

Implementasi Metode K-Means Clustering Untuk Pengelompokan Data Stok Produk Toko Online Perdagangan Kaos

Sukri Illaihi Wahyudi^{1*}, Arief Wibowo²

^{1,2}Fakultas Teknologi Informasi, Teknik Informatika, Universitas Budi Luhur, Jakarta, Indonesia

Email: ^{1*}sukriillaihi@gmail.com, ²arief.wibowo@gmail.com
(* : corresponding author)

Abstrak- Perdagangan Kaos merupakan sebuah toko yang bergerak dibidang retail kaos. Dimana pada proses yang di alami perdagangan kaos masih memiliki suatu masalah seperti tidak bisa menetapkan suatu produk terlaris untuk promosi toko kedepannya. Salah satu permasalahan yang di alami toko perdagangan kaos yaitu sering terjadi suatu masalah untuk stok yang akan di produksi yang paling diminati konsumen. Untuk mencegah suatu permasalahan tersebut maka dibutuhkan sebuah sistem dapat mampu menetapkan suatu produk terlaris bukan hanya satu produk saja namun kategori variasi produk terlaris. Adapun tujuan yang ingin dicapai dari penelitian ini adalah untuk mengetahui penjualan stok produk pada toko Perdagangan Kaos. Menggunakan metode K-Means dapat menghasilkan pengelompokan produk yang akan di stok oleh toko. Pengkategorian variasi produk terlaris bermaksud untuk di perhitungkan maka digunakan metode K-Means Clustering, K-Means Clustering dapat menghitung secara akurat dan tepat. Maka program diharapkan dapat mengelompokkan suatu variasi kaos terlaris dengan kompleks dan tepat yang akan membantu mengatasi permasalahan yang ada pada Perdagangan Kaos. Algoritme K-Means digunakan untuk program yang akan dibuat dan dijalankan toko Perdagangan Kaos untuk mendapatkan barang dari kaos yang mana termasuk laku, sedikit laku dan tidak laku. Implementasi algoritme K-Means Clustering pada toko Perdagangan Kaos ini yaitu dengan mengumpulkan stok variasi kaos. Lalu langsung di selektif 3 group secara acak sebagai awal centroid proses. Terlihat 20 data uji produk pada hasil akhir program berbasis web ada 5 data variasi kaos laku keras, 5 data variasi kaos laris, dan 10 data variasi kaos sedikit penjualan, Lalu penerapan algoritme K-Means pada program berbasis web maka lakukan melalui data stok variasi produk merupakan stok awal dan stok akhir, bahwa untuk diubah menjadi database di Microsoft Excel, data tersebut akan terhubung ke Program berbasis web dan dijalankan. dan merupakan algoritme clustering. Setelah itu, Program akan memproses hasil produksi mana yang permintaannya tinggi, permintaannya sedang, dan permintaannya kurang. Sedikit penjelasan mengenai K-Means Clustering, yaitu merupakan satu algoritme clustering yang berbentuk non tingkatan yang mempartisi data ke dalam cluster sehingga data yang memiliki karakteristik sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok lain. Dari kesimpulan di atas merupakan penjelasan algoritme K Means Clustering ini dapat membantu sistem berjalan dengan baik.

Kata kunci : *perdagangan kaos, produk, stok, data mining, algoritme k-means clustering*

Implementation of K-Means Clustering Method for Online Store Product Stock Data Grouping T-shirt Trading

Abstract- *T-shirt trading is a shop engaged in the retail of t-shirts. Where the process experienced by the t-shirt trade still has a problem such as not being able to determine a best-selling product for future store promotions. One of the problems experienced by the t-shirt trading shop is that there is often a problem for the stock that will be produced which is most in demand by consumers. The goal to be achieved from this research is to find out the sale of product stock at the T-shirt Trading store. Using the K-Means method can produce product groupings that will be stocked by the store. The categorization of the best-selling product variations is intended to be taken into account, so the K-Means Clustering method is used, K-Means Clustering can calculate accurately and precisely. So the program is expected to be able to classify a variety of best-selling t-shirts in a complex and precise manner that will help overcome the problems that exist in the T-shirt Trade. The K-Means algorithm is used for a program that will be created and run by the T-shirt Trading store to get goods from T-shirts which include selling, selling a little and not selling. The implementation of the K-Means Clustering algorithm at this T-shirt Trading store is by collecting stock of variations of t-shirts. Then immediately selectively 3 groups randomly as the beginning of the centroid process. There are 20 product test data in the final results of the web-based program, there are 5 data on variations of T-shirts selling well, 5 data on variations of best-selling t-shirts, and 10 data on variations of T-shirts selling slightly is the starting stock and ending stock, that to be converted into a database in Microsoft Excel, the data will be connected to a web-based program and executed. and is a clustering algorithm. After that, the Program will process which production results are in high demand, medium demand, and low demand. A little explanation about K-Means Clustering, which is a non-level clustering algorithm that partitions data into clusters so that data with the same characteristics are grouped into the same cluster and data with different characteristics are grouped into other groups. From the conclusion above, an explanation of the K Means Clustering algorithm can help the system run well.*

Keywords: *t-shirt trading, product, stock, data mining, k-means clustering algorithm*

1. PENDAHULUAN

Perdagangan Kaos merupakan toko yang bergerak di bidang industri fashion, kaos polos dengan berbagai model seperti turtleneck, V-neck, dan O-neck. Banyak toko lain juga berada dikategori yang sama. Hal ini tentu saja menimbulkan kompetisi komersial dengan toko lain. Oleh karena itu, diperlukan suatu program untuk menentukan produk terbaik di suatu toko agar penjualannya dapat bersaing dengan toko lain. Setiap transaksi penjualan di toko perdagangan kaos masih menggunakan sistem manual. Pengelompokan data penjualan selalu rekap dan disimpan dalam database pencatatan note. Data Mining adalah mengekstraksi bentuk yang menarik dari data dalam ukuran besar dan data yang akan di proses oleh program. Strategi yang diharapkan mudah dipahami oleh pengguna, *valid* untuk data penjualan yang akan diprediksi dengan bagian spesifik, menguntungkan, dan baru. Data mining memiliki beberapa nama alternative, meskipun definisi eksaknya berbeda, seperti *Knowledge Discovery in Databases* (KDD), analisis pola, arkeologi data, pemanenan informasi, dan intelegensia bisnis [1]. Clustering adalah *algoritme* yang menganalisa data bertujuan untuk mengkategorikan keterangan dengan nilai yang mirip ke suatu wilayah yang mirip. Program digunakan dalam mengembangkan metode clustering yaitu metode K-Means, dimana algoritme ini merupakan salah satu metode pengkategorian data non tingkat(sekatan) yang berusaha mempartisi data ke dalam bentuk dua atau lebih kelompok (kluster) yang berkarakteristik sama dimasukkan ke dalam satu kategori yang mirip[2]. Mencatat tiap-tiap transaksi sangat berguna untuk menganalisis pemasaran. Untuk permasalahan yang berada di toko ini peningkatan yang terjadi masih memiliki sedikit kendala pada penentuan stok produksi, dimana dalam penempatan stok produksi hanya dapat diterima melalui data berada di toko. Data mining bertujuan untuk mengungkap solusi nyata dalam pengambilan ketentuan metode di atas dengan menggunakan teknik pengklasifikasian data, K-Means merupakan sebuah *algoritme* clustering. Metode *algoritme* K-Means sangat populer karena kemudahan dan kemampuannya untuk mengklarifikasi data banyak dan outlier dengan sangat cepat. Jumlah permintaan dari konsumen merupakan turun dan naik mengakibatkan stok yang harus disiapkan Perdagangan Kaos menjadi tidak akurat. Penunjuk algoritme K-Means disebabkan algoritme ini harus menggunakan data fisik tidak abstrak dan bersifat kompleks, hal ini sesuai dengan data digunakan pada permasalahan di dalam pengelompokan penjualan di toko perdagangan kaos. hanya itu, algoritme ini pada program bersifat fleksibel sebab pengguna dapat menentukan jumlah cluster yang akan dibuat[3]. Terkadang dikarenakan sesuatu yang kurang di ingin terjadi kekurangan variasi produk atau produk tertentu pada saat permintaan konsumen dalam jumlah besar, peneliti pada toko mengambil langkah yaitu melakukan pemesanan barang produk tertentu lebih besar dari pada sebelumnya. Hal ini dapat mengatasi kekurangan persediaan stok produk tertentu dari toko perdagangan kaos juga tidak perlu melakukan pemesanan berulang-ulang ke pengrajin, tetapi mengakibatkan biaya simpan yang tinggi dan tidak ekonomis Program ini dapat menyelesaikan permasalahan di atas, maka penelitian ini membutuhkan program web data mining untuk mengkategorikan produk mana saja yang laku dan tidak. Data mining merupakan analisa terhadap data untuk menemukan hubungan yang jelas serta menyimpulkannya yang belum diketahui sebelumnya dengan cara terkini dipahami dan berguna bagi pemilik data tersebut[4]. Maka dari itu dengan pengkategorian data ini bagian dari perdagangan kaos dapat menyadari barang paling laris, laris dan tidak laris, sehingga barang yang ada digudang tidak menumpuk[5]. Hasil penelitian ini bermanfaat bagi bagian yang terkait dengan perencanaan kebutuhan stok barang pada toko, yang membutuhkan data jenis-jenis stok barang mana sajakah yang penjualannya rendah, sedang, dan tinggi setiap tahun nya, sehingga dapat digunakan sebagai acuan perencanaan kebutuhan stok barang untuk kedepannya[6]. Diharapkan adanya penelitian ini dan kajian lebih lanjut mengenai pengembangan program aplikasi clustering yang menggunakan metode K-Means Clustering dimasa yang akan datang. Serta melakukan evaluasi secara berkala terhadap kinerja program sebagai bahan masukan untuk keperluan pengembangan program selanjutnya[7].

2. METODE PENELITIAN

2.1 *Algoritme K-Means Clustering*

K-Means yakni merupakan algoritme klustering dengan menentukan sejumlah data untuk di klaster dalam kesamaan karakteristik dan memaksimalkan perbedaan antar klaster. K-Means algoritma salah satu algoritme clustering yang paling banyak orang pakai untuk mengkategorikan sebuah produk. Yang dimaksud K sebagai konstanta merupakan total kluster yang diperlukan, Means dalam hal ini berarti nilai rata-rata dari suatu kategori data yang dalam hal ini didefinisikan sebagai cluster, sehingga K-Means Clustering adalah suatu metode menganalisa data atau metode data mining yang melakukan proses pemodelan tanpa supervisi (unsupervised) dan merupakan salah satu metode yang melakukan pengkategorian data dengan sistem partisi[8]. Algoritme ini partisi data ke dalam golongan maka dari itu data berkategori berbagai dapat disertakan ke dalam golongan merupakan lain, K-Means termasuk dalam cluster splitting. Semua data harus berada di cluster tertentu, dan semua data yang ada di kluster tertentu pada satu tahap proses dapat dipindah ke kluster lain pada tahap selanjutnya yang saya miliki.

Algoritma K-Means dikenal karena kesederhanaan dan kemampuannya untuk mengklasifikasikan sejumlah besar data dan outlier dengan sangat cepat. [9]

Berikutlah proses *algoritme K Means* :

- a. Membuktikan nilai kluster menjadi jumlah kluster yang dibentuk.
- b. Membuktikan titik awal (*centroid*) kluster.

Dalam menentukan buah pusat cluster awal dilakukan pemilihan bilangan random yang mempresentasikan urutan data input. Pusat awal kluster didapatkan dari data sendiri bukan dengan menentukan titik baru, yaitu dengan mengacak (random) pusat awal dari data. Kemudian untuk menghitung *centroid cluster* ke-I dengan menggunakan persamaan rumus sebagai berikut :

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

Keterangan untuk rumus :

v : *Centroid* pada kluster

x_i : suatu rumus ke-j

n : merupakan objek/jumlah anggota kluster

- c. Perdiksi jarak dengan pusat cluster

Buat mengukur suatu jarak setiap stok data dari titik awal cluster digunakan rumus Euclidian Distance. Selisih perhitungan akan dibandingkan dan dipilih pusat terdekat antara data dengan pusat cluster, pusat ini menunjukkan bahwa data tersebut berada dalam suatu group dengan pusat cluster terdekat, berikut ini adalah Rumus dari proses:

$$d(x, y) = || x - y || = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}; i = 1, 2, 3, \dots n \quad (2)$$

Keterangan untuk rumus :

X_i : pusat titik kluster

Y_i : jumlah data

d(x,y) : Euclidian Distance untuk menghitung jarak antara data pada titik x dan titik y

- d. Mengelompokkan data

Setelah sekumpulan populasi data ditemukan dekat dengan salah satu *centroid* yang ada, populasi data tersebut secara otomatis dimasukkan ke dalam kelas dengan *centroid* yang sesuai.

- e. ulangi dan gunakan rumus untuk menempatkan centroid baru

- f. Jika masih ada kelompok pergeseran data, jika perubahan nilai centroid melebihi ambang batas yang ditentukan, jika ambang batas yang ditentukan berubah, atau jika nilai fungsi tujuan yang digunakan adalah Ulangi langkah 3 meski

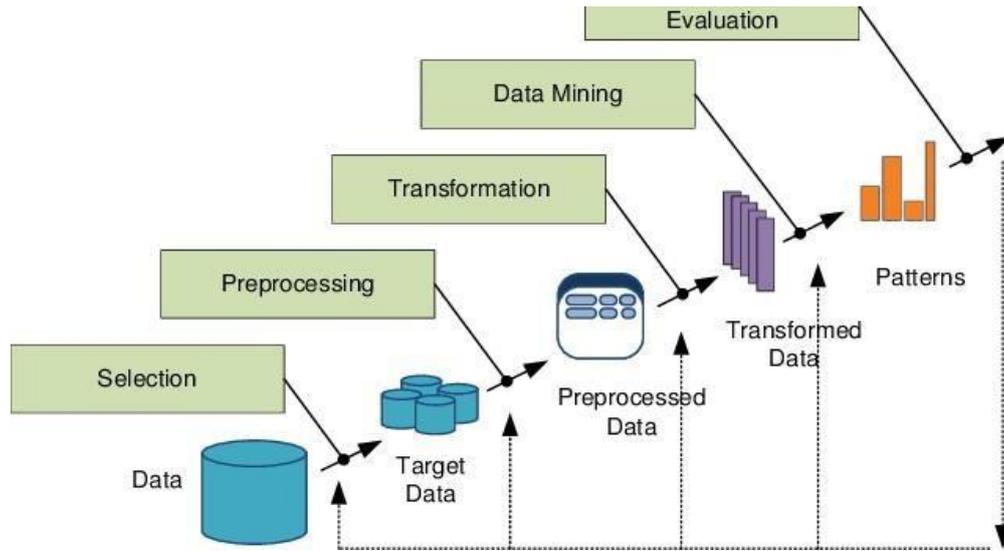
2.2 Tahapan Data Mining

Data Mining termasuk dalam kumpulan *Knowledge discovery in databases* (KDD), yang bersambung dengan seni mengintegrasikan dan penemuan ilmiah, interpretasi dan visualisasi pola dalam kumpulan data. Urutan proses ini mempunyai fase-fase berikut:

- a. Pembersihan data (untuk menghilangkan data dan noise yang tidak konsisten).
- b. Integrasi data (penggabungan data dari berbagai sumber).
- c. Transformasi data (data diubah menjadi format yang sesuai untuk penambangan).
- d. Menerapkan teknik data mining, proses penggalian pola dari data yang ada.
- e. Evaluasi pola yang ditemukan (proses menafsirkan pola pengetahuan yang dapat digunakan untuk mendukung pengambil keputusan).
- f. Penyajian pengetahuan (menggunakan teknik visualisasi).

2.3 Proses Knowledge Discovery in Databases (KDD)

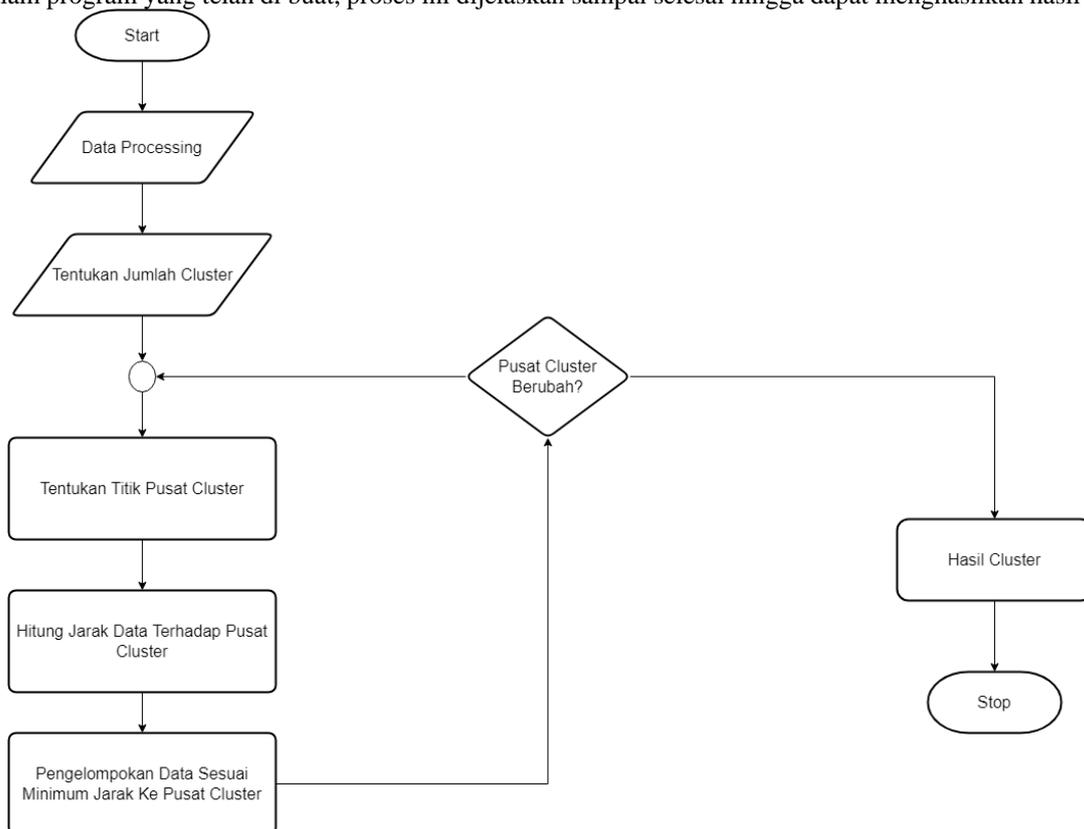
Proses Tahapan ini dimulai bagian dari proses pencarian pengetahuan yang memeriksa untuk melihat apakah ada pola atau informasi yang ditemukan berdampak dengan fakta atau hipotesis yang ada [10].



Gambar 1. Proses KDD

2.4 Flowchart Tampilan Proses K-Means Pada Program

Flowchart ini menjelaskan alur untuk proses k-means yang ada pada program setelah data sudah di importkan ke dalam program yang telah di buat, proses ini dijelaskan sampai selesai hingga dapat menghasilkan hasil cluster



Gambar 2. Proses K-Means

3. HASIL DAN PEMBAHASAN

Pada pembahasan bab ini mengenai analisa metode, implementasi ataupun pengujian serta pembahasan dari topik penelitian, yang bisa dibuat terlebih dahulu metodologi algoritme penelitian. Bagian ini juga merepresentasikan penjelasan yang berupa penjelasan, gambar, tabel dan lainnya.

3.1 Data Uji K-Means Clustering

Berikut data asli yang akan di proses menggunakan k-means clustering.

Tabel 1. Data Asli

Kode Barang	Stok Awal	Stok Akhir
K-001	150	145
K-002	150	145
K-003	150	144
K-004	150	143
K-005	150	147
K-006	150	135
K-007	150	138
K-008	150	138
K-009	150	135
K-010	150	142
K-011	100	95
K-012	100	92
K-013	100	97
K-014	100	89
K-015	100	98
K-016	120	109
K-017	120	105
K-018	120	112
K-019	120	115
K-020	120	117

Berikut data yang sudah di preprocessing atau di normalisasikan menggunakan codingan program yang telah dibuat.

Tabel 2. Data Normalisasi

Kode Barang	Stok Awal	Stok Akhir
K-001	1	0.965517241
K-002	1	0.965517241
K-003	1	0.948275862
K-004	1	0.931034483
K-005	1	1
K-006	1	0.793103448
K-007	1	0.844827586
K-008	1	0.844827586
K-009	1	0.793103448
K-010	1	0.913793103
K-011	0	0.103448276
K-012	0	0.051724138
K-013	0	0.137931034
K-014	0	0
K-015	0	0.155172414
K-016	0.4	0.344827586
K-017	0.4	0.275862069
K-018	0.4	0.396551724
K-019	0.4	0.448275862
K-020	0.4	0.482758621

3.2 Proses Dari Algoritme Perhitungan Clustering

Pada perhitungan ini dilakukan mengkategorikan data pemasaran untuk dijalankan melalui data Excel yang disediakan oleh peneliti menggunakan teknik clustering K-Means. Di bawah ini adalah proses dari algoritma K-Means. di mana proses input adalah jumlah catatan data dan inisialisasi centroid melompat $C=3$, menurut penelitian ini. Dari banyaknya data penjualan yang diperoleh, peneliti menerapkan algoritma K-Means pada lima jenis produk sebagai sampel. Pengujian dilakukan dengan menggunakan parameter berikut:

Nilai Kluster : 3

Data Contoh : 20

Berikut data atribut dan tabel yang di gunakan peneliti , untuk data yang di gunakan peneliti hanya data stok awal dan stok akhir saja berguna untuk menentukan penjualan pada toko tersebut. Stok awal dan stok akhir untuk menentukan clustering.

Iterasi ke-1

a. Penetapan Awal Pusat Kluster

Pusat awal kluster atau centroid didapatkan secara random, untuk penentuan awal kluster adalah :

Pusat cluster 1 : (1, 0.9)

Pusat cluster 2 : (0.4,0.389655)

Pusat cluster 3 : (0 , 0.089655)

b. Perhitungan jarak pusat kluster percobaan 1

Untuk mengukur jarak antara data dengan pusat cluster digunakan Euclidian distance, kemudian akan didapatkan matrik jarak sebagai berikut :

Rumus Euclidian Distance :

$$d = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$

X_i = Pusat kluster

Y_i = data

$d(x,y)$ = Euclidian Distance merupakan jarak antara data pada titik x dan titik y.

Dari 200 data dijadikan contoh telah dicari pusat awal cluster yaitu : C1(1,0.9),C2(0.4,0.389655),dan C3(0,0.089655). Kemudian dilakukan perhitungan jarak dari sisa percontohan data dengan pusat cluster yang dimisalkan dengan M(a,b) dimana a bagian dari stok awal , dan b stok akhir.

Hitung Hasil Perhitungan *Euclidean distance* dari semua data ketiap titik pusat pertama :

$$d = \sqrt{(Mx - Cx)^2 + (My - Cy)^2} = \sqrt{(100 - 1)^2 + (99 - 0.9)^2} = 0.06551724$$

Hitung Hasil Perhitungan *Euclidean distance* dari semua data ketiap titik pusat kedua :

$$d = \sqrt{(Mx - Cx)^2 + (My - Cy)^2} = \sqrt{(100 - 0.4)^2 + (99 - 0.389655)^2} = 0.83163521$$

Hitung Hasil Perhitungan *Euclidean distance* dari semua data ketiap titik pusat ketiga :

$$d = \sqrt{(Mx - Cx)^2 + (My - Cy)^2} = \sqrt{(100 - 0)^2 + (99 - 0.089655)^2} = 1.32933606$$

Hasil perhitungan *Euclidean distance* :

Tabel 3. Hasil Perhitungan Euclidian Distance

Kode Barang	C1	C2	C3
K-001	0.06551724	0.83163521	1.32933606
K-002	0.06551724	0.83163521	1.32933606
K-003	0.04827586	0.81979087	1.31804002
K-004	0.03103448	0.8081408	1.30687381
K-005	0.1	0.85587429	1.35230459
K-006	0.10689655	0.72302871	1.22263628
K-007	0.05517241	0.75311482	1.25311028
K-008	0.05517241	0.75311482	1.25311028
K-009	0.10689655	0.72302871	1.22263628
K-010	0.0137931	0.79669352	1.29584078
K-011	1.27847356	0.49184793	0.0137931
K-012	1.3113245	0.5236386	0.03793103
K-013	1.25727845	0.47261511	0.04827586
K-014	1.3453624	0.55841844	0.08965517
K-015	1.24690342	0.46366169	0.06551724
K-016	0.81744505	0.04482759	0.47446071
K-017	0.86576449	0.1137931	0.44121764
K-018	0.78323698	0.00689655	0.50416812

Berikut ini akan ditampilkan data matriks pengelompokan golongan iterasi 3 , nilai 1 berarti data tersebut berada dalam golongan (kelompok data).

Tabel 4. Hasil Pengelompokkan Clustering

Kode Barang	C1	C2	C3
K-001	1		
K-002	1		
K-003	1		
K-004	1		
K-005	1		

Kode Barang	C1	C2	C3
K-006	1		
K-007	1		
K-008	1		
K-009	1		
K-010	1		
K-011			1
K-012			1
K-013			1
K-014			1
K-015			1
K-016		1	
K-017		1	
K-018		1	

Anggota C1 = {K-001,K-002,K-003,K-004,K-005,K-006,K-007,K-008,K-009,K-010}

Anggota C2 = {K-016,K-017,K-018,K-019,K-020}

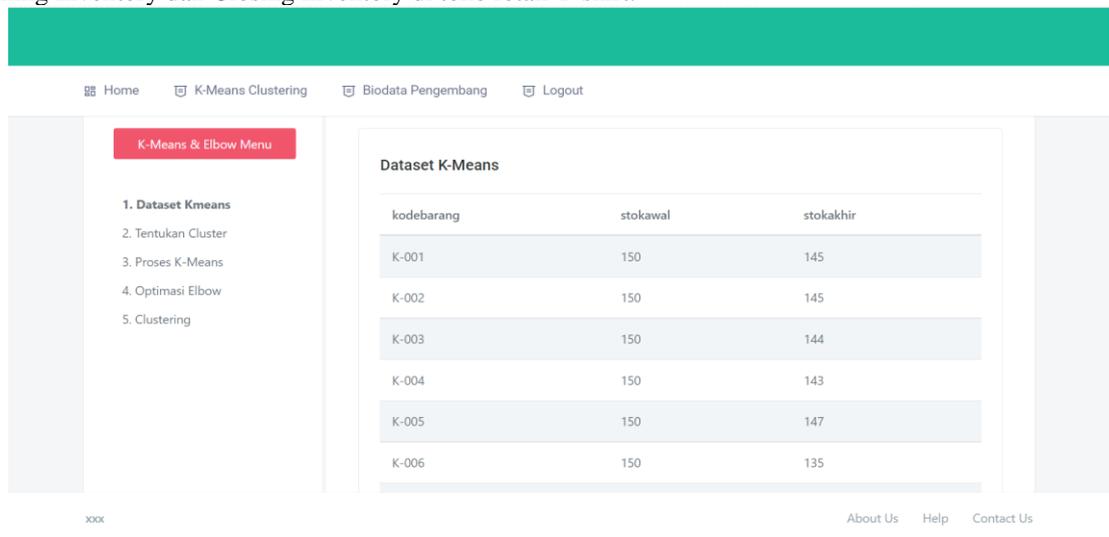
Anggota C3 = {K-011,K-012,K-013,K-014,K-015}

Kesimpulan Dari Proses :

Tingkat keuntungan juga rendah untuk produk yang anggota C1 dan di bawahnya memiliki sedikit aktivitas penjualan. Kaos C2 tergolong kaos profit dan penjualan sedang, dan kaos C3 tergolong kaos tingkat penjualan dan profit keuntungan tinggi.

3.3 Hasil Program

Implementasi Hasil Program adalah langkah menuju pengoperasian sistem yang sedang dibangun. Penelitian ini menjelaskan bagaimana sistem bekerja. Di bawah ini adalah tampilan implementasi Data Mining Group Opening Inventory dan Closing Inventory di toko retail T-shirt.

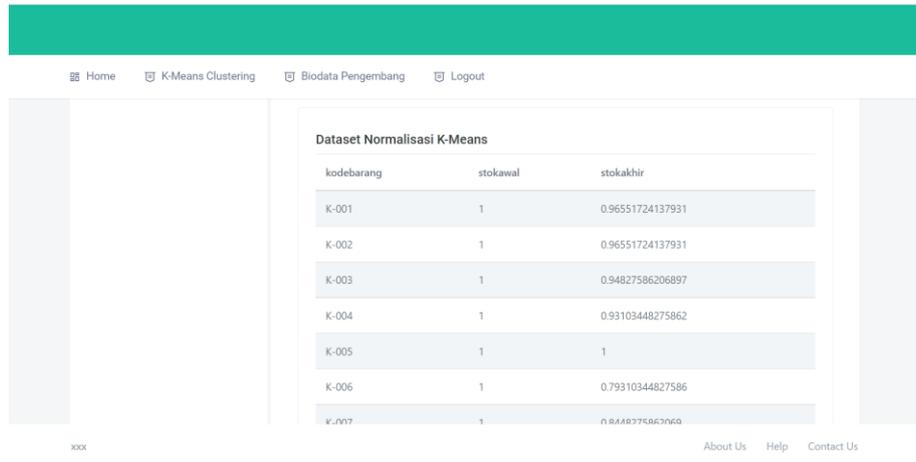


The screenshot shows a web application interface for K-Means Clustering. The main content area displays a table titled 'Dataset K-Means' with the following data:

kodebarang	stokawal	stokakhir
K-001	150	145
K-002	150	145
K-003	150	144
K-004	150	143
K-005	150	147
K-006	150	135

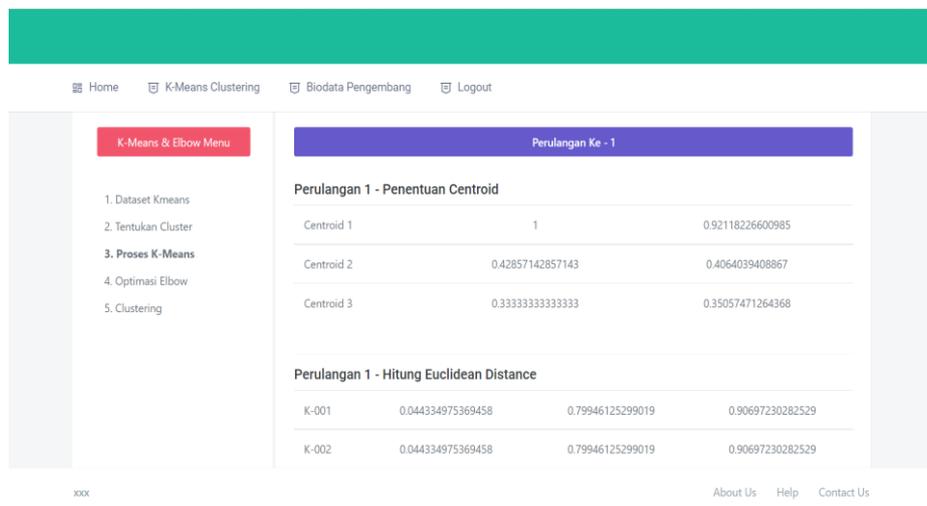
The interface also includes a navigation menu with 'Home', 'K-Means Clustering', 'Biodata Pengembang', and 'Logout'. A sidebar on the left contains a 'K-Means & Elbow Menu' with steps: 1. Dataset Kmeans, 2. Tentukan Cluster, 3. Proses K-Means, 4. Optimasi Elbow, and 5. Clustering. The footer contains 'About Us', 'Help', and 'Contact Us' links.

Gambar 3. Tampilan Data Set



kodebarang	stokawal	stokakhir
K-001	1	0.96551724137931
K-002	1	0.96551724137931
K-003	1	0.94827586206897
K-004	1	0.93103448275862
K-005	1	1
K-006	1	0.79310344827586
K-007	1	0.84482758620690

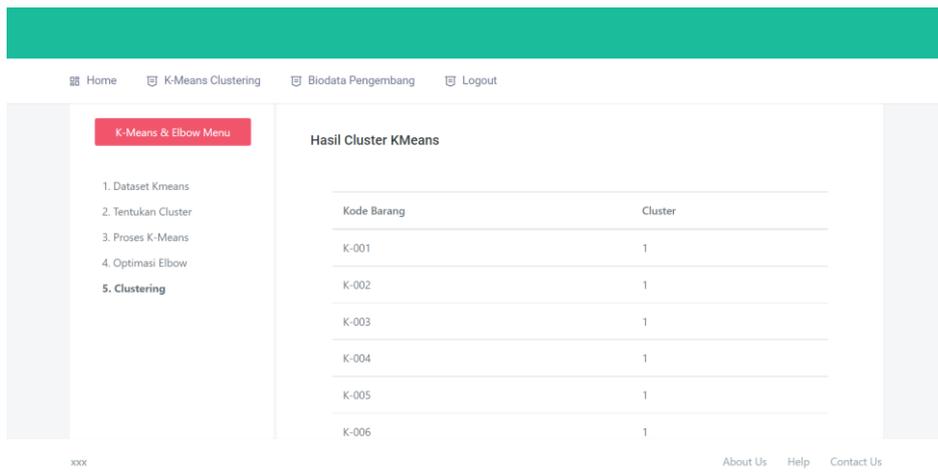
Gambar 4. Tampilan dataset normalisasi pada program.



Centroid	1	2	3
Centroid 1	1	0.92118226600985	
Centroid 2	0.42857142857143	0.4064039408867	
Centroid 3	0.33333333333333	0.35057471264368	

Kode	0.044334975369458	0.79946125299019	0.90697230282529
K-001	0.044334975369458	0.79946125299019	0.90697230282529
K-002	0.044334975369458	0.79946125299019	0.90697230282529

Gambar 5. Tampilan Proses K-Means



Kode Barang	Cluster
K-001	1
K-002	1
K-003	1
K-004	1
K-005	1
K-006	1

Gambar 6. Tampilan Hasil Clustering

4. KESIMPULAN

Dari hasil semua pembahasan tentang klasifikasi produk data mining pada toko khusus T-shirt. Dengan menggunakan algoritme K-Means, peneliti ingin menyampaikan kesimpulan sebagai berikut: Diperlukan sistem data mining dengan menggunakan metode K-Means untuk menyelesaikan permasalahan terkait klasifikasi produk pada industri kaos. Perancangan sistem klasifikasi pola penjualan produk di toko. Diperlukan aplikasi yang dapat

membangun sistem untuk jual beli kaos. Untuk menganalisis klasifikasi produk toko jual. Penelitian ini menemukan bahwa masih banyak kekurangan dalam penelitian ini yang perlu pengembang untuk mencapai kinerja yang lebih baik. Kriteria penelitian ini belum sepenuhnya sesuai untuk sistem kategorisasi untuk menentukan persediaan menggunakan K-Means clustering. Penetapan centroid awal dapat menghasilkan centroid yang berbeda, jadi disarankan untuk menggunakan meningkatkan akurasi proses clustering, penelitian dapat membandingkannya dengan algoritma lain atau mengembangkan algoritma clustering K-Means untuk membuat proses lebih akurat.

DAFTAR PUSTAKA

- [1] D. Sariapura And M. Zunaidi, “J-Sisko Tech Jurnal Teknologi Sistem Informasi Dan Sistem Komputer Tgd Analisis Data Mining Pada Strategi Penjualan Produk Pt Aquasolve Sanaria Dengan Menggunakan Metode K-Means Clustering,” Vol. 4, No. 1, Pp. 127–136, 2021.
- [2] Y. Darmi, A. Setiawan, J. Bali, K. Kampung Bali, K. Teluk Segara, And K. Bengkulu, “Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk,” 2016.
- [3] P. Studi, S. Informasi, R. Kisaran, And J. M. Yamin, “Pemanfaatan Rapid Miner Studio 8.2 Untuk Pengelompokan Data Penjualan Aksesoris Menggunakan Algoritma K-Means Mardalius,” Vol. Iv, No. 2, Pp. 123–132, 2018.
- [4] H. Samosir, M. Amin, And I. R. Harahap, “Penerapan Data Mining Untuk Klasifikasi Produk Merk Bata Menggunakan Algoritma K-Means,” *Jutsi (Jurnal Teknologi Dan Sistem Informasi)*, Vol. 2, No. 1, Pp. 161–166, Jun. 2021, Doi: 10.33330/Jutsi.V2i1.1163.
- [5] S. Parsaoran Tamba And F. Toknady Kesuma, “Penerapan Data Mining Untuk Menentukan Penjualan Sparepart Toyota Dengan Metode K-Means Clustering,” *Jurnal Sistem Informasi Ilmu Komputer Prima (Jusikom Prima)*, Vol. 2, No. 2, 2019.
- [6] G. Gustientiedina, M. H. Adiya, And Y. Desnelita, “Penerapan Algoritma K-Means Untuk Clustering Data Obat-Obatan,” *Jurnal Nasional Teknologi Dan Sistem Informasi*, Vol. 5, No. 1, Pp. 17–24, Apr. 2019, Doi: 10.25077/Teknosi.V5i1.2019.17-24.
- [7] M. Hariyanto And R. Tahara Shita, “17 Clustering Pada Data Mining Untuk Mengetahui Potensi Penyebaran Penyakit Dbd Menggunakan Metode Algoritma K-Means Dan Metode Perhitungan Jarak Euclidean Distance Clustering Pada Data Mining Untuk Mengetahui Potensi Penyebaran Penyakit Dbd Menggunakan Metode Algoritma K-Means Dan Metode Perhitungan Jarak Euclidean Distance,” 2018.
- [8] A. Triningsih And H. Supriyono, “Aplikasi Data Mining Berbasis Web Menggunakan Metode K-Means Clustering Untuk Pengelompokan Penjualan Terlaris Produk Kacamata.”
- [9] M. Dahria, R. Gunawan, And Z. Lubis, *Seminar Nasional Sains & Teknologi Informasi (Sensasi) Implementasi K-Means Untuk Pengelompokan Produk Terbaik Pt. Koko Pelli*. [Online]. Available: [Http://Prosiding.Seminar-Id.Com/Index.Php/Sensasi/Issue/Archivepage|495](http://Prosiding.Seminar-Id.Com/Index.Php/Sensasi/Issue/Archivepage|495), 2019.
- [10] I. Nawangsih And R. Puspita, “Pelita Teknologi Implementasi Algoritma K-Means Dalam Mengkategorikan Produk Terlaris Dan Kurang Laris Pada Toko Alfamart Cikarang,” *Jurnal Pelita Teknologi*, Vol. 16, No. 1, Pp. 79–87, 2021.