

PENERAPAN METODE *K-MEANS CLUSTERING* UNTUK STOK PENJUALAN SPAREPART SEPEDA MOTOR

Akbar Firdaus¹, Sriani², Ali Darta³

Universitas Islam Negeri Sumatera Utara, Medan

e-mail: ¹azhariakbar668@gmail.com, ²sriani@uinsu.ac.id, ³alidarta@uinsu.ac.id

Abstract: *Availability of motorcycle spare parts in CV. Solid Mandiri Cemerlang must be monitored to avoid product shortages. The problem that occurs in reporting regarding Most of the items purchased by Most of the customers is under control. Processing incoming and outgoing goods that are not processed by the system requires goods management techniques. The more complete the types of spare parts, customer needs will be met. The collection of available spare parts will be divided into several groups to get the spare parts that customers have purchased the most for each transaction. Data mining is sourced from raw database. This causes problems in databases which tend to be dynamic, complete and large. The K-means Clustering algorithm is capable and effective for finding clusters in data. This calculation will determine the number of clusters at the calculation center and the maximum iteration of data that has been entered into the system. The purpose of implementing the K-means algorithm is to find the value of the goods purchased by the majority of customers so that it makes it easier to provide spare parts. The results of the k-means calculation: C1 (high) has 9 items, C2 (low) has 1 items.*

Keyword: *Motorcycle Parts, K-Means, Cluster*

Abstrak: Ketersediaan suku cadang sepeda motor di CV. Solid Mitra Cemerlang harus dimonitor untuk menghindari kekosongan barang. Masalah yang terjadi dalam pelaporan mengenai Sebagian besar barang yang dibeli oleh Sebagian besar pelanggan menjadi kendali. Mengolah barang masuk dan keluar yang tidak diproses dengan sistem membutuhkan teknik mengelola barang. Semakin lengkap jenis-jenis suku cadang, kebutuhan pelanggan akan terpenuhi. Pengumpulan suku cadang yang tersedia akan dibagi menjadi beberapa kelompok untuk mendapatkan suku cadang yang paling banyak dibeli pelanggan untuk setiap transaksi. Penambangan data bersumber dari basis data mentah. Hal ini menyebabkan masalah dalam database yang cenderung dinamis, lengkap dan besar. Algoritma *K-means Clustering* mampu dan efektif untuk menemukan *cluster* dalam data. Pada perhitungan ini akan menentukan jumlah *cluster* pada pusat perhitungan dan iterasi maksimum data yang telah dimasukkan kedalam sistem. Tujuan dari penerapan algoritma *K-means* adalah untuk menemukan nilai dari barang yang dibeli oleh Sebagian besar pelanggan sehingga memudahkan untuk menyediakan suku cadang. Hasil perhitungan k-means: C1 (tinggi) ada 9 barang, C2 (rendah) ada 1 barang.

Kata kunci: Suku Cadang Sepeda Motor, *K-Means, Cluster*

PENDAHULUAN

Pengelola perusahaan harus memperhatikan jumlah suku cadang yang dibeli oleh setiap pelanggan. Ketika permintaan pembelian suku cadang meningkat, perusahaan dapat mengambil penentuan untuk menambah stok barang

agar permintaan pelanggan terpenuhi. Semakin lengkap jenis suku cadang yang ada maka kebutuhan dari pelanggan akan terpenuhi. Koleksi dari suku cadang yang tersedia akan dibagi menjadi beberapa *cluster* sehingga diperoleh suku cadang yang paling banyak dibeli pelanggan pada setiap transaksi. Pengelolaan data

penjualan sparepart yang paling banyak dibeli pelanggan pada CV Solid Mitra Cemerlang. Data yang ada sewaktu-waktu dapat digunakan sebagai pendukung keputusan dalam meningkatkan penjualan (Chaerunisa et al., 2021).

Stok barang terdiri dari beberapa komponen yang membentuk suatu kesatuan dan mempunyai fungsi tertentu. Setiap suku cadang mempunyai fungsi tersendiri dan dapat terkait atau terpisah dengan suku cadang lainnya. Adapun item yang terdapat dalam suku cadang yaitu katup/klep (*valve*), c, dan pen piston (*pin clip*). Semua item yang terdapat didalam suku cadang digunakan didalam mesin sepeda motor. Pelanggan dapat membeli suku cadang dengan berbagai item sesuai dengan kebutuhan pelanggan. Item merupakan jenis dari suku cadang dalam setiap transaksi pembelian (Murni et al., 2022).

Dari item akan ditemukan hubungan/kombinasi dari beberapa pembelian sparepart dalam setiap transaksi. Pada tahun 2022 CV. Solid Mitra Cemerlang mengalami penurunan penjualan sampai 30% dari yang biasanya 10234 item menjadi 7163 item sparepart yang terjual diberbagai kota di Sumatera Utara akibat stok yang digudang menumpuk atau cara tata letak barang tidak tersusun dengan baik. CV. Solid Mitra Cemerlang merupakan perusahaan yang bergerak dibidang penjualan suku cadang sepeda motor, dalam data penjualan sparepart pada tahun 2022 perusahaan tersebut belum maksimal dalam melakukan pemenuhan stok barang atau produk sehingga mengalami penurunan penjualan sehingga jumlah permintaan dari konsumen yang fluktuatif mengakibatkan stok yang harus disiapkan menjadi tidak stabil.

Untuk menyelesaikan permasalahan pada penjelasan diatas adalah dengan pemanfaatan algoritma *K-Means Clustering* ini yang dianggap sesuai karena memiliki kelebihan dalam efisiensi, kecepatan dan kemudahan dalam penggunaannya (Amin et al., 2022). *K-means Clustering* mendukung

pengelompokan data dalam jumlah besar yang dilakukan dengan mengalokasikan data pada suatu *cluster* berdasarkan kedekatan jaraknya dengan titik pusat data pada masing-masing *cluster*.

Algoritma K-Means Clustering adalah metode data *clustering* non-hirarki yang mengelompokan data dalam bentuk satu atau lebih cluster atau kelompok. Data-data yang memiliki karakteristik yang sama akan dikelompokan ke dalam satu *cluster* dan data tersebut yang memiliki karakteristik yang berbeda akan dikelompokan ke dalam *cluster* lain sehingga data yang berada dalam satu kelompok tersebut memiliki tingkat variasi yang kecil (Fahmi et al., 2020).

Menurut (Amin et al., 2022) dengan judul “Penerapan Metode *K-Means* dalam penjualan produk Souq.Com” dijelaskan data yang diambil dari Swalayan MM.Tika, menghasilkan dua jenis kelompok data penjualan rendah dan data penjualan tinggi. Sehingga dengan adanya pengelompokan data dapat diketahui jenis barang yang laris terjual dan barang yang di gudang tidak menumpuk.

METODE

Teknik pengumpulan data terbagi dua yaitu data pendukung dan data utama atau primier. Data pendukung terdiri dari penelitian kepustakaan atau studi literatur dan untuk data utama yaitu data yang akan digunakan dengan wawancara dan observasi (Rijal Fadli, 2021). Teknik pengumpulan data yang dilakukan dalam penelitian ini sebagai berikut:

Observasi

Yaitu dengan pengumpulan data dengan pengamatan secara langsung pada dalam penentuan stok penjualan *sparepart* sepeda motor menggunakan *k-means clustering* (studi kasus: CV. Solid Mitra Cemerlang) dengan segala aspek yang berhubungan langsung dengan penelitian.

Wawancara

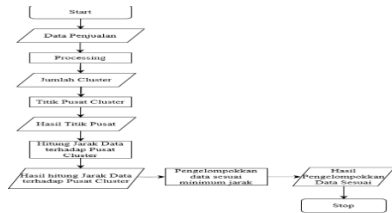
Yaitu dengan melakukan kegiatan

wawancara untuk mencari informasi dengan menggunakan tanya jawab dengan pihak yang berwenang mengenai stok barang sparepart. Data yang digunakan untuk dilakukan metode *k-means clustering* data akan di lampirkan di lampiran.

Studi Pustaka

Melakukan studi pustaka mengenai teori-teori dan konsep yang berhubungan dengan penelitian, seperti teori mengenai keamanan data dengan menggunakan metode *k-means clustering*. Referensi yang digunakan penelitian adalah buku, jurnal ilmiah *online*, dan situs.

Flowchart Sistem



Gambar 1 Flowchart Sistem dengan K-Means

Tabel 1 Dataset

No	Nama Barang	Stok Awal	Barang Masuk	Barang Keluar	Stok Akhir
1	KNZ BLOCK SEHER+RING MIO	2091	50	234	1907
2	KNZ BLOCK SEHER+RING MIO J	2053	12	23	2042
3	KNZ BLOCK SEHER+RING MIO M3	2044	3	5	2042
4	KNZ BLOCK SEHER+RING NEOTECH	2253	212	40	2425
5	KNZ BLOCK SEHER+RING NMAX	2263	222	156	2329
...
10	KNZ BLOCK SEHER+RING BEAT F1	2042	1	12	2031

Terdapat rentang data yang besar di column Barang Masuk, Barang Keluar, Stok Akhir dimana pada *column* Barang Masuk nilai minimumnya adalah 1 dan nilai maksimumnya adalah 2321, *column* Barang Keluar nilai minimumnya adalah 1 dan nilai maksimumnya adalah 2049 dan *column* data Stok Akhir nilai minimumnya adalah 18 dan nilai maksimumnya adalah 6673. Rentang data yang besar mempengaruhi hasil akhir *cluster* dimana data dengan rentang besar akan mendominasi *cluster* akhir. Solusi untuk permasalahan tersebut dapat menggunakan *minmaxscaler*. *Minmaxscaler* adalah normalisasi data

HASIL DAN PEMBAHASAN

Preprocessing

Preprocessing adalah tahap pengolahan data mentah menjadi data yang lebih relevan dan mudah dipahami dimana hasilnya dapat berdampak pada proses data mining (Chrisinta & Simarmata, 2024).

Preprocessing pada tahap ini yaitu melakukan normalisasi data (Nur Tiara Shanty et al., 2020). Normalisasi data adalah mengubah fitur data ke dalam skala yang sama tanpa mendistorsi perbedaan rentang nilai data sehingga hasilnya lebih akurat (Ramadhana et al., 2023). *Column* data yang akan digunakan pada penelitian ini yaitu Stok Awal, Barang Masuk, Barang Akhir, Stok Akhir.

dimana data akan ditempatkan pada antara nilai minimum 0 dan nilai maksimum 1.

$$X_{normalisasi} = \frac{X_{awal} - X_{min}}{X_{max} - X_{min}}$$

$X_{normalisasi}$ = Nilai normalisasi data

X_{awal} = Nilai awal sebelum normalisasi

X_{max} = Nilai maksimal pada fitur data

X_{min} = Nilai minimum pada fitur data

Penerapan *minmaxscaler* pada data dapat

dilihat berikut ini.

$$X_{normalisasi} = \frac{2091 - 2042}{4362 - 2042} = \frac{49}{2320}$$

$$X_{normalisasi} = \frac{2091 - 2042}{2320}$$

$$X_{normalisasi} = 0,0211$$

Tabel 2 Hasil Normalisasi Dataset

No.	Nama Barang	Stok Awal	Barang Masuk	Barang Keluar	Stok Akhir
1	KNZ BLOCK SEHER+RING MIO	0,0211	0,0211	0,1138	0,2838
2	KNZ BLOCK SEHER+RING MIO J	0,0047	0,0047	0,0107	0,3041
3	KNZ BLOCK SEHER+RING MIO M3	0,0009	0,0009	0,002	0,3041
4	KNZ BLOCK SEHER+RING NEOTECH	0,0909	0,0909	0,019	0,3617
5	KNZ BLOCK SEHER+RING NMAX	0,0953	0,0953	0,0757	0,3473
...
10	KNZ BLOCK SEHER+RING BEAT F1	0	0	0,0054	0,3025

Perhitungan Algoritma Metode K-Means Clustering

Dalam penelitian ini, peneliti menggunakan salah satu metode dalam Data Mining yaitu Penentuan Stok Penjualan Sparepart Sepeda Motor Menggunakan *K-Means Clustering* (Studi Kasus: CV Solid Mandiri Cemerlang) agar didapat hasil Keputusan Stok Penjualan *Sparepart* Sepeda Motor yang sesuai dengan kemampuan nilai yang dihasilkan.

Tahap pertama yaitu penentuan jumlah *cluster*. Jumlah *cluster* yang ditentukan pada penelitian sebanyak 2, yaitu C1 dan C2. Proses selanjutnya yaitu memilih *centroid awal* atau titik pusat *cluster* awal secara acak (Dcc et al., 2020).

Tabel 3 Data Centroid CENTROID DIPILIH SECARA ACAK

	Stok Awal	Barang Masuk	Barang Keluar	Stok Akhir
C1	0,0953	0,0953	0,0757	0,3473

C2	1	1	0,0044	1
----	---	---	--------	---

Tahap selanjutnya menghitung jarak objek data dengan *centroid*. Penghitungan jarak objek data dengan *centroid* menggunakan rumus *Euclidean distance*.

$$[(x, y), (a, b)] = \sqrt{(x - a)^2 + (y - b)^2}$$

Keterangan:

x: Data pertama (diambil dari atribut pertama)

a: Titik pusat *cluster / centroid* untuk atribut pertama

y: Data kedua (diambil dari atribut kedua)

b: Titik pusat *cluster / centroid* untuk atribut kedua

Setelah didapat titik pusat *cluster* awal, kemudian dilakukan perhitungan *Euclidean distance* dan mengekompokkan berdasarkan jarak. Berikut ini hasil perhitungan jarak *Euclidean Distance* pada data 1. Berikut adalah tabel hasil perhitungan tiap objek data dengan masing-masing *centroid* pada iterasi pertama.

Tabel 4 Hasil Perhitungan K-Means Iterasi Pertama

No.	Nama Barang	C1	C2	Cluster
1	KNZ BLOCK SEHER+RING MIO	0,1284	1,5625	1
2	KNZ BLOCK SEHER+RING MIO J	0,15	1,5702	1
3	KNZ BLOCK SEHER+RING MIO M3	0,1585	1,575	1
4	KNZ BLOCK SEHER+RING NEOTECH	0,0588	1,4355	1
5	KNZ BLOCK SEHER+RING NMAX	0	1,4381	1
...
10	KNZ BLOCK SEHER+RING BEAT F1	0,1585	1,5769	1

Kemudian tentukan kembali *centroid* baru berdasarkan data yang berada di tiap *cluster* yang sama. Proses perhitungan *centroid* baru menggunakan rumus rata-rata.

$$\begin{aligned} \text{Stok Awal}_{(1)} &= (0,0211 + 0,0047 \\ &+ 0,0009 + 0,0909 \\ &+ 0,0953 + 0,0004 \\ &+ 0,0052 + \\ &0,0483 + 0)/9 \end{aligned}$$

$$\text{Stok Awal}_{(1)} = 0,0296$$

$$\begin{aligned} \text{Barang Masuk}_{(1)} &= (0,0211 + 0,0047 \\ &+ 0,0009 + 0,0909 \\ &+ 0,0953 + 0,0004 \\ &+ 0,0052 + \\ &0,0483 + 0)/9 \end{aligned}$$

$$\text{Barang Masuk}_{(1)} = 0,0296$$

$$\begin{aligned} \text{Barang keluar}_{(1)} &= (0,1138 + \\ &0,0107 + 0,002 + 0,019 + 0,0757 + \\ &0 + 1 + 0,0054 \\ &+ 0,0054)/9 \end{aligned}$$

$$\text{Barang Keluar}_{(1)} = 0,1369$$

$$\begin{aligned} \text{Stok Akhir}_{(1)} &= (0,2838 + 0,3041 \\ &+ 0,3041 + 0,3617 \end{aligned}$$

$$\begin{aligned} &+ 0,3473 + 0,3044 + 0 + \\ &0,3361 + \\ &0,3025)/9 \end{aligned}$$

$$\text{Stok Akhir}_{(1)} = 0,2827$$

$$\text{Stok Awal}_{(2)} = 1$$

$$\text{Barang Masuk}_{(2)} = 1$$

$$\text{Barang Keluar}_{(2)} = 0,0044$$

$$\text{Barang Keluar}_{(2)} = 1$$

Setelah melakukan perhitungan rata-rata tiap data yang berada di *cluster* yang sama, didapatkan *centroid* baru. Berikut ini *centroidnya*.

Tabel 5. *Centroid* Baru Setelah Iterasi Pertama

CENTROID BARU

	Stok Awal	Barang Masuk	Barang Keluar	Stok Akhir
C1	0,0296	0,0296	0,1369	0,2827
C2	1	1	0,0044	1

Lakukan kembali perhitungan jarak objek data dengan *centroid* baru pada tabel 4.6 dengan rumus *Euclidean distance*. Berikut adalah hasil perhitungan jarak tiap objek data dengan *centroid* baru menggunakan rumus *Euclidean distance* dipada iterasi kedua.

Tabel 6 Hasil Perhitungan *K-Means* Iterasi Kedua

No.	Nama Barang	C1	C2	Cluster
1	KNZ BLOCK SEHER+RING MIO	0,0261	1,5625	1
2	KNZ BLOCK SEHER+RING MIO J	0,1328	1,5702	1
3	KNZ BLOCK SEHER+RING MIO M3	0,1425	1,575	1
4	KNZ BLOCK SEHER+RING NEOTECH	0,1663	1,4355	1
5	KNZ BLOCK SEHER+RING NMAX	0,1287	1,4381	1
...
10	KNZ BLOCK SEHER+RING BEAT F1	0.01684	1.01971	1

Setelah perhitungan jarak tiap objek data dengan *centroid* baru, didapatkan hasil *cluster* pada tabel 4. Dan tabel 6. tidak ada perbedaan *cluster* data. Jika *cluster* tidak berubah lagi, maka proses sudah selesai dengan 2 iterasi.

Tabel 7 Total Data Tiap Cluster

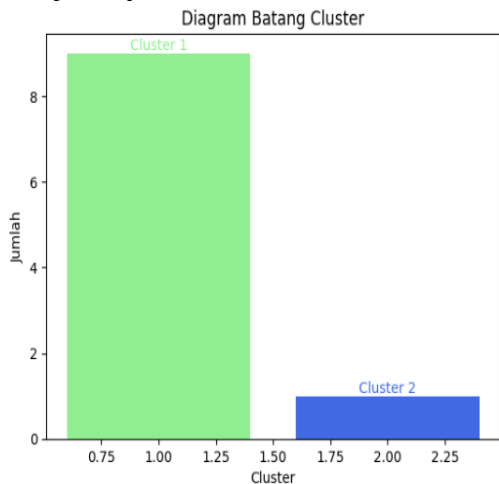
Cluster	Total
1	9
2	1

Berikut ini adalah hasil akhir dari perhitungan yang telah didapatkan untuk kategori data C1 sering dibeli sehingga dibutuhkan stok yang cukup banyak untuk memenuhi permintaan, untuk C2 kategori penjualan jarang dibeli sehingga stok yang dibutuhkan sedikit. Dengan kode barang sebagai berikut:

1. Kode barang C1 [1, 2, 3, 4, 5, 6, 8, 9, 10].
2. Kode barang C2 [7].

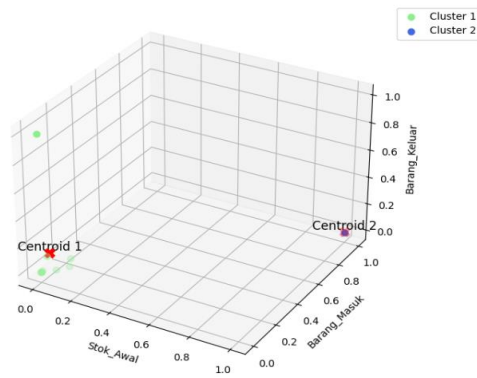
Hasil

Selanjutnya menampilkan jumlah data pada tiap *cluster* dalam bentuk diagram batang. Berikut ini adalah tampilannya.



Gambar 2 Diagram Batang Jumlah Data Tiap Cluster

Tahap terakhir adalah memvisualisasikan data dalam bentuk plot 3D dimana sumbu x adalah Stok Awal, sumbu y adalah Barang Masuk, dan sumbu z adalah Barang Keluar.



Gambar 3. Visualisasi K-means Clustering Pada data_normalisasi

Pada gambar 3. dapat dilihat bahwa titik berwarna hijau mewakili *cluster* 1 dan titik berwarna biru mewakili *cluster* 2 dengan tanda x berwarna merah mewakili *centroid* tiap *cluster*.

SIMPULAN

Berdasarkan penjelasan dan

pembahasan penelitian yang sudah dilakukan maka dapat diambil kesimpulan sebagai berikut.

1. Dari hasil penelitian ini dapat disimpulkan bahwa pengembangan *prototype* untuk mendapatkan barang sparepart sepeda motor yang sangat sering dibeli dan jarang dibeli. Berdasarkan pengelompokan dari berbagai atribut dengan menggunakan metode *algoritma k-means* maka dapat diperoleh C1 terdapat 9 barang *sparepart* yang sering dibeli, C2 terdapat 1 barang jarang dibeli dari 10 data yang digunakan. Dengan adanya pengelompokan data penjualan ini dapat membantu pihak Cv. Solid Mitra Cemerlang.
2. Hasil pemodelan *k-means clustering* berhasil mengidentifikasi pola pembelian konsumen terhadap stok barang. *Cluster* 1 menunjukkan barang yang sering dibeli dengan konsumen memiliki minat yang tinggi terhadap kategori barang tersebut. *Cluster* 2 mewakili barang yang jarang dibeli dan dapat menjadi fokus terhadap pemasaran barang lebih lanjut.
3. Aplikasi ini telah berhasil dibangun dan dapat berfungsi sesuai tujuan, yaitu Tujuan permasalahan yang akan dibahas adalah untuk mengetahui Penentuan Stok Penjualan Sparepart Sepeda Motor.

DAFTAR PUSTAKA

- Amin, F., Anggraeni, D. S., & Aini, Q. (2022). Penerapan Metode K-Means dalam Penjualan Produk Souq.Com. *Applied Information System and Management (AISM)*, 5(1), 7–14. <https://doi.org/10.15408/aism.v5i1.22534>
- Chaerunisa, D. R., Rahaningsih, N., Basysyar, F. M., Purnamasari, A. I., & Suarna, N. (2021). KOPERTIP: Jurnal Ilmiah Manajemen Informatika dan Komputer

- Pengelompokan Penjualan Madu Menggunakan Algoritma K-Means. *KOPERTIP: Jurnal Ilmiah Manajemen Informatika Dan Komputer*, 5, .23-28. <http://jurnal.kopertipindonesia.or.id/23>
- Dcc, A., Lampung, B., Susanto, F., & Marisa, N. (2020). SISTEM PENDUKUNG KEPUTUSAN PENILAIAN KINERJA KARYAWAN DENGAN METODE SIMPLE ADDITIVE WEIGHTING (SAW) (Studi Kasus STMIK Surya Intan Kotabumi). *Jurnal Cendikia*, 405–409.
- Fahmi, M., Daniati, E., Firliana, R., Kunci -SPK, K., & Kinerja Karyawan, P. (2020). Sistem Pendukung Keputusan Penilaian Karyawan Terbaik Menggunakan Metode SAW (Simple Additive Weighting) Abstrak-Best Employee Assessment Decision Support Systems Using Methods SAW (Simple Additive Weighting). *Seminar Nasional Inovasi Teknologi*.
- Murni, D., Efendi, B., Rahmadani, N., Informasi, S., & Tinggi Manajemen Informatika dan Komputer Royal Kisan, S. (2022). IMPLEMENTATION OF EMPLOYEE DISCIPLINE CLUSTERING AT GOTTING SIDODADI VILLAGE OFFICE BANDAR PASIR MANDOGUE USING K-MEANS ALGORITHM. *Jurnal Teknik Informatika (JUTIF)*, 3(2), 295–304. <https://doi.org/10.20884/1.jutif.2022.3.2.236>
- Nur Tiara Shanty, R., Nurhidayat, R., Cahyono, D., & Kunci, K. (2020). Perbandingan Hasil Analisis Klasterisasi Data Bimbingan Konseling Peserta Didik Tingkat Menengah Pertama sebagai Identifikasi Kinerja Akademik STATUS ARTIKEL. *Surabaya Jurnal Sistem Cerdas Dan Rekayasa (JSCR)*, 2(1), 1–10.
- Rijal Fadli, M. (2021). Memahami desain metode penelitian kualitatif. *Humanika, Kajian Ilmiah Mata Kuliah Umum*, 21(1), 33–54. <https://doi.org/10.21831/hum.v21i1>