

PENERAPAN ALGORITMA X-MEAN MENENTUKAN PENERIMA BANTUAN SOSIAL

Ahmad Jaelani, Ade Irma Purnamasari, Irfan Ali

Teknik Informatika, STMIK IKMI CIREBON

Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45135

Jayganteng1402@gmail.com

ABSTRAK

Pertambahan jumlah data keluarga tidak mampu di Desa Jatipancur dapat berpotensi mengakibatkan dampak buruk akibat peningkatan jumlah penduduk. Masalah ketidakseimbangan keluarga tidak mampu dapat menyebabkan masalah kemiskinan dan tekanan penduduk. Tujuan penelitian ini adalah Mendapatkan hasil cluster dari penerima bantuan Sosial dalam pengaplikasian metode x-mean, mengetahui kelompok penerima bantuan sosial dan cluster terbaik. Penelitian menggunakan pendekatan data mining dengan menerapkan Algoritma X-Means Clustering. Algoritma X-Means diimplementasikan menggunakan software Ms. Excel dan Rapidminer. Penelitian ini menggunakan data dari Badan Pusat Statistik (BPS) Jatipancur, Data yang digunakan melibatkan 7 atribut dan mencakup periode tahun 2019-2023. Hasil pengelompokan menggunakan X-Means menghasilkan 10 cluster, Cluster 1 mendekati nilai sempurna dengan Davies Bouldin Index sebesar 0.002, Jarak antara cluster terbaik adalah pada cluster 1 dengan jumlah 0.000, serta Performance terdekat ada pada cluster 0 dengan nilai 5.672.239.884.760. Peneliti berharap penelitian ini dapat membantu Desa Jatipancur dalam proses penyaluran bantuan dan Optimalisasi dalam penyaluran bantuan yang lebih tepat sasaran.

Kata kunci: RapidMiner, data mining, pengelompokan, algoritma X-means, bantuan sosial

1. PENDAHULUAN

Meningkatnya angka pertumbuhan penduduk yang tidak seimbang dengan kesempatan kerja, berakibat pada peningkatan jumlah pengangguran. data keluarga tidak mampu di Desa Jatipancur. Apabila tidak diatasi secara efektif dan sistematis, hal tersebut akan menimbulkan konsekuensi yang merugikan. Ketidakmampuan dalam data keluarga juga dapat mempengaruhi pertumbuhan ekonomi dan kesejahteraan masyarakat. diperlukan adanya pengelompokan data keluarga tidak mampu sebagai salah satu upaya untuk mencegah kemiskinan penduduk. Maka dari itu, dirancanglah pengolahan data yang dapat menangani permasalahan tersebut, yakni melalui penerapan teknik data mining dengan menggunakan algoritma X-Means. terhadap data keluarga tidak mampu di Desa Jatipancur.

Data Mining adalah suatu proses penyusunan untuk mencari nilai tambah berupa informasi yang tidak dapat ditemukan secara manual dari sebuah basis data. Proses ini umumnya digunakan untuk mengekstraksi pengetahuan atau pola dari data yang berjumlah besar, sehingga sering disebut sebagai Knowledge Discovery Database (KDD). Salah satu teknik dalam Data Mining adalah Clustering (Muhamad Khandava Mulyadien et al 2022). Dalam penelitian ini, metode Clustering yang diterapkan adalah menggunakan Algoritma X-Means.

Algoritma X-Means merupakan metode Clustering non-hirarki yang berupaya mempartisi data ke dalam satu atau lebih Cluster, di mana data yang memiliki karakteristik serupa dikelompokkan ke dalam Cluster yang sama. Dalam penelitian ini, Algoritma X-Means akan digunakan untuk mengelompokkan jumlah penduduk berdasarkan

karakteristiknya. Namun, terdapat kelemahan pada X-Means terkait penentuan matriks jarak, yang merupakan faktor kritis yang memengaruhi kinerja Algoritma. Nilai matriks jarak yang dihasilkan akan mempengaruhi performa dari Algoritma ini [1]. Pemilihan Data Mining didasarkan pada kemampuannya untuk mengekstraksi informasi prediktif dari database yang besar (Siti Amaliyah et al 2023), Pengelompokan dilaksanakan sebagai solusi untuk mendorong perhatian pemerintah terhadap kelompok jumlah penduduk, dengan tujuan mencegah kepadatan penduduk dan dampak buruk lainnya, sehingga kesejahteraan masyarakat dapat lebih terjamin.

Berdasarkan latar belakang yang telah diuraikan, penelitian ini difokuskan pada “Penerapan Algoritma X-Means Menentukan Penerima Bantuan” Hasil penelitian ini diharapkan dapat membantu pihak Desa Jatipancur dalam proses penyaluran bantuan, sehingga pemberian bantuan ini menjadi lebih optimal dan ditujukan secara langsung kepada masyarakat yang benar-benar membutuhkan.

2. LITERATURE REVIEW

Pencarian literatur yang relevan dengan topik penelitian. Langkah ini dapat memberikan gambaran mengenai Penerapan Data Mining Clustering Menggunakan Algoritma X-Means Proses pencarian dan pengumpulan artikel dengan menggunakan tools publish or perish, yang diambil dari database akademik yaitu *Google Scholar*, *Shinta* dan *Scopus* dengan kata kunci “Data Mining” dan “Clustering X-Means”.

Hasil literature review yang telah dilakukan pada jurnal-jurnal penelitian terkait topik Penerapan

Metode Optimasi pada Jaringan Komputer untuk Meningkatkan Efisiensi Penggunaan Sumber Daya dapat dijabarkan sebagai berikut:

[2] dengan judul “Penerapan Algoritma K-Means Untuk Menentukan Prioritas Penerimaan Bantuan PKH” membahas mengenai Implementasi yang dilakukan adalah penerapan Algoritma K-Means Clustering terhadap data Penduduk dengan jumlah data yang akan diolah oleh penulis sebanyak 1.011. Data ini akan di klasterisasikan kedalam kategori masyarakat “Mampu” dan “Kurang Mampu”.

[3] dengan judul “Pemetaan Penduduk Penerima Bantuan Sosial Desa Waru Jaya Menggunakan Algoritma K-Means Clustering” membahas mengenai Penerapan data mining bertujuan untuk memastikan penyaluran bantuan ini tepat sasaran. Oleh karena itu, diperlukan metode klasterisasi (clustering) sebagai teknik pengelompokan data. Ada banyak metode klasterisasi yang diajukan oleh para ahli, salah satunya adalah metode k-means, yang termasuk dalam kategori berbasis partisi. Metode ini membagi data menjadi sejumlah kelompok untuk mencapai tujuan tersebut.

[4] dengan judul “Implementasi Metode K-means Clustering Untuk Menganalisa Penerima Bantuan di Desa Palasah” membahas mengenai Penelitian ini bertujuan untuk mengoptimalkan penyebaran akurasi penerima bantuan di Desa Palasah, Kabupaten Majalengka. Metode yang diterapkan dalam penelitian ini adalah menggunakan metode K-Means Clustering.

[5] dengan judul “Analisis Pengelompokan Data set Komputer Menggunakan Algoritma X-Means” Artikel ini membahas implementasi metode Clustering X-Means dalam mengelompokkan dataset komputer berdasarkan harga dan spesifikasinya. Penjelasan diberikan mengenai proses pengklasifikasian atau clusterisasi data set komputer dengan menggunakan Algoritma X-Means. Pengelompokan ini telah diuji melalui proses pengujian menggunakan machine learning menggunakan Software RapidMiner Studio.

[6] dengan judul “Penggunaan X-Means Clustering Method untuk Mengelompokkan Potensi Sekolah Menengah Unggul di Kabupaten Banyumas” membahas mengenai Penggunaan metode X-Means dalam mengelompokkan SMA dan SMK di Banyumas menghasilkan hasil yang memuaskan dengan empat cluster, yang dibuktikan dengan nilai Davies-Bouldin Index (DBI) sebesar 0.846. Keempat cluster tersebut adalah cluster unggul dengan 28 anggota, cluster baik dengan 26 anggota, cluster cukup dengan 43 anggota, dan cluster kurang sebanyak 21 anggota. Dari hasil tersebut, dapat disimpulkan bahwa tingkat pendidikan menengah di Kabupaten Banyumas masih tergolong dalam cluster cukup.

[7] dengan judul “Analisis Rumah Tidak Layak Huni Menggunakan Algoritma X-Means Analyzing

Uninhabitable Houses Using X-Means Algorith” membahas mengenai Pengelompokan rumah tidak layak huni dilakukan dengan mempertimbangkan struktur bangunan dan fasilitas rumah pada setiap rumah tangga di empat wilayah, yaitu: Batu Layang, Siantan Hilir, Siantan Tengah, dan Siantan Hulu. Salah satu metode yang dapat digunakan untuk pengelompokan ini adalah metode X-Means clustering. Dalam proses memetakan atau mengelompokkan rumah tidak layak huni menggunakan algoritma X-Means, hasilnya diorganisir menjadi beberapa cluster, dan nilai akurasi terbaik diperoleh melalui uji Davies-Bouldin Index (DBI).

[8] dengan judul “Penerapan Algoritma Clustering K-Means Untuk Menentukan Prioritas Penerima Bantuan Dana Sosial PKH Di Kelurahan Kampung Singkep” Artikel ini membahas perhitungan Clustering K-Means untuk menetapkan penerima bantuan berdasarkan data penduduk Kelurahan Kampung Singkep. Tujuannya adalah untuk menentukan jenis bantuan yang sesuai dengan kondisi mereka, memungkinkan tim penyeleksi memberikan bantuan sesuai dengan kriteria yang telah ditetapkan, apakah layak atau tidak layak menerima bantuan tersebut.

[9] dengan judul “Data Mining Menentukan Cluster Penerima Program Bantuan dengan Metode K-Means” membahas pemanfaatan data mining untuk melakukan pengelompokan data dengan menggunakan metode K-Means, dan pengolahan data dilakukan melalui aplikasi RapidMiner. Dalam konteks data mining, pada saat penelitian, dilakukan pengklasifikasian data, pembuatan prediksi, serta estimasi untuk menghasilkan informasi berharga dari kumpulan data yang akan diolah.

[10] dengan judul “Pengelompokkan Tingkat Pemahaman Guru PAUD Terhadap Pembelajaran Berbasis STEAM Menggunakan Metode X-Means Clustering” membahas penerapan algoritma X-Means clustering sebagai salah satu algoritma data mining untuk mengelompokkan data tingkat pemahaman guru PAUD terhadap pembelajaran berbasis STEAM. Hasil pengelompokkan menggunakan algoritma X-Means Clustering menunjukkan tingkat pemahaman guru PAUD di Kecamatan Medan Area, dengan 12 guru yang sangat paham, 7 guru yang paham, 7 guru yang cukup paham, dan 4 guru yang tidak paham terhadap pembelajaran berbasis STEAM.

[11] dengan judul “Data Mining Manhattan Distance dan Euclidean Distance pada Algoritma X-Means dalam Klasifikasi Minat dan Bakat Siswa” membahas mengenai Sistem data mining yang dikembangkan dapat memberikan bantuan dan kemudahan dalam proses klasifikasi minat dan bakat siswa. Penilaian variabel dilakukan melalui sistem, memungkinkan admin untuk mengupdate nilai maksimal dan minimal variabel tanpa perlu merubah program secara manual.

3. METODE PENELITIAN

Penelitian ini dilaksanakan dengan merujuk pada tahapan-tahapan penting yang diarahkan pada indikator keberhasilan dalam pengelompokan data miskin di Kecamatan Banjarmasin Tengah, guna mengatasi permasalahan multi-objektif. Untuk mencapai indikator tersebut, penelitian ini mengikuti tahapan-tahapan berikut:

- Analisis masalah, fokus pada kebutuhan untuk menganalisis permasalahan terkait pengelompokan data masyarakat miskin.
- Analisis kebutuhan, melibatkan identifikasi dan pengumpulan semua kebutuhan penelitian, termasuk referensi dari jurnal, buku, literatur, serta alat dan perangkat lunak yang dibutuhkan.
- Merancang sistem pengelompokan berdasarkan algoritma K-Means.
- Melakukan analisis menggunakan alat bantu RapidMiner.
- Menguji data dengan menerapkan algoritma K-Means yang telah dirancang.
- Menyusun laporan dan merumuskan hasil penelitian."

4. HASIL DAN PEMBAHASAN

Pada penelitian ini dataset yang digunakan yaitu penerima bantuan pangan non tunai desa jatipancur sebanyak 567 data dari tahun 2022 sampai 2023 menggunakan atribut-atribut sebagai berikut, RT, RW, NIK, Nama, Pekerjaan, Penghasilan dan Tanggungan. Bisa dilihat pada Gambar berikut ini:

No	RT	RW	NIK	NAMA	PEKERJAAN	PEGHASILAN	TANGGUNGAN
1	1.000	1.000	33280242059...	HUROKHATU...	PEGAWAI SW...	2000000.000	2
2	1.000	1.000	32093855088...	NIHA SANIAH	MENGURUS...	1300000.000	3
3	1.000	1.000	32093843047...	SUNTIAH	MENGURUS...	2300000.000	4
4	1.000	1.000	32093845077...	IKHA HARTIKA	MENGURUS...	1550000.000	2
5	1.000	1.000	32090641074...	SANAM	MENGURUS...	2500000.000	1
6	1.000	1.000	32093852105...	SADIAH	MENGURUS...	1500000.000	1
7	1.000	1.000	32093853086...	YUSTALINA	MENGURUS...	1500000.000	3
8	1.000	1.000	32093855089...	SUSI SUSILA...	MENGURUS...	1000000.000	4
9	1.000	1.000	32093855068...	TETI SETIAWA...	MENGURUS...	1250000.000	1
10	1.000	1.000	32093866106...	ANI	MENGURUS...	1250000.000	2
11	1.000	1.000	32093855065...	UMIYATI	MENGURUS...	1320000.000	4
12	1.000	1.000	32092249078...	MISRI	MENGURUS...	1280000.000	2
13	1.000	1.000	32093841014...	SETIASIH	MENGURUS...	1250000.000	1
14	1.000	1.000					

Gambar 1. Dataset Penerima Bantuan

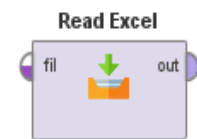
4.1. Data Selection

Pada tahap ini, dilakukan pemilihan data dengan tujuan untuk memilih data yang akan digunakan dalam proses pengelompokan melalui tools Rapidminer, dalam tahap seleksi yang dilakukan peneliti dilakukan di Microsoft Excel pada proses data mining pengelompokan bantuan social. Hasil dari seleksi data yang dilakukan peneliti dapat diketahui pada data penerima bantuan sosial terdapat 7 atribut seperti pada table berikut ini:

Tabel 1. Atribut Data Dari Dataset

No	Atribut	Type
1	RT	Integer
2	RW	Integer
3	NIK	Real
4	Nama	Polunominal
5	Pekerjaan	Polynomial
6	Penghasilan	Integer
7	Tanggungan	Integer

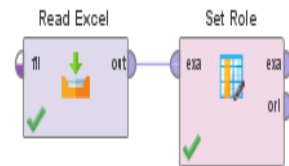
Agar dataset tersebut bisa dikelola di RapidMiner, maka perlu dilakukan Import dataset penerima bantuan BPNT.xlsx dengan menggunakan operator *Read Excel* pada RapidMiner seperti gambar berikut:



Gambar 2. Operator Read Excel

4.2. Preprocessing

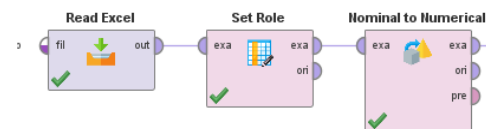
Proses digunakan untuk mengidentifikasi variabel yang terdapat pada dataset tersebut yang kemudian hal ini memungkinkan RapidMiner melakukan pemrosesan data yang sesuai seperti variabel kategoris. Yang mana dapat dilihat di bawah ini Operator set rol yang digunakan:



Gambar 3. Operator Set role pada RapidMiner

4.3. Transformation

Langkah selanjutnya yaitu Transformasi, pada langkah transformation dilakukan untuk data diolah menjadi bersifat numeric. Untuk mengonversi data Polynomial menjadi numerik, dapat menggunakan operator Nominal to Numerical yang tersedia di RapidMiner, sebagaimana terlihat pada contoh di bawah ini:



Gambar 4. Nominal to Numerical pada RadpidMiner

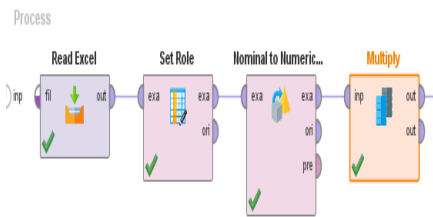
Untuk Parameter yang di gunakan dalam operator Nominal to Numerical peneli menggunakan 3 atribut yang dapat dilihat pada tabel berikut:

Tabel 2. Parameter Nominal to Numerical

Parameter	Keterangan
Attribute Filter Type	subset
Attributes	Select Attributes
Coding type	Unique integers

4.4. Data Mining

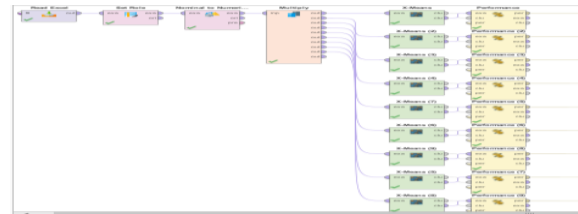
Dalam Proses data mining penulis menerapkan operator Multiply dalam pengoprasian klaster karena penulis menggunakan 10 klaster dalam kesempatan kali ini. Operator multiply dapat dilihat pada gambar 4:



Gambar 5. Operator Multiply

Pada langkah data mining, dilakukan pengelompokan data menggunakan algoritma x-means. Menggunakan 10 klaster yang dimana hasil yang di dapatkan dari 10 klaster tersebut di ambil

mana nilai klaster yang terbaik. Atribut X-means dan performance yang bisa di lihat pada gambar berikut ini :



Gambar 6. X-means dan Performance

Parameter yang digunakan dalam metode X-Means kali ini bisa di lihat pada tabel-tabel berikut ini:

Tabel 2. Parameter X-Means

Cluster	K min	K max	Measure type	Numerical measure	Cluster algoritm
K1	2	60	NumericalMeasurement	EuclideanDistance	Kmeans
K2	3	60	NumericalMeasurement	EuclideanDistance	Kmeans
K3	4	60	NumericalMeasurement	EuclideanDistance	Kmeans
K4	5	60	NumericalMeasurement	EuclideanDistance	Kmeans
K5	6	60	NumericalMeasurement	EuclideanDistance	Kmeans
K6	7	60	NumericalMeasurement	EuclideanDistance	Kmeans
K7	8	60	NumericalMeasurement	EuclideanDistance	Kmeans
K8	9	60	NumericalMeasurement	EuclideanDistance	Kmeans
K9	10	60	NumericalMeasurement	EuclideanDistance	Kmeans

Dari 9 klaster yang dilakukan peneliti kali ini di dapatkan klaster terbaik ada pada klaster K1 dengan Davies Bouldin terbaik dengan index yang di hasilkan dari algoritma X-Means sebesar 0.002.

4.5. Interpretasi Data

Hasil dari data mining yang diperoleh melalui penerapan algoritma X-Means adalah sebagai berikut:

- Hasil pengelompokan dengan menggunakan algoritma X-Means menunjukkan bahwa cluster terbaik terdapat pada klaster K1, dengan 2 klaster optimal, yaitu cluster 0 sebanyak 566 item dan cluster 1 dengan 1 item.

Cluster Model

```
Cluster 0: 566 items
Cluster 1: 1 items
Total number of items: 567
```

Gambar 7. Cluster Model

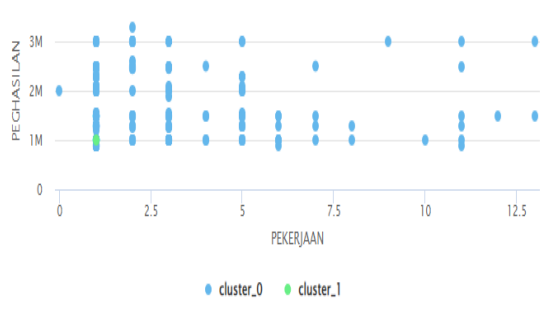
- Dari Cluster K1, nilai Davies Bouldin Index yang dihasilkan oleh algoritma X-Means adalah 0.002. Dengan menghitung jarak antara rata-rata cluster, dapat disimpulkan bahwa cluster 0 merupakan kelompok terbaik dengan jarak terdekat sebesar 5.672.239.884.760.000.

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: -32104877747755506000000000.000
Avg. within centroid distance_cluster 0: -32161600146603130000000000.000
Avg. within centroid distance_cluster 1: -0.000
Davies Bouldin: -0.002
```

Gambar 8. PerformanceVecctor pada metode X-Means

- Plot
Grafik plot yang di hasilkan dari metode algoritma X-means clustering dapat dilihat pada gambar berikut ini dimana cluster 0 dengan hasil nilai terbesar nya yaitu pada x = 2 dan Y = 3.300.00 dan nilai terkecilnya yaitu x = 11 dan Y = 900.000 dan dalam cluster 1 terdapat satu nilai saja yaitu dengan nilai X = 1 dan Y = 1.000.000. dapat dilihat pada gambar berikut ini:



Gambar 9. Grafik Plot

Memuat hasil, Pengujian dan pembahasan tentang skripsi yang telah dilakukan.

5. KESIMPULAN DAN SARAN

Berdasarkan penelitian ini, dapat menyimpulkan sebagai berikut: Hasil pengelompokan menggunakan X-Means, Dari 10 kluster atau klasifikasi bahwa kluster pertama sangat mendekati nilai sempurna pada klasifikasi tersebut. Nilai davies bouldin index sebesar 0.002. dengan menghitung jarak antara kluster, kluster 1 adalah yang terbaik dengan jumlah 0.000 Dengan nilai performance terdekat ada pada cliuster 0 dengan nilai 5.672.239.884.760.

Dalam rangka memperbaiki kekurangan serta guna pelengkapan penelitian kali ini maka beberapa saran sebagai berikut: penelitian ini dapat dikembangkan dengan menggunakan algoritma yang lain seperti naïve bayes, KNN, decision tree, dan lain-lain. Pada penelitian selanjutnya diusahakan menggunakan dataset dengan atribut lebih lengkap dan lebih banyak agar hasil yang di dapatkan bisa lebih jelas dan akurat lagi. Peneliti berharap agar penelitian yang dilakukan dapat dijadikan acuan untuk pemerintah atau aparat desa dalam melakukan verifikasi data warga yang tidak mampu dan mampu agar tidak ada lagi salah sasaran dalam pembagian bantuan sosial.

DAFTAR PUSTAKA

- [1] M. Mughnyanti and S. Hafiz Nanda Ginting, "Data Mining Manhattan Distance dan Euclidean Distance Pada Algoritma X-Means Dalam Klasifikasi Minat dan Bakat Siswa," *remik*, vol. 7, no. 1, pp. 835–842, Jan. 2023, doi: 10.33395/remik.v7i1.12162.
- [2] B. Fibriyanti and S. Informasi, "Penerapan Algoritma K-Means Untuk Menentukan Prioritas Penerimaan Bantuan PKH."
- [3] D. Bahtiar *et al.*, "PEMETAAN PENDUDUK PENERIMA BANTUAN SOSIAL DESA WARU JAYA MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING," 2023. [Online]. Available: <http://pijarpemikiran.com/index.php/Scientia>
- [4] O. Mar, atun Sholihah, N. Suarna, and G. Dwilestari, "Implementasi Metode K-means Clustering Untuk Menganalisa Penerima Bantuan di Desa Palasah," *Jurnal Informatika dan Teknologi Informasi*, vol. 1, no. 2, 2023, doi: 10.56854/jt.v1i2.121.
- [5] R. R. Khaerullah, N. Suarna, and O. Nurdian, "Analisa Pengelompokan Dataset Komputer Menggunakan Algoritma X-Means," *Jurnal Informatika dan Teknologi Informasi*, vol. 1, no. 2, 2023, doi: 10.56854/jt.v1i2.135.
- [6] A. Wijayanto and R. Adhitama, "Penggunaan X-Means Clustering Method untuk Mengelompokkan Potensi Sekolah Menengah Unggul di Kabupaten Banyumas," vol. 2, no. 1, pp. 80–088, 2019, doi: 10.20895/INISTA.V2I1.
- [7] G. F. Rohendi, N. Suarna, and G. D. Lestari, "Analisis Rumah Tidak Layak Huni Menggunakan Algoritma X-Means," *Jurnal Janitra Informatika dan Sistem Informasi*, vol. 3, no. 1, pp. 18–29, May 2023, doi: 10.25008/janitra.v3i1.151.
- [8] P. Penerima Bantuan Dana Sosial PKH Di Kelurahan Kampung Singkep Sari Ufriani and Y. Arviita SKom, "Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM) Penerapan Algoritma Clustering K-Means Untuk Menentukan," JAKAKOM, 2023. [Online]. Available: <http://ejournal.unama.ac.id/index.php/jakakom>
- [9] R. R. A. Aria, S. Susilowati, and I. R. Rahadjeng, "Data Mining Menentukan Cluster Penerima Program Bantuan dengan Metode K-Means," *remik*, vol. 7, no. 1, pp. 291–300, Jan. 2023, doi: 10.33395/remik.v7i1.12030.
- [10] S. Fatimah and A. Usman, "Pengelompokan Tingkat Pemahaman Guru PAUD Terhadap Pembelajaran Berbasis STEAM Menggunakan Metode X-Means Clustering," 2022.
- [11] M. Mughnyanti and S. Hafiz Nanda Ginting, "Data Mining Manhattan Distance dan Euclidean Distance Pada Algoritma X-Means Dalam Klasifikasi Minat dan Bakat Siswa," *remik*, vol. 7, no. 1, pp. 835–842, Jan. 2023, doi: 10.33395/remik.v7i1.12162.