskin Di Provinsi Banten. INTI Nusa Mandiri, 18(1), 93–99. https://doi.org/10.33480/inti.v18i1.4399

Hari Utami, F. (2022). Aplikasi Pelayanan Antrian Pasien Menggunakan Metode FCFS Menggunakan PHP dan MySQL. Jurnal Media Infotama, 18(1), 341139.

Hermiati, R., Asnawati, A., & Kanedi, I. (2021). Pembuatan E-Commerce Pada Raja Komputer Menggunakan Bahasa Pemrograman Php Dan Database Mysql. Jurnal Media Infotama, 17(1), 54–66. https://doi.org/10.37676/jmi.v17i1.1317

Hidayat, R. (2022). Pemanfaatan Data Mining Untuk Melihat Minat Siswa Setelah Menyelesaikan Pendidikan Sekolah Menengah Atas (Sma) Dengan Algoritma K-Means Clustering. 01(02), 85–97.

Hilmi, Marumu, M. N. H. D., Ramlawati, & Peuru, C. D. (2022). Pengaruh Jumlah Penduduk Dan Pengangguran Terhadap Tingkat Kemiskinan Di Kabupaten Tolitoli. Growth: Jurnal Ilmiah Ekonomi Pembangunan, 1(1), 20–27.

Nasril, Rustanto, A. E., Setiawan, & Fadli, A. (2019). PERANCANGAN SISTEM INFORMASI PEMBAYARAN BARBERSHOP DI ARFA BARBERSHOP DEPOK. JSM STMIK Mikroskil, 13(1), 63–71.

Rady Putra, L. G., & Anggrawan, A. (2021). Pengelompokan Penerima Bantuan Sosial Masyarakat dengan Metode K-Means. MATRIK : Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer, 21(1), 205–214. https://doi.org/10.30812/matrik.v21i1.1554

Senang, H. G. (2023). PENERAPAN METODE K-MEDOIDS UNTUK PENGELOMPOKAN PENDUDUK KURANG MAMPU (STUDI KASUS KANTOR CAMAT AFULU).

Sukmayadi, F., Firdaus, A., & Juliane, C. (2022). Klasterisasi Konsentrasi Keahlian Siswa SMK Berdasarkan Kurikulum Merdeka. Jurnal Media Informatika Budidarma, 6(4), 2410. https://doi.org/10.30865/mib.v6i4.4848

Tarigan, P. M. S., Hardinata, J. T., Qurniawan, H., Safii, M., & Winanjaya, R. (2022). Implementasi Data Mining Menggunakan Algoritma Apriori Dalam Menentukan Persediaan Barang. Jurnal Janitra Informatika Dan Sistem Informasi, 2(1), 9–19. https://doi.org/10.25008/janitra.v2i1.142

Triyana, M., Juita, R., & Suhendra, C. D. (2022). Penerapan Metode K-Means dalam Pengelompokan Data Penduduk Tidak Mampu di Distrik Oransbari. INFORMAL: Informatics Journal, 7(3), 220. https://doi.org/10.19184/isj.v7i3.34722.