ANALISIS EFEKTIVITAS PELAYANAN PUBLIK MENGGUNAKAN K-MEANS CLUSTERING DI KECAMATAN SUKAGUMIWANG

Siti Oop Sofiyah[1], Nining R.[2], Raditya Danar Dana[3]
[1] Program Studi Teknik Informatika S1, STMIK IKMI Cirebon
[2] Program Studi Komputer Akutansi, STMIK IKMI Cirebon
[3] Program Studi Manajemen Informatika, STMIK IKMI Cirebon
oopsofiyah1724@gmail.com

ABSTRAK

Dalam memberikan pelayanan kepada masyarakat selama ini sering banyak masukan dari masyarakat terutama tentang pelayanan pengurusuan surat-surat di kantor Kecamatan Sukagumiwang Indramayu. Sehingga perlu adanya pelaksanaan survey pelayanan kepada masyarakat. Permasalahan dalam penyelenggaraan pelayanan kepada masyarakat di Kecamatan Sukagumiwang yang perlu diatasi. Seperti dalam pengurusan surat pengantar kartu keluarga, surat pengantar ijin usaha dan lain-lain dalam pelayanannya masih lambat dan masih ada beberapa formulir yang masih diisi secara manual yaitu diketik menggunakan mesin tik. Tujuan penelitian ini untuk menganalisis bagaimana mengklusterkan kepuasan pelayanan yang terbaik kepada masyarakat di Kecamatan Sukagumiwang. Penelitian yang digunakan dalam metode ini yaitu menggunakan metode k-means clustering dengan teknik KDD (Knowledge Discovery in Database). Kemudian data yang digunakan untuk analisis dengan metode K-means Clustering untuk mengidentifikasi cluster-cluster yang ada pada data dan mengukur efektivitas pelayanan kepada masyarakat. Penelitian ini memberikan sebuah kontribusi untuk pemerintah setempat untuk meningkatkan efektivitas pelayanan kepada masyarakat Kecamatan Sukagumiwang, sehingga bisa membantu kebutuhan masyarakat dengan lebih optimal. Dan menghasilkan akurasi data Nilai k optimal yaitu k = 15 dengan nilai dbi = 0,984. Jenis cluster distance yang menghasilkan nilai k yang optimal tersebut adalah ChebychevDistance. Nilai max run yang menghasilkan nilai k yang optimal tersebut adalah 29.

Kata kunci: *Analisis efektivitas pelayanan public, K-means Clustering, ecamatan Sukagumiwang.*

1. PENDAHULUAN

Pelayanan publik yang efektif adalah salah satu aspek penting dalam menjaga kepercayaan dan kepuasan masyarakat kepada pemerintah. Analisis efektivitas pelayanan publik merupakan suatu cara untuk mengukur kinerja pemerintah dalam memberikan pelayanan publik kepada masyarakat. Metode analisis yang digunakan dalam analisis efektivitas pelayanan publik dapat bervariasi, seperti survei kepuasan masyarakat, analisis data, atau pengukuran kinerja. Hasil analisis efektivitas pelayanan masyarakat bisa digunakan sebagai dasar untuk melakukan perbaikan atau peningkatan pelayanan publik yang lebih baik di masa depan. Algoritma k-means clustering ini dapat dipakai untuk mengelompokkan data wilayah atau daerah ke dalam beberapa klaster, bagaimana memanfaatan algoritma k-Means untuk mengklaster atau mengelompokkan data wilayah berdasarkan beberapa indikator pelayanan kesehatan kecamatan-kecamatan di Kabupaten Blora [1].

Hasil evaluasi efektivitas pelayanan masyarakat di Lubuk Pakam menggunakan metode K-Means Clustering dari penelitian menunjukkan bahwa terdapat beberapa kelompok pelayanan publik yang memiliki tingkat efektivitas yang rendah dan perlu ditingkatkan [2]. Pendapat lain yang mengkaji efektivitas pelayanan masyarakat di Kota Pekanbaru dengan metode fuzzy C-Means Clustering. Hasil penelitian menunjukkan bahwa terdapat perbedaan dalam efektivitas pelayanan publik antara kelurahan di Kota Pekanbaru, serta memberikan rekomendasi untuk pemerintah agar meningkatkan efektivitas pelayanan publik di kelurahan yang memiliki tingkat efektivitas rendah [3].

Tujuan utama dari penelitian adalah untuk menganalisis efektivitas pelayanan publik di Kecamatan Sukagumiwang dengan menggunakan metode K-Means Clustering. Dengan melakukan analisis ini, diharapkan dapat memberikan informasi tentang kualitas pelayanan masyarakat yang diberikan oleh instansi-instansi terkait di daerah tersebut.

Alasan pemilihan judul "Analisis Efektivitas Pelayanan Publik Menggunakan K-Means Clustering di Kecamatan Sukagumiwang" dipilih karena pentingnya kualitas dan efektivitas pelayanan publik dalam memberikan layanan kepada masyarakat dan relevannya dengan kebutuhan saat ini dalam meningkatkan kualitas pelayanan publik di Kecamatan Sukagumiwang.

2. TINJAUAN PUSTAKA

2.1. Algoritma K-Means

Algoritma K-Means merupakan algoritma clustering yang paling sederhana dibandingkan dengan algoritma yang lain. Algoritma clustering ini mudah diterapkan dan dijalankan, relative cepat, mudah untuk diadaptasi dan banyak digunakan dalam tugas data mining. Algoritma clustering termasuk salah satu

algoritma yang sering di gunakan dalam data mining [4].

2.2. Klasterisasi

Clustering adalah alat analisis data yang memecahkan masalah clustering. Tujuannya adalah untuk distribusi dalam kelompok sedemikian rupa sehingga anggota dari cluster yang sama sangat terhubung dan anggota dari cluster yang berbeda kurang terhubung [5][6].

2.3. Data Mining

Data mining adalah proses menganalisis dan mengekstraksi pengetahuan secara otomatis menggunakan satu atau lebih teknik pembelajaran komputer. Penambangan data dapat melibatkan pencarian tren dan pola minat dalam basis data yang sangat besar untuk membuat keputusan di masa mendatang [7]. Sementara algoritma K-means lebih efisien untuk data kecil, Selain itu, algoritma K-means memiliki dua kelemahan: hasil pengelompokan dipengaruhi oleh posisi pengelompokan pertama. Perbedaannya terletak pada pilihan konsep cluster: pusat K-means didasarkan pada rata-rata koordinat semua titik dalam cluster, sedangkan pusat K-word harus menjadi sampel. titik dalam kelompok. Semua pusat cluster antara, dan itu Jarak cluster antara titik terkecil [8].

2.4. Pelayanan Masyarakat

Bidang Pelayanan Publik merupakan bidang kerja nonstruktural yang menyelenggarakan kegiatan pelayanan publik secara langsung kepada masyarakat. Setiap tahun, Kementerian Pendayagunaan Aparatur Negara dan Reformasi Birokrasi [9]. Yang dimaksud dengan "pelayanan publik" adalah setiap kegiatan yang dilakukan oleh pemerintah atas nama suatu kelompok atau badan yang melakukan suatu kegiatan yang mendatangkan manfaat dan kepuasan, sekalipun hasilnya tidak terikat secara fisik dengan produk. Dalam bukunya Reformasi yang berjudul "Reformasi Pelayanan Publik" Seminar Nasional Informatika Medis [1].

3. METODE PENELITIAN

Dalam penelitian ini, kami menggunakan data berupa angka sebagai alat untuk menganalisis informasi tentang apa yang ingin ketahui. Penelitian ini bersifat deskriptif analisis karena dalam penelitian ini dapat mengklusterkan tingkat kepuasan masyarakat. Pada populasi atau sampel tertentu sesuai dengan banyaknya populasi atau sample yang akan di uji [10]. Penelitian Penelitian kuantitatif adalah penelitian yang menggambarkan atau menjelaskan suatu masalah yang hasilnya dapat digeneralisasikan. Jenis penelitian yang digunakan dalam penelitian ini didasarkan pada pendekatan penelitian kuantitatif. Metode penelitiannya sendiri menggunakan algoritma k-means sebagai berikut: Algoritma pengelompokan yang digunakan dalam artikel ini adalah K-Means. Ini adalah algoritme yang dapat menentukan objek di dalam pusat grup data (biasanya pusat grup data) [11].

K-Means adalah metode pengelompokan data non-hierarkis yang dapat mengelompokkan data ke dalam kelompok berdasarkan kesamaan data, sehingga data dengan karakteristik yang baik dikelompokkan dalam satu kelompok dan file dengan karakteristik yang berbeda dikelompokkan menjadi satu. Metode K-Means dapat diterapkan pada berbagai macam data, termasuk clustering di wilayah studi [12].

Algoritma ini berulang untuk menemukan pusat terbaik. Metode K-means ini membagi klaster sesuai dengan jumlah klaster yang awalnya diidentifikasi atau dimulai saat algoritma dijalankan. Algoritma K-Means mengikuti jalur ini:

1. Tentukan nilai k sebagai jumlah cluster yang akan dibentuk.
2. Mulai dengan k sebagai centroid yang dibuat secara acak.
3. Hitung jarak dari setiap titik data ke setiap centroid menggunakan persamaan jarak Euclidean.
4. Kelompokkan semua data berdasarkan jarak terpendek antara data dengan centroid.
5. Tentukan letak titik berat baru (k).
6. Jika lokasi centroid baru dan lama tidak sama, kembali ke langkah 3. Saat menghitung jarak antara data dan centroid, jarak Euclidean digunakan dalam rumus berikut:

$$d(P,Q) = \sqrt{\sum_{j=1}^{p} \left(x_j(\mathrm{p}) - x_j(\mathrm{Q})\right) \dot{2}(1)}$$

Kemudian, pada langkah 5, temukan centroid berikutnya menggunakan yang berikut ini:

Type equation here.

$$C(I) = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{\sum x} \ (2)$$

3.1. CanberraDistance

CanberraDistance adalah metrik yang digunakan dalam pengelompokan dan penambangan data untuk mengukur jarak antara dua titik atau vektor. Ini sering digunakan saat membandingkan istilah dengan atribut yang dapat dipengaruhi oleh perbedaan kecil atau saat menangani data yang berbeda. Dalam Clustering, jarak Canberra digunakan untuk mengukur perbedaan antara titik-titik data untuk mengelompokkan titik-titik yang serupa ke dalam kelompok-kelompok berdasarkan kesamaannya pada lokasi tertentu.

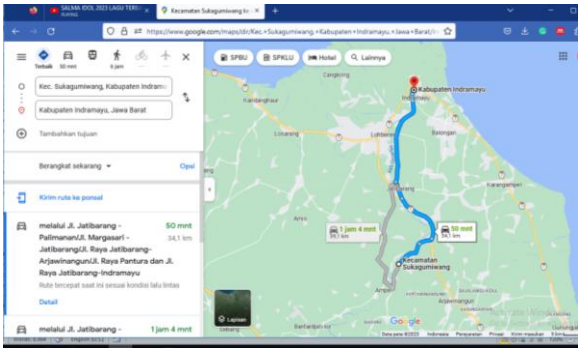
3.2. ChebyshevDistance

ChebyshevDistance adallah Jarak Chebyshev (juga disebut "maksimum" atau "maksimum") adalah ukuran jarak antara dua titik dalam ruang, dengan mempertimbangkan perbedaan maksimum antara

setiap panjang dari dua titik. Dengan kata lain, jarak Chebyshev menghitung selisih maksimum antara setiap elemen dalam dua vektor. Perbedaan antara keduanya yaitu antara luas yang dimiliki oleh EuclideanDistance sedangkan pada kasus ChebyshevDistance jarak maksimum

3.3. Lokasi Penelitian

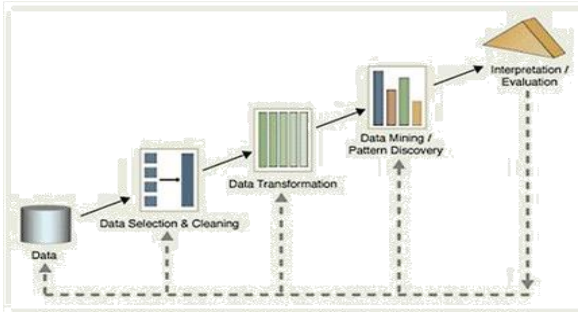
Tempat lokasi penelitian yang dilakukan oeh peneliti yaitu penelitian secara deskripsi karena sumber data dilakukan dari Kantor Kecamatan Sukagumiwang Adapun alamat sebagai berikut:



Gambar 1. Lokasi Penelitian

3.4. Teknik Analisi Data

Teknik yang digunakan untuk menganalisis data Penambangan ini menggunakan tahapan proses KDD (Knowledge Discovery in Databases) yang terdiri dari data, *pembersihan data*, *transformasi data*, *Data mining*, *Evolusi*, seperti terlihat gambar 2:



Gambar 2. Knowledge Discovery In Database

a) Darabase
Data analisa survaey kepuasan masyarakat didapat dari dataset hasil survey kepuasan masyarakat yang bersumber dari Pemerintahan Kecamatan Sukagumiwang dengan menggunakan data pada penelitian ini ialah sebanyak 88 dataset dan 10 Atribut yaitu: No. Resp, u1, u2, u3, u4, u5, u6, u7, u8, u9.

b) Data Cleaning
Secara umum, data yang diperoleh baik dari database maupun survei berisi entri yang salah seperti: B. Data hilang, data tidak valid, atau kesalahan ketik sederhana. Disarankan juga untuk membuang data yang tidak relevan, karena keberadaannya nantinya dapat mempengaruhi kualitas dan keakuratan hasil data mining. Pembersihan data juga memengaruhi kinerja sistem penambangan data dengan mengurangi jumlah dan kompleksitas pemrosesan data.

c) Data Transformation
Tahap ini adalah Proses transformasi data terpilih agar sesuai untuk proses data mining.

d) Data Mining
Proses pemeriksaan dan analisis data dalam jumlah besar, dengan menggunakan teknik dan metode tertentu, dengan tujuan menemukan pola dan informasi yang menarik dalam data yang disimpan dalam jumlah besar. Teknik, metode, dan algoritme yang tepat bergantung pada tujuan Anda dan keseluruhan proses KDD (penemuan pengetahuan dalam basis data).

e) Evaluation
Penerjemah alur yang dihasilkan oleh data mining harus mewakili aliran informasi yang dihasilkan oleh suatu proses dalam bentuk yang mudah dipahami oleh pihak yang berkepentingan. Proses ini adalah proses KDD-nya yang melibatkan pemeriksaan apakah aliran atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada.

4. HASIL DAN PEMBAHASAN

Hasil penelitian yang dilakukan dalam pembahasan ini yaitu akan menguraikan proses bagaimana pengklasifikasian atau clusteriasi dataset Survei Kepuasan Masyarakat dengan metode Algoritma K-Mean. pengelompokkan dengan menggunakan proses pengujian Machine Learning yaitu Software RapidMiner Studio.

4.1. Data

Data dalam bentuk file excel yang bernama dataset Survei Kepuasan Masyarakat diperoleh dari hasil kuesioner.

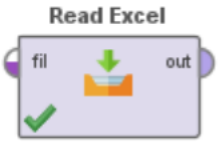
Jumlah record sebanyak 88 record dan terdiri dari 10 atribut seperti tampak pada tabel dibawah ini.

Tabel 1. Atribut default dari file electronic item price

No	Atribut	Type Data	Keterangan
1	No.RESP	integer	Nomor urut responden
2	u1	Integer	Parameter persyaratan
3	u2	Integer	Parameter prosedur
4	u3	Integer	Parameter waktu pelayanan
5	u4	Integer	Parameter biaya atau tarif
6	u5	Integer	Parameter produk/ layanan
7	u6	Integer	Parameter pelaksana
8	u7	Integer	Parameter perilaku pelaksana
9	u8	Integer	Parameter sarana prasana
10	u9	integer	Parameter penangana pengaduan

4.2. Selection

Untuk membaca dataset dalam bentuk *file* excel, menggunakan operator *Read Excel* seperti pada gambar berikut:



Gambar 3. Operator Read Excel pada Rapidminer

Dari hasil pembacaan Operator Read Excel didapat informasi sebagai berikut.

Tabel 2. Operator Read Excel

No.	*Uraian*	*Keterangan*
1.	Record	88
2.	Special Attribute	0
3.	Reguler Atribute	10
4.	Attribute :	
	NO. RESP	Integer
	U1	Integer
	U2	Integer
	U3	Integer
	U4	Integer
	U5	Integer
	U6	Integer
	U7	Integer
	U8	Integer
	U9	Integer

Karena dataset tidak memiliki id, digunakan operator Set Role seperti pada gambar berikut:

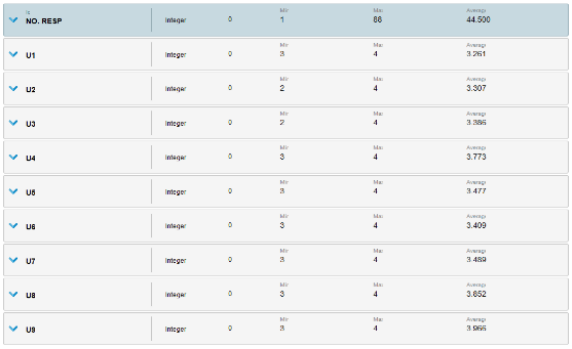
Gambar 4. Operator Set Role

4.3. Preprocessing

Proses *cleansing* atau pembersihan data yang *missing* atau memiliki nilai yang tidak konsisten pada langkah *preprocessing*. Sebelum melakukan proses ini, dilakukan analisa terlebih dahulu apakah atribut pada dataset yang dipilih memiliki nilai *missing* atau tidak serta memiliki data yang konsisten atau tidak.
Dari hasil *result* dari statistik dataset seperti tampak pada gambar dibawah ini, diketahui bahwa tidak ada atribut yang memiliki nilai *missing*. Untuk memeriksa konsisten atau tidak konsistennya dataset yang digunakan diperiksa per-*record* secara langsung dan menunjukan bahwa dataset memiliki data yang konsisten terhadap nilainya.

Result dari uraian statistik dataset tersebut diatas, maka proses Preprocessing tidak dilakukan karena dapat dianggap bahwa dataset tidak memiliki nilai missing dan dataset sudah konsisten.

Model proses pada rapidminer sampai dengan langkah Preprocessing sama dengan gambar sebelumnya.



Gambar 5. Model proses langkah Selection di Rapidminer

4.4. Transformasi

Karena tipe data pada dataset sudah berjenis integer, maka tidak perlu melakukan proses transformasi.

Model proses pada rapidminer sampai dengan langkah *Tranformation* sama dengan gambar sebelumnya.

4.5. Data Mining

Pada langkah data mining karena menggunakan operator *K-Means* pada gambar di bawah ini.

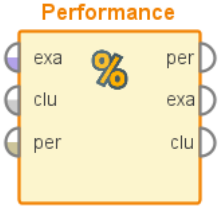
Gambar 6. Operator K-Means Rapidminer

Berikut adalah parameter pada operator *K-Means* yang digunakan tampak pada tabel dibawah ini.

Tabel 3. Operator K-Means

No	**Parameter**	**Isi**
1	K	2-15
2	Measure types	NumericalMeasure
3	Numerical measure	EuclideanDistance
		EuclideanDistance
		CanberraDistance
		ChebychevDistance
		CosineSimilarity
		KernelEuclideanDistance
		ManhattanDistance

Untuk mengeukur performance dari hasil pengelompokan menggunakan operator Cluster Distance Performance seperti gambar berikut.



Gambar 7. Operator Cluster Distance Performance

Parameter pada operator Cluster Distance Performance yang digunakan tampak pada tabel dibawah ini.

Tabel 4. operator Cluster Distance Performance

No.	Parameter	Isi
1	main criterion	davies_bouldin:

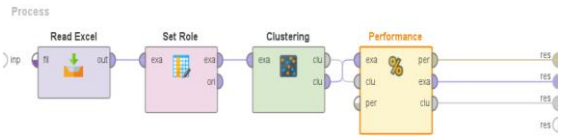
Davies-Bouldin adalah ukuran yang digunakan dalam analisis data untuk mengevaluasi kualitas partisi. Analisis klaster merupakan metode pengelompokan (clustering) data berdasarkan kesamaan fitur atau fitur. Semakin kecil nilai Davies-Bouldin, semakin baik distribusinya.

Fungsi indeks Davies-Bouldin: indeks Davies-Bouldin digunakan untuk mengukur kualitas klasifikasi dari algoritma clustering. Beberapa keunggulannya adalah:

1. Validasi Cluster: Indeks ini membantu memastikan bahwa cluster yang dibuat memiliki grup yang bermakna dan terpisah satu sama lain.
2. Evaluasi Metode Pengelompokan: Indeks Davies-Bouldin dapat digunakan untuk membandingkan kinerja berbagai metode. Prosedur yang memberikan nilai Davies-Bouldin lebih rendah dianggap menghasilkan cluster yang lebih baik.

Menentukan distribusi statistik terbaik: Indeks dapat membantu menentukan distribusi statistik yang paling sesuai dengan data dengan menguji beberapa kelompok berbeda.

Model proses pada rapidminer sampai dengan langkah Data mining pada gambar berikut ini.



Gambar 8. Data Mining

4.6. Evaluation

Evaluasi yang dilakukan terhadap hasil kegiatan eksperimen terhadap dataset yang diperoleh hasil sebagai berikut.

Tabel 5. CamberraDistance, dan k : 2

No.	K	Cluster distance	Max-run	dbi
1	2	CamberraDistance	1	1.673
2	2	CamberraDistance	2	1.807
3	2	CamberraDistance	3	1.797
4	2	CamberraDistance	4	1.764
5	2	CamberraDistance	5	1.721
....	...	.....	.....	.....
100	2	CamberraDistance	100	1.764

Tabel pada evaluasi: Cluster distance keseluruhan setiap table terdapat 100 record, namun jumlah record yang ditampilkan hanya 5 record dari 100 record mengingat table tersebut terlalu banyak sehingga ada range antara 1-5…100.

Pada analisa cluster distance: ManhattanDistance dan k: 15 menghasilkan:

Tabel 6. ManhattanDistance dan k : 15

No.	K	Cluster distance	Max-run	dbi
1	2	ChebychevDistance	1	1.799
2	2	ChebychevDistance	2	1.815
3	2	ChebychevDistance	3	1.815
4	2	ChebychevDistance	4	1.801
5	2	ChebychevDistance	5	1.815
....	...	.....	.....	…….
100	2	ChebychevDistance	100	2.073

Setelah dilakukan analisa menggunakan algoritma k-means klustering maka dapat di simpulkan oleh penulis yaitu dapat mengetahui Nilai k optimal yaitu k = 15 dengan nilai dbi = 0,984, selain itu dapat mengetahui Jenis cluster distance yang menghasilkan nilai k yang optimal tersebut adalah ChebychevDistance. Untuk nilai max run menghasilkan nilai k yang optimal adalah 29.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian klastering data survey kepuasan masyarakat menggunakan algoritma K-Means Clustering, maka dapat disimpulkan sebagai berikut : Dapat mengetahui Nilai k optimal yaitu k = 15 dengan nilai dbi = 0,984.

1) Dapat mengetahui Jenis cluster distance yang menghasilkan nilai k yang optimal tersebut adalah ChebychevDistance.
2) Dapat mengetahui nilai max run yang menghasilkan nilai k yang optimal tersebut adalah 29.

DAFTAR PUSTAKA

[1] N. H. Cahyana, “Metode Data Mining K-Means Untuk Klasterisasi Data Penanganan Dan Pelayanan Kesehatan Masyarakat,” pp. 24–31, 2019.

[2] Y. Syahra, D. R. Habibie, M. Nasution, H. N. Nasution, and A. H. Nasyuha, “Klasterisasi Data Penanganan dan Pelayanan Kesehatan Masyarakat dengan Algoritma K-Means,” vol. 9, no. 5, pp. 1423–1433, 2022, doi: 10.30865/jurikom.v9i5.4888.

[3] E. H. Wisanta and Y. N. Marlim, “Analisis Algoritma K-Means Untuk Clustering Kepuasan Pelayanan : Mall Pelayanan Publik Pekanbaru,” 2021.

[4] D. K. Sitinjak, B. A. Pangestu, B. N. Sari, T. Informatika, and U. S. Karawang, “Clustering Jumlah Tenaga Kesehatan Berdasarkan Kecamatan di Kabupaten Karawang Menggunakan Algoritma K-Means,” vol. 6, no. 1, pp. 47–54, 2022.

[5] P. E. ASRONI, FITRI HIDAYATUL, “Penerapan Metode Clustering dengan

Algoritma K-Means pada Pengelompokkan Data Calon Mahasiswa Baru di Universitas Muhammadiyah Yogyakarta (Studi Kasus : Fakultas Kedokteran dan Ilmu Kesehatan , dan Fakultas Ilmu Sosial dan Ilmu Politik),” vol. 21, no. 1, pp. 60–64, 2018, doi: 10.18196/st.211211.

[6] B. Parlambang, J. S. Informasi, P. Minggu, and J. Selatan, “IMPLEMENTASI ALGORITMA K-MEANS DALAM PROSES PENILAIAN KUESIONER KEPADA DOSEN GUNA MENDUKUNG KEPUASAN MAHASISWA TERHADAP DOSEN,” 2020.

[7] P. D. Lestari, D. Kecomberan, K. Talun, M. Clustering, A. K. Clustering, and N. Pengguna, “DATAMINING PADA PENJUALAN AIR BERSIH DI SPAM AKIDAH MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING,” vol. 7, no. 1, pp. 412–416, 2023.

[8] S. Zhang and C. Duan, “Clustering Optimization Algorithm for Data Mining Based on Artificial Intelligence Neural Network,” vol. 2022, 2022.

[9] D. Arkham, D. Swanjaya, T. Informatika, F. Teknik, U. Nusantara, and P. Kediri, “K-Means Method For Clustering Public Service Assessment of Goverment Organization In Kediri City,” pp. 155–160, 2020.

[10] E. E. Lubis, N. Rimayanti, And N. Yohana, “Pengaruh Penggunaan Media Sosial Terhadap,” Pp. 163–172, 2015.

[11] F. P. Hidayat, R. Bagus, F. Hakim, And U. I. Indonesia, “Implementasi Metode Clustering K-Medoids Dalam,” Pp. 106–114, 2021.

[12] I. Dan *Et Al.*, “Systematic Literature Review : Penggunaan Algoritma K-Means Untuk Clustering Di Indonesia Dalam Bidang Pendidikan,” Vol. 2, No. 1, Pp. 19–24, 2021.