

ANALISIS DATA PENJUALAN MENGGUNAKAN ALGORITMA K-MEANS CLUSTERING PADA TOKO KECANTIKAN PUTRI

Agung prayogo bagustio¹, Ade Irma Purnamasari², Irfan Ali³

^{1,2}Teknik Informatika, STMIK IKMI Cirebon

³Rekayasa Perangkat Lunak, STMIK IKMI Cirebon

Jln Perjuangan No 10B Majasem Kesambi Kota Cirebon

E-mail: Saji4897@gmail.com1 , irma2974@yahoo.com2 , irfanaali0.0@gmail.com3

Abstrak - Industri kecantikan mengalami pertumbuhan pesat dalam beberapa tahun terakhir, mendorong tingginya permintaan produk kecantikan. Toko Kecantikan Putri merupakan salah satu toko kecantikan yang berkembang pesat di wilayah Cirebon. Untuk meningkatkan strategi penjualan dan memahami pola pembelian pelanggan, Toko Kecantikan Putri perlu menganalisis data penjualan secara efektif. Analisis data penjualan tradisional tidak dapat memberikan *insights* yang mendalam mengenai pola pembelian pelanggan. Hal ini dapat menghambat Toko Kecantikan Putri dalam mengembangkan strategi penjualan yang tepat dan efektif. Penelitian ini menerapkan algoritma K-Means clustering untuk menganalisis data penjualan Toko Kecantikan Putri selama bulan Januari 2024. Data yang dianalisis terdiri dari 122 baris dengan 22 atribut. Data penjualan meliputi informasi produk, kategori produk, dan jumlah penjualan. Algoritma K-Means clustering digunakan untuk mengelompokkan data penjualan berdasarkan karakteristik produk. Teknik analisis menggunakan *Knowledge Discovery in Database* (KDD) merupakan proses untuk menemukan pengetahuan baru dari data yang dikumpulkan. Hasil analisis menunjukkan bahwa data penjualan Toko Kecantikan Putri dapat dilakukan menjadi menjadi 9 *cluster*, dari *cluster* k=2 sampai k=9. Masing-masing *cluster* memiliki karakteristik pola pembelian yang berbeda. *Cluster* terbaik terdapat pada *cluster* K=8 dengan Nilai DBI sebesar 0,021 yang berasal dari *measure Types Mixed Measures* nilai *davies bouldin index* 0.077, *Numerical Measures* nilai *davies bouldin index* 0.114, *Bregman Divergences* nilai *davies bouldin index* 0.114. Analisis data penjualan Toko Kecantikan Putri dengan Algoritma K-Means Clustering menghasilkan *insights* berharga mengenai pola pembelian pelanggan. Evaluasi menyeluruh terhadap berbagai parameter, jumlah *cluster*, dan nilai *Davies Bouldin Index* optimal, menunjukkan bahwa Algoritma K-Means Clustering menghasilkan pengelompokan data yang paling optimal dan informatif.

Kata Kunci: Algoritma K-Means *clustering*, Analisis Data Penjualan, Toko Kecantikan Putri, Pola Pembelian Pelanggan, Strategi Penjualan

I. PENDAHULUAN

Kemajuan teknologi, orang dapat menggunakan aplikasi penjualan atau mencari lokasi toko, meningkatkan penjualan produk atau UMKM. Dengan demikian, ada banyak peluang bagi para pengusaha untuk memperluas lapangan usahanya untuk menarik lebih banyak peminat, baik di kota maupun di dalam kota. Persaingan antara toko adalah hal yang biasa di dunia bisnis. Toko harus mengembangkan dan menggunakan peluang di berbagai bidang dan kebutuhan agar dapat bertahan dari persaingan. Di tengah perkembangan teknologi dan sistem informasi, ada peluang terbesar yang dapat membantu bisnis saat ini. Kebutuhan bisnis tetap penting, dan perusahaan dapat memenuhi tiga kebutuhan ini: profitabilitas, pengurangan biaya operasional, peningkatan kapasitas produk, dan pemasaran.

Toko kecantikan Putri, menjual produk kecantikan, tetapi tidak semua dari berbagai jenis produk kecantikan yang dijual sangat laris atau kurang laris. Toko putri ini memiliki data yang tidak terorganisir tentang pengeluaran tidak terduga, penjualan, dan barang yang dibeli. Akibatnya, data ini hanya dapat digunakan sebagai catatan toko dan tidak dapat digunakan untuk membangun strategi

pemasaran. Data yang tersedia seharusnya dapat digunakan sebagai sistem pengambilan keputusan untuk solusi bisnis dan membutuhkan dukungan infrastruktur teknologi.

Toko kecantikan putri menghadapi masalah ketika mereka ingin membuat keputusan tentang pemilihan produk tren dan laku dengan hanya menggunakan hipotesis. Selain itu, toko kecantikan putri tidak memiliki sistem IT yang memungkinkan pengolahan data yang efisien dan efektif. Keputusan dibuat karenanya, yang memiliki resiko tinggi dan berdampak besar pada kelangsungan bisnis. Contoh konsekuensi yang akan dihadapi adalah kehilangan biaya untuk produk baru yang tidak memiliki pangsa pasar yang diperlukan, dan pemesanan produk yang tidak perlu atau tidak berguna.

Dalam penelitian ini, penulis akan menganalisis data transaksi penjualan di Toko Kecantikan Putri untuk mengidentifikasi produk yang diminati sesuai dengan tren masyarakat saat ini dan untuk menentukan produk mana yang paling diminati pelanggan untuk dijual di toko tersebut.

Analisis data adalah proses mengubah data menjadi informasi yang dapat dipahami dan digunakan untuk menyelesaikan masalah dan kesimpulan. *Clustering* adalah suatu metode analisis data yang bertujuan untuk mengelompokkan data

dengan atribut yang sama atau berbeda. Dari berbagai metode analisis data yang ada, metode K-means clustering adalah salah satu yang dapat memenuhi dan memenuhi kebutuhan dan kebutuhan teknologi dan sistem informasi perusahaan dan toko.

K-means clustering adalah salah satu algoritma unsupervised learning yang termasuk ke dalam analisis kluster (*cluster analysis*) non hirarki yang digunakan untuk mengelompokkan data berdasarkan variabel atau *feature*.

Analisis data mining biasanya dikombinasikan dengan analisis gudang data di gudang untuk menemukan atau memisahkan hubungan dari sejumlah besar data yang sebelumnya tidak diketahui. Tiga pendekatan umum untuk pengolahan data adalah regresi, klasifikasi, dan klastering. Regresi adalah metode analisis yang digunakan untuk menentukan hubungan antara dua atau lebih variabel. Klasifikasi juga sama dengan regresi, tetapi konsep yang digunakan berbeda. Salah satu subkategori pengolahan data adalah klasifikasi, yang merupakan proses membagi sampel yang sama menjadi kelompok-kelompok yang disebut *cluster*. Setiap cluster terdiri dari sampel yang memiliki karakteristik yang sama dan berbeda dari sampel yang tersedia dari kelompok lain.

Toko Kecantikan Putri akan menggunakan data mining untuk mengolah data transaksi penjualan produk yang dijual oleh toko berdasarkan masalah yang dihadapi. Diharapkan penelitian ini akan membantu Toko Kecantikan Putri membuat keputusan. Dalam penelitian mereka yang berjudul "Analisis Data Penjualan Menggunakan Algoritma K-Means Clustering Pada Toko Kecantikan Putri", penulis membahas masalah penelitian yang disebutkan sebagai "Bagaimana menerapkan Data Mining pada penjualan untuk mengetahui produk yang diminati oleh masyarakat menggunakan metode K-Means Clustering Studi Kasus Toko Kecantikan Putri". Hasil penelitian menunjukkan bahwa metode K-Means Clustering membantu dalam pengelompokan pola penjualan selama Untuk membedakan merek.

II. TINJAUAN PUSTAKA

Tinjauan pustaka berisi kajian literatur terdahulu yang dilakukan oleh peneliti lain yang relevan dengan penelitian yang dilakukan untuk menunjukkan kebaruan ilmiah artikel tersebut. Pada bagian ini diuraikan *research gap* atau *state of the art* sebagai dasar pernyataan kebaruan ilmiah dari artikel.

III. METODE PENELITIAN

Penelitian ini menggunakan metode eksperimen kuantitatif untuk memproses dan menganalisis data secara sistematis dan objektif, menghasilkan kesimpulan yang akurat.

Pengumpulan data dilakukan melalui interaksi langsung dengan responden melalui wawancara dan observasi lapangan. Metode eksperimen memungkinkan pengumpulan data melalui proses pengujian terkontrol. Data penjualan Toko Putri yang diperoleh kemudian akan dianalisis menggunakan algoritma K-Means Clustering. Data yang digunakan dalam penelitian ini adalah data penjualan produk dengan total 426 item.



Gambar 1. Diagram Alur Metode Penelitian

Berikut ini adalah ringkasan metode penelitian yang digunakan dalam penelitian ini:

1. Identifikasi Masalah

Data penjualan diolah dan dibersihkan terlebih dahulu, kemudian dianalisis menggunakan algoritma K-Means Clustering. Algoritma ini mengelompokkan produk baju berdasarkan pola penjualan, dengan nilai K optimal ditentukan berdasarkan *Davies Bouldin Index* (DBI). Mengidentifikasi produk baju terlaris, laris, dan kurang laris di Toko Putri Kosmetik. Hal ini dilakukan dengan menganalisis data penjualan baju selama periode tertentu. Data ini akan diolah dan dibersihkan terlebih dahulu, kemudian dianalisis menggunakan algoritma K-Means Clustering. Algoritma ini akan mengelompokkan produk baju berdasarkan pola penjualan, dengan nilai K optimal yang ditentukan berdasarkan *Davies Bouldin Index* (DBI). Hasil analisis ini akan memberikan informasi penting untuk menentukan strategi persediaan stok penjualan yang tepat, sehingga Toko Putri Kosmetik dapat memaksimalkan keuntungan dan meminimalkan kerugian.

2. Pengumpulan Data

Proses pengumpulan data dilakukan dengan meminta izin data penjualan produk berdasarkan data stok penjualan. Data yang dikumpulkan terdiri dari 326 record dan berasal dari data penjualan toko kecantikan Toko Putri selama bulan Januari 2024. Proses pengumpulan data dilakukan dengan meminta izin data penjualan produk berdasarkan data stok penjualan. Data yang dikumpulkan terdiri dari 326 record dan berasal dari data penjualan toko kecantikan Toko Putri selama bulan Januari 2024.

3. Analisis Data

Proses analisis dilakukan untuk mengidentifikasi permasalahan penjualan di Toko Putri Kosmetik. Analisis ini dilakukan dengan mengevaluasi data penjualan selama periode tertentu. Data penjualan tersebut

kemudian dianalisis menggunakan algoritma K-Means Clustering. Proses analisis dilakukan untuk mengidentifikasi permasalahan penjualan di Toko Putri Kosmetik. Analisis ini dilakukan dengan mengevaluasi data penjualan selama periode tertentu. Data penjualan tersebut kemudian dianalisis menggunakan algoritma K-Means Clustering. Algoritma digunakan untuk mengelompokkan produk berdasarkan karakteristik data penjualan, sehingga dapat membantu Toko Putri Kosmetik.

4. Pengolahan Data

Data penjualan produk kecantikan Toko Putri diolah dan dianalisis menggunakan metode K-Means Clustering dalam software RapidMiner. Proses ini bertujuan untuk mengelompokkan produk berdasarkan pola penjualan. Data diolah secara sistematis dengan software RapidMiner, menghasilkan pengelompokan produk yang menjadi dasar strategi persediaan stok yang tepat dan memaksimalkan keuntungan Toko Putri. Data penjualan produk kecantikan Toko Putri diolah dan dianalisis menggunakan metode K-Means Clustering dalam software RapidMiner. Proses ini bertujuan untuk mengelompokkan produk berdasarkan pola penjualan. Data diolah secara sistematis dengan software RapidMiner.

5. Pengujian Data Menggunakan Software RapidMiner

Peneliti melakukan pengujian data penjualan kosmetik yang telah diolah dengan software RapidMiner. Pengujian menggunakan algoritma K-Means Clustering. Peneliti melakukan pengujian data penjualan kosmetik yang telah diolah dengan software RapidMiner. Pengujian ini menggunakan algoritma K-Means Clustering untuk mengelompokkan produk berdasarkan pola penjualan. Tujuannya adalah untuk mengidentifikasi produk terlaris, laris, dan kurang laris, sehingga Toko Putri Kosmetik dapat menentukan strategi persediaan stok yang tepat dan memaksimalkan keuntungan.

6. Pengambilan Kesimpulan

Dalam proses KDD, tahap interpretasi memegang peran penting. Pada tahap ini, pola informasi hasil penambangan data perlu diinterpretasikan dan ditampilkan dalam format yang mudah dipahami. Fokus utamanya adalah menarik kesimpulan dari hasil pengolahan data menggunakan software RapidMiner. Pola informasi hasil penambangan data perlu diinterpretasikan dan ditampilkan dalam format yang mudah dipahami. Tahap ini, yang dikenal ssebagai interpretasi dalam proses KDD, bertujuan untuk menarik kesimpulan dari hasil perhitungan dan pengolahan data sistem

menggunakan software RapidMiner. Informasi yang dihasilkan perlu dikomunikasikan secara efektif kepada pihak-pihak yang berkepentingan, sehingga mereka dapat memahami pola dan tren yang teridentifikasi, dan mengambil keputusan berdasarkan wawasan tersebut.

Sumber Data

Penelitian ini memanfaatkan data primer yang bersumber dari transaksi penjualan Toko Kecantikan Putri Rt.33/Rw.7 Desa Bakung Lor Kecamatan Jamblang Kabupaten Cirebon selama bulan Januari 2024. Data yang dianalisis terdiri dari 122 baris dengan 22 atribut, merepresentasikan informasi terkait produk, kategori produk, jumlah penjualan, dan karakteristik lainnya. Data transaksi ini menjadi sumber informasi berharga untuk memahami pola pembelian pelanggan dan merumuskan strategi bisnis yang lebih efektif bagi Toko Kecantikan Putri.

Tabel 1. Statistik Data Penjualan

Nama Field	Type	Keterangan
Paket lengkap kezia beauty ii gratis parfum & tas cantik	Numeric	Merupakan rangkaian perawatan wajah untuk kulit normal, kering dan sensitif
Paket lengkap cream luminous ii anti kusam & flek hitam ii glowing	Numeric	Memberikan kelembapan yang maksimal dari kandungan Anastatica dan regenerasi sel-sel kulit.
Toner glowing ii anti kusam ii flek hitam ii mencerahkan kulit wajah	Numeric	Mencerahkan kulit karena dapat memudahkan noda hitam pada kulit
Paket lengkap cream glowing beauty glow ii original ii glowingkan wajah	Numeric	Melembapkan kulit wajah
Cream sp hitam super uv whitening original satuan	Numeric	Membuat kulit halus, warna kulit merata wajah lebih alami.
Face toner ii brightening toner fresh ii original ii bpom	Numeric	Mengangkat sisa kotoran dari wajah
Serum glowing ii serum brightening ii	Numeric	Mencerahkan kulit, menyamarkan noda hitam, melembapkan kulit, melindungi

Nama Field	Type	Keterangan
bpom ii original		kulit dari radikal bebas, dan mencegah penuaan dini
Salep flek super quality (dosis tinggi) ii salep flek ii melasma	Numeric	Mengobati peradangan dan gatal pada kulit akibat berbagai penyebab seperti reaksi alergi, eksim dan psoriasis
Serum express glowing permanen dalam 7 hari ii serum glowing vit c	Numeric	Membantu melembapkan, mencerahkan, meratakan tekstur kulit, dan memberikan kilau alami pada kulit
Paket 5in 1 Theraskin II Paket Lengkap II BPOM II original	Numeric	Membuat kulit wajah menjadi lebih bersih cerah, mengurangi kekusaman karena hiperpigmentasi, dan mengatasi kerutan halus

Populasi dan Sampel

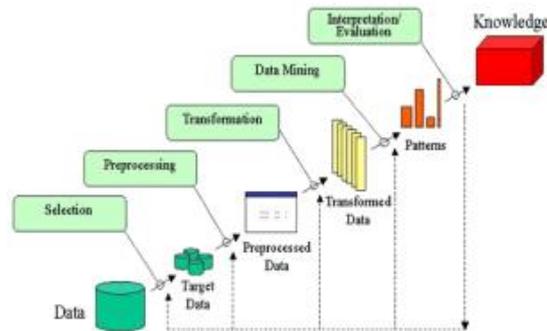
Populasi dalam penelitian ini adalah keseluruhan data penjualan pada toko kecantikan Putri, yang terdiri dari 326 record. Data ini mewakili semua transaksi penjualan yang pernah terjadi di toko tersebut. Di sisi lain, adalah bagian kecil dari populasi yang dipilih untuk mewakili keseluruhan. Sampel harus dipilih dengan cara yang representatif agar dapat mencerminkan karakteristik populasi secara akurat.

Teknik Pengumpulan Data

Dalam penelitian ini, teknik pengumpulan data yang digunakan adalah data sekunder. Data sekunder diperoleh dari sumber yang sudah ada, dalam hal data yang dikumpulkan berupa data dengan 326 record. Sampel data dari data penjualan di toko kecantikan Putri. Hasil dari pengumpulan data disimpan dalam bentuk format excel kemudian data di-import ke dalam Rapidminer.

Teknik Analisis Data

Dalam penelitian ini, metode analisis data yang digunakan adalah Knowledge Discovery in Database (KDD).



Gambar 2. Tahapan pada KDD

KDD merupakan sebuah proses yang bertujuan untuk menemukan pengetahuan baru dari data yang dikumpulkan. Proses KDD terdiri dari beberapa tahapan, yaitu:

1. Seleksi Data
Pada tahap ini, data yang akan dianalisis dipilih dan disiapkan. Peneliti perlu menentukan data mana yang relevan dengan penelitian dan data mana yang perlu dihapuskan.
2. Data Preprocessing
Data yang dipilih kemudian dibersihkan dan dipersiapkan untuk analisis. Hal ini termasuk menangani nilai yang hilang, inkonsistensi data, dan outlier.
3. Transformasi Data
Data yang telah dibersihkan kemudian ditransformasikan ke dalam format yang sesuai untuk analisis. Hal ini termasuk mengubah skala data, normalisasi data, dan reduksi dimensi.
4. Data Mining
Pada tahap ini, algoritma data mining digunakan untuk menemukan pola dan wawasan dari data.expand_more Algoritma yang digunakan dapat bervariasi tergantung pada jenis data dan tujuan penelitian.
5. Evaluasi
Hasil dari proses data mining kemudian diinterpretasikan dan dievaluasi. Peneliti perlu menentukan apakah pola yang ditemukan relevan dengan penelitian dan apakah pola tersebut dapat diaplikasikan dalam konteks nyata.

IV. HASIL DAN PEMBAHASAN

Hasil Penelitian Data

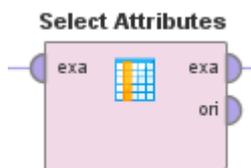
Penelitian ini memanfaatkan data sekunder yang diperoleh dari toko kecantikan putri. Data tersebut mencakup informasi mengenai penjualan produk di toko kecantikan putri dari bulan Desember 2023 hingga Februari 2024. Data yang dikumpulkan sebanyak 122 record dengan 22 atribut yang relevan. Data ini disimpan dalam format Excel dan diimpor ke dalam aplikasi RapidMiner. Atribut pada dataset dapat dilihat pada tabel 2 di bawah ini.

Tabel 2. Attribute Dataset

No	Attribute	Type
1.	Transaction Date	Nominal
2	Transaction Type	Nominal
3	Fee Name	Nominal
4	Transaction Number	Real
5	Details	Nominal
6	Seller SKU	Nominal
7	Lazada SKU	Nominal
8	Amount	Real
9	VAT in Amount	Real
10	WHT Amount	Real
11	WHT included in Amount	Nominal
12	Statement	Nominal
13	Paid Status Order No.	Nominal
14	Order Item No.	Real
15	Order Item Status	Real
16	Shipping Provider	Nominal
17	Shipping Speed	Nominal
18	Shipment Type	Nominal
19	Reference	Real
20	Comment	Nominal
21	PaymentRefId	Nominal

Data Selection

Pada tahap *data selection*, data yang telah dikumpulkan akan dipilih untuk kebutuhan analisis. *Dataset* awal memiliki 22 atribut. Untuk menyaring atribut yang akan digunakan, *operator Select Attributes* digunakan pada aplikasi RapidMiner. Dari 22 atribut dalam dataset penjualan Toko Kecantikan Putri, hanya 3 atribut yang dipilih melalui proses filtering ini: Details, Order No., dan Paid Status. Tampilan operator *Select Attributes* dapat dilihat pada gambar 3.



Gambar 3. Operator *Select Attributes*

Pada operator *Select Attributes* terdapat parameter yang perlu di sesuaikan. Parameter yang digunakan dapat dilihat pada tabel 3.

Tabel 3. Parameter *Select Attributes*

Parameter	Isi
Type	Include Attributes
Attribute filter type	A Subset
Select Attributes	Details, Order No, Paid Status

Data Preprocessing

Pada tahap *Data Preprocessing* dalam KDD, *missing value* pada data yang akan digunakan perlu

dihilangkan untuk mencegah kesalahan pada proses selanjutnya. *Missing value* dapat mengganggu proses *modeling* data mining selanjutnya, sehingga perlu ditangani dengan tepat. Pada *dataset* penjualan toko putri terdapat *missing value*, pada gambar 4 terdapat tampilan *missing value*.

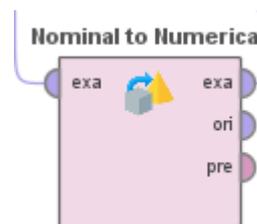
Gambar 4. *Missing Value Dataset*

Data penjualan Toko Putri memiliki 204 *missing value* yang perlu dihilangkan untuk mencegah terganggunya proses *modeling*. Gambar 5 menunjukkan hasil *dataset* setelah *missing value* dihilangkan.

Gambar 5. Dataset setelah dihilangkan *Missing Value*

Data Transformation

Proses *transformation* diperlukan pada data mining untuk mengubah tipe data agar sesuai dengan analisis yang diperlukan. dalam penelitian ini digunakan operator *nominal to numerical* untuk mengubah tipe data *nominal* menjadi *numerical*. tampilan operator *nominal to numerical* dapat dilihat pada gambar 6.



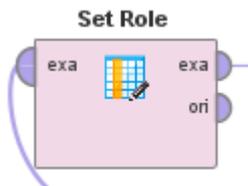
Gambar 6. Operator *Nominal to Numerical*

Pada operator *Nominal to Numerical* terdapat beberapa parameter, parameter yang digunakan dapat dilihat pada tabel 3.

Tabel 1. Parameter *Operator Nominal to Numerical*

Parameter	Isi
Attribute filter type	Subset : Details, Paid Status
Coding type	Unique Integers

Kategorisasi atribut dalam *dataset* dapat dilakukan menggunakan operator *Set Role* pada aplikasi RapidMiner, sesuai dengan kebutuhan analisis. Dalam penelitian ini, atribut "Order No." dikategorikan sebagai "ID". Tampilan operator *Set Role* dapat dilihat pada gambar 7.



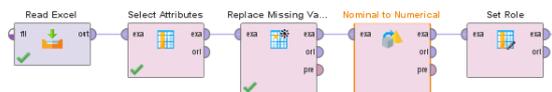
Gambar 7. Operator *Set Role*.

Pada operator *set role* terdapat beberapa parameter, pada tabel 4 ditunjukkan parameter yang digunakan.

Tabel 4. Operator *Set Role*

Parameter	Isi	
Set roles	Edit List :	
	Attribute Name :	Target role :
	Order No.	id

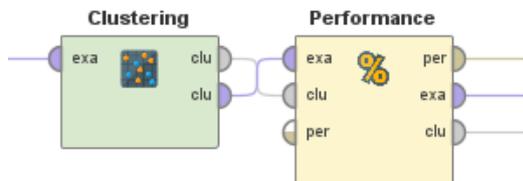
Proses permodelan pada tahap transformation di RapidMiner dapat dilihat pada gambar 7.



Gambar 7. Proses *Model Transformation*

Data Mining

Pada tahap data mining, teknik *clustering* yang diimplementasikan adalah algoritma *K-Means Clustering* menggunakan operator *Clustering*. Operator ini merupakan operator utama pemodelan untuk menghasilkan pengklasteran *dataset*. Tampilan operator *clustering* dapat dilihat pada gambar 8.



Gambar 8. Operator *Clustering*

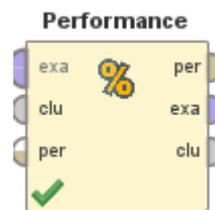
Pada operator *Clustering* menggunakan *K-Means* terdapat parameter yang perlu disesuaikan. Parameter yang digunakan dalam operator ini tampak pada tabel 5.

Tabel 5. Parameter Algoritma *K-Means*

Parameter	Isi
K	2-9
Measure Types	Numerical Measures, Bregman Divergences, Mixed Measures
Max Optimization	100

Selanjutnya ditambahkan operator *Performance* untuk mengukur *Cluster Distance Performance* dengan metode *Davies Bouldin Index* (DBI). Operator ini bertujuan untuk mengetahui nilai DBI dari proses *clustering* yang dilakukan.

Tampilan operator *Performance* dapat dilihat pada gambar 9.



Gambar 9. Tampilan Operator *Performance*

Dari hasil pembacaan operator *K-Means clustering* dengan menggunakan Parameter Measure Type Numerical Measures, Bregman Divergences, Mixed Measures, serta pembacaan operator *Cluster Distance Performance*, kemudian pada bagian *Main Criterion* pilih menggunakan *Davies Bouldin Index* (DBI), maka didapatkan informasi hasil DBI sebagai berikut:

Tabel 6. Parameter *Measure Types Numerical Measures*

Cluster	Measure Type	Davies Bouldin Index
2	Numerical Measures	0.162
3		0.225
4		0.207
5		0.184
6		0.146
7		0.104
8		0.077
9		0.092

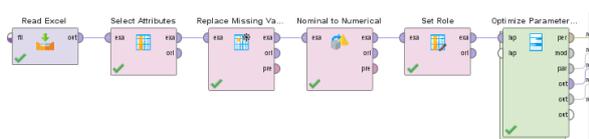
Tabel 7. Parameter *Measure Types Mixed Measures*

Cluster	Measure Type	Davies Bouldin Index
2	Mixed Measures	0.162
3		0.225
4		0.207
5		0.188
6		0.130
7		0.104
8		0.114
9		0.092

Tabel 8. Parameter *Measure Types Bregman Divergences*

Cluster	Measure Type	Davies Bouldin Index
2	Bregman Divergences	0.162
3		0.225
4		0.243
5		0.206
6		0.156
7		0.104
8		0.114
9		0.056

Proses pemodelan pada tahap data mining di RapidMiner, dapat dilihat pada gambar 10.



Gambar 10. Pemodelan Data Mining

Evaluasi

Berdasarkan hasil implementasi algoritma K-Means maka didapatkan hasil *performance* yang peneliti lakukan dengan Avg. within centroid distance: 0.164, Avg. within centroid distance_cluster_0: 0.187, Avg. within centroid distance_cluster_1: 0.121, Avg. within centroid distance_cluster_2: 0.346, Avg. within centroid distance_cluster_3: 0.111, Davies Bouldin Index: 0.243, *Measure type* yang digunakan *Numerical Measures*.

```

Performance:
PerformanceVector [
----Avg. within centroid distance: 0.071
----Avg. within centroid distance_cluster_0: 0.034
----Avg. within centroid distance_cluster_1: 0.346
----Avg. within centroid distance_cluster_2: 0.187
----Avg. within centroid distance_cluster_3: 0.111
****Davies Bouldin: 0.226
]
Clustering.k = 4
Clustering.measure_types = NumericalMeasures
Performance.main_criterion = Davies Bouldin
    
```

Gambar 11. Hasil *Performance Vektor*

Setelah model dibuat, langkah selanjutnya adalah mengeksekusi data dengan menggunakan RapidMiner. Hasil pengolahan K-Means Clustering dapat dilihat pada gambar 11. Gambar 11 menunjukkan bahwa terdapat 122 *record* data dalam analisis ini. Data tersebut dikelompokkan menjadi 4 *cluster*, *Cluster 0*: 24 *item*, *Cluster 1*: 44 *item*, *Cluster 2*: 18 *item*, *Cluster 3*: 36 *item*. Penentuan *cluster* dalam penelitian ini dilakukan berdasarkan karakteristik barang yang ada pada data penjualan

produk *skincare*. Pengelompokan ini bertujuan untuk membantu Toko Putri Kosmetik dalam data penjualan produk. *Cluster* dengan jumlah item terbanyak (*Cluster 3*).

Cluster Model

```

Cluster 0: 24 items
Cluster 1: 44 items
Cluster 2: 18 items
Cluster 3: 36 items
Total number of items: 122
    
```

Gambar 12. Anggota Cluster Model

Pembahasan

Pada tahap ini, setelah melalui proses analisis metode data mining menggunakan KDD dengan melakukan pengujian menggunakan algoritma K-means *clustering*, peneliti kemudian membahas secara menyeluruh hasil penelitian yang telah diuraikan sebelumnya. Hasil pengelompokan data Penjualan dengan model *Clustering* yang digunakan dievaluasi menggunakan Nilai DBI untuk mendapatkan nilai K yang paling optimal. Dengan demikian, dapat disimpulkan bahwa Nilai DBI yang optimal dapat dicapai dengan mencari nilai DBI yang mendekati nol. Analisis ini dilakukan dengan tujuan menjawab rumusan masalah yang telah ditetapkan pada awal penelitian.

Hasil analisis Measure Type, Mixed Measures, Bregman Divergences, dan Numerical Measures berdasarkan Davies Bouldin Index Optimal

Hasil analisis menentukan nilai K optimal dalam algoritma pengelompokan data. Dalam penelitian ini, peneliti melakukan analisis terhadap berbagai parameter pengukuran, termasuk *Mixed Measures*, *Bregman Divergences*, dan *Numerical Measures*, untuk mengidentifikasi parameter yang paling optimal dalam menentukan nilai K optimal berdasarkan *Davies Bouldin Index*, Temuan ini disajikan dalam tabel 9.

Tabel 9. Hasil analisis pengujian *cluster Measure Type*

Cluster	Measure Type	Davies Bouldin Index
8	MixedMeasures	0.077
	NumericalMeasures	0.114
	BregmanDivergences	0.114

Hasil Analisis Nilai K Berdasarkan Nilai Davies Bouldin Index Optimal

Hasil dari pengujian nilai K=2 hingga K=9, berdasarkan Davies Bouldin Index (DBI), diperoleh melalui eksperimen dengan rentang K=2 hingga K=9. Nilai DBI yang paling rendah, mendekati 0, menunjukkan kualitas cluster yang lebih baik dibandingkan dengan *cluster* yang memiliki nilai DBI yang lebih tinggi. Oleh karena itu, hasil

penelitian menunjukkan bahwa nilai $K=8$ menghasilkan DBI sebesar 0,077, yang dianggap sebagai nilai optimal. Dari hasil pengujian pengelompokan data dengan variasi jumlah *cluster* dari $k=2$ hingga $k=9$ DBI sebesar 0,077. Rincian hasil pengujian dapat ditemukan dalam tabel 10 yang terlampir.

Tabel 10. Hasil Analisis Davies Bouldin Index Optimal $k=2$ sampai $k=9$

Cluster	Measure Type	Davies Bouldin Index
2	MixedMeasures	0.162
	NumericalMeasures	0.162
	BregmanDivergences	0.162
3	MixedMeasures	0.225
	NumericalMeasures	0.225
	BregmanDivergences	0.225
4	BregmanDivergences	0.243
	NumericalMeasures	0.198
	MixedMeasures	0.207
5	MixedMeasures	0.184
	BregmanDivergences	0.188
	NumericalMeasures	0.184
6	NumericalMeasures	0.146
	MixedMeasures	0.146
	BregmanDivergences	0.178
7	NumericalMeasures	0.104
	MixedMeasures	0.104
	BregmanDivergences	0.153
8	MixedMeasures	0.077
	NumericalMeasures	0.114
	BregmanDivergences	0.114
9	MixedMeasures	0.092
	NumericalMeasures	0.092
	BregmanDivergences	0.092

V. KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil penelitian di atas dapat disimpulkan bahwa data penelitian berasal dari data transaksi penjualan Toko Kecantikan Putri Rt.33/Rw.7 Desa Bakung Lor Kecamatan Jamblang Kabupaten Cirebon.

1. Dalam penelitian ini, model algoritma k-means digunakan untuk mengakses dan mengelompokkan data tersebut. Pengujian dilakukan menggunakan model *clustering* k-means dan *cluster distance performance* untuk mengevaluasi hasil pengelompokan. Teridentifikasi 4 kelompok produk, yaitu

Cluster 0: 24 item, Cluster 1: 44 item, Cluster 2: 18 item, dan Cluster 3: 36 item, dengan *Cluster 3* memiliki jumlah item terbanyak dan nilai *Davies Bouldin Index* sebesar 0,243. Penentuan kelompok dilakukan berdasarkan karakteristik barang pada data penjualan produk *skincare*, yang memungkinkan pemilik untuk mengembangkan strategi penjualan dan pembelian ulang. Hasil pengujian dari nilai $K=2$ hingga $K=9$ menunjukkan bahwa nilai DBI terendah, mendekati 0, diperoleh saat $K=8$, menunjukkan kualitas *cluster* yang lebih baik. *Cluster* terbaik terdapat pada $K=8$ dengan nilai DBI sebesar 0,021, berasal dari berbagai jenis pengukuran seperti *Mixed Measures* (0,077), *Numerical Measures* (0,114), dan *Bregman Divergences* (0,114).

2. Berdasarkan hasil implementasi algoritma K-Means, peneliti berhasil mencapai hasil performa yang mengembirakan. Dalam proses ini, rata-rata jarak *centroid* dalam (*Avg. within centroid distance*) menunjukkan nilai yang cukup baik, dengan nilai keseluruhan sebesar 0.164. Lebih lanjut, evaluasi terhadap performa masing-masing *cluster* juga dilakukan, dengan hasil yang menarik seperti nilai rata-rata jarak *centroid* dalam untuk masing-masing *cluster*: *Cluster 0*: 0.187, *Cluster 1*: 0.121, *Cluster 2*: 0.346, dan *Cluster 3*: 0.111. Selain itu, *Davies Bouldin Index* (DBI) yang digunakan untuk mengukur kualitas *clustering* juga memberikan nilai yang memuaskan, yaitu sebesar 0.243. Penggunaan metode *Numerical Measures* dalam proses analisis metode data mining menggunakan KDD menambah keakuratan hasil penelitian. Dengan demikian, penelitian ini memberikan kontribusi yang berarti dalam pengembangan pemahaman terhadap model *clustering* dalam konteks penjualan produk *skincare* di Toko Kecantikan Putri Rt.33/Rw.7 Desa Bakung Lor Kecamatan Jamblang Kabupaten Cirebon.

Saran

Berdasarkan hasil penelitian ini, penulis merekomendasikan beberapa saran untuk penelitian lanjutan:

1. Melakukan penelitian serupa pada toko kecantikan lain atau jenis bisnis retail lainnya sangat dianjurkan untuk memperkuat temuan penelitian ini dan mengembangkan wawasan yang lebih luas tentang pola pembelian pelanggan di industri kecantikan dan retail. Penelitian ini dapat membantu mengidentifikasi pola dan tren yang mungkin tidak terlihat dalam satu toko saja, dan memungkinkan generalisasi temuan yang dapat diterapkan pada industri yang lebih luas. Selain itu, penelitian ini dapat meningkatkan pemahaman tentang faktor-faktor yang

mendorong pembelian pelanggan di berbagai jenis bisnis retail, yang dapat bermanfaat bagi bisnis dalam mengembangkan strategi pemasaran dan penjualan yang lebih efektif.

2. Selain algoritma K-means *clustering*, penelitian selanjutnya dapat mempertimbangkan penggunaan algoritma lain seperti pohon keputusan, jaringan saraf tiruan, dan algoritma regresi untuk menganalisis data penjualan Toko Kecantikan Putri. Masing-masing algoritma memiliki kelebihan dan kekurangannya sendiri, sehingga pemilihan algoritma yang tepat bergantung pada tujuan penelitian dan sifat data yang dianalisis. Penggunaan algoritma yang berbeda dapat memberikan wawasan yang lebih komprehensif tentang pola pembelian pelanggan dan membantu Toko Kecantikan Putri dalam meningkatkan strategi pemasaran dan penjualan mereka.

DAFTAR PUSTAKA

- Afiasari, Nur, et al. "Implementasi Data Mining Transaksi Penjualan Menggunakan Algoritma Clustering Dengan Metode K-Means." *Jurnal SAINTEKOM*, vol. 13, no. 1, 2023, pp. 100–10, <https://doi.org/10.33020/saintekom.v13i1.402>.
- Amin, Fadli, et al. "Penerapan Metode K-Means Dalam Penjualan Produk Souq.Com." *Applied Information System and Management (AISM)*, vol. 5, no. 1, 2022, pp. 7–14, <https://doi.org/10.15408/aism.v5i1.22534>.
- Anam, Khaerul, et al. "Analisis Segmentasi Pelanggan Menggunakan Metode K-Means Clustering." *Jurnal ICT: Information Communication & Technology*, vol. 21, 2022, pp. 273–78.
- Andi Cuhwanto, Yahya Novi, and Dewi Agushinta R. "Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K-Means." *Petir*, vol. 15, no. 1, 2021, pp. 48–56, <https://doi.org/10.33322/petir.v15i1.1358>.
- Aprilianur, Gita, and Elvin Leander Hadisaputro. "Penerapan Data Mining Menggunakan Metode K-Means Untuk Analisa Penjualan Toko Myam Hijab." *Jupiter*, vol. 14, no. 1, 2022, pp. 161–70.
- Astuti, Nur, et al. "Penerapan Data Mining Pada Penjualan Produk Digital Konter Leppangeng Cell Menggunakan Metode K-Means Clustering." *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 3, 2022, p. 754, <https://doi.org/10.30865/jurikom.v9i3.4351>.
- Dharma Putra, Yogiswara, et al. "Clustering History Data Penjualan Menggunakan Algoritma K-Means." *Majalah Ilmiah Teknologi Elektro*, vol. 20, no. 2, 2021, p. 195, <https://doi.org/10.24843/mite.2021.v20i02.p03>.
- Fahlevi, Muhammad Reza, et al. "Analisis Pengelompokan Data Pelelangan Barang Dengan Metode K-Means Clustering." *Jurasik (Jurnal Riset Sistem Informasi Dan Teknik Informatika)*, vol. 8, no. 1, 2023, pp. 53–61, <http://tunasbangsa.ac.id/ejurnal/index.php/jurasik/article/view/541>.
- Indriyani, Fintri, and Ani Irfiani. "Clustering Data Penjualan Pada Toko Perlengkapan Outdoor Menggunakan Metode K-Means." *JUITA: Jurnal Informatika*, vol. 7, no. 2, 2019, p. 109, <https://doi.org/10.30595/juita.v7i2.5529>.
- Mangku Negara, Iis Setyawan, et al. "Analisa Cluster Data Transaksi Penjualan Minimarket Selama Pandemi Covid-19 Dengan Algoritma K-Means." *JOINTECS (Journal of Information Technology and Computer Science)*, vol. 6, no. 3, 2021, p. 153, <https://doi.org/10.31328/jointecs.v6i3.2693>.
- Normah, et al. "Analisa Sentimen Perkembangan Vtuber Dengan Metode Support Vector Machine Berbasis SMOTE." *Jurnal Teknik Komputer AMIK BSI*, vol. 8, no. 2, 2022, pp. 174–80, <https://doi.org/10.31294/jtk.v4i2>.
- Noviati, Noviati, et al. "Clustering Data Penjualan Produk Makanan Pada Toko Toserba Yogya Siliwangi Dengan Menggunakan Metode K-Means." *MEANS (Media Informasi Analisa Dan Sistem)*, vol. 7, no. 1, 2022, pp. 77–84, <https://doi.org/10.54367/means.v7i1.1850>.
- Purnama, Chandra, et al. "Klasterisasi Penjualan Pakaian Untuk Meningkatkan Strategi Penjualan Barang Menggunakan K-Means." *Journal of Information Technology*, vol. 4, no. 1, 2022, pp. 35–38, <https://doi.org/10.47292/joint.v4i1.79>.
- Ramdhan, Dadan, et al. "Clustering Data Persediaan Barang Dengan Menggunakan Metode K-Means." *MEANS (Media Informasi Analisa Dan Sistem)*, vol. 7, no. 1, 2022, pp. 1–9, <https://doi.org/10.54367/means.v7i1.1826>.
- Sudrajat, Wahyu, et al. "Penerapan Algoritma K-Means Clustering Untuk Pengelompokan UMKM Menggunakan Rapidminer." *Jurnal JUPITER*, vol. 14, no. 1, 2022, pp. 27–36.
- Putri Trisnawati et al. "Penerapan Pengelompokan Produktivitas Hasil Pertanian Menggunakan Algoritma K-Means." *Infotek: Jurnal Informatika Dan Teknologi*, vol. 6, no. 2, 2023, pp. 249–57, <https://doi.org/10.29408/jit.v6i2.10198>.