

Prediksi Barang Sering dan Jarang Terjual Dengan Menggunakan Algoritma *K-Mean Clustering* (Studi Kasus Toko Bina Mulia)

Muhammad Khaerudin ^{1,*}, Imam Zaenuddin ², Tukino ³

¹ Fakultas Ilmu Komputer; Universitas Bhayangkara Jakarta Raya; Jl.Perjuangan No.81, Marga Mulya, Kec.Bekasi Utara, Kota Bekasi, Jawa Barat 17143; e-mail:

muhammad.khaerudin@dsn.ubharajaya.ac.id

² Manajemen Informatika; STMIK Pranata Indonesia; Jl Cut Mutia No 28, Kec. Margahayu, Kota Bekasi, e-mail: imamzaenuddin@gmail.com

³ Fakultas Ilmu Komputer; Universitas Buana Perjuangan, Jl. H.S. Ronggowaluyo Telukjambe Timur Karawang 41361; e-mail: tukino@ubpkarawang.ac.id

* Korespondensi: e-mail: muhammad.khaerudin@dsn.ubharajaya.ac.id

Diterima: 21 Juni 2022 ; Review: 25 Juli 2022; Disetujui: 29 Juli 2022; Diterbitkan: 29 Juli 2022

Abstract

The rapid increase in the population in the buffer zone of the capital city has an effect on the lifestyle of the people in the area, including in the Bekasi district. Likewise, the growth of small and medium community businesses, one of which is the Bina Mulia cooperative shop. This study was made to determine which types of goods are often sold and which types of goods are rarely sold. The algorithm used is K-Means Clustering, where data grouped based on the same characteristics will be included in the same group and the data sets entered into the groups do not overlap. The information displayed is in the form of groups of product names and the amount sold in one week for two months, namely April and May as a sample. The results of this study will help the store in analyzing which types of goods are often sold and which are rarely sold. The software used to help this grouping is Rapid Miner.

Keywords: *Clustering, Item Type, K-Means, Rapid Miner*

Abstrak

Peningkatan jumlah penduduk di wilayah penyangga ibu kota yang pesat berpengaruh terhadap pola hidup masyarakat di wilayah tersebut termasuk di kabupaten Bekasi. Demikian juga dengan pertumbuhan usaha masyarakat kecil dan menengah, salah satunya adalah toko koperasi bina mulia. Penelitian ini dibuat untuk menentukan jenis barang mana saja yang sering terjual dan jenis barang mana saja yang jarang terjual. Algoritma yang digunakan yaitu *K-Means Clustering*, di mana data dikelompokkan berdasarkan karakteristik yang sama akan dimasukkan ke dalam kelompok yang sama dan set data yang dimasukkan ke dalam kelompok tidak tumpang tindih. Informasi yang ditampilkan berupa kelompok – kelompok nama produk dan jumlah terjual dalam satu minggu selama dua bulan yaitu bulan April dan Mei sebagai *sample*. Hasil penelitian ini akan membantu pihak Toko dalam menganalisa jenis barang mana yang sering dan yang jarang terjual. *Software* yang digunakan untuk membantu pengelompokan ini adalah Rapid Miner.

Kata kunci: *Item Type, K-Means, Clustering, Rapid Miner*

1. Pendahuluan

Banyaknya penggunaan barang-barang alat tulis kantor menghasilkan banyaknya juga permintaan pembelian barang ke Toko Bina Mulia, yang membuat banyaknya juga data barang yang terinput dan tak sering membuat data yang diinput tidak terkontrol. Dalam penghitungan jumlah transaksi penjualan yang manual terkadang menghasilkan terjadinya penumpukan jenis barang yang berlebih, terjadinya salah perkiraan antara jumlah penyediaan barang yang akan dijual dengan jumlah barang yang terjual.

Data Mining adalah menganalisis data yang prespektif dan menyimpulkan menjadi informasi penting yang dapat digunakan untuk meningkatkan keuntungan, biaya pengeluaran diperkecil, atau juga bisa keduanya merupakan pengertian dari *Data Mining*. *Data Mining* adalah analisis pada data untuk mencari hubungan yang jelas dan menyimpulkan belum diketahui sebelumnya dengan cara terkini agar mudah dipahami dan bermanfaat (Ardiansyah et al., 2020)

Metode *K-Means*, dan *Clustering* digunakan dalam memecahkan masalah penumpukan penumpukan barang dan kemas dan untuk mengelompokkan barang yang jarang dan sering terjual. *K-Means* adalah algoritma *Clustering* partisi sederhana berbasis prototipe yang mencoba menemukan *Cluster* yang tidak tumpang tindih. *Cluster* ini diwakili dengan sentroid mereka (sentroid *Cluster* biasanya merupakan rata-rata titik dalam *Cluster* itu). Pada algoritma *K-Means* komputer mengelompokkan sendiri data-data yang menjadi masukannya tanpa mengetahui terlebih dahulu target kelasnya. Dengan menggunakan metode ini, data yang telah didapatkan dapat dikelompokkan ke dalam beberapa *Cluster* berdasarkan kemiripan dari data-data tersebut, sehingga data-data yang memiliki karakteristik yang sama dikelompokkan dalam satu *Cluster* dan yang memiliki karakteristik yang berbeda dikelompokkan dalam *Cluster* yang lain yang memiliki karakteristik yang sama (Setiawan, 2016).

Pengelompokan (*Clustering*) adalah suatu metode *Data Mining* yang berfungsi untuk mengelompokkan beberapa data sehingga menghasilkan satu data yang mirip hingga membentuk satu data yang baru. Pengelompokan (*Clustering*) untuk barang yang sering dan jarang terjual perlu dilakukan untuk memudahkan dalam merinci setiap barang yang di pesan kembali. Yang sering terjadi pada penghitungan barang yang terjual stok barang yang berlebih, terjadi kesalahan perkiraan antara jumlah barang yang disediakan dengan yang terjual sehingga bisa menghasilkan kerugian. Tidak tertata data barang yang disediakan setiap bulan.

Untuk mengatasi kesalahan perhitungan perkiraan tersebut maka dapat menggunakan beberapa metode salah satunya metode *K-Means Clustering*. Metode ini mengelompokkan setiap barang yang sering terjual dan yang jarang terjual. Dengan metode ini maka bisa meminimalisir persediaan barang yang berlebih. *Output Data Mining* dapat digunakan untuk membantu pengambilan keputusan di masa depan. *K-Means* adalah metode *Clustering* non-hierarchical yang mencoba untuk mempartisi data yang ada menjadi satu atau lebih *Cluster* atau mungkin bertujuan untuk mempartisi data menjadi beberapa *Cluster* (Aprilianur & Hadisaputro, 2022).

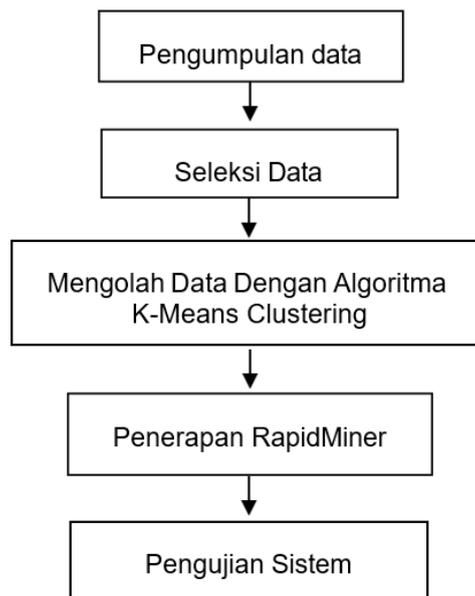
2. Metode Penelitian

Menurut Rohmawati et al (2017) dalam Srisulistiwati et al, prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi (Srisulistiwati et al., 2021)

Menurut Wijaya dan Dwiasnati (2020) dalam Darmi and Setiawan, *Data Mining* merupakan suatu langkah dalam melakukan *Knowledge discovery in Databases (KDD)*. *Knowledge discovery* sebagai suatu proses terdiri atas pembersihan data (data cleaning), integrasi data (data integration), pemilihan data (data selection), transformasi data (data transformation), *Data Mining*, evaluasi pola (pattern evaluation) dan penyajian pengetahuan (knowledge presentation). *Data Mining* mengacu pada proses untuk menambang (mining) pengetahuan dari sekumpulan data yang sangat besar untuk menghasilkan sebuah pengetahuan baru dari bidang tertentu (Darmi & Setiawan, 2017).

Kebutuhan bisnis dikategorikan menjadi kebutuhan fungsional dan kebutuhan non-fungsional Analisis kebutuhan sistem dilakukan berdasarkan literatur dan wawancara terhadap pihak, Sedangkan kebutuhan fungsional merupakan deskripsi dari aktivitas-aktivitas dan layanan-layanan yang dapat diberikan oleh sistem, yang penelitian ini adalah sistem prediksi barang (Govindaraju et al., 2009).

Rancangan penelitian yang digunakan dalam penelitian ini adalah deskriptif kualitatif yang bersifat menggambarkan, memaparkan, dan menguraikan objek yang di teliti, berikut ini adalah tahapan penelitian yang digunakan sebagaimana pada Gambar 1.



Sumber: Hasil Penelitian (2022)

Gambar 1. Tahapan Penelitian

a. Pengumpulan Data

Clustering merupakan Analisis Pengelompokan / *Clustering* merupakan proses membagi data dalam suatu himpunan ke dalam beberapa kelompok yang kesamaan datanya dalam suatu kelompok lebih besar dari pada kesamaan data tersebut dengan data dalam kelompok lain. Potensi *Clustering* adalah dapat digunakan untuk mengetahui struktur dalam data yang dapat dipakai lebih lanjut dalam berbagai aplikasi secara luas seperti klasifikasi, pengolahan gambar, dan pengenalan pola (Sumadikarta & Abeiza, 2014).

Ada beberapa tahapan dalam pemilihan kesesuaian data diantaranya adalah Fase Pemahaman Bisnis yaitu tahapan menentukan tujuan penelitian dan ruang lingkup penelitian, dimana tujuan penelitian adalah untuk memperoleh model atau pola Ruang lingkup dalam penelitian ini adalah menggunakan teknik *Clustering Data Mining*, sedangkan data yang dibutuhkan adalah data barang ATK yang sering terjual kemudian Fase pemahaman data adalah memahami data yang diperlukan dalam melakukan pengelompokan data menggunakan teknik *Clustering Data Mining*. *Data Mining* adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual (Rivanthio & Ramdhani, 2020).

Data yang digunakan dalam penelitian adalah data penjualan barang ATK pada bulan Bulan April sampai bulan Mei 2022, berikut data awal dari *data set* pada Tabel 1:

Tabel 1. Data Awal Bulan April 2022

No	Tgl	No. Transaksi	Jenis Barang	Jumlah
1	01/04/22	001/4	Buku Tulis' 18	2
			Gunting K 200	1
			Pulpen Faster C6	2
2	01/04/22	002/4	Buku Tulis' 18	1
			Buku Tulis' 32	2
			Kertas HVS	1
3	01/04/22	003/4	Amplop Coklat Air Mail Bertali	4
4	01/04/22	004/4	Box File'1	1
			Buku Kwitansi'k	2
5	01/04/22	005/4	Buku Tulis' 18	2
			Mistar Plastik'30	1
			Mistar Segitiga'8	1
			Pensil 2b Faber Castell'pcs	2
6	01/04/22	006/4	Pensil Warna Big Isi 12 Pendek	3
			Spidol 12 warna	1
			Stip 2B Hitam' Snowman	2
			Kertas HVS'pcs	30

No	Tgl	No. Transaksi	Jenis Barang	Jumlah
7	01/04/22	007/4	Buku Tulis' 20	3
			Map Biasa	2
			Pulpen Standart AE7	2
			Pulpen JOYKO Gel Pen - Pena jel 330	1
			Binder Clip No. 107	5
8	01/04/22	008/4	Gunting K 200	2
			Amplop Line Putih Panjang Perekat	1
			Box File'2	1
			Kertas Bergaris	2
			Kertas HVS	1
9	01/04/22	009/4	Correction Pen	2
			Pulpen Faster C6	3
			Kertas Bergaris	1

Sumber: Hasil Penelitian (2022)

Tabel 2. Data Awal Bulan Mei 2022

No	Tgl	No. Transaksi	Jenis Barang	Jumlah
14	02/05/22	014/5	Gunting K 400	1
			Correction Pen	2
			Map Batik Kain	1
			Pulpen Standart AE7	2
15	02/05/22	015/5	Mistar Segitiga	1
			Pensil 2b Lyra 2324	2
16	02/05/22	016/5	Stip 2B putih	1
			Kertas HVS	1
			Amplop Tali Merah Biru	2
17	02/05/22	017/5	Buku Tulis' 18	4
			Spidol 12 warna	2
18	02/05/22	018/5	Buku Tulis' 18	5
			Buku Tulis'32	2
			Mistar Plastik'60	1
			Pensil 2b Faber Castell'pcs	2
19	02/05/22	019/5	Pensil 2b Lyra 2324'pcs	1
			Stip 2B putih	1

No	Tgl	No. Transaksi	Jenis Barang	Jumlah
20	02/05/22	020/5	Stip Hitam/Merah'Snowman	1
			Spidol 12 warna'Melody	2
			Pulpen Zebra size 0.5 'pcs	2
			Mistar Plastik' 30 butterfly	1
21	02/05/22	021/5	Buku Tulis' 18	2
			Spidol Whiteboard'Snowman	2
			Box File'Maspion	2
			Binder Clip No. 155'Jocko	1
22	02/05/22	022/5	Amplop Line Putih Panjang Perekat'dz	1
			Buku Kwitansi'k	2
			Buku Tulis'20	5
			Buku Tulis'32	2
23	02/05/22	023/5	Buku Tulis'18	4
			Kertas Foto'pcs	2
			Buku Tulis' 18	2
			Correction Pen	2

Sumber: Hasil Penelitian (2022)

b. Seleksi data

Setelah di dapat data awal, perlu untuk melakukan penyeleksian data, dengan tujuan untuk memperkecil tahapan pengelompokan data. Dalam tahap ini tidak semua atribut dari data digunakan, hanya atribut-atribut yang dianggap penting yang digunakan.

Tabel 3. Seleksi Data

No	Jenis Barang	Merk / Model	Satuan	Jumlah Penjualan	
				Prediksi	Actual
1	Amplop Coklat Air Mail Bertali	Ukuran A3	pcs	30	22
2	Amplop Line Putih Panjang	Dus	Pak	35	20
	Perekat				
3	Amplop Line Putih Pendek	100 lbr	Pak	20	12
	Perekat				
4	Amplop Tali Merah Biru	Ukuran 250/350	pcs	50	32
5	Binder Clip No. 107	Codwell	Pak	30	10
6	Binder Clip No. 107	Codwell	pcs	45	0
7	Binder Clip No. 107	Jocko	Pak	25	20
8	Binder Clip No. 107	Jocko	pcs	30	14
9	Binder Clip No. 155	Codwell	Pak	25	21
10	Binder Clip No. 155	Codwell	pcs	20	0

No	Jenis Barang	Merk / Model	Satuan	Jumlah Penjualan	
				Prediksi	Actual
11	Binder Clip No. 155	Jocko	Pak	15	9
12	Binder Clip No. 155	Jocko	pcs	25	0
13	Binder Clip No. 200	Codwell	Pak	40	21
14	Binder Clip No. 200	Jocko	Pak	30	26
15	Box File	Codwell	pcs	15	9
16	Box File	Maspion	pcs	15	10
17	Buku Kwitansi	Ukuran sedang	pcs	30	15
18	Buku Kwitansi	Kecil	pcs	50	22
19	Buku Tulis	Isi 18 Lembar	pcs	150	100
20	Buku Tulis	Isi 20 Lembar	pcs	200	120
21	Buku Tulis	Isi 32 Lembar	pcs	75	62
22	Buku Tulis	Isi 34 Lembar	pcs	70	66
23	Buku Tulis	Isi 70 Lembar	pcs	60	47
24	Buku Tulis	Isi 100 Lembar	pcs	30	11
25	Correction Pen	Pentel	pcs	75	68
26	Gunting K 200	Gurindo	pcs	20	12
27	Gunting K 400	Gurindo	pcs	35	33
28	Gunting K 500	Gurindo	pcs	15	10
Double Folio 100					
29	Kertas Bergaris	lembar	Rim	40	27
Era Print 200 lembar					
30	Kertas Foto	A4	lembar	150	112
Era Print 200 lembar					
31	Kertas Foto	A4	Pak	15	10
32	Kertas HVS	Folio 60 gram	Rim	15	11
33	Kertas HVS	Folio 60 gram	lembar	300	255
34	Kertas HVS Biasa	Folio 80 gram	Rim	30	26
35	Kertas HVS Biasa	Folio 80 gram	lembar	200	145
36	Kertas HVS	Kwarto 60 gram	Rim	45	33
37	Kertas HVS	Kwarto 60 gram	lembar	200	187
38	Kertas HVS	Kwarto 70 gram A4	Rim	20	10
39	Kertas HVS	Kwarto 80 gram A4	Rim	40	35
40	Map Batik Kain		pcs	50	32

Sumber: Hasil Penelitian (2022)

c. Transformasi data

Data disederhanakan agar mempermudah dalam pembacaan data sesuai dengan variable

pada nilai bobot yang ditentukan. Data yang sudah ditransformasikan pada Tabel 4.

Tabel 4. Transfromasi Data

No	Jenis Barang	Merk / Model	Satuan	Jumlah Penjualan	
				Prediksi	Actual
1	Amplop Coklat Air Mail Bertali	Ukuran A3	pcs	40	21
2	Amplop Line Putih Panjang Perekat	Dus	Pak	30	26
3	Amplop Line Putih Pendek Perekat	100 lbr	Pak	15	9
4	Amplop Tali Merah Biru	Ukuran 250/350	pcs	15	10
5	Binder Clip No. 107	Codwell	Pak	30	15
6	Binder Clip No. 107	Codwell	pcs	50	22
7	Binder Clip No. 107	Jocko	Pak	30	22
8	Binder Clip No. 107	Jocko	pcs	35	20
9	Binder Clip No. 155	Codwell	Pak	20	12
10	Binder Clip No. 155	Codwell	pcs	50	32
11	Binder Clip No. 155	Jocko	Pak	30	10
12	Binder Clip No. 155	Jocko	pcs	45	0
13	Binder Clip No. 200	Codwell	Pak	30	11
14	Binder Clip No. 200	Jocko	Pak	75	68
15	Box File	Codwell	pcs	20	12
16	Box File	Maspion	pcs	35	33
17	Buku Kwitansi	Ukuran sedang	pcs	15	10
18	Buku Kwitansi	Kecil	pcs	40	27

Sumber: Hasil Penelitian (2022)

Perhitungan Algoritma *K-Means Clustering*

Salah satu algoritma *Clustering* adalah *K-Means* yang digunakan untuk mengelompokkan data menjadi beberapa kelompok dengan beberapa *Cluster*. Data – data dipilih menjadi beberapa kelompok dengan kriteria yang telah ditentukan lalu dikumpulkan menjadi satu dalam sebuah *Cluster*. Di mana setiap *Cluster* memiliki titik pusat yang disebut Centroid.

Algoritma *Clustering* yang masuk dalam kelompok Unsupervised learning yang digunakan untuk mengelompokkan data kedalam beberapa kelompok dengan sistem partisi. Algoritma ini menerima masukan berupa data tanpa label kelas. Pada algoritma *K-Means*, komputer menerima data-data yang tidak diketahui kelasnya terlebih dahulu kemudian mengelompokkannya. *Input* yang diterima ialah data dan jumlah kelompok (*Cluster*) yang diinginkan. Algoritma ini akan memasukan data ke dalam beberapa kelompok tersebut (Dwitri et al., 2020). Berikut adalah proses perhitungan manual dengan menggunakan Algoritma *K-Means*

Clustering berdasarkan rumus 1. Algoritma Clustering yang sudah dijabarkan:

- Menentukan data awal yang akan di proses menggunakan metode *K-Means Clustering*.
- Menentukan Jumlah *Cluster* Pada tahap ini menggunakan 2 (dua) *Cluster* yang akan diterapkan dalam perhitungan manual *K-Means* yaitu *Cluster* tinggi dan *Cluster* rendah.
- Menentukan nilai Centroid.

Tabel 5. Nilai Centroid

Centroid		
C1 (Max)	120	59
C2 (Min)	6	0

Sumber: Hasil Penelitian (2022)

Menghitung Jarak Centroid

Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke *Cluster* tertentu ditentukan jarak antara data ke tiap pusat *Cluster*. Jarak paling dekat antara satu data dengan satu *Cluster* tertentu akan menentukan suatu data masuk dalam *Cluster* mana. Untuk menghitung jarak semua data ke setiap titik pusat *Cluster* dapat menggunakan jarak Euclidean yang dirumuskan sebagai berikut: a) Tentukan jumlah *Cluster*, b) Menentukan titik awal *Cluster* secara random, c) Hitung jarak ke masing-masing centroid, d) Kelompokkan objek berdasarkan jarak minimum, dan e) Kembali ke Step 2, apabila masih ada data yang berpindah *Cluster* atau apabila perubahan nilai centroid, ada yang di atas nilai threshold yang ditentukan atau apabila perubahan nilai pada objective function yang digunakan di atas nilai threshold yang ditentukan. $x_1 - x_2$ (Aulia & Saepudin, 2021).

Menggunakan Metode Euclidean Distance

$$De = \sqrt{(xi - si)^2 + (yi - ti)^2} \quad (1)$$

Menghitung jarak Centrid 1 (C1) dan Centroid 2 (C2)

$$DA1C1 = \sqrt{(22 - 120)^2 + (21 - 59)^2} = 9642$$

$$DA1C2 = \sqrt{(22 - 6)^2 + (21 - 0)^2} = 277$$

$$DA2C1 = \sqrt{(20 - 120)^2 + (26 - 59)^2} = 10033$$

$$DA2C2 = \sqrt{(20 - 6)^2 + (26 - 0)^2} = 222$$

$$DA3C1 = \sqrt{(12 - 120)^2 + (9 - 59)^2} = 11714$$

$$DA3C2 = \sqrt{(12 - 6)^2 + (9 - 59)^2} = 10,81665$$

dan seterusnya.

Table 6. Hasil Perhitungan

No	Jenis Barang	Model	Satuan	C1	C2	Jarak
1	Amplop Coklat Air Mail Bertali	Ukuran A3	pcs	9642	277	277
2	Amplop Line Putih Panjang Perekat	Dus	Pak	10033	222	222
3	Amplop Line Putih Pendek Perekat	100 lbr	Pak	11714	10,82	10,82

Sumber: Hasil Penelitian (2022)

Menentukan Posisi Cluster

Algoritma *K-Means* ini dimulai dengan menentukan terlebih dahulu berapa *Cluster* yang ingin dibuat, selanjutnya menentukan nilai awal untuk masing-masing *Cluster*, lalu menghitung jarak dari masing-masing data dengan nilai awal yang telah ditentukan. Setelah jarak dari tiap data dihitung, data dimasukkan dalam *Cluster* terdekat. Selanjutnya menghitung rata-rata dari tiap *Cluster*, rata-rata tersebut nantinya akan digunakan sebagai nilai awal yang baru dalam menghitung jarak dari masing-masing data, pusat *Cluster* atau centroid dipilih pada tahap awal secara acak dari sekumpulan koleksi (populasi) data. *K-Means* menguji masing-masing komponen didalam populasi data dan menandai komponen tersebut ke salah satu centroid yang telah didefinisikan sebelumnya berdasarkan jarak minimum antara komponen (data) dengan masing- masing centroid (Cuhwanto & Agushinta R, 2022).

Menentukan posisi *Cluster* dari Tabel 5 mengikuti pernyataan berikut: “Jika nilai jarak terpendek berada di kolom C1 maka pada tabel posisi *Cluster*, kolom C1 diberi nilai 1” dan “Jika jarak terpendek berada di kolom C2 maka pada tabel posisi *Cluster*, kolom C2 diberi nilai 1”. Nilai 1 hanya untuk simbol atau pertanda bahwa pada kolom tersebut terdapat nilai jarak terpendek mewakili kolom tersebut.

Tabel 7. Posisi Cluster

Jenis Barang	C1	C2
Amplop Coklat Air Mail Bertali		1
Amplop Line Putih Panjang Perekat		1
Amplop Line Putih Pendek Perekat		1

Sumber: Hasil Penelitian (2022)

3. Hasil dan Pembahasan

Pada tahap ini data hasil perhitungan diterapkan untuk menemukan pengetahuan tersembunyi dari data. Metode yang digunakan adalah klasifikasi dengan algoritma *K-Means*. Berikut penerapan algoritma *K-Means* memakai *tool RapidMiner*.



Sumber: Hasil Penelitian (2022)

Gambar 2. Hasil Proses *K-Means Clustering*

Hasil dari proses *K-Means Clustering* di dapat *Cluster* model, lalu di dapat Centroid table.

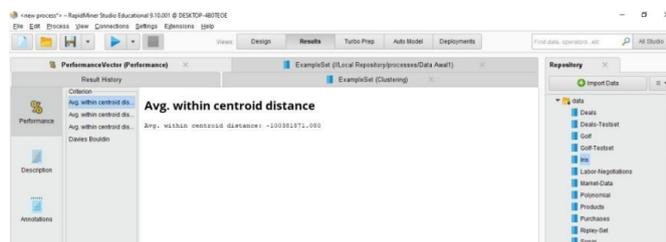


Attribute	cluster_0	cluster_1
NO	10.887	6.750
September	8476.800	25106.750
October	7420.733	29397.250

Sumber: Hasil Penelitian (2022)

Gambar 3. Centroid Table

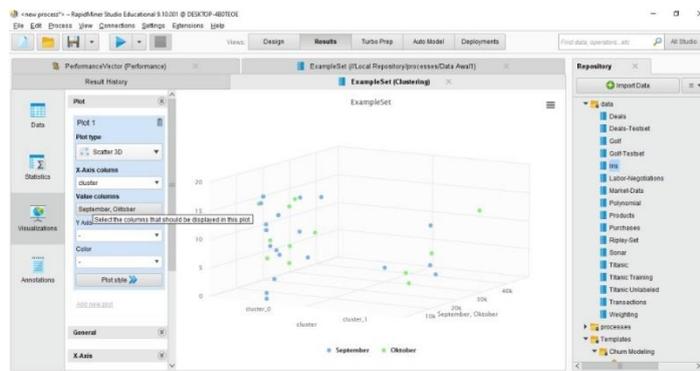
Hasil proses performance dari K-Means sebagaimana pada Gambar 4.



Sumber: Hasil Penelitian (2022)

Gambar 4. Hasil Proses Avg. dengan Centroid Distance

Dari hasil proses Performance maka di dapat contoh Clustering dalam bentuk scatter 3D pada Gambar 5 berikut.



Sumber: Hasil Penelitian (2022)

Gambar 5. Clustering Scatter 3D

4. Kesimpulan

Pengklasteran dengan menggunakan metode K-Means didapatkan untuk menentukan jenis barang yang sering di terjual pada bulan April dan Mei Cluster 1 dan Cluster 2. Menerapkan metode K-Means pada Rapidminer dilakukan dengan memasukkan data produksi bulan April dan Mei dari data Ms.Excel, data tersebut kemudian dikoneksikan ke dalam Tools

Rapidminer, dan akan diolah dan dibentuk *K-Means*. Setelah itu, *Rapidminer* akan menghasilkan jenis barang mana yang sering (banyak) terjual dan yang jarang (sedikit) terjual.

Daftar Pustaka

- Aprilianur, G., & Hadisaputro, E. L. (2022). Penerapan Data Mining Menggunakan Metode K-Means Clustering Untuk Analisa Penjualan Toko Myam Hijab Penajam. *Jurnal Jupiter*, 14(1), 161–170.
- Ardiansyah, A. H., Nugroho, W., Alfiyah, N. H., Handoko, R. A., & Bakhtiar, M. A. (2020). Penerapan Data Mining Menggunakan Metode Clustering untuk Menentukan Status Provinsi di Indonesia 2020. *Seminar Nasional Inovasi Teknologi*, 329–333.
- Aulia, P. U. F., & Saepudin, S. (2021). Penerapan Data Mining K-Means Clustering Untuk Mengelompokkan Berbagai Jenis Merk Laptop. *SEMANTIK (Seminar Nasional Sistem Informasi dan Manajemen Informatika)*, 209–217.
- Cuhwanto, Y. N. A., & Agushinta R, D. (2022). Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means. *Jurnal Petir*, 15(1), 48–56. <https://doi.org/10.33322/petir.v15i1.1358>
- Darmi, Y. D., & Setiawan, A. (2017). Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk. *Jurnal Media Infotama*, 12(2). <https://doi.org/10.37676/jmi.v12i2.418>
- Dwitri, N., Tampubolon, J. A., Prayoga, S., Zer, F. I. R. ., & Hartama, D. (2020). Penerapan Algoritma K-Means Dalam Menentukan Tingkat Penyebaran Pandemi Covid-19 di Indonesia. *Jurnal Teknologi Informasi*, 4(1), 128–132. <https://doi.org/10.36294/jurti.v4i1.1266>
- Govindaraju, R., Simatupang, T., & Samadhi, T. A. (2009). Perancangan Sistem Prediksi Churn Pelanggan PT. Telekomunikasi Seluler Dengan Memanfaatkan Proses Data Mining. *Jurnal Informatika*, 9(1), 33–42. <https://doi.org/10.9744/informatika.9.1.33-42>
- Rivanthio, T. R., & Ramdhani, M. (2020). Penerapan Teknik Clustering Data Mining untuk Memprediksi Kesesuaian Jurusan Siswa (Studi Kasus SMA PGRI 1 Subang). *Faktor Exacta*, 13(2), 125–131.
- Setiawan, R. (2016). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru (Studi Kasus : Politeknik LP3I Jakarta). *Jurnal Lentera ICT*, 3(1), 76–92.
- Srisulistiowati, D. B., Khaerudin, M., & Rejeki, S. (2021). Sistem Informasi Prediksi Penjualan Alat Tulis Kantor Dengan Metode FP-Growth (Studi Kasus Toko Koperasi Sekolah Bina Mulia). *JSI (Jurnal Sistem Informasi)*, 8(2), 243–256. <https://doi.org/10.35968/jsi.v8i2.739>
- Sumadikarta, I., & Abeiza, E. (2014). Penerapan Algoritma K-Means Pada Data Mining Untuk Memilih Produk dan Pelanggan Potensial (Studi Kasus : PT Mega Arvia Utama). *Jurnal Satya Informatika*, 1(1), 12–22.