

CLUSTERING STOK MATERIAL DI PDAM KOTA MAKASSAR WILAYAH PELAYANAN VI MENGGUNAKAN ALGORITMA K-MEANS

Satria Gunawan Zain¹, Renisa Amalia Tahir², Andi Akram Nur Risal^{*3}

^{1,2,3}Universitas Negeri Makassar, Makassar

Email: ¹satria.gunawan.zain@unm.ac.id, ²renisaamalia1@gmail.com, ³akramandi@unm.ac.id

^{*}Penulis Korespondensi

(Naskah masuk: 12 Maret 2025, diterima untuk diterbitkan: 17 Juni 2025)

Abstrak

Perusahaan Daerah Air Minum (PDAM) Kota Makassar Wilayah Pelayanan VI mengalami kesulitan dalam manajemen stok materialnya terutama dalam hal pengelompokan material berdasarkan rendah atau tingginya penggunaan material. Hal tersebut mengakibatkan seringnya terjadi kekurangan dan kelebihan stok yang dapat menghambat kegiatan dan meningkatkan biaya operasional di PDAM. Oleh karena itu, diterapkan metode *clustering* menggunakan algoritma *k-means* untuk mengelompokkan jenis material berdasarkan tingkat penggunaannya. Penelitian ini bertujuan membantu PDAM dalam mengelola stok material dengan lebih baik dengan mengidentifikasi tingkat kebutuhan berdasarkan pola penggunaan sebelumnya. Penelitian ini meliputi tahap identifikasi masalah, pengumpulan data, *preprocessing* data mencakup data *selection*, data *reduction* dan data *integration*, data *transformation*, dan standarisasi data, penerapan algoritma *clustering*, evaluasi hasil *clustering*, dan visualisasi hasil *clustering*. Hasil *clustering* menunjukkan bahwa jumlah *cluster* optimal yang diperoleh adalah 2, yaitu Cluster_0 (penggunaan rendah) dengan 123 jenis material dan Cluster_1 (penggunaan tinggi) dengan 2 jenis material. Kualitas *cluster* berdasarkan nilai *silhouette* menunjukkan hasil yang cukup baik dimana Cluster_0 sebesar 0.939 dan Cluster_1 sebesar 0.816, dan nilai *silhouette score* yaitu 0.937. Hasil *clustering* sebagai rekomendasi dalam menentukan kebutuhan stok material di masa depan sehingga dapat meningkatkan efisiensi pengelolaan stok material di PDAM Kota Makassar Wilayah Pelayanan VI.

Kata kunci: *Clustering, K-Means, Material, Silhouette Coefficient*

MATERIAL STOCK CLUSTERING AT PDAM MAKASSAR CITY SERVICE AREA VI USING K-MEANS ALGORITHM

Abstract

Regional Drinking Water Company (PDAM) Makassar City Service Region VI has difficulty managing its material stock, especially in grouping materials based on low or high material usage. This results in frequent shortages and excess stock, which can hamper activities and increase operational costs at the PDAM. Therefore, a clustering method using the *k-means* algorithm is applied to group material types based on their level of use. This research aims to assist the PDAM in managing material stocks better by identifying the level of need based on previous usage patterns. This research includes problem identification, data collection, data preprocessing, data selection, data reduction and integration, data transformation, data standardization, the application of clustering algorithms, evaluation of clustering results, and visualization of clustering results. The clustering results show that the optimal number of clusters obtained is 2, namely Cluster_0 (low usage) with 123 types of materials and Cluster_1 (high usage) with two types of materials. Cluster Quality Based on the silhouette value shows quite good results where Cluster_0 is 0.939, Cluster_1 is 0.816, and the silhouette score is 0.937. The clustering results are a recommendation for determining future material stock needs to improve the efficiency of material stock management at PDAM Makassar City Service Area VI.

Keywords: *Clustering, K-Means, Elbow Method, Material, Silhouette Coefficient*

1. PENDAHULUAN

Perusahaan Daerah Air Minum (PDAM) adalah instansi milik pemerintah yang bertanggung jawab atas penyediaan, pengembangan, serta

pendistribusian air bersih kepada masyarakat, salah satunya PDAM Kota Makassar Wilayah Pelayanan VI yang dikelola oleh pemerintah daerah Kota Makassar, Sulawesi Selatan (Annisa, Ginting and Syar, 2022). Pada sebuah perusahaan, termasuk

PDAM, manajemen stok sangat penting untuk mengatur persediaan barang yang bertujuan untuk mencegah terjadinya kekurangan dan kelebihan stok material (Ramdhan *et al.*, 2022), (Mulkan *et al.*, 2023).

Saat ini, PDAM Kota Makassar Wilayah pelayanan VI menghadapi kesulitan dalam mengelola stok material, terutama dalam mengelompokkan material berdasarkan tingkat penggunaan sebelumnya guna mengidentifikasi tingkat kebutuhan stok material. Berdasarkan rekapitulasi data penggunaan material tahun 2023, rata – rata dari stok material yang tersisa di setiap akhir bulan yaitu sebesar 26%. Hal ini menunjukkan adanya ketidakefisienan dalam perencanaan dan pengelolaan stok material. Akibatnya, PDAM sering menghadapi ketidakseimbangan stok material, baik berupa kelebihan atau kekurangan stok. Misalnya, dengan nilai total material bulanan sebesar Rp.100.000.000, rata – rata kelebihan stok sebesar 26% menunjukkan adanya material senilai Rp.26.000.000 yang belum digunakan dan berpotensi meningkatkan biaya penyimpanan, serta menurunkan kualitas material. Di sisi lain, distribusi material yang tidak efisien juga memicu kekurangan material yang lain, sehingga dapat menghambat kegiatan operasional di PDAM dan memicu biaya darurat untuk memenuhi kebutuhan yang mendesak.

Tantangan ini dapat diatasi dengan melakukan pengelompokan stok material berdasarkan pola penggunaan sebelumnya, sehingga tingkat kebutuhan material dapat diidentifikasi dengan baik. Adapun metode yang dapat diterapkan yaitu dengan menggunakan teknik data mining yaitu *clustering*. *Clustering* merupakan metode pengelompokan yang berbasis densitas berdasarkan kemiripan data. Tujuannya adalah untuk menemukan pola – pola tersembunyi yang sebelumnya tidak diketahui berdasarkan informasi yang terdapat pada kelompok data besar tersebut (Nahjan, Heryana and Voutama, 2023)(Aldino *et al.*, 2021).

Terdapat beberapa penelitian terkait penggunaan algoritma *clustering* yaitu *K-Means* yang telah dilakukan sebelumnya. Penelitian oleh Imron *et al.* (2020) yang memanfaatkan algoritma *K-Means* dengan perhitungan manual untuk mengklasterisasi produk di Toko Rizki Barokah berdasarkan popularitasnya dengan menggunakan data penjualan produk selama 3 bulan. Berdasarkan pengujian yang dilakukan menghasilkan dua *cluster* yaitu produk populer dan produk kurang populer. Namun, pada penelitian ini belum dilakukan pengujian atau evaluasi terhadap hasil *clustering*nya (Imron, Hasanah and Humaidi, 2020). Kemudian penelitian Tambunan *et al.* (2020) yang menganalisis data historis beban puncak minimum sistem kelistrikan Jawa Bali menggunakan algoritma *k-means*. Hasil analisis menunjukkan bahwa data dapat diklasifikasikan ke dalam 3 *cluster* yang mewakili

tingkat beban tinggi, menengah, dan rendah (Tambunan *et al.*, 2020).

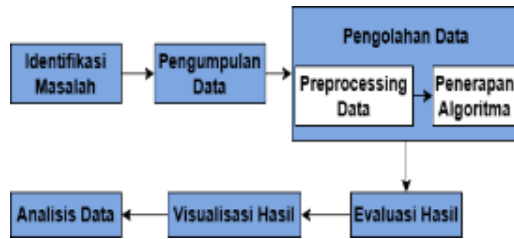
Penelitian oleh Faisal *et al.* (2022) menggunakan algoritma *k-means* untuk mengelompokkan kompetensi TIK siswa SMKN 3 Lhokseumawe berdasarkan 154 data nilai uji kompetensi yang menghasilkan 3 *cluster*, yaitu *cluster* 0 (sangat kompeten) dengan 80 siswa, *cluster* 1 (kompeten) dengan 64 siswa, dan *cluster* 2 (kurang kompeten) dengan 10 siswa (Faisal *et al.*, 2022). Sementara itu, Arifiyanti *et al.* (2022) menggunakan *k-means* untuk perencanaan pemerataan penduduk di DKI Jakarta berdasarkan kepadatan penduduk yang menghasilkan 3 *cluster*, yaitu *cluster* 0 (tidak padat), *cluster* 1 (kompeten), dan *cluster* 2 (kurang kompeten) dengan nilai *silhouette coefficient* sebesar 0.596 (Arifiyanti, Darusman and Trenggono, 2022).

Selanjutnya yaitu penelitian oleh Abdullah *et al.* (2024) yang menerapkan algoritma *K-Means* untuk mengklasterisasi produk di supermarket berdasarkan tingkat kelarisannya dengan menggunakan 18 data produk. Hasil yang diperoleh terdiri dari dua *cluster* yaitu produk tidak laris (C1) dengan 6 produk dan produk laris (C2) dengan 12 produk. Namun, penelitian ini juga belum menerapkan evaluasi terhadap hasil *clustering*nya (Abdullah and Aldisa, 2024).

Berdasarkan hasil yang diperoleh dari beberapa penelitian terkait penerapan algoritma *K-Means Clustering* yang telah dipaparkan di atas, secara keseluruhan, metode ini dapat digunakan untuk melakukan pengelompokan terhadap suatu data. Namun, belum ada penelitian yang secara khusus menerapkan algoritma *k-means* dalam pengelolaan stok material di PDAM Kota Makassar Wilayah Pelayanan VI.

Maka dari itu, dilakukan penelitian dengan judul “*Clustering* Stok Material di PDAM Kota Makassar Wilayah Pelayanan VI Menggunakan Algoritma *K-Means*”. Hal ini bertujuan untuk membantu PDAM memajemen stok material dengan lebih baik, khususnya dalam mengidentifikasi tingkat kebutuhan stok material untuk pemasangan dan perbaikan distribusi air bersih. Penelitian ini tidak hanya berfokus pada penerapan algoritma *K-Means* dalam mengelompokkan stok material, tetapi juga untuk meningkatkan kualitas *cluster* algoritma *k-means*, dilakukan proses *preprocessing* data yang lebih komprehensif termasuk standarisasi data, dan iterasi sebanyak 300 kali dengan perulangan 10 kali menggunakan *centroid* yang berbeda dalam proses pengelompokannya.

2. METODE PENELITIAN



Gambar 1. Tahapan Penelitian

Penelitian ini terdiri dari beberapa tahap seperti yang ditunjukkan pada Gambar 1, yaitu identifikasi masalah, pengumpulan data, pengolahan data yang terdiri dari *preprocessing data* dan penerapan algoritma, evaluasi hasil, visualisasi hasil, dan analisis data.

2.1 Identifikasi Masalah

Penelitian ini berfokus pada permasalahan di PDAM Kota Makassar Wilayah Pelayanan VI terkait kesulitannya dalam mengelola stok material, terutama dalam pengelompokan material berdasarkan rendah atau tingginya penggunaan yang menyebabkan terjadinya kekurangan dan kelebihan stok. Sehingga diperlukan metode untuk mengelompokkan stok material berdasarkan pola penggunaan sebelumnya untuk menentukan tingkat kebutuhan material.

2.2 Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh dari PDAM Kota Makassar Wilayah Pelayanan VI berupa data historis penggunaan material bagian distribusi dan kehilangan air sebanyak 1.120 SPK (Surat Perintah Kerja) selama bulan Januari 2023 hingga Desember 2023.

2.3 Preprocessing Data

Sebelum diimplementasikan, dilakukan preprocessing untuk meningkatkan kualitas hasil clustering.

Tahap pertama yaitu data *selection* merupakan tahap pemilihan data yang sesuai dengan tujuan analisis dari dataset awal penelitian ini agar dimensi data yang akan diolah pada tahap selanjutnya tidak terlalu besar sehingga proses pengolahan data berikutnya menjadi lebih efisien (Putry and Betha Nurina Sari, 2022). Tahap selanjutnya mencakup dua proses yaitu data *reduction*, yaitu menyederhanakan data untuk mengurangi kompleksitas dan volume data tanpa mengurangi integritas data asli dan informasi yang terkandung dalam data tersebut dan data *integration* merupakan proses penggabungan data dari berbagai sumber (Jayakumar and Saravanan, 2022)(Çetin and Yıldız, 2022)(Alvianatinova *et al.*, 2024).

Selanjutnya adalah tahap data *transformation* di mana pada tahap ini dilakukan perubahan terhadap bentuk data menjadi data yang sesuai untuk diproses dengan menggunakan algoritma data mining (Putry and Betha Nurina Sari, 2022). Setelah tahap data *transformation*, dilakukan standarisasi data dengan StandardScaler, yaitu mengurangi nilai atribut numerik dengan rata – ratanya datanya dan membaginya dengan standar deviasinya sehingga menghasilkan distribusi dengan rata – rata (*mean*) 0 dan standar deviasi 1. Metode ini diterapkan untuk memastikan nilai fitur dalam dataset berada dalam skala yang serupa sehingga algoritma tidak cenderung pada fitur dengan rentang nilai yang lebih besar, dan menghasilkan *cluster* yang lebih akurat (Zuo, 2023). Perbandingan hasil sebelum dan setelah dilakukan standarisasi data dapat dilihat pada Tabel 6 dan 7.

2.4 Penerapan Algoritma

Pada penelitian ini digunakan algoritma K-Means untuk mengelompokkan dan mengidentifikasi pola yang terdapat pada data. Adapun tahapannya yaitu sebagai berikut (Izzah and Jananto, 2022).

- Menentukan jumlah *k cluster* yang akan dibentuk. Penelitian ini menggunakan metode *elbow* untuk menentukan jumlah *k cluster* optimal dengan mengidentifikasi titik di mana penurunan nilai *Sum of Squared Errors* (SSE) melambat dan membentuk sudut siku dengan menggunakan persamaan (1) sebagai berikut.

$$SSE = \sum_{k=1}^k \sum_{x_i \in C_k} ||x_i - c_k||^2 \quad (1)$$

Pada persamaan (1) di atas, C_k adalah himpunan titik data dalam *cluster k*, c_k adalah *centroid* dari *cluster k*, dan $||x_i - c_k||$ adalah jarak *Euclidean* antara titik x_i dan *centroid* c_k (Kurniawan *et al.*, 2023) (Maravillas, 2023).

- Menentukan nilai ambang batas dan jumlah batas iterasi maksimum.
- Menentukan titik pusat *cluster* secara acak sebagai *centroid* awal sejumlah *k cluster*.
- Menghitung jarak dari setiap objek terhadap setiap *centroid* di setiap *cluster* dengan menggunakan perhitungan metrik jarak *Euclidean Distance* menggunakan persamaan (2) berikut.

$$d(x_i, y_i) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}, \text{ dimana } i = 1, 2, 3, \dots, n. \quad (2)$$

Pada persamaan (2), d adalah jarak, i adalah jumlah data, y adalah *centroid*, dan x adalah *centroid*.

- Menentukan *centroid* yang paling dekat untuk setiap objek menggunakan persamaan (3) berikut dan alokasikan objek tersebut ke *centroid* tersebut.

$$v = \frac{\sum_{i=1}^n x_i}{n}, \text{ dimana } i = 1, 2, 3, \dots n. \quad (3)$$

Pada persamaan (3), v adalah *centroid* pada *cluster*, x_i adalah objek ke- i , dan n adalah jumlah objek yang menjadi anggota *cluster*.

- f. Melakukan iterasi untuk menentukan posisi *centroid* baru hingga posisi *centroid* tidak ada lagi yang berubah atau jumlah batas iterasi maksimum yang telah ditentukan atau konvergensi tercapai.

2.5 Evaluasi Hasil

Evaluasi hasil dilakukan dengan metode *silhouette coefficient* karena metode tersebut tidak hanya mempertimbangkan jarak antar *cluster*, tetapi juga kedekatan data dalam *cluster*. Selain itu, metode ini mudah diinterpretasikan karena memiliki rentang nilai yang jelas, yaitu dari -1 hingga 1, di mana semakin mendekati nilai 1 maka semakin baik kualitas *clustering* yang dihasilkan (Shahapure and Nicholas, 2020). Metode *silhouette coefficient* yang terdiri dari beberapa tahap, yaitu sebagai berikut (Simanjuntak and Khaira, 2021).

- a. Hitung rata – rata jarak antara sebuah titik data (i) dengan semua titik data lain dalam *cluster* yang sama $a(i)$ dengan persamaan (4) berikut.

$$a(i) = \frac{1}{|A|-1} \sum_{j \in A, j \neq i} d(i, j) \quad (4)$$

Pada persamaan (4) terdapat j yaitu titik data lain dalam sebuah *cluster* A dan $d(i, j)$ yaitu jarak titik data i dan titik data j .

- b. Hitung rata – rata jarak antara titik data (i) tersebut dengan semua titik data dalam *cluster* terdekat dan mengambil nilai yang paling kecil menggunakan persamaan (5).

$$b(i) = \frac{1}{|B|} \sum_{j \in B} d(i, j) \quad (5)$$

Pada persamaan (5) terdapat $b(i)$ yaitu rata – rata jarak antara titik data i dengan semua titik data dalam *cluster* lain, $|B|$ adalah jumlah titik data dalam *cluster* B, dan $d(i, j)$ adalah jarak antara titik data i dan titik data j .

- c. Hitung *silhouette coefficient* $S(i)$ dengan menggunakan persamaan (6) berikut.

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (6)$$

Pada persamaan (6) terdapat $a(i)$ yaitu rata – rata jarak antara titik data i dengan semua titik data dalam *cluster* yang sama, $b(i)$ yaitu rata – rata jarak antara titik data i dengan semua titik data dalam *cluster* lain, dan $\max(a(i), b(i))$ yaitu nilai terbesar antara $a(i)$ dan $b(i)$.

2.6 Visualisasi Hasil

Visualisasi terhadap *clustering* dilakukan menggunakan perangkat lunak Tabelau dengan

mengimport data hasil *clustering* yang berbentuk file *excel* yang kemudian divisualisasikan dalam bentuk diagram.

2.7 Analisis Data

Pada penelitian ini, dilakukan analisis data untuk mengidentifikasi material apa saja yang secara konsisten digunakan dalam pemasangan dan perbaikan pelanggan di PDAM Kota Makassar Wilayah Pelayanan VI berdasarkan hasil *clustering* dan evaluasi hasilnya yang dilakukan menggunakan bahasa pemrograman Python pada perangkat lunak Visual Studio Code yang kemudian divisualisasikan dengan perangkat lunak Tableau. Hasil analisis tersebut akan digunakan untuk menentukan jenis material yang perlu diperbanyak atau dikurangi stoknya.

3. HASIL DAN PEMBAHASAN

Pada penelitian ini, digunakan data histori penggunaan material di PDAM sebanyak 1.120 SPK (Surat Perintah Kerja) selama bulan Januari hingga Desember 2023 yang berbentuk tabel file *excel* dengan format file *xlsx*. Dataset yang akan digunakan memiliki beberapa atribut seperti yang ditunjukkan pada Tabel 1 berikut.

Pada Tabel 1 dataset awal tersebut memiliki 7 atribut yaitu No, Hari/Tgl, BPP, SPK (Surat Perintah Kerja) Kegiatan, Material yang memiliki 3 atribut yaitu Jenis, Jml, dan St, Petugas, dan Ket.

Pada dataset awal dilakukan *preprocessing* data, yaitu data *selection*. Pada data *selection* dilakukan penyeleksian dan penghapusan beberapa atribut yang tidak digunakan yang hasilnya dapat dilihat pada tabel 2.

Hasil data *selection* pada Tabel 2 menunjukkan 4 atribut yang akan digunakan pada penelitian ini, yaitu No., Hari/Tgl, Jenis Material, dan Jml.

Kemudian dilakukan data *reduction* dan data *integration*. Pada data *reduction* dilakukan penggabungan jenis material yang digunakan berulang kali dalam setiap bulannya sehingga dalam satu bulan hanya tersisa satu saja dalam setiap jenisnya, sedangkan pada data *integration* dilakukan penjumlahan terhadap semua jumlah penggunaan material dalam setiap jenisnya dalam satu bulan.

Jadi, data yang sebelumnya berbentuk penggunaan material per pekerjaan disederhanakan menjadi penggunaan material perbulan yang hasilnya dapat dilihat pada Tabel 3, Tabel 4, dan Tabel 5.

Tabel 3 merupakan hasil data *reduction* dan data *integration* dari bulan Januari di mana pada tabel tersebut dapat dilihat perubahan dari Tabel 2 yang sebelumnya berbentuk data penggunaan material dari setiap pekerjaan menjadi data penggunaan material per bulan dari setiap jenis material. Begitupun pada Tabel 4 dan Tabel 5.

Tabel 1. Dataset Awal Januari

No	Hari/Tgl	BPP	SPK Kegiatan	Material			Petugas	Ket.
				Jenis	Jml	St		
1	02 Jan 2023	42/BPP/V/2023	Kebocoran	Pipa HDPE ¾ inch	5	Mtr	Arif Fauzi	-Saenal P
				Double Nepple GI 3/4	2	Bh		
				Repair Clamp Steel SNI 4	1	Bh		
				Socket GI 3/4	2	Bh		
				Cleam Saddle.B (HDPE) 4x1 1/4	2	Bh		
				Ferrule.B (HDPE) 1 1/4x3/4	2	Bh		
				Knie HDPE 3/4	6	Bh		
134	30 Jan 2023	727/BPP/V/2023	Kebocoran	Double Nepple GI 1/2	2	Bh	Muhtar Kolle Hamja Ilman	Muh. Mahir
				Stop Kran 1/2	1	Bh		
				Varlop Knie GI 1/2 x 3/4	2	Bh		
				Isolasi Pipa	2	Roll		

Tabel 2. Hasil Data *Selection* Januari

No	Hari/Tgl	Jenis Material	Jml
1	02 Jan 2023	Pipa HDPE 3/4 inch	5
		Double Nepple GI 3/4	2
		Repair Clamp Steel SNI 4	1
		Socket GI 3/4	2
		Cleam Saddle.B (HDPE) 4x1 1/4	2
		Ferrule.B (HDPE) 1 1/4x3/4	2
		Knie HDPE 3/4	6
134	30 Jan 2023	Double Nepple GI 1/2	2
		Stop Kran 1/2	1
		Varlop Knie GI 1/2 x 3/4	2
		Isolasi Pipa	2

Tabel 3. Hasil Data *Reduction* dan *Integration* Januari

Januari		
No.	Jenis Material	Jml
1	Brass Ball Valve With Lock GI 3/4 x 1/2	60
2	Check Valve 1/2	17
3	Cleam Saddle.B (HDPE) 2X1 1/4	17
...	...	...
55	Water_Mour 3/4	9

Tabel 4. Hasil Data *Reduction* dan *Integration* Februari

Februari		
No.	Jenis Material	Jml
1	Brass Ball Valve With Lock GI 3/4 x 1/2	55
2	Check Valve 1/2	28
3	Cleam Saddle.B (HDPE) 2X1 1/4	22
...	...	...
32	Water_Mour 3/4	1

Tabel 5. Hasil Data *Reduction* dan *Integration* Maret

Maret		
No.	Jenis Material	Jml
1	Brass Ball Valve With Lock GI 3/4 x 1/2	58
2	Check Valve 1/2	58
3	Cleam Saddle.B (HDPE) 2X1 1/4	52
...	...	...
43	Water_Mtr Air GI 1/2	2

perubahan terhadap format atau bentuk tabel data yang akan diolah, yaitu data penggunaan material bulan Januari hingga Desember dari tahap sebelumnya digabung dan mengganti atribut Jml dengan nama bulan penggunaan material yang berisikan jumlah penggunaan material dari setiap jenisnya dengan hasil seperti pada tabel 6.

Tabel 6 menampilkan hasil dari tahap data *transformation* di mana pada tabel tersebut dapat dilihat bahwa semua jenis material dari bulan Januari hingga Desember digabungkan dalam satu kolom atau atribut Jml yang sebelumnya berisikan jumlah penggunaan dari setiap jenis material per bulan diganti menjadi nama bulan dari penggunaan materialnya, yaitu atribut Januari yang berisikan jumlah penggunaan jenis material dari bulan Januari, Februari yang bersikan jumlah penggunaan jenis material dari bulan Februari, begitupun hingga atribut Desember. Dari tahap ini diperoleh sebanyak 125 jenis material yang digunakan selama bulan Januari 2023 hingga Desember 2023.

Setelah itu, data kemudian disimpan dengan format csv untuk dapat diimplementasikan ke dalam algoritma *K-Means Clustering* dengan bahasa pemrograman *Python*. Namun sebelum itu, dilakukan standarisasi numerik agar atribut dalam dataset memiliki nilai dalam skala yang seragam. Standarisasi data ini dilakukan dengan menghitung nilai rata – rata dari setiap atribut mulai dari jumlah penggunaan material bulan Januari hingga Desember. Kemudian dilanjutkan dengan menghitung standar deviasi dari masing – masing atribut tersebut dengan mengurangi nilai dari setiap titik data dengan rata – rata dari masing – masing atribut lalu hasilnya dikuadratkan dan di rata – ratakan lalu di akar kuadratkan. Kemudian, nilai asli dari setiap titik data pada masing – masing atribut dikurangkan dengan nilai rata – ratanya dan dibagi dengan standar deviasi dari atribut tersebut. Hasil dari standarisasi data ini dapat dilihat pada Tabel 7 .

Selanjutnya dilakukan tahap data *transformation* di mana pada tahap ini dilakukan

Tabel 6. Hasil Data *Transformation*

No	Jenis Material	Jan	Feb	Maret	April	Mei	Jun	Jul	Agust	Sep	Okt	Nov	Des
1	Air Valve GI 1	0	0	0	0	3	0	0	0	12	1	0	0
2	Bend 90° FOR HDPE 2	0	0	0	0	0	1	0	0	0	0	4	0
3	Bend 90° FOR HDPE 3	0	0	0	0	4	0	2	0	0	0	0	0
4	Bend 90° FOR HDPE 4	0	0	0	0	0	1	0	0	0	0	0	0
5	Bend 90° FOR HDPE 6	0	0	0	0	4	0	0	0	0	0	0	0
6	Brass Ball Valve With Lock GI 3/4 x 1/2	60	55	58	25	34	35	48	31	13	5	5	4
...	...	...	...	...	...	...	...	...	...	...	...	...	...
125	Water_Mour 3/4	9	1	0	2	3	12	11	4	11	9	11	5

Tabel 7. Hasil Standarisasi Data

No.	Jan	Feb	Mar	Apr	Mei	Jun	Juli	Agust	Sep	Okt	Nov	Des
1	-0.270	-0.250	-0.289	-0.258	-0.302	-0.288	-0.253	-0.269	0.114	-0.198	-0.237	-0.203
2	-0.270	-0.250	-0.289	-0.258	-0.371	-0.269	-0.253	-0.269	-0.246	-0.235	-0.134	-0.203
3	-0.270	-0.250	-0.289	-0.258	-0.279	-0.288	-0.227	-0.269	-0.246	-0.235	-0.237	-0.203
4	-0.270	-0.250	-0.289	-0.258	-0.371	-0.269	-0.253	-0.269	-0.246	-0.235	-0.237	-0.203
5	-0.270	-0.250	-0.289	-0.258	-0.279	-0.288	-0.253	-0.269	-0.246	-0.235	-0.237	-0.203
...	...	...	...	...	...	...	...	...	...	...	...	...
125	-0.055	-0.250	-0.289	-0.186	-0.302	-0.064	-0.111	-0.191	0.084	0.105	0.045	-0.035

Setelah itu, dilakukan iterasi *k-means* untuk menentukan jumlah *cluster* dengan metode *elbow* dengan menghitung nilai *Sum of Square Errors* (SSE). Iterasi dilakukan mulai dari jumlah *cluster* dalam rentang 1 hingga 10 *cluster* dengan maksimal iterasi atau konvergensi sebanyak 300 iterasi, serta dengan 10 kali pengulangan inisialisasi dari masing – masing *cluster* dengan *centroid* yang berbeda. Adapun nilai SSE untuk setiap jumlah *cluster* dapat dilihat pada Tabel 8 berikut dan grafiknya pada Gambar 2.

Berdasarkan hasil perhitungan nilai *Sum of Squared Errors* (SSE) yang dapat dilihat pada Tabel 8 dan grafik *Elbow* pada Gambar 2, menunjukkan bahwa jumlah *cluster* yang paling optimal adalah saat *k cluster* berjumlah 2. Hal itu dikarenakan pada saat *k cluster* berjumlah 2, penurunan nilai SSE mulai melambat dan grafik SSE yang paling membentuk siku.

Tabel 8. Hasil Perhitungan Nilai SSE

Jumlah Cluster	Nilai SSE
1	1500.000
2	339.959
3	164.630
4	113.482
5	79.427
6	64.712
7	55.817
8	46.465
9	41.790
10	36.830

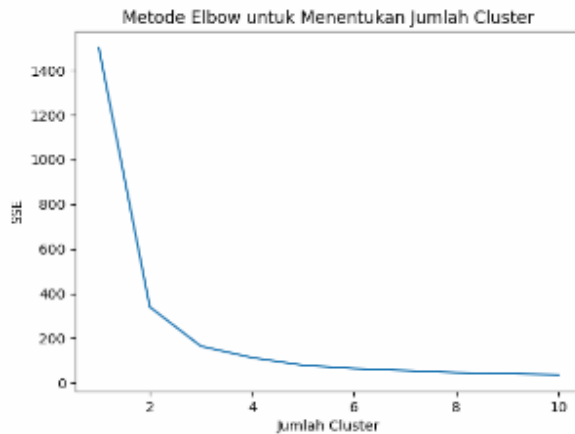
Di mana pada saat jumlah *k cluster* sama dengan 1, nilai SSE-nya tinggi karena titik data tergabung dalam satu kelompok yang besar sehingga jarak antara titik data dengan titik pusat *centroid* *clusternya* menjadi jauh. Dengan menambah jumlah *cluster*, titik data dapat dikelompokkan menjadi *cluster* yang lebih kecil yang mengakibatkan jarak antara titik data dengan titik pusat *centroidnya* menjadi lebih dekat dan nilai SSE-nya menurun.

Namun, setelah jumlah *k cluster* sebanyak 2, penurunan nilai SSE antara 1 *cluster* dan 2 *cluster*. Hal ini menunjukkan bahwa menambah jumlah *k cluster* tidak menghasilkan peningkatan yang signifikan lagi dalam kualitas *clusteringnya*.

Selanjutnya, dilakukan inisialisasi algoritma *K-Means* dengan jumlah *cluster* sebanyak 2 dengan 10 kali pengulangan inisialisasi dari masing – masing *cluster* dengan *centroid* yang berbeda untuk memastikan hasil *clustering* yang konsisten dan tidak bergantung pada pemilihan *centroid* awal. Proses iterasi dilakukan dengan batas maksimal 300 iterasi untuk menjaga efisiensi serta memastikan konvergensi tercapai dengan posisi *centroid* yang stabil dengan batas konvergensi (tol) sebesar 0.0001 sehingga menghasilkan *cluster* dengan pengelompokan yang optimal. Adapun hasil *clustering* yang diperoleh dengan jumlah *k cluster* sama dengan 2 dapat dilihat pada tabel di bawah ini.

Adapun pada Tabel 9 terdapat 123 jenis material yang termasuk dalam *Cluster 0*. Material yang termasuk dalam *cluster* ini memiliki tingkat penggunaan yang rendah, artinya material ini tidak selalu digunakan dalam kegiatan operasional sehari – hari.

Misalnya Air Valve GI 1 yang sering digunakan dalam pekerjaan “tidak dapat air” untuk mengontrol dan mengeluarkan udara yang terperangkap dalam pipa, sehingga memastikan aliran air kembali normal dan tekanan dalam pipa tetap stabil. Namun, pengaduan pelanggan terkait masalah “tidak dapat air” tergolong sedikit sehingga jenis material ini tidak perlu ditingkatkan jumlahnya, cukup memastikan ketersediaannya saat dibutuhkan tanpa harus menyimpan dalam jumlah besar. Begitupun dengan material yang lain yang termasuk dalam *cluster* ini.



Gambar 2. Grafik Elbow

Tabel 9. Hasil *Clustering* Material *Cluster 0*

No	Jenis Material <i>Cluster 0</i>
1	Air Valve GI 1
2	Bend 90° FOR HDPE 2
3	Bend 90° FOR HDPE 3
4	Bend 90° FOR HDPE 4
5	Bend 90° FOR HDPE 6
6	Brass Ball Valve With Lock GI 3/4 x 1/2
7	Check Valve 1/2
8	Cleam Saddle 2x3/4
9	Cleam Saddle 3x3/4
10	Cleam Saddle 4x3/4
...	...
123	Water_Mour 3/4

Hal ini bertujuan untuk menghindari kelebihan stok yang bisa menyebabkan pemborosan, baik dari sisi ruang penyimpanan maupun biaya yang dikeluarkan untuk penyimpanan tambahan jika terdapat material yang berlebih. Selain itu, kelebihan stok untuk jenis material dengan tingkat penggunaan yang rendah juga meningkatkan risiko kerusakan material karena terlalu lama disimpan dan tidak digunakan.

Tabel 10. Hasil *Clustering Cluster 1*

No	Jenis Material <i>Cluster 1</i>
1	Isolasi Pipa
2	Knie HDPE 3/4

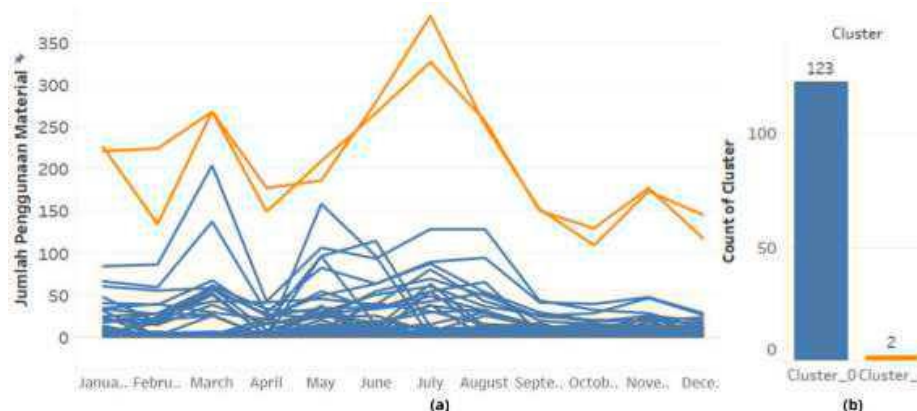
Tabel 10 menunjukkan bahwa hanya terdapat 2 jenis material yang termasuk dalam *Cluster 1*, yaitu Isolasi Pipa dan Knie HDPE 3/4, dimana terdapat

perbedaan yang signifikan terhadap jumlah material yang termasuk dalam *Cluster 0* dan *Cluster 1*. Hal ini disebabkan oleh tingkat penggunaan material di *Cluster 1* yang jauh lebih tinggi, sehingga algoritma *k-means* secara otomatis memisahkan keduanya berdasarkan pola penggunaan yang jelas. Dalam proses pemasangan dan perbaikan di PDAM terutama pada kebocoran dan rehabilitasi, isolasi pipa dan knie HDPE 3/4 sangat diperlukan. Isolasi pipa berfungsi untuk memperkuat sambungan antara dua pipa atau komponen untuk mencegah kebocoran, serta memastikan sambungan tahan terhadap tekanan air sehingga aliran air tetap lancar. Sementara itu, knie HDPE 3/4 digunakan untuk menyambungkan pipa dalam instalasi pipa distribusi air atau jaringan pipa PDAM, terutama pada titik rawan kebocoran seperti sambungan rumah tangga atau distribusi kecil. Kedua material ini menjadi kebutuhan utama karena pekerjaan terkait kebocoran dan rehabilitasi merupakan aktivitas yang paling sering dilakukan di PDAM.

Setelah itu, hasil *clustering* tersebut akan dievaluasi dengan menghitung nilai *silhouette coefficient* dan *silhouette score*nya sehingga diperoleh hasil yang dapat dilihat pada tabel di bawah ini.

Tabel 11. Nilai Evaluasi Hasil *Clustering Cluster 0*

No	Jenis Material	<i>Silhouette Coefficient Cluster 0</i>
1	Air Valve GI 1	0.954
2	Bend 90° FOR HDPE 2	0.964
3	Bend 90° FOR HDPE 3	0.966
4	Bend 90° FOR HDPE 4	0.966
5	Bend 90° FOR HDPE 6	0.966
6	Brass Ball Valve With Lock GI 3/4 x 1/2	0.825
7	Check Valve 1/2	0.895
8	Cleam Saddle 2x3/4	0.966
9	Cleam Saddle 3x3/4	0.965
10	Cleam Saddle 4x3/4	0.966
...	...	...
123	Water_Mour 3/4	0.942
Silhouette Score Cluster_0		0.939



Gambar 3. (a) Grafik Distribusi Material (b) Jumlah Material Setiap Cluster

Tabel 12. Nilai Evaluasi Hasil *Clustering Cluster 1*

No	Jenis Material	<i>Silhouette Coefficient Cluster 1</i>
1	Isolasi Pipa	0.810
2	Knie HDPE 3/4	0.823
Silhouette Score Cluster_1		0.816

Terkait hasil evaluasi *cluster* pada Tabel 11, diperoleh nilai *silhouette score* untuk *Cluster 0* yaitu sebesar 0.939. Hasil tersebut menunjukkan bahwa setiap titik data yang ada pada *Cluster 0* memiliki kesesuaian yang tinggi dengan *clusternya* sendiri. Artinya, titik – titik data dalam *cluster* tersebut memiliki tingkat kerapatan yang tinggi di dalam *clusternya* karena lebih dekat ke pusat *cluster* mereka sendiri daripada ke pusat *cluster* yang lain. Sementara itu, nilai *silhouette score* untuk *Cluster 1* yang dapat dilihat pada Tabel 12, juga menunjukkan kesesuaian yang baik meskipun nilainya sedikit lebih rendah daripada *Cluster 0* yaitu sebesar 0.816. Hal ini dikarenakan perbedaan jumlah titik data antara titik – titik data di *Cluster 1* yang lebih sedikit dibandingkan dengan jumlah titik data pada *Cluster 0*, serta tingkat kerapatan antara titik – titik data di *Cluster 1* yang lebih rendah. Adapun nilai *silhouette score* hasil *clustering* tersebut secara keseluruhan menunjukkan hasil yang cukup baik yaitu sebesar 0.937.

Kemudian visualisasi dari hasil *clustering* tersebut dapat dilihat pada gambar 3 di atas. Gambar 3 (a) menunjukkan grafik distribusi material di setiap *cluster*. *Cluster 0* yang diwakili oleh garis yang berwarna biru mencakup jenis material dengan tingkat penggunaan yang lebih rendah. Sementara itu, *Cluster 1*, yaitu jenis material dengan tingkat penggunaan rendah ditunjukkan oleh garis berwarna oranye. *Cluster* ini mencerminkan kebutuhan material yang dominan, meskipun hanya terdiri dari dua jenis material. Penggunaan kedua material ini mengalami peningkatan yang sangat signifikan pada bulan Agustus yang dipicu oleh peningkatan laporan atau pengaduan yang masuk terkait aktivitas perbaikan di PDAM.

Berdasarkan gambar 3 (b) di atas, dapat dilihat bahwa dari 123 jenis material yang ada, terdapat 123

jenis material atau 98,4% dari total material termasuk dalam *Cluster 0* atau jenis material dengan tingkat penggunaan yang rendah dan 2 jenis material atau 1,6% dari total jenis material termasuk dalam *Cluster 1* atau jenis material dengan tingkat penggunaan yang tinggi.

Berdasarkan hasil pengelompokan material, dapat diketahui bahwa jenis material yang paling sering digunakan atau memiliki frekuensi penggunaan yang tinggi adalah jenis material yang berada dalam *Cluster 1*, yaitu Isolasi Pipa dan Knie HDPE 3/4. Kedua jenis material tersebut memiliki frekuensi penggunaan yang tinggi terutama dalam pekerjaan yang berkaitan kebocoran dan rehab. Oleh karena itu, sangat penting untuk meningkatkan stok material tersebut untuk mencegah terjadinya kekurangan stok saat material tersebut dibutuhkan. Karena ketergantungan penggunaan material juga dipengaruhi oleh jenis pekerjaan yang dilakukan, dan berdasarkan pada dataset yang digunakan, kebocoran dan rehabilitasi merupakan pengaduan pelanggan yang sering terjadi dalam beberapa bulan terakhir sehingga menjadikan kedua material tersebut penting untuk selalu tersedia.

Penelitian ini berhasil mencapai tujuan awal, yaitu mengelompokkan material berdasarkan tingkat penggunaannya agar dapat dikelola dengan lebih efisien di PDAM. Hasil ini sejalan dengan penelitian sebelumnya seperti Imron et al. (2020) dan Abdullah et al. (2024) yang menunjukkan bahwa algoritma *clustering k-means* efektif dalam meningkatkan efisiensi pengelolaan stok. Selain itu, nilai *silhouette* dalam penelitian ini juga berhasil ditingkatkan, di mana pada penelitian Arifiyati et al. (2022) memiliki nilai *silhouette coefficient* sebesar 0.596, sementara pada penelitian ini berhasil ditingkatkan menjadi 0.937. Berdasarkan hasil pengelompokan ini, dapat dibuat sistem informasi yang secara otomatis mengelompokkan material. Meskipun rancangan sistem tersebut belum ditentukan secara spesifik dalam penelitian ini, namun hasilnya dapat menjadi dasar bagi pengembangan sistem manajemen pengelolaan stok yang dilakukan secara otomatis.

Namun, metode dalam penelitian ini memiliki keterbatasan, yaitu sangat bergantung pada jumlah

cluster yang ditentukan sebelumnya. Pemilihan jumlah *cluster* yang tidak tepat dapat mengurangi kualitas pengelompokan. algoritma ini juga sensitif terhadap skala data dan *outlier* yang dapat memengaruhi kualitas hasil *clustering*. Oleh karena itu, *preprocessing* data dan metode penentuan jumlah *cluster* optimal, seperti metode *elbow*, sangat penting untuk memastikan hasil yang representatif.

Penelitian selanjutnya dapat difokuskan pada perancangan dan implementasi sistem tersebut, baik itu melalui aplikasi berbasis web, *cloud*, maupun teknologi lainnya yang sesuai dengan kebutuhan operasional PDAM. Dengan adanya sistem ini, pengelolaan stok dapat dilakukan secara lebih efisien, terintegrasi, dan adaptif terhadap perubahan kebutuhan operasional. Hal ini diharapkan dapat mengurangi risiko kekurangan atau kelebihan stok, sekaligus mendukung kelancaran penyedia layanan air bersih.

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, diperoleh jumlah *k cluster* berdasarkan rekomendasi dari metode *elbow* untuk inisialisasi pada algoritma *K-Means* yaitu sebanyak 2 cluster pada pengelompokan stok material di PDAM Kota Makassar Wilayah Pelayanan VI dengan nilai evaluasi metode *silhouette coefficient* secara keseluruhan yaitu sebesar 0.937, dimana nilai ini mendekati 1 sehingga termasuk cukup baik dan membuktikan bahwa algoritma *K-Means* dapat digunakan untuk mengelompokkan jenis material di PDAM berdasarkan tingkat penggunaannya.

Secara keseluruhan, penelitian ini memberikan kontribusi ilmiah dalam pengelompokan material berdasarkan tingkat penggunaannya, yang dapat digunakan untuk meningkatkan efisiensi pengelolaan stok material di PDAM. Berdasarkan hasil pengelompokan yang telah dilakukan, maka dapat diketahui bahwa jenis material yang direkomendasikan untuk ditingkatkan stoknya adalah jenis material yang tergolong ke dalam *Cluster* 1, yaitu Isolasi Pipa dan Knie HDPE 3/4. Rekomendasi tersebut dapat membantu PDAM dalam mengoptimalkan ketersediaan material yang dibutuhkan, mengurangi risiko kekurangan stok, dan mendukung kelancaran operasional.

5. SARAN

Pada penelitian ini, penulis menyadari bahwa masih terdapat beberapa kekurangan sehingga diperlukan penelitian yang lebih lanjut. Adapun saran untuk penelitian selanjutnya yaitu melakukan penambahan jumlah dataset untuk meningkatkan kualitas dari hasil *clustering* dan representasi pola penggunaan material. Selain itu, peneliti selanjutnya juga dapat mengembangkan penelitian untuk diimplementasikan dalam sebuah sistem manajemen stok berbasis web atau *cloud* agar dapat digunakan

secara otomatis dalam pengelolaan stok material di PDAM Kota Makassar Wilayah Pelayanan VI.

DAFTAR PUSTAKA

- ABDULLAH, M.A. & ALDISA, R.T. 2024. Grouping Products In Supermarkets Using The K-Means. *International Journal of Society Reviews (INJOSER)*, 2(4), pp. 912–922. Available at: <https://injoser.joln.org/index.php/123/article/view/125>.
- ALDINO, A.A. *et al.* 2021. Implementation of K-Means Algorithm for Clustering Corn Planting Feasibility Area in South Lampung Regency', *Journal of Physics: Conference Series*, 1751(1). Available at: <https://doi.org/10.1088/1742-6596/1751/1/012038>.
- ALVIANATINOVA, V. *et al.* 2024. Penerapan Algoritma K-Means Clustering Dalam Pengelompokan Data Penjualan Supermarket Berdasarkan Cabang (BRANCH). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(2), pp. 1529–1535. Available at: <https://doi.org/https://doi.org/10.36040/jati.v8i2.8993>.
- ANNISA, K., GINTING, B.S. AND SYAR, M.A. 2022. Penerapan Data Mining Pengelompokan Data Pengguna Air Bersih Berdasarkan Keluhannya Menggunakan Metode Clustering Pada Pdam Langkat. *ALGORITMA: Jurnal Ilmu Komputer dan Informatika*, 6(1), pp. 112–129. Available at: <https://doi.org/http://dx.doi.org/10.30829/algoritma.v6i1.11624>.
- ARIFIYANTI, A.A., DARUSMAN, F.S. AND TRENGGONO, B.W. 2022. Population Density Cluster Analysis in DKI Jakarta Province Using K-Means Algorithm. *Journal of Information System and Informatics*, 4(3), pp. 772–783. Available at: <https://doi.org/http://dx.doi.org/10.51519/journalisi.v4i3.315>.
- ÇETIN, V. AND YILDIZ, O. 2022. A comprehensive review on data preprocessing techniques in data analysis. *Pamukkale University Journal of Engineering Sciences*, 28(2), pp. 299–312. Available at: <https://doi.org/10.5505/pajes.2021.62687>.
- FAISAL, M. *ET al.* 2022. Information and Communication Technology Competencies Clustering for students for Vocational High School Students Using K-Means Clustering Algorithm. *International Journal of Engineering, Science & Information Technology (IJESTY)*, 2(3), pp. 111–120. Available at: <https://doi.org/http://dx.doi.org/10.52088/ijesty.v2i3.318>.
- IMRON, M., HASANAH, U. AND HUMAIDI, B.

2020. Analysis of Data Mining Using K-Means Clustering Algorithm for Product Grouping. *IJIS: International Journal of Informatics and Information Systems*, 3(1), pp. 12–22. Available at: <https://doi.org/10.47738/ijis.v3i1.3>.
- IZZAH, L. AND JANANTO, A. 2022. Penerapan Algoritma K-Means Clustering Untuk Perencanaan Kebutuhan Obat Di Klinik Citra Medika. *Jurnal Ilmiah Komputer*, 18(1), pp. 69–76. Available at: <https://doi.org/http://dx.doi.org/10.35889/progressif.v18i1.769>.
- JAYAKUMAR, R. AND SARAVANAN, D.R.A. 2022. Weather Data Analysis Data Preprocessing. *Journal of Pharmaceutical Negative Results*, 13(8), pp. 3149–3158. Available at: <https://doi.org/https://doi.org/10.47750/pnr.2022.13.S08.390>.
- KURNIAWAN, Y.I. *et al.* 2023. Pengelompokan Prioritas Negara Yang Membutuhkan Bantuan Menggunakan Clustering K-Means dengan Elbow dan Silhouette. *Jurnal Pendidikan dan Teknologi Indonesia (JPTI)*, 3(10), pp. 445–453. Available at: <https://doi.org/https://doi.org/10.52436/1.jpti.343>.
- MARAVILLAS, A.B. 2023. Integration of K-means Algorithm and Elbow Method in Clustering the Bivalve Species. *International Journal of Advance Research in Computer Science and Management Studies*, 11(1). Available at: https://www.researchgate.net/publication/368719963_Integration_of_Kmeans_Algorithm_and_Elbow_Method_in_Clustering_the_Bivalve_Species.
- MULKAN, M.B.I. *et al.* 2023. Aplikasi Peramalan Stok Barang Pada Cv. Sarana Usaha Mandiri Dengan Metode Single Moving Average. *Jurnal Sains Komputer Dan Sistem Informasi*, 1(2), pp. 38–44. Available at: <https://doi.org/https://doi.org/10.61674/jursakoms.v1i2.120>.
- NAHJAN, M.R., HERYANA, N. AND VOUTAMA, A. 2023. Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), pp. 101–104. Available at: <https://doi.org/https://doi.org/10.36040/jati.v7i1.6094>.
- PUTRY, N.M. AND BETHA NURINA SARI, M.K. 2022. Komparasi Algoritma Knn Dan Naïve Bayes Untuk Klasifikasi Diagnosis Penyakit Diabetes Melitus. *Jurnal Sains dan Manajemen*, 10(1), pp. 45–57. Available at: <https://doi.org/https://doi.org/10.31294/evolusi.v10i1.12514>.
- RAMDHAN, D. *et al.* 2022. Clustering Data Persediaan Barang Dengan Menggunakan Metode K-Means. *MEANS (Media Informasi Analisa dan Sistem)*, 7(1), pp. 1–9. Available at: <https://doi.org/10.54367/means.v7i1.1826>.
- SHAHAPURE, K.R. AND NICHOLAS, C. 2020. Cluster quality analysis using silhouette score. *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 747–748. Available at: <https://doi.org/10.1109/DSAA49011.2020.00096>.
- SIMANJUNTAK, K.P. AND KHAIRA, U. 2021. Pengelompokan Titik Api di Provinsi Jambi dengan Algoritma Agglomerative Hierarchical Clustering. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 1(1), pp. 7–16. Available at: <https://doi.org/10.57152/malcom.v1i1.6>.
- TAMBUNAN, H.B. *et al.* 2020. Electrical Peak Load Clustering Analysis Using K-Means Algorithm and Silhouette Coefficient. *2020 International Conference on Technology and Policy in Electric Power & Energy (ICT-PEP)*, pp. 258–262. Available at: <https://doi.org/10.1109/ICTPEP50916.2020.9249773>.
- ZUO, Y. 2023. Investigation on the Impact of Preprocessing Methods and Parameter Selection in Acoustic Scene Classification Based on K-means Clustering Algorithm', in. Atlantis Press International BV, pp. 300–306. Available at: https://doi.org/10.2991/978-94-6463-300-9_30.