PENERAPAN ALGORITMA *K-MEANS CLUSTERING* UNTUK PERENCANAAN PERSEDIAAN STOK SEPATU DI TOKO DIOMCLOTHING

Fitri Rizki Ani, Rudi Kurniawan, Tati Suprapti
Teknik Informatika, STMIK IKMI Cirebon
Jalan Perjuangan No.10B, Karyamulya Cirebon, Indonesia
vitrivitir@gmail.com

ABSTRAK

Toko DIOMclothing merupakan sebuah perusahaan dagang yang menawarkan berbagai produk sepatu dengan merek seperti converse, vans, asics, ventela hingga produk sepatu lainnya. Dalam mengelola usahanya, toko DIOMclothing menghadapi tantangan strategis terkait penentuan produk yang paling diminati oleh pelanggan untuk direncanakan persediaannya. Kesalahan dalam menentukan persediaan produk dapat mengakibatkan kerugian finansial karena persediaan yang berlebihan atau kehilangan peluang penjualan karena persediaan yang kurang. Penelitian ini bertujuan mengklaster data penjualan sepatu untuk mengetahui tren penjualan produk sepatu sesuai periode waktu tertentu menggunakan penerapan metode K-means Clustering. Metode ini memungkinkan pengelompokan data penjualan berdasarkan karakteristik kemiripan dari datanya. Penerapan metode ini menghasilkan 3 klaster terbaik dengan karakteristik Klaster 0 (Sangat diminati) terdapat 469 data dengan tren penjualan dari Oktober hingga Desember, Klaster 1 (Cukup diminati) terdiri dari 581 data dengan tren penjualan dari Juli hingga September, Klaster 2 (Kurang diminati) terdiri dari 263 data dengan tren penjualan dari Agustus hingga Desember. Evaluasi menggunakan Davies Bouldin Index (DBI) menunjukkan nilai sebesar 0,165, menandakan kinerja klaster yang baik karena nilai DBI yang dihasilkan mendekati nol dan tidak negatif. Hasil penelitian ini dapat dimanfaatkan oleh DIOMclothing untuk mengoptimalkan pengadaan stok pada periode waktu tertentu, meningkatkan efisiensi manajemen persediaan, dan merancang strategi pemasaran yang lebih tepat sasaran.

Kata kunci : *Clustering, Davies Bouldin Index, K-Means, Perencanaan Persediaan Stok, Tren Penjualan*

1. PENDAHULUAN

Perkembangan era globalisasi saat ini telah memberikan dampak signifikan pada sektor bisnis, terutama dalam bisnis distribusi alas kaki seperti sepatu. Di indonesia, terdapat sebanyak 18.687 industri alas kaki, yang terdiri dari 18.091 usaha skala kecil, 441 usaha skala menengah, dan 155 usaha skala besar (Sumber: kemenperin.go.id). Keadaan tersebut mendorong pelaku usaha untuk mengoptimalkan potensi, mengatasi berbagai ancaman yang timbul, serta berupaya meningkatkan pendekatan dengan para pelanggan untuk bisa unggul dalam persaingan [1]. Salah satu pelaku usaha yang perlu meningkatkan potensi dan mempertahankan pangsa pasarnya adalah Toko DIOMclothing. Toko ini adalah sebuah perusahaan dagang yang menawarkan berbagai jenis sepatu seperti Converse, Nike, Asics, Vans, dan sepatu lainnya. Berlokasi di Kabupaten Cirebon, tepatnya di Jalan Pangeran Cakrabuana, DIOMclothing telah berhasil menjalankan bisnisnya sejak awal berdiri dan telah melakukan sejumlah transaksi penjualan yang signifikan. Meski telah meraih kesuksesan dalam bisnisnya, Toko DIOMclothing masih dihadapkan pada beberapa tantangan yang perlu diatasi.

Berdasarkan observasi, untuk memastikan persediaan yang sesuai dengan kebutuhan masyarakat, toko DIOMclothing dihadapkan pada permasalahan dalam menentukan produk alas kaki yang saat ini dibutuhkan pelanggan. Meskipun proses pencatatan transaksi penjualan produk di toko tersebut sudah menggunakan sistem komputerisasi, namun pengadaan stok produk masih terbatas pada pencarian data sepatu yang persediaannya rendah, tanpa mempertimbangkan tingkat kebutuhan masyarakat terhadap produk tersebut. Praktik ini dapat menyebabkan masalah penumpukan stok atau bahkan kelangkaan stok di toko DIOMclothing. Menurut [2] Kesalahan dalam menentuan persediaan stok dapat menimbulkan berbagai masalah, seperti kelebihan stok yang dapat menghambat efisiensi gudang, meningkatkan risiko penumpukan barang yang dapat mengakibatkan kerusakan barang sehingga barang sulit terjual. Di sisi lain, memiliki persediaan barang yang terlalu sedikit dapat menghadirkan risiko kehabisan stok. Keadaan ini dapat berpotensi menimbulkan risiko kehilangan penjualan atau lost sales akibat ketidaktersediaan stok produk yang diinginkan oleh pelanggan [3]. Jika hal ini terus dibiarkan maka toko akan terancam merugi.

Salah satu solusi untuk mengatasi masalah ini adalah dengan menggunakan aplikasi data mining untuk pengolahan data guna mengekstrak informasi berharga dari data yang tersedia. Data mining dikenal sebagai proses penggalian informasi dari kumpulan data atau dataset berskala besar [4]. Ada berbagai macam metode dalam *data mining* salah satunya adalah metode pengelompokan atau biasa dikenal dengan *clustering*. Menurut [5] Clustering adalah suatu teknik yang digunakan dalam data mining yang bertujuan untuk menggabungkan objek-objek menjadi beberapa kelompok (cluster). Beberapa algoritma yang dapat diterapkan pada teknik ini antara lain *K-*

Nearest Neighbors (KNN), Fuzzy C-Means, dan K-Means [6]. Algoritma *Clustering* yang seringkali diaplikasikan secara luas dalam proses penggolongan data adalah K-Means, algoritma ini dikenal sebagai algoritma yang sangat sederhana dalam mengelompokan data [4]. Menurut [7] *Algoritma K-Means* tergolong dalam algoritma pembelajaran tanpa pengawasan (*unlabeled*). Algoritma ini digunakan untuk membagi data menjadi beberapa kelompok dengan menggunakan pendekatan sistem partisi. Algoritma ini merupakan pilihan yang tepat untuk mengelompokkan informasi mengenai inventaris, strategi penjualan, dan mengidentifikasi produk alas kaki yang populer, kurang populer, dan tidak populer[8].

Penelitian yang berkaitan dengan penerapan *K-means Clustering* dalam memecahkan masalah persediaan stok produk pernah dilakukan oleh beberapa peneliti terdahulu. Seperti oleh [9] dalam jurnalnya yang berjudul "Penerapan Algoritma K-Means Dalam Prediksi Penjualan Karoseri" Menerapkan algoritma *K-means Clustering* pada data penjualan karoseri untuk memprediksi tipe karoseri yang banyak dibutuhkan masyarakat guna membantu pihak produksi dalam menentukan tipe karoseri yang akan di produksi di tahun berikutnya. Penerapan ini menghasilkan 2 cluster dengan karakteristik kurang laris, dan laris. Penelitian lain oleh [10] dalam jurnalnya yang berjudul "Penerapan *Data Mining Metode K-Means Clustering* Untuk *Analisis* Penjualan Pada Toko Fashion Hijab Banten" mengelompokan data stok baju menggunakan metode *K-means Clustering* untuk mengembangkan strategi persediaan stoknya. Penerapan tersebut sukses menghasilkan 3 cluster dengan kriteria produk sangat laris, laris dan kurang laris. Diperoleh nilai yang *konvergen* pada iterasi ke 2. Sementara penelitian yang dilakukan oleh [11] dengan jurnalnya yang berjudul "*Analisis Cluster K-means* dengan *Metode Elbow* untuk Menentukan Pola Penjualan Produk 9Traffic Room Summarecon Mal Bekasi". Dalam penelitiannya, Penulis menggunakan pengelompokan K-means untuk mengelompokkan data guna mengidentifikasi item yang harus ditambahkan atau dikurangi dari produksi. Penelitian ini menghasilkan 4 klaster optimal dengan nilai 594.366,733 atau persentase 65,5% dari pengujian *Sum of Square Error (SSE)* menggunakan perhitungan dari google collab.

Berdasarkan permasalahan dan penjelasan di atas, Penulis akan melakukan *Clusterisasi* menggunakan *algoritma K-means clustering* terhadap data penjualan sepatu di toko DIOMclothing dengan tujuan untuk mengetahui *trend* penjualan produk sepatu sesuai periode waktu tertentu guna mempermudah pemilik toko dalam merencanakan jenis sepatu yang perlu diadakan persediaanya.

Metode analisis data yang digunakan dalam penelitian ini adalah *KDD* (knowledge Discovery in database). *KDD* merupakan suatu metode penggalian pola atau aturan yang terkandung dalam informasi dalam jumlah besar [6]. *KDD* mendukung identifikasi aspek-aspek yang diinginkan, serta mengelola data dengan tujuan menghasilkan informasi yang saling terkait dan berkorelasi [12]. Metode ini merupakan suatu proses yang perlu dilakukan secara sistematis mulai dari pemilihan data, pembersihan data, transformasi data, penggalian data dan evaluasi. Dalam tahap *Data Mining* inilah *algoritma K-means* akan di terapkan untuk mencari nilai klaster yang optimal serta melakukan evaluasi (pengujian) terhadap kinerja klaster berdasarkan nilai *Davies Bouldin Index (DBI)*.

Implikasi dari hasil penelitian ini, memberikan kontribusi yang berharga di bidang Informatika dengan fokus pada manajemen persediaan di toko DIOMclothing. Temuan ini memberikan solusi praktis bagi pemilik toko dalam meningkatkan efisiensi operasiona, membantu menghindari risiko overstock atau kekurangan stok, serta meningkatkan penjualan. Selain itu, hasil temuan ini juga membuka peluang untuk penelitian lebih lanjut dalam bidang Informatika, terutama terkait pengembangan teknologi informasi untuk analisis data dalam konteks manajemen persediaan dengan mengembangkan metode lain serta menerapkannya pada sektor lain yang memilik tantangan serupa. Dengan demikian, temuan ini memberikan manfaat praktis sekaligus menjadi landasan inspiratif untuk penelitian lanjutan dalam pemahaman dan penerapan teknologi informasi dalam bisnis.

2. TINJAUAN PUSTAKA

2.1. Metode Literature Review

Penelitian ini mengadopsi metode tinjauan pustaka yang melibatkan analisis artikel-artikel yang relevan dengan ruang lingkup pengelompokan data penjualan dalam memecahkan permasalahan terkait perencanaan persediaan barang. Metodenya terdiri dari empat tahap, yaitu (1) perumusan pertanyaan, (2) pencarian literatur, (3) evaluasi data, dan (4) analisis dan interpretasi. [13].

Dalam fase perumusan permasalahan, penelitian ini menitikberatkan topiknya pada perencanaan persediaan stok barang.. Pencarian literatur dilakukan menggunakan alat "Publish or Perish" berbasis Google Scholar dengan kata kunci "pengelompokan dengan *K-Means*" dan "persediaan stok barang," menghasilkan 50 artikel jurnal. Tahap evaluasi data melibatkan penyaringan berdasarkan keterbaruan (lima tahun terakhir) sementara kesesuaian dinilai berdasarkan judul, masalah yang dibahas, indeks SINTA, atau aspek lain yang dianggap memiliki reputasi. Sebanyak 5 artikel jurnal terpilih untuk direview. Selanjutnya, langkah tinjauan pustaka dilaksanakan dengan menerapkan teknik perbandingan, kontrast, kritik, sintesis, dan ringkasan [13].

2.2. Hasil Literature Review

Hasil literature review pada jurnal-jurnal penelitian terkait pengelompokan data transaksi menggunakan algoritma K-means dalam memecahkan masalah persediaan stok barang dapat diuraikan sebagai berikut.

Penelitian oleh [14], yang berjudul "Pengelompokan Penjualan Madu Menggunakan Algoritma K-Means," berhasil menentukan kelompok jenis madu yang paling diminati konsumen. Dalam konteks penelitian ini, metode pengukuran jarak yang diterapkan adalah *Numerical Measure* dengan menerapkan *Euclidean distance*. Meskipun menghasilkan 6 klaster optimal, nilai *Davies Bouldin Index (DBI)* yang sebesar -0.365 dan Avg sebesar -0.094 menunjukkan kinerja klasterisasi yang kurang baik.

Sebaliknya, penelitian oleh [15], berjudul "Penerapan Algoritma K-Means Clustering Untuk Menganalisis Transaksi Penjualan Jasa Cetak Pada Unit Print on Demand (Pod) Percetakan Gramedia" juga menggunakan teknik *Numerical Measure* dan *Euclidean Distance*. Namun, nilai DBI yang dihasilkan lebih baik, yaitu 0.209, ketika parameter *maximize* dan *normalize* diaktifkan pada Rapidminer.

Penelitian lain oleh [16], berjudul "Penerapan Algoritma K-Means Clustering Berdasarkan Hasil Penjualan Produk Inez Cosmetic Di Toko Rosalia," menghasilkan 3 kelompok terbaik dengan nilai DBI sebesar 0.128, tanpa mengaktifkan parameter *maximize dan normalize*.

Di sisi lain, penelitian oleh [17], berjudul "Analisa Penerapan Metode Clustering K-Means Untuk Pengelompokan Data Transaksi Konsumen," menggunakan teknik pengukuran jarak *Bregman Divergences* dengan parameter *Squared Euclidean Distance*, menghasilkan 2 klaster optimal dengan nilai DBI sebesar 0.261.

Terakhir, penelitian oleh [18], berjudul "Implementasi Data Mining Pada Stok Penggunaan Barang Di Gmf Aeroasia Menggunakan Algoritma K-Means Clustering," menggunakan alat pengukuran jarak *mixed measures* untuk mengelompokkan data dengan atribut yang berbeda. Penelitian ini menghasilkan 3 kelompok terbaik dengan nilai DBI sebesar 0.494.

Berdasarkan hasil literatur di atas, teknik pengukuran jarak yang paling baik dalam mengelompokkan data penjualan adalah *Numerical Measures* dengan nilai DBI mendekati nol dan tidak negatif. Maka, dalam rangka penelitian ini, penulis memilih untuk menggunakan *Numerical Measures* dengan mempertimbangkan penggunaan parameter *maximize dan normalize* pada Rapidminer

2.3. Landasan Teori

2.3.1. Clustering

Clustering merupakan suatu proses pengelompokan objek berdasarkan informasi yang diperoleh dari data, dengan tujuan maksimalkan tingkat kesamaan di antara anggota dalam satu kelompok dan meminimalkan tingkat kesamaan antara kelompok atau kluster [18].

2.3.2. K-Means

K-Means adalah Algoritma yang dapat mengelompokkan data ke dalam beberapa kelompok melalui pendekatan partisi. Algoritma ini dirancang khusus untuk menangani data yang tidak memiliki label kelas di mana komputer secara otomatis mengelompokkan data tanpa memiliki pengetahuan sebelumnya terkait target kelasnya [15].

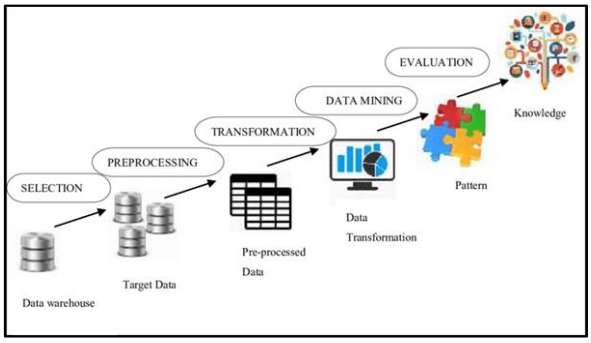
Pengelompokan data menggunakan algoritma K-Means dilakukan dengan langkah-langkah yang cukup sederhana. Pada tahap awal, tentukan jumlah kelompok data yang akan dibentuk. Setelah penetapan jumlah kelompok, data kemudian dipilih sebagai titik sentroid cluster. Langkah selanjutnya melibatkan iterasi hingga mencapai nilai yang konvergen [4].

2.3.3. Davies Bouldin Index (DBI)

Davies-Bouldin Index (DBI) merupakan metode evaluasi yang digunakan untuk mengukur kinerja hasil pengelompokan dan termasuk dalam kategori evaluasi internal. Metode ini didasarkan pada nilai kohesi dan separasi. Dalam konteks pengelompokan, kohesi mencakup jumlah kedekatan data terhadap centroid kelompok, sedangkan separasi berkaitan dengan jarak antara centroid dari masing-masing klaster[7]. Performa klasterisasi dengan K-Means dianggap semakin optimal jika nilai Davies-Bouldin Index (DBI) yang dihasilkan semakin rendah atau sama dengan 0 dan tidak negatif (non-negatif >= 0) [6], [7], [16].

3. METODE PENELITIAN

Metodologi penelitian yang diterapkan dalam kajian ini adalah Knowledge Discovery in Database (KDD). Tahap-tahap dari pendekatan ini tergambar pada Gambar 1 berikut ini.



Gambar 1. Tahapan KDD [19]

a) *Selection*
Pada fase seleksi ini, dilakukan pemilihan data yang relevan dengan kebutuhan penelitian.

b) Pre-processing
Pada tahap ini, dilakukan kegiatan pemeriksaan terhadap data yang mengalami inkonsistensi,

data yang terdapat kesalahan ketik, penghapusan data *duplikat*, dan data yang tidak memiliki relevansi dengan penelitian.

c) Transformation
Pada langkah ini, terjadi proses pengubahan tipe data ke dalam format yang sesuai dengan keperluan penelitian. Selain itu, pada tahap ini juga dilakukan normalisasi data untuk mengatur data ke dalam rentang yang spesifik.

d) Data Mining
Proses Data Mining ini melibatkan penerapan algoritma untuk memodelkan data sesuai dengan kebutuhan dan tujuan penelitian. Algoritma yang diterapkan dalam penelitian ini adalah *K-Means Clustering*.

e) Evaluation
Pada tahap evaluasi ini, dilakukan penilaian terhadap klaster yang dihasilkan dari tahap Data Mining (penerapan algoritma *K-Means Clustering*). Metode evaluasi yang digunakan adalah *Davies Bouldin Index (DBI)*.

f) Knowledge
Dalam langkah ini, hasil klasterisasi akan dipresentasikan melalui diagram atau visualisasi data yang mudah dipahami.

3.1. Sumber Data

Sumber data yang digunakan dalam penelitian ini bersumber dari data primer. Data primer merujuk pada informasi yang diperoleh peneliti secara langsung dari sumber aslinya. Data tersebut berupa data transaksi penjualan dari toko DIOMclothing selama periode waktu 6 bulan dari bulan Juli 2022 - Desember 2022

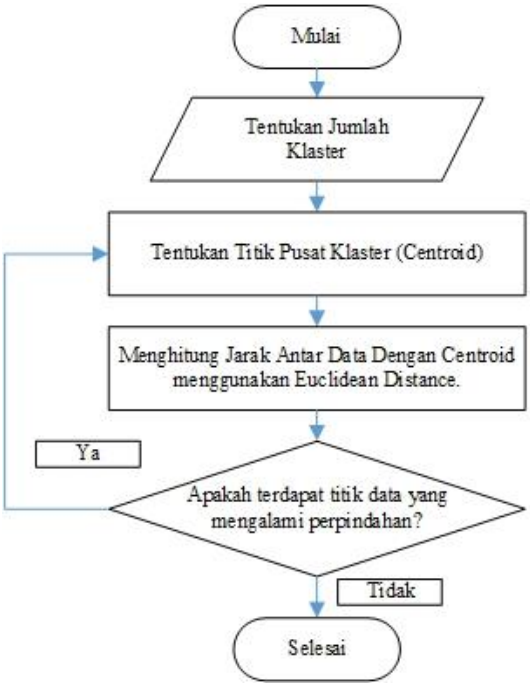
3.2. Teknik Pengumpulan Data

Dalam konteks penelitian ini, penulis memanfaatkan dua metode pengumpulan data untuk memperoleh data yang akan dijadikan sebagai bahan analisis. Beberapa metode pengumpulan data mencakup hal-hal berikut:

1. Observasi
   Dilakukan langsung di toko sepatu DIOMclothing. Tujuan dilakukannya observasi agar dapat mengamati secara langsung keadaan yang terjadi di toko DIOMclothing
2. Wawancara.
   Proses wawancara dilakukan bersamaan dengan kegiatan observasi. Dengan menyampaikan maksud dan tujuan kedatangan kepada pemilik toko serta menanyakan mengenai persoalan yang masih menjadi tantangan dari proses operasional toko DIOMclothing.

3.3. Teknik Analisis Data

Metode analisis data yang diterapkan dalam penelitian ini melibatkan penerapan algoritma K-Means Clustering. Algoritma ini dipergunakan untuk melakukan pengelompokan data berdasarkan kesamaan karakteristik data. Algoritma ini akan diterapkan pada data transaksi penjualan sepatu, dengan tujuan mengidentifikasi kelompo dengan tren penjualan sepatu yang sangat diminati, cukup diminati dan kurang diminati pada periode waktu tertentu menggunakan perangkat lunak Rapidminer. Ilustrasi diagram alur yang menggambarkan proses algoritma K-Means Clustering dapat dilihat pada Gambar 2 berikutt ini.



Gambar 2. Flowchart K-means Clustering

Diagram flowchart yang tertera di atas mengilustrasikan rangkaian langkah dari algoritma K-Means sebagai berikut:

1. Tentukan Jumlah Klaster
   Jumlah kelompok yang akan dibentuk dari 1313 data ialah 3 klaster, yaitu kelompok dengan tren penjualan produk sangat diminati, cukup diminati, dan kurang diminati sesuai periode waktu tertentu.
2. Tentukan Titik Pusat Klaster
   Menentukan titik pusat klaster sesuai dengan jumlah klaster yaitu sebanyak 3 centroid.
3. Menghitung Jarak Antar Data dengan Centroid
   Setiap data dari jumlah transaksi penjualan sepatu dilakukan perhitungan jarak terhadap masing-masing pusat klaster yang telah ditentukan menggunakan persamaan Euclidean Distance sebagai berikut.

$$M_e = \sqrt{(xj - sj)^2 + (yj - tj)^2} \quad (1)$$

   Keterangan:
   Me : *Euclidean Distance*
   j : Banyaknya Data
   (x,y) : Koordinat Data
   (s,t) : Koordinat *Centroid*

4. Ulangi langkah 2 dan 3 apabila masih terdapat perubahan letak dari titik data.

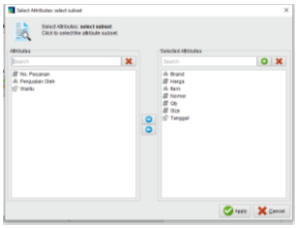
Sementara itu, untuk menilai performa klasterisasi yang dihasilkan, digunakan metode evaluasi dalam bentuk Davies-Bouldin Index (DBI). DBI digunakan untuk mengukur validitas atau kualitas klaster yang dihasilkan. Semakin kecil nilai DBI dan non-negatif menunjukkan kualitas klaster yang lebih baik.

4. HASIL DAN PEMBAHASAN

4.1. Data Selection

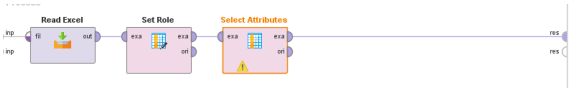
Dalam langkah ini, dilakukan pemilihan data yang akan digunakan dalam proses klasterisasi menggunakan perangkat RapidMiner. Langkah pertama yang perlu dilakukan yaitu Import data ke *Repository RapidMiner* dengan menggunakan operator *Read Excel* tujuannya adalah untuk membaca dataset bertipe Excel.

Langkah selanjutnya yaitu memasukan operator *set role* untuk mengubah peran atribut nomor menjadi id. Setelah itu menyeleksi atribut yang akan digunakan dalam pengklasteran menggunakan operator select atribut. Proses pemilihan atribut dapat dilihat pada Gambar 3 berikut ini.



Gambar 3. Proses Select Atribut

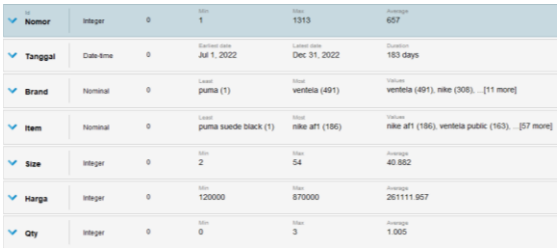
Berdasarkan gambar di atas, terdapat 3 atribut yang tidak digunakan dan 7 atribut yang digunakan dalam pengolahan. Model proses pada tahap seleksi data ini dapat dilihat pada Gambar 4 di bawah ini.



Gambar 4. Model Proses Data Selection

4.2. Data Preprocessing

Pada Tahap Preprocessing data ini dilakukan proses pengecekan melalui statistik data untuk melihat apakah ada data yang missing. Berikut ini adalah hasil pengecekan dapat dilihat pada Gambar 5 di bawah ini.



Gambar 5. Statistik Data

Dikarenakan pada statistik data tidak ditemukan nilai missing maka data siap untuk diolah ke tahap berikutnya.

4.3. Data Transformation

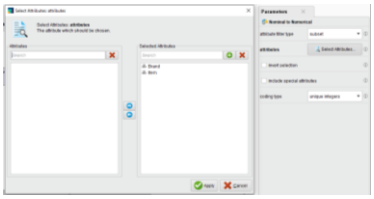
Dalam tahap ini dilakukan proses pengubahan tipe data ke dalam bentuk data yang sesuai dengan kebutuhan klasterisasi. Algoritma *K-Means* membutuhkan tipe data *numeric* (angka) untuk dapat dilakukan pengolahan. Berikut ini adalah data atribut yang akan dilakukan perubahan dapat dilihat pada Tabel 1 yang ada di bawah ini.

Tabel 1. Data Atribut dan Tipe Data

No.	Atribut	Tipe Data	Target Tipe data
2.	Item	Nominal	Numeric
3.	Brand	Nominal	Numeric
4.	Tanggal	Date time	Numeric

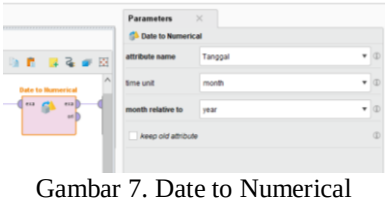
Berdasarkan tabel di atas, terdapat 2 atribut dengan tipe data nominal dan 1 atribut dengan tipe data date time. Nominal adalah tipe data yang digunakan untuk mengelompokkan atau mengkategorikan objek berdasarkan label atau nama tertentu. Sedangkan tipe data date time merupakan tipe data yang menunjukan kategori tanggal. Dikarenakan ketiga atribut tersebut bukan numeric maka perlu dilakukan perubahan ke tipe data numeric (angka).

Untuk atribut item dan brand digunakan operator *nominal to numerical* yang dapat mengubah tipe data nominal ke tipe data numeric. Proses transfomasi data nominal ke numerical dapat dilihat pada Gambar 6 berikut ini.



Gambar 6. Nominal to Numerical

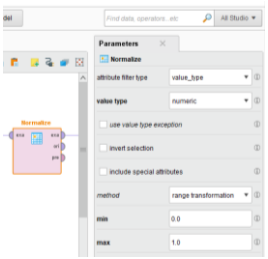
Sedangkan untuk atribut tanggal digunakan operator *date to numerical* dengan parameter time unit = month, month relative to = year (satuan yang digunakan pada atribut tanggal berupa bulan) berikut ini prosesnya dapat dilihat pada Gambar 7 di bawah ini.



Gambar 7. Date to Numerical

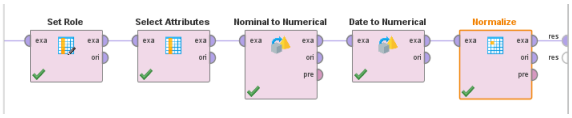
Selanjutnya melakukan normalisasi data menggunakan operator *normalize* untuk mengubah nilai dari semua atribut ke rentang nilai 0 hingga 1 guna membuat skala yang konsisten. Berikut ini adalah

proses normalisasi data menggunakan operator normalize dapat dilihat pada Gambar 8 di bawah ini.



Gambar 8. Proses Normalisasi Data

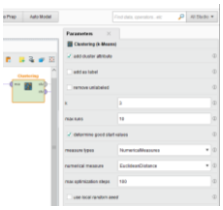
Model proses sampai tahap transformasi data dapat dilihat pada Gambar 9 yang tertera di bawah.



Gambar 9. Model Proses Sampai Tahap Transformation

4.4. Data Mining

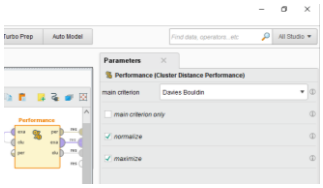
Setelah melalui tahap transformasi data, langkah berikutnya melibatkan penerapan algoritma K-Means Clustering. Ilustrasi proses klasterisasi menggunakan Operator Clustering K-Means dapat ditemukan dalam Gambar 10 di bawah ini..



Gambar 10. Proses Klasterisasi K-Means

Dalam pengaturan parameter K, nilai K mencerminkan jumlah klaster yang akan dibentuk. Dalam penelitian ini, peneliti memilih untuk membentuk 3 kelompok (k=3) guna mengidentifikasi tren penjualan produk sepatu yang sangat diminati, cukup diminati, dan kurang diminati pada periode waktu tertentu. Proses pembentukan kelompok ini dijalankan sebanyak 10 kali running (Max Run= 10), dengan menerapkan teknik pengukuran jarak berupa metode pengukuran numerik (*Numerical Measures).*

Langkah selanjutnya adalah menambahkan operator *Cluster Distance Performance* untuk melakukan evaluasi performa dari klaster yang telah ditentukan menggunakan *Davies Bouldin* dengan mengaktifkan parameter *normalize* dan *maximize*, dapat dilihat pada Gambar 11 berikut ini.



Gambar 11. Evaluasi menggunakan Operator Cluster Distance Performance

Model proses pada tahap data mining dapat dilihat pada Gambar 3 yang tertera di bawah ini.

Gambar 12. Model Proses Sampai Tahap Data Mining

Dalam menentukan klaster optimal, telah dilakukan pengujian nilai K sebanyak K 2 3 4 5. Hasil dari pengujian sejumlah K dari K 2 sampai K 5 , dapat dilihat pada Tabel berikut ini.

Tabel 2. Hasil Pengujian K

No.	Nilai Klaster	DBI
1.	2	0.177
2.	3	0.165
3.	4	0.170
4.	5	0.180

Berdasarkan Tabel 2 tersebut, dapat diketahui bahwa klaster yang menghasilkan nilai DBI terkecil dan mendekati nol berada pada klaster 3 dengan nilai DBI sebesar 0,165.

Setelah diketahui klaster yang paling *optimal*, selanjutnya dilakukan proses *running* untuk mengetahui *output* dari klaster tersebut. Berikut ini adalah hasil *running* dari k 3 menghasilkan *output* sebagai berikut :

Cluster Model

```
Cluster 0: 469 items
Cluster 1: 581 items
Cluster 2: 263 items
Total number of items: 1313
```

Gambar 13. Hasil Clusterisasi Model

Berdasarkan Gambar 13, cluster model yang dihasilkan dalam penelitian ini adalah *Cluster 0* berisi 469 data, *Cluster 1* berisi 581 data dan *Cluster 2* berisi 263 data dengan total data sebanyak 1313 data.

4.5. Evaluasi

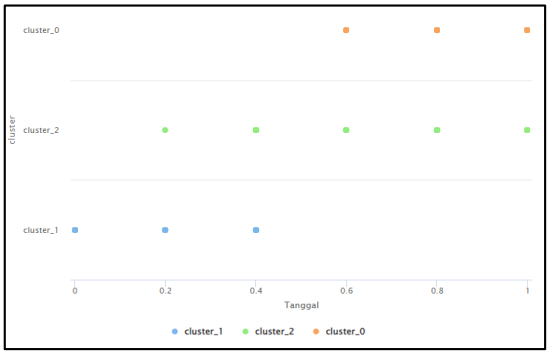
Hasil evaluasi kinerja klaster berdasarkan nilai Davies Bouldin pada data transaksi penjualan sepatu di toko DIOMclothing diperoleh senilai 0,165. Nilai tersebut menunjukan bahwa klasterisasi yang dihasilkan memiliki performa yang baik karena nilai DBI yang dihasilkan kecil dan mendekati nol serta bukan negatif. Ini berarti penerapan algoritma *K-Means Clustering* dapat dengan baik mengelompokan

data penjualan sepatu dari toko DIOMclothing sesuai tren penjualan pada periode waktu tertentu. nilai Davies Bouldin tersebut dapat dilihat pada Gambar 14 berikut ini3699

Gambar 14. Hasil Evaluasi Kinerja Klaster

4.6. Knowledge

Untuk dapat memahami informasi dari hasil klasterisasi yang telah dilakukan di atas. Berikut ini visualisasi hasil klasterisasi menggunakan scatter bubble nampak pada Gambar 15 di bawah ini.



Gambar 15. Scatter Cluster Berdasarkan Tanggal

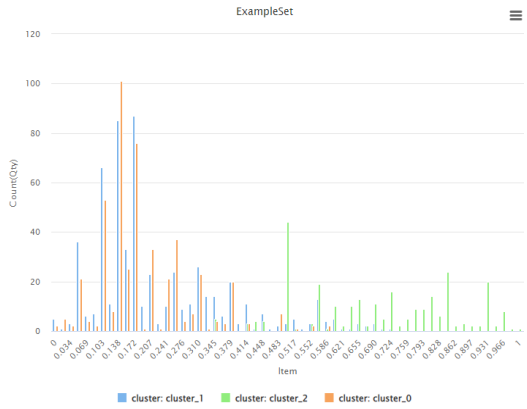
Gambar 10 di atas adalah scatter buble dengan ketentuan sumbu x merupakan tanggal penjualan dalam bulan yang dimuat dalam nilai dengan range dari 0 hingga 1. Berikut ini adalah keterangan dari masing masing nilai yang ada pada sumbu x dapat disimak melalui Tabel 3 berikut.

Tabel 3. Keterangan Nilai Sumbu X Pada Bulan

Angka	0	0,2	0,4	0,6	0,8	1
Bulan	7	8	9	10	11	12

Berdasarkan Gambar 10 dan tabel 4 di atas dapat di jelaskan bahwa cluster 1 berisi item dengan tren penjualan yang terjadi dari bulan 0 hingga bulan 0,4 (Juli hingga September), cluster 2 berisi item dengan tren penjualan dari bulan 02 hingga bulan 1 (Agustus hingga Desember) selama 5 bulan, sedangkan cluster 0 berisi item dengan tren penjualan yang terjadi dari bulan 0,6 hingga bulan 1 atau dari Oktober hingga Desember.

Selain itu berikut ini adalah visualisasi data dalam bentuk grafik yang menampilkan cluster berdasarkan atribut Qty (jumlah penjualan) dan Item (tipe model sepatu). Penggunaan visualisasi ini dapat mengidentifikasi tingkat penjualan dari tiap *cluste*r. Visualisasi tersebut dapat dilihat pada Gambar 16 di bawah ini.



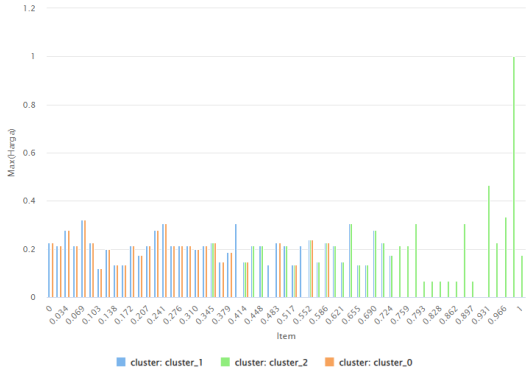
Gambar 16. Hasil Cluster berdasarkan Item dan Qty

Menurut Gambar 16 dari grafik di atas, dapat diidentifikasi bahwa cluster 0 dengan garis berwarna orange merupakan klaster dengan tingkat penjualan tertinggi. Jumlah penjualan terbanyak dari *cluster* 0 ini mencapai 101 pasang, item tersebut adalah Nike Afi.

Sementara itu cluster 1 dengan garis warna biru menempati tingkat penjualan tinggi kedua (sedang), dimana jumlah penjualan terbanyak dari *cluster* 1 ini mencapai 87 pasang yaitu item Ventela Public dan 86 pasang yaitu Nike Af1.

Sedangkan *cluster* 2 dengan garis warna hijau menempati tingkat penjualan paling rendah dengan jumlah penjualan terbanyaknya hanya sejumlah 44 pasang yaitu item NB 997.

Selanjutnya berikut ini adalah grafik *visualisasi* hasil klasterisasi berdasarkan atribut item dan atribut harga dapat dijelaskan pada Gambar 12 di bawah.



Gambar 17. Grafik Cluster Berdasarkan Item dan Harga

Gambar 17 menunjukan tingkat harga setiap *Cluster* dapat diidentifikasi sebagai berikut. *Cluster 0* dan *Cluster 1* merupakan kelompok dengan tingkat harga yang relatif sedang (tidak terlalu rendah maupun terlalu tinggi), sementara *Cluster 1* merupakan kelompok dengan tingkat harga yang relativ bervariatif yaitu dari harga yang sangat rendah hingga harga yang sangat tinggi.

5. KESIMPULAN DAN SARAN

Penerapan algoritma K-Means pada data transaksi penjualan sepatu di toko DIOMclothing melibatkan tahapan KDD yang mencaku*p seleksi data*, *preprocessing*, *transformasi*, dan *evaluasi* untuk mengidentifikasi tren penjualan. Hasil klasterisasi menunjukkan bahwa jumlah Klaster optimal adalah 3, dengan karakteristik Klaster 0 berisi kumpulan item sangat diminati dengan jumlah 469 data, tren penjualannya terjadi dari Oktober hingga Desember. Klaster 1 cukup diminati memiliki 581 data dengan tren penjualan terjadi dari Juli hingga September, Klaster 2 kurang diminati memiliki 263 item dengan tren penjualan yang terjadi dari Agustus hingga Desember.

Evaluasi kinerja klaster berdasarkan *Davies Bouldin Index* menghasilkan nilai 0,165, menandakan klasterisasi yang baik. Berdasarkan pembahasan, dapat disimpulkan juga item yang sangat diminati dan memiliki penjualan tertinggi adalah Nike AF1, dengan penjualan mencapai 101 pasang dari Juli hingga September, dan 86 pasang dari Oktober hingga Desember. Ventela Public cukup diminati, terjual 87 pasang dari Oktober hingga Desember. Jadi Rencana persediaan sebaiknya difokuskan pada Nike AF1 untuk Juli-Desember dan Ventela Public untuk Oktober-Desember.

DAFTAR PUSTAKA

[1] S. Avriyanti, "Peran E-Commerce untuk Meningkatkan Keunggulan Kompetitif di Era Industri 4.0 (Studi pada UKM yang terdaftar pada Dinas Koperasi, Usaha Kecil dan Menengah Kabupaten Tabalong)," *J. PubBis*, vol. 4, no. 1, pp. 82–99, 2020.

[2] N. Sari, "Perencanaan Dan Pengendalian Persediaan Barang Dalam Upaya Meningkatkan Efektivitas Gudang," *J. Bisnis, Logistik dan Supply Chain*, vol. 2, no. 2, pp. 85–91, 2022, doi: 10.55122/blogchain.v2i2.542.

[3] I. F. P. Ginting, D. Saripurna, and E. Fitriani, "Penerapan Data Mining Dalam Menentukan Pola Ketersediaan Stok Barang Berdasarkan Permintaan Konsumen Di Chykes Minimarket Menggunakan Algoritma Apriori," *J. SAINTIKOM (Jurnal Sains Manaj. Inform. dan Komputer)*, vol. 20, no. 1, p. 28, 2021, doi: 10.53513/jis.v20i1.2504.

[4] Y. Hartati, S. Defit, and G. W. Nurcahyo, "Klasterisasi Bibit Terbaik Menggunakan Algoritma K-Means dalam Meningkatkan Penjualan," *J. Inform. Ekon. Bisnis*, vol. 3, pp. 4–10, 2021, doi: 10.37034/infeb.v3i1.56.

[5] N. Y. Aswad, "Clustering Algoritma K-Means Pengadaan Barang Non Medis Di Rumah Sakit Jantung Hasna Medika Cirebon," *J. Data Sci. dan Inform.*, vol. 2, no. 1, pp. 6–14, 2022.

[6] T. A. Aria, Yuliadi, M. Julkarnain, and F. Hamdani, "Penerapan Algoritma K-Means Untuk Mengklasifikasi Data Obat," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 12, no. 1, pp. 53–62, 2023, doi: 10.32736/sisfokom.v12i1.1461.

[7] Z. Nabila, A. Rahman Isnain, and Z. Abidin, "Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means," *J. Teknol. dan Sist. Inf.*, vol. 2, no. 2, p. 100, 2021, [Online]. Available: http://jim.teknokrat.ac.id/index.php/JTSI

[8] N. Afiasari, N. Suarna, and N. Rahaningsih, "Implementasi Data Mining Transaksi Penjualan Menggunakan Algoritma Clustering dengan Metode K-Means E-commerce K-Means melakukan analisis penerapan Data Mining dalam mengelompokkan jumlah," *SAINTEKOM*, vol. 9, pp. 100–110, 2023, [Online]. Available: https://ojs.stmikplk.ac.id/index.php/saintekom/article/view/402

[9] D. Anggarwati, O. Nurdiawan, I. Ali, and D. A. Kurnia, "Penerapan Algoritma K-Means Dalam Prediksi Penjualan Karoseri," *J. Data Sci. Inform.*, vol. 1, no. 2, pp. 58–62, 2021.

[10] Normah, B. Rifai, S. Vambudi, and R. Maulana, "Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Pada Toko Fashion Hijab Banten," *J. Tek. Komput. AMIK BSI*, vol. 8, no. 2, pp. 174–180, 2022, doi: 10.31294/jtk.v4i2.

[11] A. Prasetya, R. Salkiawati, and A. D. Alexander, "Analisis Cluster K-Means dengan Metode Elbow untuk Menentukan Pola Penjualan Produk Traffic Room Summarecon Mal Bekasi," *J. Students' Res. Comput. Sci.*, vol. 4, no. 1, pp. 105–118, 2023, doi: 10.31599/jsrcs.v4i1.2480.

[12] D. Winarti, M. Kom, E. Revita, and M. Kom, "Penerapan Data Mining untuk Analisa Tingkat Kriminalitas Dengan Algoritma Association Rule Metode FP-Growth," *J. SIMTIKA*, vol. 4, no. 3, pp. 8–22, 2021.

[13] T. Yuniati and M. F. Sidiq, "Literature Review : Legalisasi Dokumen Elektronik Menggunakan Tanda," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 4, no. 6, pp. 1058–1069, 2020, [Online]. Available: https://doaj.org/article/159b35f326d7453eb754b389eb4214c7%0Ahttps://repository.ittelkom-pwt.ac.id/5976/%0Ahttps://garuda.kemdikbud.go.id/documents/detail/1980051

[14] Danya Rizki Chaerunisa, Nining Rahaningsih, Fadhil Muhammad Basysyar, Ade Irma Purnamasari, and Nana Suarna, "Pengelompokan Penjualan Madu Menggunakan Algoritma K-Means," *KOPERTIP J. Ilm. Manaj. Inform. dan Komput.*, vol. 5, no. 1, pp. 23–28, 2021, doi: 10.32485/kopertip.v5i1.144.

[15] G. Gunadi, "Penerapan Algoritma K-Means Clustering Untuk Menganalisa Transaksi Penjualan Jasa Cetak Pada Unit Print on Demand (Pod) Percetakan Gramedia," *Infotech J. Technol. Inf.*, vol. 8, no. 2, pp. 117–126, 2022, doi: 10.37365/jti.v8i2.148.

[16] N. KUMALA and R. D. Dana, "Penerapan Algoritma K-Means Clustering Berdasarkan Hasil Penjualan Produk Inez Cosmetic Di Toko Rosalia," *E-Link J. Tek. Elektro dan Inform.*, vol. 18, no. 1, p. 50, 2023, doi: 10.30587/e-link.v18i1.5343.

[17] Masjunedi, N. Suarna, and Y. A. Wijaya, "ANALISA PENERAPAN METODE CLUSTERING K-MEANS UNTUK PENGELOMPOKAN DATA TRANSAKSI KONSUMEN (Studi Kasus : Cv . Mitra Indexindo Pratama)," vol. 7, no. 2, pp. 1322–1328, 2023.

[18] T. B. Pamungkas, S. Maesaroh, and P. Ardiansyah, "Implementasi Data Mining Pada Stok Penggunaan Barang Di Gmf Aeroasia Menggunakan Algoritma K-Means Clustering," *J. Ilm. Sains dan Teknol.*, vol. 7, no. 2, pp. 112–123, 2023, doi: 10.47080/saintek.v7i2.2697.

[19] P. R. Wulandhari, N. Rahaningsih, I. Ali, and C. L. Rohmat, "IMPLEMENTASI DATA MINING DALAM PERENCANAAN PERSEDIAAN OBAT," *JATI (jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, 2023, [Online]. Available: https://ejournal.itn.ac.id/index.php/jati/article/view/6404