

PEMODELAN POLA PENJUALAN SEPATU IMPOR MELALUI PENDEKATAN ALGORITMA K-MEANS

Rosnia Harum¹, Fathurrohman², Ade Rizki Rinaldi³

¹ Teknik Informatika, STMIK IKMI Cirebon

^{2,3} Rekayasa Perangkat Lunak, STMIK IKMI Cirebon

Jl Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45135

rosniaharum25@gmail.com

ABSTRAK

Toko Myi Store merupakan toko yang bergerak dibidang penjualan Sepatu impor, adapun permasalahan yang dihadapi oleh toko myi store yaitu owner tidak tahu data penjualan sepatunya dalam jangka tiga tahun terakhir itu meningkat atau menurun dan juga memastikan ketersediaan barang, stok barang, dan kelengkapan barang agar menghindari penumpukan barang yang sama dan kurang di minati oleh konsumen. Tujuan penelitian ini untuk mengetahui produk yang diminati oleh konsumen dan bisa mendapatkan analisis data penjualan yang paling laku, laku dan kurang laku untuk dijual. Hasil dari penelitian ini melakukan 4 kali percobaan dengan k yang berbeda, dan diketahui dari 4 kali percobaan, nilai k yang paling optimal terdapat pada $k = 4$ dengan hasil DBI nya adalah 0.189, Cluster yang dihasilkan memperoleh sebanyak 4 Cluster Setiap klaster memiliki karakteristiknya sendiri dalam tingkat penjualan yang berbeda-beda tergantung merek sepatu yang dijual. Klaster tersebut menunjukkan sejumlah pola yang paling banyak diminati oleh pembeli *myi store* terhadap sepatu yang dijual, Hasil dari cluster Sepatu yang paling laku adalah Converse dan Nike, sepatu yang laku yaitu Adidas dan New Balance, sedangkan sepatu yang kurang laku dan kurang diminati adalah Puma. Penelitian ini diharapkan dapat memberi wawasan yang lebih luas bagi industry sepatu untuk mengembangkan strategi pemasaran yang lebih terfokus dan efektif.

Kata Kunci : Data mining, Algoritma K-Means Clustering, Sepatu Impor

1. PENDAHULUAN

Industri sepatu adalah sektor yang besar dan berkembang pesat di seluruh dunia. Produksi sepatu terjadi di banyak negara, dengan berbagai tingkat teknologi, biaya tenaga kerja, dan standar kualitas. Negara-negara seperti Tiongkok, Vietnam, Indonesia, dan Italia adalah produsen sepatu utama di dunia. Globalisasi telah memfasilitasi perdagangan internasional yang lebih besar dalam industri sepatu. Perusahaan sepatu dapat dengan mudah mengakses pasar global untuk memasarkan produk mereka dan memperluas jangkauan mereka. Ini juga memungkinkan konsumen untuk memiliki akses lebih besar terhadap sepatu dari berbagai negara. Perkembangan teknologi juga mempengaruhi industri sepatu, termasuk produksi, desain, dan pemasaran. Teknologi baru memungkinkan produsen untuk meningkatkan efisiensi produksi, mengurangi biaya, dan menghasilkan produk yang lebih inovatif dan berkualitas tinggi, yang dapat menarik minat konsumen untuk membeli sepatu impor.

Toko Myi Store adalah sebuah Toko yang bergerak dibidang retail khususnya seperti sepatu. Dengan adanya suatu data kita dapat melakukan analisis terhadap penjualan, data penjualan tersebut kita manfaatkan untuk mengetahui barang yang paling laku, laku dan kurang laku. Ketersediaan barang, stock barang dan kelengkapan barang pada suatu toko adalah elemen yang sangat penting. Sehingga proses manajemen untuk mengatur ketersediaan persediaan barang sangat diperlukan

untuk menghindari penumpukan barang yang sama dan kurang diminati oleh konsumen.

Permasalahan yang dihadapi toko sepatu *myi store* yaitu owner tidak tahu data penjualan sepatunya dalam jangka tiga tahun terakhir itu meningkat atau menurun dan juga memastikan ketersediaan barang, stok barang, dan kelengkapan barang agar menghindari penumpukan barang yang sama dan kurang di minati oleh konsumen karena belum adanya cluster yang bisa mengelompokkan berapa jumlah Sepatu yang telah terjual. Pihak toko *Myi Store* mengelola data penjualan secara manual, sehingga menyebabkan pengelolaan data menjadi tidak efisien. Selain itu sering terjadi ketidakakuratan dalam menghitung jumlah karena masih belum diclusterkan.

Penelitian ini menggunakan teknik klasifikasi dengan K-Means untuk mengelompokkan data kategorikal hingga menghasilkan klaster yang lebih stabil. Algoritma K-means Clustering adalah suatu metode penganalisaan data atau metode data mining yang melakukan proses pemodelan tanpa supervisi (unsupervised) dan merupakan salah satu metode yang melakukan pengelompokan data dengan sistem partisi. Metode K-Means berusaha mengelompokkan data yang ada kedalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik.

Penelitian ini bertujuan untuk mengetahui produk yang diminati oleh konsumen dan menentukan *clustering* dari penjualan sepatu impor dan bisa mendapatkan data penjualan yang paling

laku, laku dan kurang laku untuk dijual dengan menggunakan metode algoritma k-means clustering.

Hasil penelitian ini dapat diharapkan memberi wawasan yang lebih luas bagi industry sepatu untuk mengembangkan strategi pemasaran yang lebih terfokus dan efektif. disimpulkan bahwa algoritma K-Means menjadi salah satu metode yang cocok digunakan untuk mengelompokan data penjualan sepatu import dari penjualan paling laku, laku dan kurang laku Dengan mengetahui pola pembelian konsumen perusahaan dapat mengoptimalkan tingkat persediaan dan produksi, ini dapat mengurangi resiko kelebihan stok atau kekurangan stok sehingga meningkatkan efesiensi operasional.

2. TINJAUAN PUSTAKA

2.1. Literature Review

Penelitian yang berjudul Implementasi Data Mining Untuk Memprediksi Kelulusan Siswa SMK Al Huda Kedungwungu Dengan Menggunakan Algoritma Naïve Bayes Classifie penerapan metode Naive Bayes menerapkan data latih untuk menghasilkan probabilitas setiap kriteria untuk kelas yang berbeda, sehingga nilai probabilitas dari kriteria tersebut dapat dioptimalkan untuk menentukan prediksi kelulusan siswa secara cepat dan efisien berdasarkan klasifikasi dilakukan dengan metode Naive Bayes, kemudian dari hasil pengujian dengan metode Naive Bayes didapatkan hasil nilai akurasi sebesar 76,25%, sehingga hasil ini memiliki akurasi yang sangat baik. Dengan begitu metode ini dapat diterapkan dalam memprediksi kelulusan siswa s[1].

Penelitian yang berjudul Sistem Informasi Gudang Berbasis Web Untuk Penyimpanan Barang Di PT Mitra Sukses Bangun Bersama Metode yang digunakan dalam penelitian ini adalah dengan melakukan studi literatur untuk memperoleh pengetahuan dan konsep dasar mengenai sistem informasi gudang dan pengembangan web. Selanjutnya, dilakukan analisis kebutuhan sistem dan desain sistem informasi gudang berbasis web. Kemudian, dilakukan implementasi sistem dan pengujian fungsionalitas sistem. Berdasarkan penelitian yang sudah dilakukan berhasil membuat sistem informasi gudang untuk penyimpanan yang dikembangkan menggunakan metode SDLC waterfall [2]

Penelitian yang berjudul Jaringan Neural Konvolusional untuk Motif Klasifikasi dan Pengaruh Dimensi Gambar metode konvolusional jaringan saraf (CNN). Hasil penelitian ini menunjukkan bahwa Multi-Layer Perceptron (MLP) merupakan deep yang sangat disukai metode pembelajaran untuk klasifikasi data, terutama pada domain yang melibatkan klasifikasi gambar. Ukuran dari gambar yang digunakan mempengaruhi keakuratan algoritma jaringan saraf komputasi (CNN). Hasilnya menunjukkan bahwa pengujian dengan menggunakan perbandingan data latih sebesar 60%, 30% dan 10%

menghasilkan kerugian sebesar 01,89% sebesar 1,18% dan Peningkatan akurasi 100% [3].

Penelitian yang berjudul Pembelajaran Mendalam Untuk Klasifikasi Sindrom Ovarium Polikistik Menggunakan Jaringan Syaraf Tahap penelitian mempunyai 3 tahap diantaranya tahap pertama mengidentifikasi masalah dan mengumpulkan dataset dari dataverse Telkom University berupa gambar dan tinjauan pustaka yang beragam sumber. Tahap kedua adalah Pra Pemrosesan data gambar, Pelatihan Data, desain pemodelan dengan pengelolaan gambar data dan klasifikasi menggunakan model deep learning Algoritma Convolutional Neural Network dan pengujian. Itu tahap ketiga adalah evaluasi hasil pengujian dan pembahasan hasil ketelitian dalam menentukan status Sindrom Ovarium Polikistik Normal atau PCOS. Hasil pelatihan dan validasi pada gambar xray ovarium dataset menggunakan arsitektur CNN yang telah dibuat, 40 iterasi (epoch), dan 4 step_per_epochs menunjukkan nilai akurasi sebesar 0,8947 atau 89,47% dan nilai loss sebesar 0,2684 [4]

Penelitian yang berjudul Rancangan Bangun Alat Pakan Ikan otomatis Menggunakan Mikrokontroler Esp8288 Dengan Berbasis Internet Of Things Budidaya ikan Teknologi yang akan diterapkan adalah alat pakan ikan otomatis yang dapat diatur jadwal pemberian pakan ikannya secara otomatis dan juga dapat memonitoring kondisi pakan melalui aplikasi Blynk pada smartphone. Komponen elektronika yang digunakan adalah ESP8266 sebagai pengolah data, Motor Servo MG996R sebagai penggerak keran pakan, Motor Servo SG90 sebagai penggerak Potensiometer, Potensiometer sebagai kontrol aktifnya Blower, Blower sebagai penebat pakan ikan, Sensor Ultrasonik HY-SRF05 sebagai monitoring kondisi pakan, Power Supply sebagai input tegangan 12 V, Buzzer sebagai alarm notifikasi dan aplikasi Blynk pada smartphone. Hasil pengujian menunjukkan alat yang dibangun dapat berfungsi dengan baik, pengaturan jadwal pakan dan monitoring pakan bekerja normal sesuai perintah yang telah diatur pada aplikasi Blynk [5]

Penelitian berjudul Analisa Kanker Paru Paru Dengan Menggunakan Algoritma K-Nearest Neighbor metode klasifikasi untuk kanker paru-paru sangat penting dalam upaya meningkatkan deteksi dini dan tepat perlakuan. Tahapan penelitian analisis kanker paru-paru menggunakan algoritma k-Nearest Neighbor (k-NN). Itu tahap awal dalam penelitian (Abdi Mangun et al., 2023) ini adalah pengumpulan data yang berkaitan dengan kanker paru-paru. Setelah data terkumpul maka langkah selanjutnya adalah memproses data terlebih dahulu untuk membersihkannya. Memilih nilai k, yaitu jumlah tetangga terdekat, merupakan hal yang krusial langkah dalam analisis kanker paru-paru. Begitu nilai knya adalah dipilih, model k-NN dilatih menggunakan pelatihan data untuk mempelajari hubungan antara atribut dan status kanker paru-paru.

Hasil penelitian ini adalah akurasi yang diperoleh sebesar 80,40 [6]

Penelitian dengan judul Rancangan Bangun Sistem Informasi Akuntansi Menggunakan Codeignter Dengan Metode Scrum Studi Kasus PT Surya Marga Sarana. Metode pengembangan yang digunakan dalam membuat sistem ini adalah metode Scrum yang memiliki keunggulan untuk merespon perubahan dengan cepat. Dengan tahapan yaitu, product backlog, sprint planning, daily scrum, sprint review, serta sprint retrospective. Hasil penelitian menunjukkan bahwa sistem informasi akuntansi yang telah dikembangkan di PT Surya Marga Sarana mampu menghasilkan informasi keuangan dengan lebih cepat dan efektif, serta data yang dihasilkan sesuai dengan kebutuhan perusahaan. Dari penilaian kuesioner, tingkat keefektifan dan kemudahan sistem informasi akuntansi PT Surya Marga Sarana mendapatkan nilai rata-rata sebesar 72%, yang menunjukkan bahwa sistem tersebut layak digunakan [7].

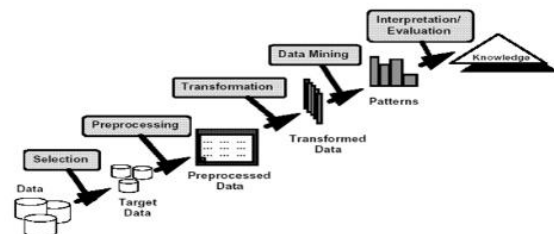
Penelitian yang dengan judul Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Yana Sport penelitian ini menetapkan 2 kelompok yaitu laris terjual dan tidak laris terjual. Hasil Data Performance menjelaskan bahwa Cluster 0 dengan nilai 110 dikategorikan sebagai Tidak Laris Terjual sedangkan 21389 dikategorikan sebagai laris terjual. Hasil rekomendasi penelitian ini mendapatkan informasi atau pola dari penerapan algoritma k-means dengan data penjualan terdapat sebanyak 99 item barang yang laris terjual dan terdapat 23 item barang yang tidak terjual sehingga pemilik dapat melakukan strategi penjualan dan pembelian ulang berdasarkan barang yang laris terjual [8]

Penelitian dengan judul Analisa Pengelompokan Dataset Komputer Menggunakan Algoritma X-Means Salah satu metode clustering yang dapat digunakan untuk mengelompokkan data adalah X-Means clustering. Hasil penelitian dari data mining pengelompokan dataset komputer dengan x-means Cluster model yang dihasilkan sebanyak 3 cluster, yaitu cluster 0 dengan anggota sebanyak 710 item, cluster 1 dengan anggota sebanyak 136 item dan cluster 2 sebanyak 15 item. (Khaerullah et al., 2023)

Penelitian dengan judul Penerapan Metode K-Means Pada Data Penjualan Untuk Mendapatkan Produk Terlaris Di PT. Titian Nusantara Boga Metode CRISP-DM yang melalui proses business understanding, data understanding, data preparation, modelling, evaluation, dan deployment. Pengujian ini dilakukan menggunakan aplikasi RapidMiner 9.10 dengan jumlah 3 atribut dan 793 record data kemudian berhasil membentuk 3 cluster, yaitu cluster pertama mendapatkan hasil 128 produk, cluster kedua mendapatkan 622 produk dan cluster ketiga mendapatkan 43 produk. Kemudian dilakukan uji validasi dengan Davies Bouldin Index dan mendapatkan nilai DBI pada cluster pertama 0,679,

cluster kedua 0,816 dan cluster ketiga 0,837. Semakin kecil nilai DBI yang dihasilkan maka kualitasnya semakin baik, maka dari nilai DBI yang dihasilkan dapat diketahui cluster pertama masuk ke dalam kategori sangat laris, yaitu dengan 128 produk dan nilai DBI 0,679. (Wahyudi et al., 2023)

3. METODE PENELITIAN



Gambar 1. Tahapan Metode KDD (Knowledge Discovery in Databases)

3.1. Selection Data

Tahapan yang pertama yaitu data *selection*. Pada tahap ini memasukkan operator *Read Excel* yang berfungsi untuk membaca data *excel* Penjualan Sepatu Import di toko myi store. Selanjutnya yaitu *import* data. Seperti pada gambar 2 Operator *Read Excel*.



Gambar 2. Operator Read Excel

Parameter operator Read Excel menggunakan parameter default. Data berhasil di import dengan Read excel menampilkan data preview. Berikut hasil Gambar 3 data Read Excel.

Tahun integer	Bulan polynomial id	Jenis Sepatu polynomial label	Merek Sepatu polynomial weight	Jumlah Terjual integer
1 2021	Januari	Formal	Converes	10
2 2021	Februari	Formal	NewBalance	11
3 2021	Maret	Formal	Addass	27
4 2021	April	Formal	Nike	11
5 2021	Mei	Formal	Puma	16
6 2021	Juni	Formal	Vans	23
7 2021	Juli	Formal	Converes	35
8 2021	Agustus	Formal	NewBalance	13
9 2021	September	Formal	Addass	19
0 2021	Oktober	Formal	Nike	11
1 2021	November	Formal	Puma	18
2 2021	Desember	Formal	Vans	22

no problems

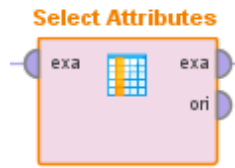
← Previous

 Finish

 Cancel

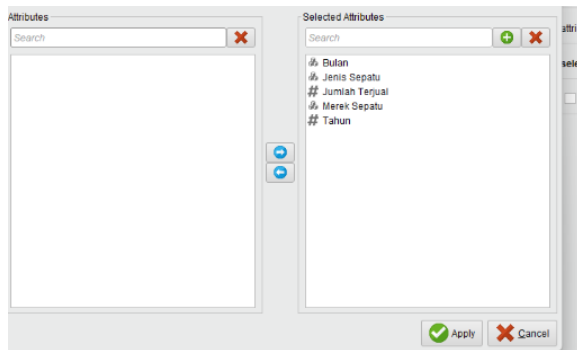
Gambar 3. Data Read excel

Langkah selanjutnya adalah masuk ke operator Select Attributes yang memiliki fitur mana yang akan dipilih atau atribut mana yang akan dianalisis. Berikut Gambar 3 Operator Select Atributes.



Gambar 4. Operator Select Attributes

Dari gambar 4 operator *Select Attributes* diatas di bagian parameter pilih *attribute filter type* dengan pilihan *a subset* dan lanjut untuk *Select Attributes* seperti yang ditampilkan pada gambar 5 dibawah ini.



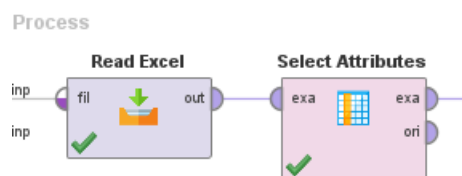
Gambar 5. Select Attributes

Pada gambar 5 diatas atribut yang akan digunakan dan tidak digunakan ada pada tabel 1 dibawah ini

Tabel 1. Hasil Operator *Select Attributes*

Atribut Yang Digunakan	Atribut Yang Tidak Digunakan
Tahun	
Jenis Sepatu	
Jumlah Terjual	
Merek Sepatu	
Bulan	
Terjual	

Dari total 5 atribut yaitu Jenis sepatu, jumlah terjual, Merek sepatu, Bulan dan Tahun. dipilih menjadi 4 atribut yaitu jenis sepatu, jumlah terjual, harga, dan merek sepatu. Pemilihan atribut ini didasarkan pada relevansi informasi terkait dengan analisis yang ingin dilakukan. Kombinasi dari keempat atribut ini dapat memberikan wawasan yang mendalam tentang penjualan produk sepatu dalam suatu konteks tertentu. yang tertera pada tabel 1 diatas.

Gambar 6. Model proses langkah *Selection* di *Rapidminer*

Parameter pada operator *replace Missing Values* menggunakan parameter bawaan, setelah dilakukan pembersihan data dengan operator *replace missing value* maka hasilnya seperti pada gambar 3.6 yaitu tidak menyebabkan perubahan pada data dan masih tetap ada. sama, 37 data dari tahun 2021-2023 karena data yang digunakan sudah lengkap, tidak ada nilai yang kosong, tumpang tindih atau hilang seperti terlihat pada Gambar 3.5 *missing values*.

3.2. Preprocessing

Pada tahap preprocessing, data disiapkan sebelum analisis dengan membersihkan data yang kosong, menghapus duplikat, dan menangani nilai yang hilang. Salah satu langkah umum adalah menggunakan operator "Replace Missing Values" untuk mengisi atau mengganti nilai yang hilang dengan nilai valid seperti mean, median, atau modus. Hal ini bertujuan untuk memastikan bahwa data yang digunakan bebas dari nilai yang hilang agar analisis menjadi lebih akurat. Gambar 3.7 menunjukkan contoh dari operator Replace Missing Values.

Gambar 7. Operator *Missing Value*

Setelah dilakukan pembersihan data oleh operator Replace Missing Values memiliki hasil seperti pada gambar 7 yang menghasilkan tidak adanya perubahan data dan masih tetap sama berjumlah 37 data dari Januari tahun 2022 sampai Desember tahun 2022 , dikarenakan data yang digunakan merupakan data lengkap tidak ada data yang kosong, ataupun missing seperti pada gambar 8

Name	Type	Missing	Statistics	Filter (Attributes)	Default Value
Tanggal	Date	0	Earliest date: Jan 1, 2022 Latest date: Dec 31, 2022 Duration: 364 days		
Jenis Sepatu	Nominal	0	Earliest: Boots (337) Latest: Pantini (375) Most: Pantini (375), Sneakers (319), (31 more)		
Merek Sepatu	Nominal	0	Earliest: Nike (242) Latest: Adidas (315) Most: Adidas (315), Converse (273), (31 more)		
Jumlah Terjual	Integer	0	Min: 3 Max: 15 Average: 6.100		

Gambar 8. Operator *Replace Missing Values*

a.) Data Transformation

Pada tahap transformasi data, atribut yang telah dipilih diubah menjadi bentuk yang cocok untuk digunakan dalam algoritme tertentu seperti K-Means. Salah satu tahap dalam proses ini adalah menggunakan operator "Nominal to Numerical" untuk mengubah atribut non-numerik menjadi tipe numerik. Ini penting karena algoritme K-Means hanya dapat memproses data numerik. Operator

tersebut mengubah atribut non-numerik menjadi representasi numerik, seperti menggunakan one-hot encoding atau pengubahan label menjadi nilai numerik sesuai dengan skema yang ditetapkan. Dengan transformasi ini, atribut yang awalnya non-numerik menjadi numerik, memungkinkan analisis klusterisasi oleh algoritme K-Means.



Gambar 9. Nominal to Numerical

Parameter pada operator *Nominal to Numerical* hanya mengganti coding type unique integers.

b.) Data Mining

Tahap selanjutnya memasukkan operator *Clustering* yang berfungsi untuk menentukan Cluster atau pengelompokan data. Penelitian ini menggunakan algoritme *K-Means* dan menggunakan 4 Cluster max perulangan 10. Seperti pada gambar 9 operator *clustering* menggunakan *k-means*.



Gambar 10. Operator Clustering menggunakan K-Means

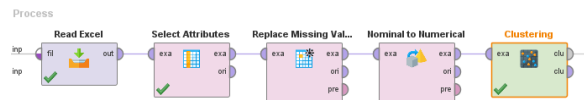
Parameter pada operator *Clustering K-Means* dimana diketahui $K = 3$, $\text{max run} = 10$ kali putaran, tipe *measure = Numerical Measure* yang *Euclidean Distance* dan $\text{max optimal steps} = 100$. Hasil operator *Clustering K-Means* diperoleh 4 Cluster yaitu Cluster 0 sebanyak 9 items, Cluster 1 sebanyak 23 items, Cluster 2 sebanyak 1 items, Cluster 3 sebanyak 3 items dengan total items sebanyak 36 items. Seperti pada gambar 11 Cluster Model

Cluster Model

```
Cluster 0: 9 items
Cluster 1: 23 items
Cluster 2: 1 items
Cluster 3: 3 items
Total number of items: 36
```

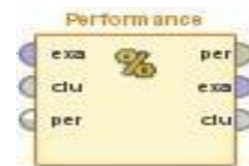
Gambar 11. Cluster Model

Model proses pada *rapidminer* dilangkah *Data Mining* tampak pada gambar 3. 12 dibawah ini.



Gambar 13. Model proses langkah Data Mining di Rapidminer

Setelah melakukan klusterisasi dengan K-Means, langkah selanjutnya adalah mengevaluasi kualitas kluster menggunakan Davies Bouldin Index (DBI). DBI mengukur jarak antar kluster, semakin mendekati 0 semakin baik. Operator ini membantu menentukan nilai k yang optimal dengan melakukan evaluasi DBI pada berbagai nilai k . Hasil evaluasi ini memberikan pemahaman tentang seberapa baik klusterisasi yang dilakukan oleh algoritme K-Means.



Gambar 14. operator Performance

Hasil dari nilai pada *Davies Bouldin Index* yang di dapat dari penelitian ini dan yang hasilnya paling mendekati 0 yaitu 0.189, menggunakan nilai $k = 4$ dengan maksimal perulangan 10 kali seperti gambar 4.1 hasil operator *performance*

Tabel 2. Hasil Nilai DBI

No	Cluster	Nilai Bouldin Indeks
1	K2	0,435
2	K3	0,210
3	K4	0,189
4	K5	0,000

3.3. Sumber Data

Sumber data yang di gunakan merupakan data yang di peroleh dari pihak toko Myi Store yang beralamat di JL.Telagasari Desa Tempel Kulon Blok Ci Rt.10 Rw.01 Kec.Lelea Kab.Indramayu Jawa Barat 45261 dan waktu pngambilan data tersebut pada bulan oktober 2023 yang telah melalui proses perizinan dari pihak kampus ke pihak Toko sepatu Myi Store tersebut, kemudian data yang di gunakan adalah data penjualan sepatu pada tahun 2021-2023 yaitu data yang berisi 37 record dan 5 atribut yaitu tahun, bulan, jenis sepatu, merek sepatu dan jumlah terjual, di bawah ini sample data penjualan sepatu sebelum melakukan pengolahan data.

Teknik pengumpulan data merupakan bagian terpenting dalam kegiatan menyusun tugas akhir, ketersediaan data akan sangat menentukan dalam proses pengimpelemntasian data. Jenis data yang digunakan ialah data primer teknik pengumpulan data yang di lakukan yaitu: Observasi, Pengamatan langsung maupun tidak langsung dalam proses menyusun tugas akhir dan dalam proses kegiatan yang ada untuk menghasilkan data penunjang.

3.4. Teknik Pengumpulan Data

Teknik Pengumpulan Data yang dilakukan yaitu observasi, Observasi adalah metode pengumpulan data primer dengan melakukan pengamatan terhadap aktivitas dan kejadian tertentu yang terjadi. Penelitian langsung datang ketempat usaha myi store untuk mengamati aktivitas yang terjadi pada toko tersebut untuk mendapatkan data atau informasi yang sesuai dengan apa yang dilihat dan sesuai dengan kenyataannya [9]

3.5. Teknik Analisis Data

K-means merupakan salah satu metode pengelompokan data nonhierarki (sekatan) yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan kedalam kelompok yang lain. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diatur dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok [5].

Langkah-langkah melakukan clustering dengan metode K-means adalah sebagai berikut :

- Pilih jumlah cluster k
- Inisialisasi ke pusat cluster ini bisa dilakukan dengan berbagai cara. Cara yang paling sering dilakukan adalah dengan random atau acak.[10]

Langkah-langkah dalam melakukan clustering dengan metode K-Means adalah sebagai berikut:

- Tentukan jumlah Kluster k
- Menentukan titik pusat awal dari setiap Kluster
- Alokasikan semua data/objek ke kluster terdekat, kedekatan kedua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga dekatkan suatu data ke kluster tertentu ditentukan jarak antara data dengan pusat cluster. Dalam tahap ini perlu dihitung jarak tiap data ke tiap cluster. Jarak antara satu data dengan satu kluster tertentu akan menentukan suatu data masuk dalam kluster mana untuk menghitung jarak semua data ke setiap titik pusat kluster dapat menggunakan teori jarak centroid yang dirumuskan sebagai berikut :
- Hitung kembali pusat kluster dengan keanggotaan cluster yang sekarang. Pusat kluster adalah rata-rata dari semua data/objek

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

dalam cluster tertentu.

4. HASIL DAN PEMBAHASAN

4.1. Hasil nilai Davies Bouldin Index(DBi) yang optimal dari penerapan algoritma K-Means terhadap Data Penjualan Sepatu di toko Myi Store

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: 1.621
Avg. within centroid distance_cluster_0: 1.235
Avg. within centroid distance_cluster_1: 0.000
Avg. within centroid distance_cluster_2: 4.667
Avg. within centroid distance_cluster_3: 1.889
Avg. within centroid distance_cluster_4: 1.389
Davies Bouldin: 0.507
```

Gambar 15. Hasil operator Performance

Penerapan algoritma K-Means pada data penjualan dan pengelompokan produk sepatu di Toko Myi Store menghasilkan nilai DBI optimal 0.189 untuk K=4, menunjukkan pemisahan kelompok produk yang baik dengan karakteristik yang berbeda secara signifikan. Ini menegaskan keberhasilan algoritma K-Means dalam mengidentifikasi pola pembelian yang konsisten dan berbeda di antara kelompok produk sepatu. Hasil ini memberikan keyakinan bahwa segmentasi produk sepatu berbasis K-Means dapat dipercaya untuk pengambilan keputusan strategis terkait pemasaran, pengelolaan stok, dan pengembangan produk di Toko Myi Store.

4.2. Hasil nilai cluster yang optimal dari penerapan algoritma K-Means terhadap Data Penjualan sepatu di Myi Store Dengan Penerapan Algoritma K-Means

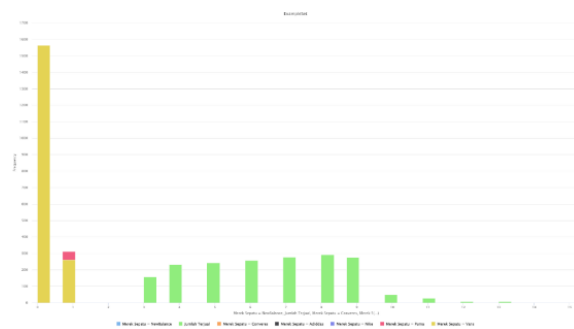
Row No.	Jenis Sepatu	label	Merck Sepa...	Merck Sepa...	Merck Sepa...	Merck Sepa...	Merck Sepa...	Harga	Jumlah Ter...
1009	Boots	cluster_1	1	0	0	0	0	300000	7
1010	Sneakers	cluster_0	0	1	0	0	0	250000	7
1011	Formal	cluster_3	0	0	1	0	0	200000	9
1012	Sport	cluster_3	0	0	0	1	0	180000	5
1013	Casual	cluster_2	0	0	0	0	1	150000	8
1014	Boots	cluster_1	0	0	0	0	1	300000	4
1015	Sneakers	cluster_0	1	0	0	0	0	250000	4
1016	Formal	cluster_3	0	1	0	0	0	200000	6
1017	Sport	cluster_3	0	0	1	0	0	180000	8
1018	Casual	cluster_2	0	0	0	1	0	150000	9
1019	Boots	cluster_1	0	0	0	0	1	300000	7
1020	Sneakers	cluster_0	0	0	0	0	1	250000	8
1021	Formal	cluster_3	1	0	0	0	0	200000	4
1022	Sport	cluster_3	0	1	0	0	0	180000	6
1023	Casual	cluster_2	0	0	1	0	0	150000	5

Gambar 16. Hasil dari nilai cluster

Cluster yang dihasilkan memperoleh sebanyak 4 Cluster Setiap kluster memiliki karakteristiknya sendiri dalam tingkat penjualan yang berbeda-beda tergantung merek sepatu yang dijual. Kluster tersebut menunjukkan sejumlah pola yang paling banyak diminati oleh pembeli myi store terhadap sepatu yang dijual Hasil dari cluster Sepatu yang paling laku adalah Converse dan Nike Sepatu yang laku sedang Adidas dan New Balance sedangkan sepatu yang kurang laku diminati adalah Puma dan vans.

4.3. Pembahasan

Dari hasil analisis data penjualan dan pengelompokan produk sepatu di Myi Store dengan algoritma k-means, algoritma K-Means dapat membantu mengidentifikasi tren konsumen terhadap produk sepatu di Myi Store dengan cara Segmentasi konsumen, Algoritma k-means bisa digunakan untuk menentukan clustering produk sepatu dengan langkah langkah pemilihan fitur, pemrosesan data, penentuan jumlah cluster, inisialisasi pusat cluster, penetapan, dan evaluasi cluster. dan hasil analisis penjualan dan pengelompokan produk kopi dapat memberikan wawasan strategis untuk meningkatkan kinerja bisnis dan profitabilitas Toko Sepatu Myi Store adalah dengan cara pemahaman konsumen, penyesuaian harga, strategi pemasaran yang lebih tepat dan efisiensi.



Gambar 17. Jumlah Data Penjualan Sepatu

Hasil penelitian dari pengelompokan ini dapat digunakan untuk menganalisis jumlah penjualan sepatu berdasarkan Cluster. Dapat diketahui, jumlah Penjualan paling tinggi pada Cluster 2, kemudian jumlah penjualan persatsi sedang pada Cluster 0 dan Cluster 3, kemudian jumlah persatsi terendah terdapat pada Cluster 1.

Tabel 3. Cluster 0

Tahun	Bulan	Merek Sepatu	Terjual
2021	Mei	Puma	16
2021	September	Adidas	19
2021	November	Puma	18
2022	Januari	Converes	19
2022	Februari	New Balance	20
2022	Maret	Adidas	15
2022	Desember	Vans	17
2023	Februari	New Balance	18
2023	Oktober	Nike	16

Dari tabel 3 diatas dengan cluster 0 penjualan sepatu dengan presentase penjualan sedang.

Tabel 4. Cluster 1

Tahun	Bulan	Merek Sepatu	Terjual
2021	Januari	Converes	10
2021	Februari	New Balance	11
2021	April	Nike	11
2021	Agustus	New Balance	13
2021	Oktober	Nike	11

Tahun	Bulan	Merek Sepatu	Terjual
2022	April	Nike	10
2022	Mei	Puma	11
2022	Juni	Vans	11
2022	Juli	Converes	14
2022	Agustus	New Balance	10
2022	September	Adidas	9
2022	Oktober	Nike	10
2022	November	Puma	10
2023	Januari	Converes	12
2023	Maret	Adidas	10
2023	April	Nike	10
2023	Mei	Puma	12
2023	Juni	Vans	10
2023	Juli	Converes	11
2023	Agustus	New Balance	10
2023	September	Adidas	14
2023	November	Puma	10
2023	Desember	Vans	12

Dari tabel 4 diatas dengan cluster 1 penjualan sepatu dengan presentase penjualan paling rendah.

Tabel 5. Cluster 2

Tahun	Bulan	Merek Sepatu	Terjual
2021	Juli	Converes	35

Dari tabel 5 diatas dengan cluster 2 penjualan sepatu dengan presentase penjualan paling tinggi.

Tabel 6. Cluster 3

Tahun	Bulan	Merek Sepatu	Terjual
2021	Maret	Adidas	27
2021	Juni	Vans	23
2021	Desember	Vans	22

Dari tabel 6 diatas dengan cluster 3 penjualan sepatu dengan presentase penjualan sedang.

5. KESIMPULAN DAN SARAN

Kesimpulannya, berdasarkan hasil analisis dan pembahasan penerapan algoritma klasterisasi k-means untuk mengelompokkan penjualan sepatu impor di Myi Store, yang bertujuan untuk membantu toko meningkatkan penjualan dan manajemen stok sesuai dengan preferensi konsumen, penulis dapat menyimpulkan: Penelitian ini melakukan 4 kali percobaan dengan k yang berbeda. Dan diketahui dari 4 kali percobaan, nilai k yang paling optimal terdapat pada k = 4 dengan hasil DBI nya adalah 0. 189. Algoritma menghasilkan empat klaster, masing-masing memiliki karakteristik penjualan yang berbeda tergantung pada merek sepatu yang dijual. Klaster tersebut menunjukkan pola preferensi pembeli terhadap sepatu yang dijual di Myi Store. Sepatu yang paling dominan di klaster adalah Converse dan Nike. Sepatu yang kurang laku dan kurang diminati adalah Puma. Hasil klasterisasi membantu Myi Store menyediakan stok barang yang dibutuhkan dan sesuai dengan minat konsumen.

Saran dari penerapan algoritma klasterisasi k-means untuk penjualan di Myi Store masih memiliki kelemahan, sehingga penulis menyarankan: Analisis Klaster Mendalam; Investigasi karakteristik tiap klaster, termasuk faktor lain seperti penjualan sepatu dan produk lainnya di *toko myi store*. Eksplorasi Variabel Lain; Tambahkan variasi jenis dan merek sepatu untuk meningkatkan penjualan di *Myi Store*. Penggunaan Algoritma K-Means; Terapkan pada pasar internasional, tradisional, dan supermarket untuk meningkatkan penjualan, manajemen stok, dan mengurangi penumpukan barang yang kurang diminati.

DAFTAR PUSTAKA

- [1] P. Putriyana dan O. Nurdiawan, "Implementasi Data Mining Untuk Memprediksi Kelulusan Siswa SMK Al Huda Kedungwungu Dengan Menggunakan Algoritma Naïve Bayes Classifier," *Exp. Student Exp.*, vol. 1, no. 1, hal. 1–7, 2023.
- [2] F. Baihaqqi, N. Suarna, dan O. Nurdiawan, "Sistem Informasi Gudang Berbasis Web Untuk Penyimpanan Barang Di Pt Mitra Sukses Bangun Bersama," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 2, hal. 1204–1211, 2023, doi: 10.36040/jati.v7i2.7317.
- [3] S. Aisyah, R. Astuti, F. M. Basysyar, dan ..., "Convolutional Neural Networks for Classification Motives and the Effect of Image Dimensions," ... (*Rekayasa Sist. dan ...*, vol. 5, no. 158, hal. 181–188, 2024, [Daring]. Tersedia pada: <http://www.jurnal.iaii.or.id/index.php/RESTI/article/view/5623%0Ahttps://www.jurnal.iaii.or.id/index.php/RESTI/article/download/5623/910>
- [4] O. Nurdiawan, H. Susana, dan A. Faqih, "Deep Learning for Polycystic Ovarian Syndrome Classification Using Convolutional Neural Network," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 9, no. 2, hal. 218–226, 2024, doi: 10.33480/jitk.v9i2.4575.
- [5] R. Z. Anzary, D. A. Kurnia, dan O. Nurdiawan, "RANCANG BANGUN ALAT PAKAN IKAN OTOMATIS MENGGUNAKAN MIKROKONTROLER ESP8266 DENGAN BERBASIS INTERNET OF THINGS," vol. 10, hal. 53–60, 2024.
- [6] T. Abdi Mangun, O. Nurdiawan, dan A. Irma Purnamasari, "Lung Cancer Analysis Using K-Nearst Neighbor Algorithm," *J. Tek. Ind. Sist. Inf. dan Tek. Inform.*, vol. 2, no. 2, hal. 58–61, 2023, [Daring]. Tersedia pada: https://ejournal.ubibanyuwangi.ac.id/index.php/jurnal_tinsika
- [7] I. Heriyawan, U. Hayati, dan O. Nurdiawan, "Rancang Bangun Sistem Informasi Akuntansi Menggunakan Codeigniter Dengan Metode Scrum," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 2, hal. 1236–1241, 2023, doi: 10.36040/jati.v7i2.6597.
- [8] D. Anggarwati, O. Nurdiawan, I. Ali, dan D. A. Kurnia, "Penerapan Algoritma K-Means Dalam Prediksi Penjualan Karoseri," *J. Data Sci. Inform.*, vol. 1, no. 2, hal. 58–62, 2021.
- [9] R. Adhitama, A. Burhanuddin, dan R. Ananda, "Penentuan Jumlah Cluster Ideal Smk Di Jawa Tengah Dengan Metode X-Means Clustering Dan K-Means Clustering Determining Vocational Ideal Cluster Number in Central Java With X-Means Clustering and K-Means Clustering Methods," *J. Inform. dan Komputer) Akreditasi KEMENRISTEKDIKTI*, vol. 3, no. 1, hal. 1–5, 2020, doi: 10.33387/jiko.
- [10] V. Miralda, M. Zarlis, dan E. Irawan, "Penerapan Metode K-Means Clustering Untuk Daging Ayam Buras," *Build. Informatics, Technol. Sci.*, vol. 2, no. 2, hal. 91–98, 2020, doi: 10.47065/bits.v2i2.493.