

# Klasterisasi Penjualan Produk Clothing Line Menggunakan Metode K-Means Clustering Untuk Penerapan Konsep Up-Selling

Muhammad Ikkal Sukirman, Tacbir Hendro Pudjiantoro, Irma Santikarama  
Jurusan Informatika, Fakultas Sains dan Informatika  
Universitas Jenderal Achmad Yani  
Jl. Terusan Sudirman, Cimahi  
muhammadikbalsukirman@gmail.com

**Abstrak**— Ketersediaan data transaksi yang cukup besar dapat digunakan sebagai analisa sehingga dapat menghasilkan suatu informasi yang sangat berguna bagi perusahaan dalam melakukan manajemen stok pada produk potensial sehingga meminimalisir terjadinya kekosongan dan kelebihan pada produk tertentu. Seperti Jhonarmy Clothing yang dalam setiap bulan memiliki data transaksi yang besar, tidak dilakukan proses analisis terhadap data tersebut sehingga informasi mengenai produk potensial untuk manajemen stok menjadi sulit untuk didapat. Penelitian ini dilakukan bertujuan untuk melakukan teknik data mining yaitu proses klasterisasi dengan metode K-Means Clustering terhadap data penjualan yang dapat menghasilkan suatu informasi mengenai produk yang sangat diminati dan kurang diminati, dimana produk yang sangat diminati dapat menjadi rekomendasi produk yang dapat diterapkan kedalam konsep up-selling. Penelitian ini menghasilkan 3 buah cluster dan hasil nilai dari Davies-Bouldin Index yaitu 0,181 dengan menghasilkan rekomendasi produk untuk diterapkan ke dalam konsep up-selling sebanyak 10 produk.

**Kata kunci**— Clustering; K-Means; Penjualan; Up-Selling

## I. PENDAHULUAN

Data Mining telah banyak digunakan sebagai salah satu teknik dalam memperoleh informasi yang dapat bermanfaat bagi perusahaan. Saat ini Data Mining telah banyak digunakan diberbagai sektor bisnis salah satunya pada bidang industri fashion. Data Mining merupakan proses untuk menemukan pola yang dapat berguna, berdasarkan data yang memiliki volume yang sangat besar. Proses Data Mining dimulai dengan mengekstrak informasi yang sebelumnya tidak diketahui sehingga dapat dipahami dan dapat di tindak lanjuti [1]. Berdasarkan data yang sangat melimpah dan kebutuhan akan pengetahuan yang dapat dijadikan acuan oleh pemangku kepentingan dalam proses pengambilan keputusan penting bagi perusahaan, merupakan salah satu penyebab yang melatarbelakangi teknologi data mining.

Industri fashion merupakan salah satu area bisnis dengan ketersediaan data yang besar dengan meningkatnya salah satu kebutuhan primer manusia utamanya permintaan akan gaya pakaian seiring dengan perkembangan fashion yang sangat berkembang pesat pada saat ini. Biasanya perusahaan akan melakukan pengembangan produk untuk meningkatkan penjualan produk [2]. Pada proses penjualannya Jhonarmy Clothing memiliki variasi penjualan yang minim dan tidak

melakukan manajemen stok terhadap produk yang seringkali terjadi dead stock didalam gudang yang akan menyebabkan biaya penyimpanan menjadi tinggi dan menjadi tidak ekonomis.

Dalam proses penjualan produk terdapat sejumlah data yang cukup besar namun proses analisis data dalam jumlah data yang cukup besar membutuhkan waktu yang cukup lama apabila dilakukan secara manual, sehingga ketersediaan informasi atau pengetahuan mendalam mengenai produk potensial menjadi sulit didapatkan, dengan menggunakan data mining khususnya dengan melakukan klasterisasi atau pengelompokan data potensi minat pelanggan terhadap produk yang paling diminati dan kurang diminati dapat dijadikan informasi guna meningkatkan penjualan dalam memberikan manfaat yang lebih untuk perusahaan. Teknik yang digunakan mengacu kepada proses pencarian informasi dari data yang cukup besar melalui suatu metode algoritma. Berkenaan dengan pengelompokan data yang merupakan sebuah permasalahan, cara untuk melakukan pengelompokan data tersebut dapat dilakukan dengan mengelompokkan region penjualan menggunakan metode clustering K-Means. Melalui metode tersebut, data dapat dikelompokkan berdasarkan pola-pola distribusi, dan dapat menemukan keterkaitan yang menarik antara atribut data [3].

Dalam beberapa studi yang telah dilakukan, dengan menggunakan algoritma K-Means yang digunakan untuk mengelompokkan data produk dan data penjualan untuk dapat mengetahui produk potensial dengan menghitung jumlah transaksi untuk digunakan sebagai informasi dalam penentuan strategi penjualan [4]. kemudian pada studi dengan judul Analisis Clustering menggunakan algoritma K-Means terhadap penjualan produk pada PT Batamas Niaga Jaya Dimana berdasarkan data transaksi penjualan produk pada perusahaan tersebut dilakukan klasterisasi sebanyak 9 kali iterasi yang terdiri dari 2 cluster yaitu produk yang tidak laris (C0) dan produk yang laris (C1) [5].

Klasterisasi penjualan produk dilakukan berdasarkan data penjualan, proses klasterisasi menghasilkan kelompok-kelompok tertentu berdasarkan data uji dengan jarak yang paling dekat dengan pusat centroid dimana hasil tersebut akan diterapkan dalam konsep Up-Selling. Dimana Up-Selling merupakan suatu strategi yang dapat meningkatkan keuntungan bagi perusahaan dengan menentukan persepsi,

pengetahuan, dan kebijaksanaan. Dalam proses pertumbuhan perusahaan, Up-Selling merupakan salah satu teknik terpenting untuk meningkatkan keuntungan dalam sektor bisnis [6]. Sehingga dengan melakukan pengelompokan data dengan menggunakan metode tersebut, perusahaan dapat memperoleh informasi mengenai jumlah penjualan Produk paling diminati dan sedikit diminati sehingga dapat menentukan produk yang harus memiliki banyak stok berdasarkan produk yang paling diminati yang selanjutnya dapat diterapkan didalam konsep Up-Selling guna meningkatkan penjualan produk pada perusahaan.

## II. METODE

### A. Pengumpulan data

Data penelitian diperoleh dari pengambilan data langsung ke instansi terkait yaitu Jhonarmy Clothing yang diperlukan untuk mendukung penelitian yaitu data penjualan produk dengan rentan waktu selama 3 tahun dari bulan Januari tahun 2019 hingga Januari tahun 2021 dengan jumlah 2400. Data transaksi penjualan ini berisi data penjualan jenis dan jumlah produk selama periode waktu tersebut.

### B. Data Pre-processing

Pre-processing data merupakan sekumpulan data operasional yang perlu dilakukan seleksi sebelum tahap penggalian informasi dalam *Knowledge Discovery Database* (KDD) dimulai. Proses *cleaning* data dilakukan untuk menghilangkan data yang tidak konsisten, atau menghapus atribut yang tidak diperlukan [11]. Kegiatan yang dilakukan didalamnya meliputi:

#### 1) Pemilihan data (Data Selection)

Data Selection merupakan tahapan untuk melakukan seleksi data berdasarkan sekumpulan data yang diperoleh sehingga dapat menciptakan kelompok data target, pemilihan kelompok, dan memfokuskan pada sampel data berdasarkan data asset yang telah dimiliki sebelumnya [8].

#### 2) Pembersihan Data (Data Cleaning)

Cleaning data merupakan tahapan data yang telah dimiliki perlu diseleksi kembali agar tidak terjadi redundansi data dan inkonsistensi data [9].

### C. Transformasi Data

Transformasi merupakan sebuah proses integrasi berdasarkan data yang sudah dipilih. Pada tahapan ini dilakukan proses normalisasi dimana Proses Data mining akan dapat menghasilkan hasil yang efektif dan tepat jika nilai yang dihasilkan memiliki keakuratan yang tinggi [10]. Dimana metode normalisasi yang digunakan yaitu teknik min-max normalization [11]. Transformasi data dilakukan oleh Persamaan (1).

$$v'_i = \frac{v_i - \min_A}{\max_A - \min_A} (\text{new\_max}_A - \text{new\_min}_A) + \text{new\_min}_A \quad (1)$$

Keterangan:

$v'_i$  : Nilai data yang baru dari hasil normalisasi min-max.

$v_i$ : Nilai data yang akan dilakukan normalisasi.

$\max_A$  : Nilai maksimum data.

$\min_A$  : Nilai minimum data.

$\text{new\_max}_A$ : Nilai maksimum yang diharapkan dari proses normalisasi.

$\text{new\_min}_A$ : Nilai minimum yang diharapkan dari proses normalisasi.

Proses transformasi tersebut dilakukan agar dapat memperoleh data yang sesuai dengan kriteria yang dibutuhkan dan agar data yang digunakan menjadi efektif. Kualitas data sangat berperan penting dalam proses klusterisasi karena data yang kotor dapat menyebabkan kesalahan pada keputusan.

### D. Clustering dengan metode K-means Clustering

Clustering merupakan tahapan yang digunakan dalam melakukan pengelompokan data yang dilakukan pada suatu data yang atribut kelas nya belum di deskripsikan, clustering merupakan suatu metode yang digunakan untuk mengelompokkan kemiripan terhadap data dari bagian antar kelas. Dimana clustering dapat menghasilkan beberapa kelompok (cluster) pada suatu data [12]. Metode ini digunakan untuk metode untuk menemukan struktur cluster yang dikarakterisasi dengan kesamaan terbesar dalam kelompok yang sama pada perbedaan terbesar antara cluster yang berbeda [13].

Algoritma *K-Means* merupakan metode pengelompokan data tanpa sekatan yang dapat di partisi dalam bentuk dua atau lebih kelompok. *K-Means* dapat menjadi sebuah tahapan pengelompokan yang melakukan partisi terhadap dataset menjadi sejumlah cluster (kelompok) yang sudah ditentukan di awal. Penggunaan metode *K-Means* adalah untuk membentuk cluster yang terpisah pisah dari sejumlah titik data [14]. Tujuannya adalah untuk membentuk kelompok yang terpisah pisah dari sejumlah titik data. Dalam lima puluh tahun terakhir, k-means merupakan algoritma pengelompokan yang banyak digunakan dalam proses pengolahan data [15]. Berikut ini merupakan penerapan Metode *K-Means Clustering*:

- a) Masukkan data yang akan di klusterisasi.
- b) Tentukan jumlah klaster.
- c) Ambil sebaran data sebanyak jumlah klaster secara acak sebagai pusat centroid.
- d) Hitung jarak antara data dengan centroid dengan menggunakan persamaan (2).

$$d(x_i - c_j) = \sqrt{\sum_{i=1}^N (X_i - C_j)^2} \quad (2)$$

Dimana :

$X_i$  = Jarak data

$C_j$  = Titik Centroid

- e) Hitung Kembali pusat klaster dengan keanggotaan klaster yang baru.
- f) Jika pusat klaster tidak mengalami perubahan maka proses klusterisasi telah selesai.

### E. Interpretasi

Berdasarkan Data Transaksi penjualan produk pada Jhonarmy Clothing yang akan dikelompokkan berdasarkan kemiripan setiap data. Pengelompokkan tersebut berdasarkan atribut yang dipilih saat pengumpulan dan preparasi data yang kemudian atribut tersebut akan diolah dengan menggunakan algoritma K-means yang akan menghasilkan cluster dan centroid. Maka berdasarkan hasil dari clustering menggunakan metode k-means sehingga dapat menghasilkan pengetahuan yang menghasilkan visualisasi dan yang dapat merepresentasikan hasil penggalan pengetahuan yang selanjutnya dapat diterapkan kedalam konsep penjualan Up-selling.

### F. Davies-Bouldin Index (DBI)

*Davies-Bouldin Index* (DBI) merupakan sebuah metode yang digunakan untuk melakukan evaluasi terhadap *cluster*. [17]. Dengan menggunakan Metode *Davies-Bouldin Index* memiliki proses evaluasi didalam internal *cluster*, dimana menghitung bagaimana baik atau tidaknya hasil *cluster* dilihat berdasarkan kedekatan dan kuantitas antar data dari hasil *cluster*. Metode *Davies-Bouldin Index* memiliki beberapa tahapan dalam perhitungannya diantaranya:

#### 1) Sum of Square Within-cluster (SSW)

Untuk mengetahui kohesi dalam sebuah *cluster* ke-*i* adalah dengan menghitung nilai dari *Sum of Square Within-cluster* (SSW). Kohesi didefinisikan sebagai jumlah dari kedekatan data terhadap titik pusat *cluster* dari sebuah *cluster* yang diikuti. Persamaan (3) digunakan untuk memperoleh nilai *Sum of Square Within cluster*.

$$SSW_i = \frac{1}{m_i} \sum_{j=i}^{m_i} d(x_j, c_j) \quad (3)$$

#### 2) Sum of Square Between-cluster (SSB)

Perhitungan *Sum of Square Between-cluster* (SSB) bertujuan untuk mengetahui separasi antar *cluster*. Persamaan (4) digunakan untuk menghitung nilai *Sum of Square Between cluster*.

$$SSB_{i,j} = d(c_i, c_j) \quad (4)$$

#### 3) Ratio (Rasio)

Rasio bertujuan untuk mengetahui nilai perbandingan antara *cluster* ke-*i* dan *cluster* ke-*j*. Untuk menghitung nilai rasio yang dimiliki oleh masing-masing *cluster*, untuk menghitung rasio digunakan persamaan (5) berikut.

$$R_{i,j} = \frac{SSW_i + SSW_j}{SSB_{i,j}} \quad (5)$$

#### 4) Davies Bouldin Index

Nilai Rasio yang telah didapat sebelumnya dapat digunakan untuk menghitung nilai *Davies Bouldin Index* (DBI) dengan menggunakan Persamaan (6) berikut:

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{i \neq j} (R_{i,j}) \quad (6)$$

Dari persamaan tersebut, *k* adalah jumlah *cluster*. Semakin kecil nilai *Davies Bouldin Index* (DBI) yang diperoleh (non-negatif  $\geq 0$ ), maka semakin baik *cluster* yang diperoleh dari pengelompokan menggunakan algoritma *clustering*.

## III. HASIL DAN DISKUSI

### A. Pengumpulan Data

Data yang digunakan dalam melakukan penelitian ini merupakan data penjualan pada Jhonarmy Clothing yang terdiri dari 2400 record data penjualan dengan rentan waktu selama 3 tahun dimulai dari bulan Januari tahun 2019 hingga Januari tahun 2021. Data Transaksi yang diperoleh memiliki atribut seperti pada Tabel 1 berikut.

TABEL 1. ATRIBUT DATA

No.	Nama Atribut
1.	Kode Penjualan
2.	Nama Pelanggan
3.	No Telepon
4.	Status
5.	Tanggal
6.	Nama Produk
7.	Nama Variasi
8.	Kategori
9.	Warna
10.	Size
11.	QTY
12.	Berat
13.	Harga Satuan
14.	Total Harga

### B. Data Pre-Processing

Sebelum melakukan perhitungan perlu melakukan Pre-processing dimana ini merupakan tahapan sebelum melakukan penggunaan data sebelum proses pengolahan yang akan dilakukan untuk melakukan pengubahan data penjualan produk yang sudah tersedia agar selanjutnya dapat di proses menjadi data yang siap untuk di olah sebagai objek penelitian. Dalam proses pre-processing terdapat beberapa tahapan diantaranya yaitu data selection, dan data cleaning.

#### 1) Data Selection

Data yang telah dimiliki selanjutnya di seleksi agar dapat menjadi sekumpulan data yang dapat digunakan untuk proses penggalan data. Pada tahap ini ditentukan attribute yang dibutuhkan diantaranya Kode Penjualan, Nama Produk, Kategori, dan QTY. Contoh dari data selection dapat dilihat pada Tabel 2.

TABEL 2. DATA SELECTION

Nama Produk	Harga Awal	Jumlah	Total Harga
Kemeja Flanel Panjang	Rp.80.000	23	Rp1.840.000
Flanel Shirt Jhonarmy 001	Rp.80.000	2	Rp160.000
Flanel Shirt Jhonarmy 002	Rp.80.000	1	Rp80.000
Jaket Parasut Jhonarmy	Rp.130.000	4	Rp520.000
Kemeja Polos Jhonarmy	Rp.80.000	5	Rp400.000

Nama Produk	Harga Awal	Jumlah	Total Harga
Long Sleeve Distortion	Rp.60.000	43	Rp2.580.000
Long Sleeve Jhonarmy	Rp.60.000	25	Rp1.500.000
Hoodie Jhonarmy Enthusy	Rp.125.000	5	Rp625.000
Hoodie Jhonarmy Enthusy II	Rp.125.000	24	Rp3.000.000
Kemeja Flanel Panjang	Rp.80.000	23	Rp1.840.000

Pemilihan atribut yang digunakan dapat mendukung pengalihan data yang sesuai dengan tujuan informasi yang akan didapatkan. Pemilihan atribut yang dibutuhkan adalah Nama Produk, Jumlah, Harga Awal. Atribut yang diambil mengacu pada penelitian sebelumnya yang relevan dengan penelitian ini yang melakukan klusterisasi produk menggunakan K-Means dengan menggunakan atribut yang sama [18].

## 2) Data Cleaning

Pada tahap Data Cleaning dilakukan pembersihan data untuk melakukan analisa terhadap data yang tidak mempunyai kelengkapan data seperti data penjualan yang tidak memiliki kelengkapan data agar dapat mempermudah proses pencarian data. Berikut ini merupakan data yang telah melalui tahap Data Cleaning dapat dilihat pada Tabel 3.

TABEL 3. DATA CLEANING

Nama Produk	Harga Awal	Jumlah	Total Harga
Kemeja Flanel Panjang	Rp.80.000	23	Rp1.840.000
Flanel Shirt Jhonarmy 001	Rp.80.000	2	Rp160.000
Flanel Shirt Jhonarmy 002	Rp.80.000	1	Rp80.000
Jaket Parasut Jhonarmy	Rp.130.000	4	Rp520.000
Kemeja Polos Jhonarmy	Rp.80.000	5	Rp400.000
Long Sleeve Distortion	Rp.60.000	43	Rp2.580.000
Long Sleeve Jhonarmy	Rp.60.000	25	Rp1.500.000
Hoodie Jhonarmy Enthusy	Rp.125.000	5	Rp625.000
Hoodie Jhonarmy Enthusy II	Rp.125.000	24	Rp3.000.000
Kemeja Flanel Panjang	Rp.80.000	23	Rp1.840.000

Hasil data yang sudah dilakukan proses pembersihan data didapatkan jumlah record data dengan jumlah data sebanyak 248 Data.

## C. Data Transformation

Data Transformation merupakan tahapan yang digunakan untuk merubah bentuk data asli menjadi data numerik untuk algoritma K-Means Clustering dan integrasi berdasarkan data yang sudah dipilih sehingga data yang dimiliki Selanjutnya melakukan normalisasi nilai setiap data dengan menggunakan

teknik data scaling yaitu *min-max normalization*. Contoh dari data Transformation dapat dilihat pada Tabel 4.

TABEL 4. DATA TRANSFORMATION

Nama Produk	Harga Awal	Jumlah	Total Harga
Kemeja Flanel Panjang	80.000	23	1.840.000
Flanel Shirt Jhonarmy 001	80.000	2	160.000
Flanel Shirt Jhonarmy 002	80.000	1	80.000
Jaket Parasut Jhonarmy	130.000	4	520.000
Kemeja Polos Jhonarmy	80.000	5	400.000

Dari Tabel 4 diatas dilakukan normalisasi menggunakan *min-max normalization* dengan menggunakan rumus 1, contoh perhitungan normalisasi tersebut dapat dilihat pada persamaan 5.

$$\frac{80000 - 55000}{270000 - 55000} (1 - 0) + 0 = 0,116 \quad (7)$$

Persamaan (7) diatas merupakan contoh perhitungan normalisasi pada produk 1 dengan harga awal 80.000 yang dinormalisasi menjadi 0,116. Setiap record data dilakukan proses normalisasi menggunakan *min-max normalization* dengan skala 0 hingga 1. Tabel 5 Berikut ini merupakan hasil dari perhitungan *min-max normalization*.

TABEL 5. HASIL NORMALISASI

Nama Produk	Harga Awal	Jumlah	Total Harga
Kemeja Flanel Panjang	0.116	0.415	0.177
Flanel Shirt Jhonarmy 001	0.116	0.018	0.009
Flanel Shirt Jhonarmy 002	0.116	0.0	0.001
Jaket Parasut Jhonarmy	0.348	0.056	0.045
Kemeja Polos Jhonarmy	0.116	0.075	0.033

## D. Clustering dengan metode K-Means

Setelah melakukan tahap Pre-Processing terhadap data yang akan digunakan maka tahap selanjutnya melakukan penerapan teknik data mining yang telah dibahas sebelumnya, teknik yang digunakan yaitu teknik klusterisasi dengan implementasi menggunakan algoritma K-Means terhadap dataset yang sebelumnya sudah melewati tahap preproses dan transformasi. Jumlah cluster dalam algoritma K-Means tidak ada ketentuan dalam menentukan jumlah cluster namun pada penelitian ini jumlah cluster ditentukan sebanyak 3. Selanjutnya akan dilakukan perhitungan dengan metode k-means berdasarkan langkah-langkah pada tinjauan pustaka sebelumnya:

### 1) Menentukan K Optimal

Selanjutnya akan masuk kedalam perhitungan klasterisasi menggunakan algoritma K-Means Clustering dimana sebelum itu harus menentukan nilai K terlebih dahulu, nilai K pada penelitian ini ditentukan sebanyak 3 karena hasil dari proses klasterisasi akan dikelompokkan menjadi 3 Cluster yaitu kelompok Sangat diminati dan kurang diminati. Berikut ini contoh perhitungan K-Means Clustering dengan contoh 5 data seperti pada Tabel 6 dibawah ini dengan centroid awal seperti pada Tabel 7.

TABEL 6. CONTOH DATA

Produk	Harga Awal	Jumlah	Total Harga
1	0.116	0.415	0.177
2	0.116	0.018	0.009
3	0.116	0.415	0.001
4	0.116	0.018	0.045
5	0.116	0.0	0.033

TABEL 7. CENTROID AWAL

Centroid	Harga Awal	Jumlah	Total Harga
C1	0.116	0.415	0.177
C2	0.116	0.018	0.009
C3	0.116	0.415	0.001

Kemudian Hitung Panjang jarak data dengan pusat centroid dengan menggunakan rumus *Euclidian Distance* Dengan rumus (8).

$$D(a, b) = \sqrt{(x_a - y_b)^2 + (x_a - y_b)^2} \quad (8)$$

Diketahui :

- Jumlah Data Produk (P) = 5
- Ditentukan Nilai K = 3
- Centroid Cluster 1 (CC-1) = {0.116, 0.415, 0.177}
- Centroid Cluster 2 (CC-2) = {0.116, 0.018, 0.009}
- Centroid Cluster 2 (CC-2) = {0.116, 0.415, 0.001}

Jarak berdasarkan data dengan pusat cluster pertama didefinisikan oleh Tabel 8.

TABEL 8. JARAK DATA DENGAN CLUSTER 1

No	Produk	Jarak ke Centroid
1	Produk 1	0
2	Produk 2	0,431
3	Produk 3	0,716
4	Produk 4	0,418
5	Produk 5	0,439

Jarak berdasarkan data dengan pusat cluster kedua didefinisikan oleh Tabel 9.

TABEL 9. JARAK DATA DENGAN CLUSTER 2

No	Produk	Jarak ke Centroid
1	Produk 1	0,431
2	Produk 2	0
3	Produk 3	0,397
4	Produk 4	0,036

5	Produk 5	0,030
---	----------	-------

Jarak berdasarkan data dengan pusat cluster ketiga didefinisikan oleh Tabel 10.

TABEL 10. JARAK DATA DENGAN CLUSTER 2

No	Produk	Jarak ke Centroid
1	Produk 1	0,176
2	Produk 2	0,397
3	Produk 3	0
4	Produk 4	0,399
5	Produk 5	0,416

Langkah berikutnya adalah menghitung centroid baru berdasarkan rata rata dari hasil perhitungan cluster yang dilakukan pada iterasi I. Yang nantinya centroid baru digunakan untuk melakukan pengulangan sampai mendapatkan hasil dengan nilai rata-rata atau *cluster* yang sudah tidak berubah atau konvergen. Hasil dari pengelompokan data berdasarkan *cluster* nya diberikan oleh Tabel 11 dan Gambar 1.

TABEL 11. HASIL KLASTERISASI

Nama Produk	Harga Awal	Jumlah	Total Harga	Cluster
Kemeja Flanel Panjang	Rp.80.000	23	Rp1.840.000	C3
Flanel Shirt Jhonarmy 001	Rp.80.000	2	Rp160.000	C1
Flanel Shirt Jhonarmy 002	Rp.80.000	1	Rp80.000	C1
Jaket Parasut Jhonarmy	Rp.130.000	4	Rp520.000	C1
Kemeja Polos Jhonarmy	Rp.80.000	5	Rp400.000	C1



Gambar 1. Visualisasi klasterisasi

### E. Interpretasi

Setelah mendapatkan hasil klasterisasi maka didapat hasil dari data mining yang dapat digunakan sebagai informasi bagi perusahaan, untuk produk yang memiliki tingkat penjualan paling tinggi sebagai rekomendasi perusahaan untuk meningkatkan penjualan, seperti diberikan oleh Tabel 12.

TABEL 12. REKOMENDASI PRODUK

Nama Produk	Harga Awal	QTY
Long Shirt Flannel JHNRMY NB	190000	53

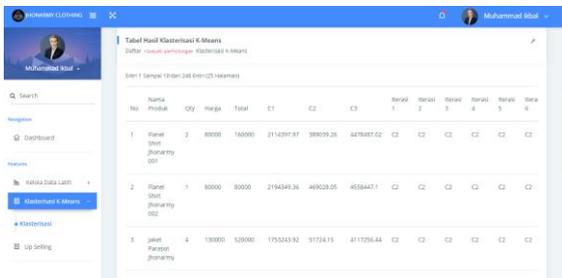
Nama Produk	Harga Awal	QTY
Hoodie Zipper Jhonarmy Skull Vol.1	260000	25
Hoodie Jhonarmy Full Up United Heaven	250000	26
Shirt Basic JHNRMY V	150000	39
Hoodie Jhonarmy Full Up Tribal Basic IV	250000	23
Hoodie Zipper Jhonarmy arm I	250000	23
Hoodie Jhonarmy Full Up Tribal Basic III	250000	20
Hoodie Jhonarmy Full Up Skate And Destroy	250000	20
Hoodie Jhonarmy Full Up Tribal Basic II	250000	19
Hoodie Jhonarmy Full Up BDG	250000	19
Long Shirt Flannel JHNRMY NB	190000	53
Hoodie Zipper Jhonarmy Skull Vol.1	260000	25

#### F. Implementasi Perangkat Lunak

Implementasi perangkat lunak memperlihatkan hasil dari perangkat lunak atau produk yang telah dibuat.

##### 1) Halaman hasil klasterisasi

Halaman hasil klasterisasi merupakan halaman yang menampilkan hasil perhitungan klasterisasi serta pengelompokan produk yang terdapat 3 cluster. Seperti yang terlihat pada Gambar 2.



Gambar 2. Halaman hasil klasterisasi

##### 2) Halaman visualisasi klasterisasi

Halaman visualisasi klasterisasi merupakan halaman yang menampilkan hasil klasterisasi serta data produk hasil klasterisasi untuk dijadikan rekomendasi konsep *Up-Selling*. Seperti yang terlihat pada Gambar 3.



Gambar 3. Implementasi perangkat lunak

#### G. Davies-Bouldin Index (DBI)

Perhitungan yang telah dilakukan menggunakan metode *K-Means Clustering* selanjutnya dilakukan uji performace yang digunakan untuk menilai bagaimana baik atau tidaknya hasil *cluster* dilihat berdasarkan kedekatan dan kuantitas antar data dari hasil *cluster*. Dimana berdasarkan perhitungan

*Davies-Bouldin Index* dengan menggunakan bantuan tools *rapidminer* dengan penggunaan 3 klaster didapatkan nilai sebesar 0,181. Dimana dengan nilai yang didapat tersebut dapat dikatakan kualitas klaster cukup baik

#### IV. KESIMPULAN

Dalam penelitian Klasterisasi Penjualan Produk Clothing Line Menggunakan Metode K-Means Clustering Untuk Penerapan Konsep *Up-Selling* yang dilakukan dapat menghasilkan beberapa rekomendasi produk berdasarkan tingkat penjualan yang tinggi yang didapat dari hasil klasterisasi yang telah dilakukan dengan menggunakan 3 cluster yang mendapatkan nilai Davies-Bouldin Index sebesar 0,181. Dengan demikian tujuan yang dilakukan pada penelitian yang dilakukan menghasilkan rekomendasi produk dimana hasil rekomendasi produk tersebut dapat digunakan oleh perusahaan dengan menggunakan konsep *Up-Selling* untuk upaya meningkatkan penjualan pada perusahaan.

#### DAFTAR PUSTAKA

- [1] k. M. Raval, "data mining techniques," *springerbriefs appl. Sci. Technol.*, vol. 2, no. 10, pp. 439–442, 2012, doi: 10.1007/978-3-319-22294-3\_3.
- [2] c. Italina and fahrurrazi, "pengaruh strategi pemasaran terhadap peningkatan volume penjualan pakaian pada toko grosir kadafi collection di kota sigli kabupaten pidie," *j. Sains ris.*, vol. 9, no. September 2019, pp. 61–67, 2019.
- [3] F., f. T. Kesuma, and s. P. Tamba, "penerapan data mining untuk menentukan penjualan sparepart toyota dengan metode k-means clustering," *j. Sist. Inf. Dan ilmu komput. Prima(jusikom prima)*, vol. 2, no. 2, pp. 67–72, 2020, doi: 10.34012/jusikom.v2i2.376.
- [4] i. Sumadikarta and e. Abeiza, "penerapan algoritma k-means pada data mining untuk memilih produk dan pelanggan potensial (studi kasus : pt mega arvia utama)," *j. Satya inform.*, vol. 1, no. 1, pp. 12–22, 2014.
- [5] m. Silalahi, "analisis clustering menggunakan algoritma k-means terhadap penjualan produk padapt batamas niaga jaya," *comput. Based inf. Syst. J.*, vol. 02, pp. 20–35, 2018.
- [6] s. M. A. M. Manchanayake et al., "potential upselling customer prediction through user behavior analysis based on cdr data," 2019 *ieee 14th int. Conf. Ind. Inf. Syst. Eng. Innov. Ind. 4.0, iciis 2019 - proc.*, pp. 46–51, 2019, doi: 10.1109/iciis47346.2019.9063278.
- [7] c. W. Randi rian putra, "implementasi data mining pemilihan pelanggan potensial menggunakan algoritma k-means," *intecom s. j. Technol. Comput. Sci.*, vol. 1, no. 1, pp. 72–77, 2018.
- [8] b. M. Metisen and h. L. Sari, "analisis clustering menggunakan metode k-means dalam pengelompokan penjualan produk pada swalayan fadhila," *j. Media infotama*, vol. 11, no. 2, pp. 110–118, 2015.
- [9] a. Asroni, h. Fitri, and e. Prasetyo, "penerapan metode clustering dengan algoritma k-means pada pengelompokan data calon mahasiswa baru di universitas muhammadiyah yogyakarta (studi kasus: fakultas kedokteran dan ilmu kesehatan, dan fakultas ilmu sosial dan ilmu politik)," *semesta tek.*, vol. 21, no. 1, pp. 60–64, 2018, doi: 10.18196/st.211211.
- [10] v. R. Patel and r. G. Mehta, "impact of outlier removal and normalization approach in modified k-means clustering algorithm," *int. J. Comput. Sci. Issues*, vol. 8, no. 5, pp. 331–336, 2011.
- [11] w. E. Nurjanah, r. S. Perdana, and m. A. Fauzi, "analisis sentimen terhadap tayangan televisi berdasarkan opini masyarakat pada media sosial twitter menggunakan metode k-nearest neighbor dan pembobotan jumlah retweet," *j. Pengemb. Teknol. Inf. Dan ilmu komput. Univ. Brawijaya*, vol. 1, no. 12, pp. 1750–1757, 2017.
- [12] r. Muliono and z. Sembiring, "data mining clustering menggunakan algoritma k-means untuk klasterisasi tingkat tridarma pengajaran dosen," *cess (journal comput. Eng. Syst. Sci.)*, vol. 4, no. 2, pp. 2502–714, 2019.

- [13] k. P. Sinaga and m. S. Yang, "unsupervised k-means clustering algorithm," *iee access*, vol. 8, pp. 80716–80727, 2020, doi: 10.1109/access.2020.2988796.
- [14] p. O. Olukanmi and b. Twala, "k-means-sharp: modified centroid update for outlier-robust k-means clustering," 2017 *pattern recognit. Assoc. South africa robot. Mechatronics int. Conf. Prasa-robmech 2017*, pp. 14–19, 2017, doi: 10.1109/robomech.2017.8261116.
- [15] j. Qi, y. Yu, l. Wang, and j. Liu, "k\*-means: an effective and efficient k-means clustering algorithm," *proc. - 2016 iee int. Conf. Big data cloud comput. Bdcloud 2016, soc. Comput. Networking, soc. 2016 sustain. Comput. Commun. Sustain. 2016*, pp. 242–249, 2016, doi: 10.1109/bdcloud-socialcom-sustaincom.2016.46.
- [16] f. Purwaningtias, "penerapan product knowledge dan strategi upselling pada chandra tech palembang," *j. Inform. J. Pengemb. It*, vol. 3, no. 2, pp. 253–258, 2018, doi: 10.30591/jpit.v3i2.826.