DATA MINING DALAM KONTEKS TRANSAKSI PENJUALAN HIJAB DENGAN MENGGUNAKAN ALGORITMA CLUSTERING K-MEANS (STUDI KASUS: HIJABER TRENDY)

Riky Maoulana, Bambang Irawan, Agus Bahtiar
Teknik Informatika, STMIK IKMI CIREBON
Jl. Perjuangan No.10B, Karyamulya Cirebon, Indonesia
rickyrocksteady07@gmail.com

ABSTRAK

Industri perdagangan hijab menghadapi tantangan dinamika pasar yang kompleks, sehingga diperlukan pendekatan analitik untuk meningkatkan efisiensi dan optimasi penjualan. Penelitian ini bertujuan menerapkan teknik data mining khususnya clustering dengan algoritma K-Means pada data transaksi penjualan hijab. Dengan pendekatan tersebut diharapkan dapat mengelompokkan pola transaksi berdasarkan kemiripan dan memahami preferensi perilaku pembelian konsumen hijab. Metode penelitian diawali dengan mengumpulkan data historis transaksi dari sebuah toko hijab, lalu menerapkan algoritma K-Means clustering untuk pengelompokan data transaksi serupa. Analisis mempertimbangkan atribut seperti jumlah pembelian, frekuensi transaksi, jenis dan variasi produk yang dibeli. Hasil clustering K-Means berhasil mengidentifikasi beberapa kelompok besar transaksi hijab dengan pola pembelian mirip. Hal ini memberikan informasi berharga terkait perilaku dan selera konsumen hijab. Selanjutnya hasil analitik dapat dimanfaatkan untuk penyesuaian strategi pemasaran agar lebih memenuhi preferensi konsumen dan meningkatkan kepuasan serta loyaltas pelanggan. Penelitian serupa dapat diterapkan pada data transaksi industri fashion Muslim lainnya.

Kata kunci : *Clustering, Penjualan Hijab, Data Mining, K-Means.*

1. PENDAHULUAN

Perkembangan teknologi dan digitalisasi telah membawa perubahan signifikan dalam dunia bisnis, terutama dalam manajemen dan analisis data. Dalam konteks ini, penerapan teknik Data Mining, khususnya dengan menggunakan algoritma clustering K-Means, menjanjikan solusi efektif untuk menganalisis pola transaksi penjualan hijab.

Dalam konteks transaksi penjualan hijab, beberapa permasalahan mendasar muncul. Pertama, terdapat kompleksitas pola pembelian pelanggan yang sulit diidentifikasi secara manual, menghambat potensi pengembangan strategi pemasaran yang efisien. Ketiga, kurangnya pemanfaatan teknologi Data Mining berpotensi mengakibatkan kurangnya wawasan mendalam terhadap perilaku pelanggan dan kecenderungan tren penjualan. Keenam, kesulitan dalam mengekstraksi nilai tambah dari data transaksi tanpa pendekatan analisis yang canggih dapat menghambat inovasi bisnis.

Tujuan penelitian ini untuk mengatasi beberapa permasalahan yang diidentifikasi dalam konteks transaksi penjualan hijab di Hijaber Trendy. Kedua, memahami secara mendalam preferensi konsumen terhadap beragam produk hijab yang ditawarkan, memungkinkan Hijaber Trendy untuk menyusun strategi pemasaran yang lebih personal dan sesuai. Keempat, meningkatkan efisiensi operasional dan penetapan harga dengan pemahaman yang lebih baik terhadap segmen pasar dan tren pembelian.

Metode penelitian ini melibatkan implementasi algoritma clustering K-Means dalam analisis data transaksi penjualan hijab pada Hijaber Trendy. Algoritma clustering K-Means kemudian diterapkan untuk mengelompokkan data menjadi beberapa klaster berdasarkan pola pembelian yang serupa. Selain itu, metode ini dapat membantu dalam penetapan harga yang lebih cerdas dan meningkatkan pengelolaan stok dengan memahami permintaan produk yang berbeda.

Hasil penelitian implementasi K-Means mampu menghasilkan kelompok-kelompok transaksi yang saling mirip, mengidentifikasi pola pembelian konsumen. Ini memberikan wawasan berharga tentang preferensi konsumen dan memungkinkan penyesuaian strategi pemasaran untuk meningkatkan efisiensi dan kepuasan pelanggan.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Paper 1 membahas mengenai "Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering"[1]. Kesimpulan pada penelitian ini adalah dengan menggunakan metode Algoritma K-Means Clustering data mining, area penjualan produk dapat diidentifikasi menjadi tinggi, sedang, dan rendah. Di daerah dengan penjualan produk yang rendah, dilakukan promosi penjualan produk. Sementara di daerah dengan penjualan tinggi, tidak dilakukan promosi penjualan. Aplikasi yang dibuat diharapkan dapat memudahkan pengelompokan (clustering) data transaksi penjualan sehingga dapat menghemat waktu dan membantu menyusun strategi untuk meningkatkan penjualan.

Paper 2 membahas mengenai "Klasterisasi Pola Penjualan Menggunakan Metode K-Means Clustering

(Study Kasus Penjualan di Tunas Ponsel Kota Langsa)"[2]. Berdasarkan data penjualan ponsel, metode clustering K-Means menghasilkan produk terlaris dan laris dengan 100 nama ponsel serta aksesorinya dalam waktu 3 bulan pengujian. Perhitungan K-Means mengelompokkan nama-nama ponsel dan aksesorisnya menjadi dua kelompok, yaitu Cluster 1 berlabel Bestseller berisi 28 nama ponsel dan aksesoris, Cluster 2 berlabel Bestseller berisi 13 nama ponsel dan aksesoris.

Paper 3 membahas mengenai "Implementasi Data Mining Untuk Menentuksn Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering"[3]. Clustering adalah pendekatan penting untuk menemukan kesamaan dalam data dan mengklasifikasikan data serupa ke dalam kelompok. Pengelompokan melibatkan pembagian kumpulan data menjadi beberapa kelompok, dengan kesamaan yang lebih besar dalam suatu kelompok dibandingkan antar kelompok (Rui Xu dan Donald. 2009). Ide pengelompokan atau pengelompokan data merupakan hal yang sederhana dan dekat dengan cara berpikir manusia dalam mengolah data dalam jumlah besar menjadi beberapa kelompok atau kategori terbatas sehingga dapat dilakukan analisis lebih lanjut.

Paper 4 membahas mengenai "Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Mengetahui Minat Customer Di Toko Hijab"[4]. Penelitian ini dilakukan untuk mengetahui minat pelanggan terhadap toko hijab dengan menggunakan data transaksi dan penjualan. Hasil yang diperoleh dengan menggunakan algoritma K-Means menunjukkan bahwa kelompok pertama terdiri dari merek hijab yang low demand yaitu Dian Pelangi, Kami Idea dan Mechanism.

Paper 5 membahas mengenai "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Yana Sport"[5]. Dalam penelitian ini dibedakan dua kelompok yaitu yang terlaris dan yang tidak terlaris. Hasil data kinerja menjelaskan bahwa cluster 0 dengan nilai 110 tergolong bad seller sedangkan 21389 tergolong good seller. Hasil dari tip pencarian ini dapat memperoleh informasi atau pola dari penerapan algoritma K-Means dengan data penjualan. Terdapat 99 item merchandise yang laku dan 23 item merchandise yang tidak laku, sehingga pemilik dapat menyusunstrategi penjualan dan pembelian kembali pada produk yang laku.

Paper 6 membahas "Implementasi Metode Algoritma K-Means Clustering pada Toko Eidelweis"[6]. Pada toko Eidelweis telah diimplementasikan algoritma clustering K-Means yang terintegrasi pada sistem melalui beberapa menu, yaitu menu login, menu input data, menu hasil dan laporan, serta menu utama. Menu utama dirancang dengan antarmuka yang user-friendly sehingga memudahkan akses ke semua menu dan melihat laporan penjualan toko. Dengan adanya implementasi algoritma clustering K-Means ini, metode pengelompokan data mining telah berhasil diaplikasikan pada data transaksi toko Eidelweis guna menunjang analisis pola penjualannya.

Paper 7 membahas "Decision Support System Metode K-Means Clustering Dalam Memprediksi Tingkat Penjualan Pada Store Ryan Mart"[7]. Aplikasi Implementasi Algoritma K-Means Clustering pada Toko Ryan Mart dibuat untuk mengelola data penjualan harian dan mengelompokkannya ke dalam 3 klaster: Klaster I (Barang Paling Laku), Klaster II (Barang Kurang Laku), dan Klaster III (Barang Tidak Laku). Pengelompokan data dilakukan berdasarkan total penjualan setiap jenis barang selama setahun yang dihitung secara otomatis.

2.2. Machine Learning

Machine Learning, salah satu cabang ilmu komputer yang berkaitan dengan kecerdasan buatan, berfokus pada pembuatan sistem yang belajar dari data. Data adalah aspek kunci dari algoritma Machine Learning. Data pelatihan yang dibagi menjadi data pelatihan dan data pengujian digunakan untuk mengembangkan model, sedangkan data pengujian digunakan untuk mengevaluasi kinerja algoritma yang dilatih pada data baru. Dengan memanfaatkan data, Machine Learning meningkatkan kemampuannya untuk memahami pola, membuat prediksi, dan belajar dari pengalaman.[8]

2.3. Data Mining

Data mining merupakan suatu proses yang kompleks dengan menggunakan statistik, matematika, kecerdasan buatan, dan machine learning guna mengekstraksi informasi yang berharga dari database berskala besar. Informasi yang terhimpun diolah dalam database untuk mendukung pengambilan keputusan.[8] menguraikan tahapan-tahapan data mining, mencakup pembersihan data, integrasi data dari beragam sumber, pemilihan data yang relevan, transformasi data, proses mining, evaluasi pola, dan presentasi pengetahuan. Tahap terakhir mencakup perumusan keputusan berdasarkan hasil analisis yang didapatkan.

2.4. Clustering K-Means

Clustering K-Means adalah salah satu metode analisis data yang digunakan untuk mengelompokkan data menjadi beberapa kelompok yang saling berhubungan berdasarkan kesamaan atribut. Metode ini bekerja dengan cara mendefinisikan pusat klaster (centroid) secara acak, kemudian menghitung jarak antara setiap data dengan centroid tersebut. Setiap data akan dikelompokkan ke dalam klaster yang memiliki centroid terdekat. Setelah semua data terkelompok, akan dihitung kembali centroid baru berdasarkan rata-rata dari data dalam satu klaster.[9]

3. METODE PENELITIAN

Dari judul "Implementasi Data Mining Dalam Konteks Transaksi Penjualan Hijab Dengan Menggunakan Algoritma Clustering K-Means (Studi

Kasus: Hijaber Trendy)," metode yang digunakan cenderung bersifat kuantitatif. Penggunaan algoritma clustering K-Means pada konteks transaksi penjualan hijab menunjukkan pendekatan analisis data yang lebih terstruktur dan berorientasi pada angka-angka. Kuantitatif mencakup pengumpulan, analisis, dan interpretasi data dalam bentuk numerik, yang sesuai dengan tujuan penelitian untuk mengidentifikasi pola penjualan dengan menggunakan algoritma K-Means.[10]

3.1. Tahapan Metode Penelitian

Tabel 1. Alur Metode Penelitian

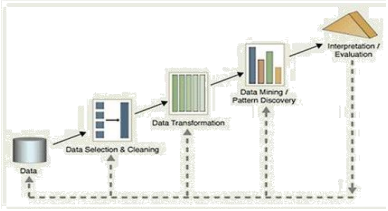
Tahapan	Aktivitas	Deskripsi Aktivitas
Inisialisasi Centroid	a. Pilih jumlah cluster (k) yang diinginkan. b. Pilih secara acak k observasi dari data sebagai centroid awal.	yaitu pemilihan jumlah cluster yang diinginkan, mempertimbangkan beberapa faktor seperti karakteristik data, tujuan analisis, dan pemahaman mendalam terhadap domain masalah.
Assign Data ke Cluster	a. Hitung jarak antara setiap data dengan centroid. b. Tentukan data masuk ke cluster dengan centroid terdekat.	Proses ini dilakukan untuk seluruh data dalam dataset. Hasilnya adalah pembentukan cluster-cluster yang mencerminkan kesamaan antar data dalam masing-masing cluster, dan setiap cluster diidentifikasi oleh centroidnya.
Update Centroid	Hitung rata-rata dari setiap cluster sebagai centroid baru.	Dalam aktivitas "Update Centroid" pada algoritma K-Means, centroid baru untuk setiap cluster dihitung dengan mengambil nilai rata-rata dari semua data yang termasuk dalam cluster tersebut. Proses ini membawa perubahan posisi centroid, sehingga mencerminkan perubahan dalam pusat massa data dalam setiap cluster.
Iterasi	Ulangi langkah 2 dan 3 sampai tidak ada perubahan atau perubahan minimal dalam penempatan data ke dalam cluster.	Dalam konteks algoritma K-Means, "Iterasi" merujuk pada langkah-langkah perulangan yang dilakukan hingga mencapai konvergensi atau titik di mana tidak ada perubahan yang signifikan dalam penempatan data ke dalam cluster.

Pada Tabel 1. menjelaskan tahapan metode K-Means.

3.2. Sumber Data

Dataset yang digunakan dalam penelitian ini terdiri dari 598 record dan terdiri dari 7 atribut periode 2021-2022. Kumpulan data berasal dari Distro Hijaber Trendy di Jl. Komplek Perumahan Keandra Blok No 10 Kalijaga Kota Cirebon. Hasil data mining berupa dokumen fleksibel berformat Microsoft Excel.

3.3. Teknik Analisis Data



Gambar 1. Alus Teknik Analisis Data

Pada Gambar 1. menjelaskan teknik analisis data dalam penerapan data mining ini menggunakan proses tahapan Knowledge Discovery in Databases (KDD) yang terdiri dari Data, Data Cleaning, Data transformation, Data mining, Pattern evolution.

3.4. Populasi dan Sampel

Populasi dalam dataset ini mencakup seluruh transaksi penjualan hijab yang tercatat oleh Hijaber Trendy selama periode tertentu. Dengan jumlah transaksi sebanyak 598 record, populasi ini mencerminkan keseluruhan aktivitas penjualan hijab yang terjadi dalam rentang waktu tersebut. Untuk penelitian ini, sampel diambil sebanyak 598 record dari populasi tersebut sebagai representasi yang dapat mewakili karakteristik umum dari seluruh transaksi penjualan hijab di Hijaber Trendy. Pemilihan sampel dilakukan secara acak untuk memastikan keberagaman dan representativitas data yang digunakan dalam analisis menggunakan algoritma clustering K-Means.

4. HASIL DAN PEMBAHASAN

4.1. Hasil

Hasil penelitian yang dilakukan dalam pembahasan ini yaitu akan menguraikan proses bagaimana pengklasifikasian atau mengemlompokan dataset hijab dengan menggunakan Algoritma K-Means.

4.1.1. Data

Tabel 2 Dataset Hijab

Kode	Nama Barang	Stok Awal	Sisa Stok	Harga Beli	Harga Jual
SP-10-01	Segi Empat Paris 110x110 Milo	20	0	23000	28000
SP-10-02	Segi Empat Paris 110x110 Black	35	4	23000	28000
SP-10-03	Segi Empat Paris 110x110 BW	18	9	23000	28000
SP-10-04	Segi Empat Paris 110x110 Sage	20	4	23000	28000

Kode	Nama Barang	Stok Awal	Sisa Stok	Harga Beli	Harga Jual
SP-10-05	Segi Empat Paris 110x110 Brown	15	0	23000	28000
SP-10-06	Segi Empat Paris 110x110 Olive	17	10	23000	28000
SP-10-07	Segi Empat Paris 110x110 Army	10	8	23000	28000
SP-10-08	Segi Empat Paris 110x110 Wardah	10	6	23000	28000
SP-15-01	Segi Empat Paris 115x115 Milo	18	3	25000	30000
SP-15-02	Segi Empat Paris 115x115 Black	15	0	25000	30000
……	……	..	…..	……	….
……	……	..	…..	……	….
……	……	..	…..	……	….
SP-15-06	Segi Empat Paris 115x115 Olive	15	5	25000	30000

Pada Tabel 2. adalah penjelasan dataset penelitian berjumlah 598 record dengan 7 atribut, bersumber dari penjualan Distro Hijaber Trendy alamat Jl. Komplek perumahan Keandra Blok N. 10 Kalijaga Kota Cirebon periode 2021-2022. Data berbentuk soft file Microsoft Excel.

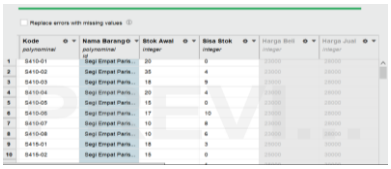
4.1.1. Data Selection

Tabel 3 Data Selection

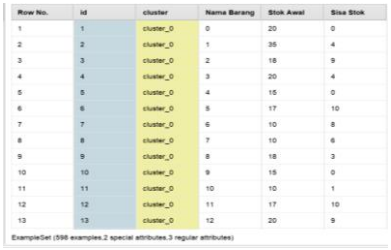
No	Nama Atribut	Type
1	No	Integer
2	Kode Barang	Integer
3	Nama Produk	Polynominal
4	Harga Beli	Nominal
5	Harga Jual	Nominal
6	Stock Awal	Polynominal
7	Sisa Stock	Polynominal

Data yang dipilih berjumlah 598 record penjualan hijab tahun 2022 dengan 7 atribut: No, Kode, Nama produk, Harga beli, Harga jual, Stock awal, Sisa stock. Setelah seleksi, hanya 4 atribut yang digunakan karena 3 atribut lain nilainya seragam.

4.1.2. Data Transformation



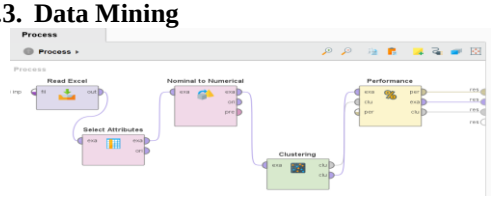
Gambar 2. Data Sebelum diTransformasi



Gambar 3. Data Sesudah diTransformasi

Transformasi data dilakukan untuk mengubah data bertipe nominal menjadi numerik dengan menggunakan operator Nominal to Numeric, seperti di sajikan pada Gambar 2. dan Gambar 3. Didapatkan hasil proses dari rapid miner untuk operator terdiri dari 598 examples, 1 spesial attribute, 4 atribut reguler diantaranya operator rows no, Id, Cluster, nama barang, stock awal dan, sisa stock.

4.1.3. Data Mining



Gambar 4. Data Mining

Pada Gambar 4. Dengan operator algoritma k-means dan untuk mengukur besarnya Davies Bouldin yang diperoleh dari algoritma k-means menggunakan operator Cluster Distance Performance, maka menghasilkan parameter seperti pada table diabawah ini:

Tabel 4. Parameter K-Means

Parameter	Keterangan
k min	3
k max	10
measure type	NumericalMeasurement
numerical measure	"EuclideanDistance
clustering algorithm	Kmeans

Tabel 4. menjelaskan hasil parameter k-means adalah k-means =3, kmax =10 adalah untuk menguji k-meas pada langkah ke 3 untuk mnghasilkan nilai terbaik davies bouldin yang mendekati angka 0.

4.2. Pembahasan

Pada proses Clustering K-Means terdapat beberapa operator yang digunakan untuk menganalisis data. Operator tersebut adalah :

1. Read Excel, digunakan untuk membaca data yang akan dianalisis.
2. Select Atribut, Untuk menyeleksi atribut yang di butuhkan yaitu : Kode,Nama Produk,Stock awal dan sisa stock.
3. Nominal to numerical, yaitu untuk merubah atribut type Nominal menjadi Numeric
4. Clustering, digunakan untuk mengelompokkan kedalam beberapa cluster.
5. Operator Cluster Performance Distance digunakan untuk evaluasi kinerja pengelompokkan berbasis centroid yang menghasilkan model cluster pusat. Model cluster centroid memiliki informasi mengenai clustering yang dilakukan. Pada Operator Cluster Performance Distance terdapat dua ukuran kinerja yang didukung yaitu rata-rata dalam jarak cluster dan indeks Davies-Bouldin. Ukuran kinerja ini dijelaskan dalam parameter. Algoritma cluster yang menghasilkan kumpulan cluster dengan indeks Davies-Bouldin terkecil dianggap sebagai algoritma terbaik berdasarkan kriteria ini. Untuk mencari nilai k optimum langkah yang harus dilakukan adalah melakukan percobaan nilai k untuk menemukan nilai indeks Davies-Bouldin terkecil. Berikut parameter yang akan diuji untuk menemukan nilai Davies-Bouldin :

Tabel 5. Parameter Clustering

Parameters	Value
K/Cluster	3 sampai 10
Max Runs	10
Measure Types	NumericalMeasures
Numerical Measure	EuclideanDistance
Main Criterion	Davies Bouldin

Dari hasil percobaan menggunakan nilai parameter diatas, berikut hasil percobaan yang dilakukan :

a) Data Cluster 3

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: -14.374
Avg. within centroid distance_cluster_0: -14.594
Avg. within centroid distance_cluster_1: -0.641
Avg. within centroid distance_cluster_2: -16.105
Davies Bouldin: -0.819
```

Gambar 5 Davies Bouldin k=3

Pada Gambar 5. k=3, max 10 diperoleh nilai Davies Bouldin : -0.819

b) Data Cluster 4

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: -6.513
Avg. within centroid distance_cluster_0: -8.833
Avg. within centroid distance_cluster_1: -0.641
Avg. within centroid distance_cluster_2: -2.249
Avg. within centroid distance_cluster_3: -8.394
Davies Bouldin: -0.543
```

Gambar 6. Davies Bouldin k=4

Pada Gambar 6. k=4 diperoleh nilai Davies Bouldin : -0.543

c) Data Cluster 5

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: -4.842
Avg. within centroid distance_cluster_0: -5.200
Avg. within centroid distance_cluster_1: -0.641
Avg. within centroid distance_cluster_2: -2.249
Avg. within centroid distance_cluster_3: -8.394
Avg. within centroid distance_cluster_4: -0.000
Davies Bouldin: -0.470
```

Gambar 7. Davies Bouldin k=5

Pada Gambar 7. k=5 diperoleh nilai Davies Bouldin : -0.470

d) Data Cluster 6

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: -3.065
Avg. within centroid distance_cluster_0: -2.249
Avg. within centroid distance_cluster_1: -5.778
Avg. within centroid distance_cluster_2: -0.641
Avg. within centroid distance_cluster_3: -4.107
Avg. within centroid distance_cluster_4: -1.778
Avg. within centroid distance_cluster_5: -2.250
Davies Bouldin: -0.604
```

Gambar 8. Davies Bouldin k=6

Pada Gambar 8. K=6 diperoleh nilai Davies Bouldin : -0.604

e) Data Cluster 7

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: -1.993
Avg. within centroid distance_cluster_0: -4.107
Avg. within centroid distance_cluster_1: -1.778
Avg. within centroid distance_cluster_2: -0.641
Avg. within centroid distance_cluster_3: -2.249
Avg. within centroid distance_cluster_4: -2.250
Avg. within centroid distance_cluster_5: -0.000
Avg. within centroid distance_cluster_6: -0.000
Davies Bouldin: -0.475
```

Gambar 9. Davies Bouldin k=7

Pada Gambar 9. k=7 diperoleh nilai Davies Bouldin : -0.475

f) Data Cluster 8

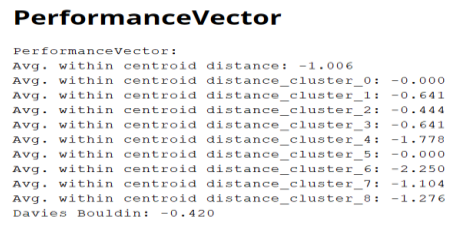
PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: -1.418
Avg. within centroid distance_cluster_0: -1.778
Avg. within centroid distance_cluster_1: -0.641
Avg. within centroid distance_cluster_2: -0.641
Avg. within centroid distance_cluster_3: -2.249
Avg. within centroid distance_cluster_4: -0.000
Avg. within centroid distance_cluster_5: -0.000
Avg. within centroid distance_cluster_6: -1.276
Avg. within centroid distance_cluster_7: -2.250
Davies Bouldin: -0.432
```

Gambar 10. Davies Bouldin k=8

Pada Gambar 10. K=8 diperoleh nilai Davies Bouldin : -0.432

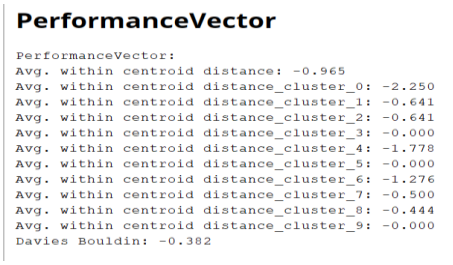
g) Data Cluster 9



Gambar 11. Davies Bouldin k=9

Pada Gambar 11. k=9 diperoleh nilai Davies Bouldin : -0.420

h) Data Cluster 10



Gambar 12. Davies Bouldin k=10

Pada Gambar 12. k=10 diperoleh nilai Davies Bouldin : -0.382.

Percobaan dilakukan dengan menggunakan k=3 sampai dengan k=10. Berikut adalah hasil yang diperoleh :

Tabel 6. Nilai davies Bouldin

Cluster	Davies Bouldin
K3	-0.819
K4	-0.543
K5	-0.470
K6	-0.604
K7	-0.475
K8	-0.432
K9	-0.420
K10	-0.382

Dari hasil percobaan dapat disimpulkan bahwa nilai k optimum adalah k=10 dengan nilai DBi = -0.382. Semakin kecil nilai DBi menunjukkan seberapa baik cluster yang diperoleh. Setelah mendapatkan nilai terbaik lalu implementasikan untuk k = 10 kedalam parameter pada operator K-Means.

5. KESIMPULAN DAN SARAN

Kesimpulan dari penelitian ini adalah penerapan algoritma K-Means pada dataset penjualan Hijab dapat menghasilkan pengelompokan data yang objektif, informatif dan akurat, dengan cluster terbaik adalah K10. Untuk pengembangan lebih lanjut, disarankan untuk membandingkan algoritma K-Means dengan algoritma lain seperti X-Means dan k-medoid untuk meningkatkan presisi clustering data. Hal ini harus mengarah pada pengembangan strategi pemasaran hijab yang lebih optimal berdasarkan model penjualan yang berasal dari pemodelan data yang lebih baik.

DAFTAR PUSTAKA

[1] S. Handoko, F. Fauziah, and E. T. E. Handayani, "Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering," *J. Ilm. Teknol. dan Rekayasa*, vol. 25, no. 1, pp. 76–88, 2020, doi: 10.35760/tr.2020.v25i1.2677.

[2] G. M. A. Siregar, J. Mauli, and R. Akram, "Klasterisasi Pola Penjualan Menggunakan Metode K-Means Clustering (Study Kasus Penjualan di Tunas Ponsel Kota Langsa)," *JITU (J. Ilm. Tek. UNIDA)*, vol. 4, no. 1, pp. 140–147, 2023.

[3] H. Prastiwi, Jeny Pricilia, and Errissya Rasywir, "Implementasi Data Mining Untuk Menentuksn Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering," *J. Inform. Dan Rekayasa Komputer(JAKAKOM)*, vol. 2, no. 1, pp. 141–148, 2022, doi: 10.33998/jakakom.2022.2.1.34.

[4] Y. Yulianti, D. Y. Utami, N. Hikmah, and F. N. Hasan, "Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Mengetahui Minat Customer Di Toko Hijab," *J. Pilar Nusa Mandiri*, vol. 15, no. 2, pp. 241–246, 2019, doi: 10.33480/pilar.v15i2.650.

[5] Agung Nugraha, Odi Nurdiawan, and Gifthera Dwilestari, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Yana Sport," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 6, no. 2, pp. 1–7, 2022.

[6] R. Supardi and I. Kanedi, "Implementasi Metode Algoritma K-Means Clustering pada Toko Eidelweis," *J. Teknol. Inf.*, vol. 4, no. 2, pp. 270–277, 2020, doi: 10.36294/jurti.v4i2.1444.

[7] M. R. Pratama and R. Supardi, "Decision Support System Metode K-Means Clustering Dalam Memprediksi Tingkat Penjualan Pada Store Ryan Mart," *J. Media Infotama*, vol. 18, no. 1, pp. 81–86, 2022, [Online]. Available: http://dx.doi.org/10.37676/jmi.v18i1.1941%0Ahttps://jurnal.unived.ac.id/index.php/jmi/article/download/1941/1672

[8] Oon Wira Yuda, Darmawan Tuti, Lim Sheih Yee, and Susanti, "Penerapan Penerapan Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Random Forest," *SATIN - Sains dan Teknol. Inf.*, vol. 8, no. 2, pp. 122–131, 2022, doi: 10.33372/stn.v8i2.885.

[9] W. N. Purba, M. Kosasih, D. Kallamas, and ..., "Penggunaan Algoritma K–Means Clustering Untuk Menentukan Penilaian Kedisiplinan Karyawan Rumah Sakit Royal Prima," *... (Teknik Inf. dan ...*, vol. 6, pp. 188–195, 2023, doi: 10.37600/tekinkom.v6i1.856.

[10] M. Syahril, K. Erwansyah, and M. Yetri,

"Penerapan Data Mining Untuk Menentukan Pola Penjualan Peralatan Sekolah Pada Brand Wigglo Dengan Menggunakan Algoritma Apriori," *J-SISKO TECH (Jurnal Teknol. Sist. Inf. dan Sist. Komput. TGD)*, vol. 3, no. 1, p. 118, 2020, doi: 10.53513/jsk.v3i1.202.