

CLUSTERING DATA PENJUALAN PRODUK KESEHATAN MENGGUNAKAN ALGORITMA K-MEANS DI PT. CITRA TEKNIK MEDICA (STUDI KASUS: PT. CITRA TEKNIK MEDICA)

Husni Mubarak, Bambang Irawan, Tati Suprpti

Program Studi Teknik Informatika, STMIK IKMI Cirebon

Jalan Perjuangan No. 10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat, Indonesia

hmrempah299@gmail.com

ABSTRAK

Penawaran produk kesehatan merupakan bagian penting dari industri kesehatan yang terus berkembang seiring meningkatnya kesadaran masyarakat akan layanan medis. PT. Citra Teknik Medica sebagai salah satu pemain utama berperan dalam menyediakan berbagai produk dan jasa untuk mendukung fasilitas kesehatan di Indonesia. Dengan permintaan yang terus meningkat dari konsumen, penting bagi PT. Citra Teknik Medica untuk mengidentifikasi produk-produk yang paling diminati untuk mengoptimalkan strategi penjualan dan pengelolaan stok barang. Penelitian ini bertujuan untuk melakukan analisis terhadap data penjualan produk kesehatan yang paling diminati di PT. Citra Teknik Medica menggunakan Algoritma K-Means Clustering. Metode penelitian mengikuti proses Knowledge Discovery in Databases (KDD), meliputi seleksi data, preprocessing, transformasi data, penerapan Algoritma K-Means Clustering, dan evaluasi hasil menggunakan Davies Bouldin Index (DBI). Data yang digunakan adalah data penjualan produk kesehatan tahun 2018-2022 yang diperoleh dari PT. Citra Teknik Medica. Dari percobaan dengan variasi jumlah cluster (K) dari 2 hingga 7, ditemukan bahwa nilai Davies Bouldin Index (DBI) teroptimal dimiliki oleh cluster 4 (K=4) dengan nilai DBI 1.566.714.125.739.650.000 yang paling mendekati 0, menunjukkan pengelompokan produk kesehatan terbaik berdasarkan kuantitas penjualan dengan menggunakan 4 cluster.

Kata kunci : *Algoritma K-means, Knowledge Discovery in Databases (KDD), Analisis Data Penjualan Produk Kesehatan, Davies Bouldin Index (DBI)*

1. PENDAHULUAN

Penawaran produk kesehatan merupakan bagian dari bisnis kesehatan yang terus berkembang seiring dengan meningkatnya kesadaran masyarakat akan pentingnya layanan medis. PT Citra Teknik Medica sebagai salah satu pemain utama dalam industri ini memainkan peran penting dalam menyediakan berbagai item dan organisasi spesialis untuk peningkatan klinik darurat di seluruh Indonesia, terutama di bidang perolehan fasilitas gas klinis fokus dan hiasan, misalnya, Argon 6 M3, Asetilin 3 KG, Karbon Dioksida 20 kg, Liquid Oksigen, Nitrogen 6 M3, Nitrous oxide 20 KG, Nitrous oxide 25 kg, Oksigen 1 M3, Oksigen 6M3, Udara Tekan 6 M3, VGL Oksigen 160 LITER, Flowmeter with Humidifier 0-15 Lpm/ Dewasa Pahsco, ALARM GASES CONTROL 2 GAS CTM, Pigtail Check Valve to Kran, Regulator O2 H25DS/ 825-25 Bar, CENTRAL OXYGEN (O2) KAP 2 X 10 CYL (20 TABUNG) SYSTEM MANUAL, Air Compressor Regulator 1/2" AW 4000, Copper Tube Tembaga dia 8 x 2 mm, Twin Flowmeter Oxygen dan lain-lain yang mendukung upaya perawatan kesehatan yang berkualitas. Untuk memaksimalkan efisiensi dalam strategi penjualan, perusahaan perlu memanfaatkan sumber daya yang ada secara optimal. Salah satu pendekatan yang dapat digunakan adalah analisis *k-means*.

Penggunaan algoritma *k-means* telah menjadi trend utama dalam berbagai industri, termasuk industri kesehatan. Data penjualan produk kesehatan yang

dihasilkan oleh PT. Citra Teknik Medica mencakup berbagai aspek, seperti bulan, tahun, nama barang, satuan, kuantitas, penjualan dan faktor-faktor lain yang memengaruhi keputusan pembelian. Analisis data ini menjadi semakin penting untuk membantu perusahaan memahami perilaku pelanggan, tren penjualan, dan peluang bisnis yang dapat ditingkatkan.

Salah satu algoritma yang dilakukan peneliti dalam menganalisis data penjualan adalah algoritma *k-means*. Algoritma ini dikenal sebagai salah satu metode klastering yang digunakan dalam analisis data untuk mengelompokkan suatu set data menjadi beberapa kelompok berdasarkan kesamaan karakteristik. Algoritma ini bekerja dengan cara mengelompokkan data ke dalam (K) kelompok (klaster) berdasarkan jarak *euclidean* antara titik data dan pusat klaster. Proses ini melibatkan inisialisasi pusat klaster secara acak, perhitungan jarak, dan pembaruan pusat klaster hingga konvergensi. Dengan menerapkan algoritma *K-Means* pada data penjualan produk kesehatan di PT. Citra Teknik Medica, perusahaan dapat mengidentifikasi segmen pelanggan berdasarkan pola pembelian produk kesehatan. Dengan memahami preferensi dan perilaku pembelian, perusahaan dapat menyusun strategi pemasaran yang lebih tepat sasaran untuk setiap kelompok pelanggan. Dalam konteks ini, penelitian ini bertujuan untuk melakukan analisis terhadap data penjualan produk kesehatan yang paling diminati di PT. Citra Teknik Medica menggunakan algoritma *K-Means*. Hasil dari

penelitian ini diharapkan dapat memberikan wawasan yang berharga bagi perusahaan dalam mengoptimalkan pengelolaan stok, distribusi produk, strategi penjualan, meningkatkan kepuasan pelanggan, dan mengidentifikasi peluang bisnis yang belum terggali.

Metode-metode yang tersedia dalam data mining untuk tujuan prediksi mencakup metode *Knowledge Discovery in Databases* (KDD) adalah suatu pendekatan sistematis untuk mengekstrak pengetahuan yang berharga dari data besar atau basis data yang kompleks. Proses *Knowledge Discovery in Databases* (KDD) melibatkan beberapa tahapan yang membentuk suatu siklus, dan sering kali terkait erat dengan penggunaan teknik analisis data, statistika, dan machine learning. Berikut adalah tahapan umum dalam proses *Knowledge Discovery in Databases* (KDD) yaitu prediktif dan deskriptif Selain itu, metode *K-Means* juga merupakan salah satu pendekatan yang relevan. merupakan metode yang digunakan untuk melakukan pengelompokan atau klasterisasi data. Tujuannya adalah untuk membagi sekumpulan data menjadi beberapa kelompok-kelompok atau klaster berdasarkan kemiripan karakteristik tertentu yang disebut klaster. Salah satu keunggulan yang dimiliki oleh metode *K-Means* adalah sederhana, mampu mengolah kumpulan data dalam jumlah besar, menghasilkan kumpulan *cluster* yang stabil, dapat melakukan generalisasi *cluster* dengan berbagai bentuk dan ukuran, dapat adaptasi data yang baru.

PT. CITRA TEKNIK MEDICA adalah organisasi yang memiliki pengalaman praktis dalam menyediakan barang-barang dan administrasi yang diperlukan dalam pengembangan klinik medis di seluruh Indonesia. Terutama di bidang akuisisi perusahaan dan perangkat keras Focal Gas Klinis. Selain itu, organisasi ini juga dinamis dalam administrasi *top off Gas Klinis* berbasis ritel di seluruh Wilayah Showcase, yang mencakup seluruh wilayah Jawa Barat dan mencakup klinik yang berbeda, mulai dari Kelas A hingga Kelas B. Melihat besarnya permintaan yang terus meningkat dari konsumen terhadap produk-produk ini, maka sangat penting untuk melakukan *clustering* terhadap produk yang paling diminati oleh konsumen. *Clustering* ini memiliki tujuan utama untuk memberikan kemudahan bagi tim pengelola stok barang di PT. Citra Teknik Medica dalam perencanaan penyediaan stok barang. Selain itu, *clustering* ini juga berfungsi untuk memberikan informasi lanjut kepada organisasi dalam hal barang yang paling sering dibeli oleh pembeli. Mengingat gambaran dasar ini, sebuah penelitian diusulkan dengan judul "*Clustering Data Penjualan Produk Kesehatan Menggunakan Algoritma K-Means Di PT. Citra Teknik Medica*".

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Dalam penelitian berjudul "*Clustering Data Penjualan pada Toko Perlengkapan Outdoor*

Menggunakan Metode K-Means," Fintri Indriyani dan Eni Irfian membahas cara membuat rekomendasi untuk barang yang diminati secara signifikan, kurang diminati, atau keduanya. Tujuan utama penelitian ini adalah memberikan panduan kepada manajemen dalam mengelola stok barang, sehingga dapat mencegah kekecewaan pelanggan karena barang yang diinginkan tidak tersedia.[1]

Dalam Penelitian berjudul "*Analisa Cluster Data Transaksi Penjualan Minimarket Selama Pandemi Covid- 19 dengan Algoritma K-means*" oleh Iis Setiawan Mangku Negara, dkk. membahas penggunaan perhitungan pengelompokan k-means untuk memecah informasi pertukaran penawaran di Pasar Berkah Abadi Tegal menjadi kelompok yang lebih kecil. Hasil perhitungan ini memberikan tiga jenis kelompok informasi transaksi, yaitu barang dengan jumlah transaksi sedikit, sedang, dan banyak, masing-masing direpresentasikan oleh kelompok 0, kelompok 1, dan kelompok 2. Kelompok 2 terdiri dari 57 produk dengan transaksi dan total penjualan terbanyak, sementara kelompok 1 terdiri dari 5 produk dengan transaksi dan total penjualan sedang. Sisanya adalah produk dengan total penjualan yang lebih rendah. Model pengelompokan k-means ini juga mencapai tingkat ketepatan sebesar 87%.[2]

Dalam penelitian berjudul "*Clustering History Data Penjualan Menggunakan Algoritma K-Means*" yang dilakukan oleh Yogiswara Dharma Putra, dkk. dipaparkan perhitungan K-Means pada 23 informasi transaksi stok barang dagangan yang telah selesai dari tahun 2019 hingga 2020. Fokus klasifikasi produk mencakup PC, PC, dan printer, dan hasilnya menghasilkan tiga kelompok yang dijelaskan sebagai sepenuhnya menarik, memikat, dan kurang memikat. Dengan mempertimbangkan centroid mendasari, khususnya C1 (800, 580, 200), C2 (500, 450, 300), dan C3 (100, 190, 95), penelitian ini menyimpulkan bahwa kelompok 1 (benar-benar memikat) terdiri dari 5 barang, kelompok 2 (menarik) terdiri dari 4 barang, dan kelompok 3 (kurang memikat) terdiri dari 14 barang dengan titik centroid terakhir masing-masing C1 (911.8, 853, 58.8), C2 (442.5, 426.25, 16.25), dan C3 (97.85, 86, 11.85).[3]

Dalam penelitian berjudul "*Klasterisasi Penjualan Paicaian Untuk Meningkatkan Strategi Penjualan Barang Menggunakan K-Means*" yang dilakukan oleh Chandra Purnama, dkk. tujuan dari penelitian ini adalah untuk memberikan suatu bentuk rekomendasi berupa nama produk, jenis produk, jumlah produk yang terjual, bulan dimana produk tersebut paling banyak terjual, dan warnanya berdasarkan data yang digunakan dalam proses clustering. Pada penelitian ini, metode K-Means digunakan untuk meningkatkan strategi penjualan.[4]

Dalam penelitian berjudul "*Penerapan Metode K-Means dalam Penjualan Produk Souq.Com*" yang dilakukan oleh Fadli Amin, dkk. menggunakan teknik k-means untuk membagi produk menjadi dua kelompok, yaitu yang laris dan yang tidak laris. Hasil

penelitian menunjukkan bahwa produk makanan dan minuman termasuk dalam kategori laris, sementara produk perawatan kecantikan termasuk dalam kategori tidak laris. Analisis tersebut melibatkan pengelompokan hingga penekanan kelima, di mana titik tengah tetap konstan dan tidak ada perubahan informasi di antara kelompok. Hasil akhirnya menghasilkan tiga kelompok, yaitu kelompok produk paling laris (C1) dengan 8 produk, kelompok produk laris lainnya (C2) dengan 26 produk, dan kelompok produk kurang laris (C3) dengan 16 produk.[5]

Dalam penelitian berjudul “Analisis Pengelompokan Data Pelelangan Barang Dengan Metode K-Means Clustering” yang dilakukan oleh Muhammad Reza Fahlevi, dkk. digunakan teknik K-Means Clustering untuk mengumpulkan informasi penjualan yang tidak sepenuhnya ditentukan oleh beberapa aturan, terutama terfokus pada Kredit Tunai dan Biaya Penjualan. Hasil akhir dari pengujian strategi K-Means Clustering dan penerapannya pada aplikasi menunjukkan keseragaman hasil dalam mengumpulkan informasi penjualan, terutama untuk jenis kredit kecil yang terdiri dari 17 informasi pelanggan, jenis kredit sedang yang terdiri dari 5 informasi pelanggan, dan jenis kredit besar yang terdiri dari 2 informasi pelanggan.[6]

Dalam penelitian berjudul “Penerapan data mining metode k-means clustering untuk analisa penjualan pada toko fashion hijab Banten” yang dilakukan oleh Normah, dkk. data yang menjadi fokus adalah data stok pakaian dari bulan Juli 2019 hingga September 2019 di toko Helai. Data tersebut mencakup aspek-aspek seperti stok awal, stok terjual, dan stok terakhir, yang akan menjadi subjek pembahasan dalam ulasan ini.[7]

Dalam penelitian berjudul “Clustering Data Penjualan Produk Makanan pada Toko Toserba Yogya Siliwangi dengan Menggunakan Metode K-Means” yang dilakukan oleh Noviaty, dkk. membahas hasil penelitian di Toserba Yogya Siliwangi dengan menggunakan data penjualan produk makanan kategori snack pada bulan Desember 2021. Penelitian ini melibatkan 470 data dan menerapkan metode clustering, khususnya K-Means melalui aplikasi RapidMiner. Hasilnya menunjukkan bahwa produk makanan kategori snack dapat dibagi menjadi tiga cluster, yaitu cluster_1 dengan jumlah item sebanyak 3, cluster_2 dengan jumlah item sebanyak 13, dan cluster_0 dengan jumlah item sebanyak 454. Davies Bouldin juga menunjukkan hasil running sebesar - 0.487, mendekati angka 1.[8]

Dalam penelitian yang berjudul “Clustering Data Persediaan Barang dengan Menggunakan Metode K-Means” yang dilakukan oleh Dadan Ramdhan, dkk.

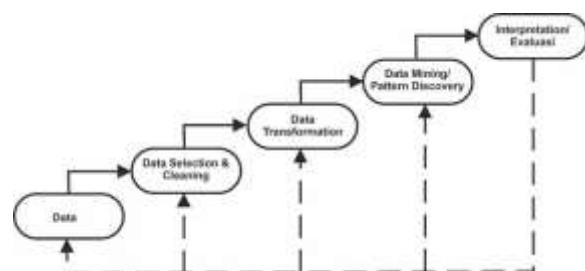
hasil penelitian diperoleh dengan membagi data menjadi dua kelompok: kelompok dengan jumlah cluster paling sedikit dan kelompok dengan jumlah cluster paling banyak. Oleh karena itu, dapat disimpulkan bahwa pengelompokan persediaan barang dengan menggunakan strategi K-Means terbukti sangat efektif dalam memperoleh nilai tipikal di dalam jarak dan kerumitan waktu.[9]

Dalam penelitian yang berjudul “Implementasi Data Mining Transaksi Penjualan Menggunakan Algoritma Clustering dengan Metode K-Means” yang dilakukan oleh Nur Afiasari, dkk. Membahas konsekuensi dari transaksi pertukaran informasi di Toko Lampu Bill di Gg. 5 Utara No.65 RW. 2 Karangampel Kidul, Karangampel, Kabupaten Indramayu, Jawa Barat 45283, dibahas secara rinci. Penelitian ini memanfaatkan data penjualan sebanyak 160 produk, dengan model algoritma K-Means digunakan untuk mengakses dataset. Pengelompokan k-implisit dataset dilakukan, dan eksekusi jarak berkelompok digunakan untuk menguji hasil pengelompokan. Hasil eksplorasi menunjukkan adanya 3 kelompok yang secara khusus dapat diidentifikasi sebagai menarik, kurang menarik, dan tidak menarik. Informasi lebih lanjut mengungkapkan bahwa Cluster_0 dianggap sebagai barang yang dapat dipasarkan dengan skor 14.131, Cluster_2 dianggap sebagai barang yang kurang menarik dengan skor 20.583, dan Cluster_1 dianggap sebagai barang yang tidak dapat dipasarkan dengan skor 34.241.[10]

3. METODE PENELITIAN

3.1. Knowledge Discovery in Database

Knowledge Discovery in Database (KDD) adalah metode yang digunakan untuk dapat memperoleh pengetahuan yang berasal dari database yang ada. Hasil pengetahuan yang diperoleh dapat dimanfaatkan untuk basis pengetahuan (knowledge base) yang digunakan dalam keperluan mengambil keputusan. Berikut adalah langkah-langkah Metode Knowledge Discovery in Database (KDD) dalam penelitian ini:



Gambar 1. Alur KDD

Tabel 1. Tahapan KDD

Tahapan	Aktivitas	Deskripsi Aktivitas
Selection	Seleksi Data	Tahapan ini dilakukan untuk meningkatkan akurasi dan kualitas hasil data mining dengan mengidentifikasi atribut yang akan digunakan dalam analisis klasterisasi data penjualan produk kesehatan di PT. Citra Teknik Medica menggunakan Algoritma K-Means.

Tahapan	Aktivitas	Deskripsi Aktivitas
<i>Preprocessing Data</i>	Pembersihan Data	Tahapan ini akan dilakukan untuk pembersihan data karena data yang diperoleh masih berupa data mentah atau data yang nilainya hilang atau tidak valid.
<i>Transformation Data</i>	Mengubah Data	Tahapan ini dilakukan untuk mengubah data agar sesuai dengan format yang mendukung dalam proses pengolahan data <i>mining</i> .
<i>Data Mining</i>	Penerapan Metode Data Mining	Tahap data <i>mining</i> adalah inti dari proses <i>Knowledge Discovery in Database</i> (KDD). Untuk penelitian ini menggunakan metode <i>Clustering</i> dengan Algoritma <i>K-Means</i> .
<i>Interpretation/ Evaluation</i>	Evaluasi	Tahapan ini akan menganalisis hasil klusterisasi guna memahami faktor-faktor yang memiliki dampak pada penelitian. Selain itu, akan dilakukan evaluasi performa model <i>Clustering</i> dengan Algoritma <i>K-Means</i> menggunakan <i>Davies Bouldin Indeks</i> (DBI), dengan tujuan untuk mengukur efektivitas pengelompokan dengan mengukur kuantitas dan kualitas data yang dikumpulkan.

3.2. Sumber Data

Penelitian ini menggunakan data sekunder yang diperoleh dari sumber selain objek atau subjek penelitian itu sendiri. Data yang digunakan merupakan informasi transaksi berupa angka-angka akurat untuk item-item Kesehatan dalam periode 2018-2022 yang diperoleh secara langsung dari PT Citra Teknik Medica guna mendapatkan informasi yang unik dan bersifat subjektif.

3.3. Populasi Dan Sampel

Dalam penelitian ini, populasi yang diteliti adalah keseluruhan data penjualan produk kesehatan di PT. Citra Teknik Medica. Sebagai sampel, diambil 735 data penjualan produk kesehatan dari populasi tersebut. Data sampel diperoleh melalui wawancara dengan pihak HRD, Akuntansi, dan Marketing PT. Citra Teknik Medica. Data sampel yang dikumpulkan mencakup informasi meliputi Bulan, Tahun, Nama Barang, Satuan, Kuantitas, dan Penjualan.

3.4. Teknik Analisis Data

Teknik analisis data yang digunakan adalah menerapkan Algoritma *K-Means Clustering* untuk mengelompokkan data penjualan produk kesehatan yang terjual paling banyak, dengan tahapan meliputi pengumpulan data penjualan, preprocessing data, seleksi atribut relevan, implementasi Algoritma *K-Means Clustering* dengan penentuan jumlah cluster yang sesuai, melakukan clustering produk berdasarkan karakteristik serupa menggunakan *K-Means*, dan mengevaluasi performa algoritma dengan menghitung nilai *Davies Bouldin Index* (DBI).

4. HASIL DAN PEMBAHASAN

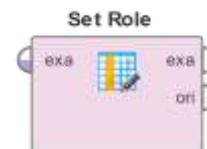
4.1. Data

Data yang digunakan ini data tahun 2018-2022 dengan jumlah 735 record dan 6 atribut diantaranya yaitu bulan, tahun, nama barang, satuan, kuantitas dan penjualan.



Gambar 2. Operator read Excel

Operator “Read Excel” pada *RapidMiner* digunakan untuk mengimpor data dari *file excel* ke dalam proses analisis *RapidMiner*. Fungsi utamanya adalah membaca data yang disimpan dalam format excel dan mengonversinya menjadi dataset yang dapat digunakan dalam proses analisis data. Dataset yang digunakan adalah dataset penjualan produk kesehatan di PT. Citra Teknik Medica.



Gambar 3. Operator Set Role

Set Role, fitur ini berfungsi untuk membedakan baris bernama atribut koordinat dan prediksi posisi yang akan dimasukkan ke dalam kategori ‘label’. Agar pada saat pengkategorian data ‘label’ tidak ikut serta terhitung dan merubah hasil. Untuk attribute yang digunakan untuk label dapat dilihat pada gambar 4. dibawah.



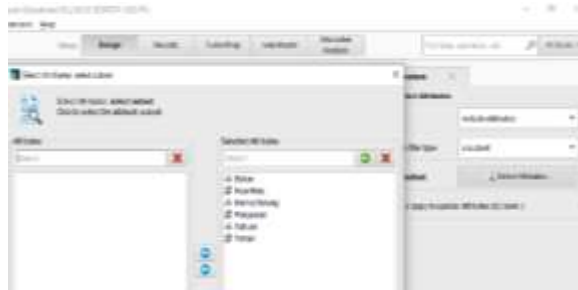
Gambar 4. Attribute Name



Gambar 5. Operator Select Attribute

Select Attributes, operator ini digunakan untuk memfilter data apa saja yang akan digunakan dalam pemrosesan data. *Attribut* dalam dataset penjualan produk kesehatan ada 6 *attribut* tidak ada *pemfilteran*

attribut, attribut yang digunakan tetap 6 *attribut* yang sudah ada. *Attribut* yang digunakan yaitu bulan, kuantitas, nama barang, penjualan, satuan dan tahun. Untuk proses pemilihan attribut dapat dilihat pada Gambar 6. berikut.



Gambar 6. Select Attribute

Setelah tahapan proses seleksi data dilakukan dapat ditentukan bahwa attribut yang digunakan untuk proses *clustering* menggunakan algoritma *K-Means* yaitu nama barang sebagai id dan bulan, satuan, tahun, kuantitas, penjualan sebagai attribut. Dapat dilihat pada Gambar 7. berikut.

No	Nama Barang	bulan	satuan	tahun	kuantitas	penjualan
1	Aspirin 0.001	cluster_0	0	2020	10	100000
2	Aspirin 0.001	cluster_0	0	2020	10	100000
3	Aspirin 0.001	cluster_0	0	2020	10	100000
4	Aspirin 0.001	cluster_0	0	2020	10	100000
5	Aspirin 0.001	cluster_0	0	2020	10	100000
6	Aspirin 0.001	cluster_0	0	2020	10	100000
7	Aspirin 0.001	cluster_0	0	2020	10	100000
8	Aspirin 0.001	cluster_0	0	2020	10	100000
9	Aspirin 0.001	cluster_0	0	2020	10	100000
10	Aspirin 0.001	cluster_0	0	2020	10	100000
11	Aspirin 0.001	cluster_0	0	2020	10	100000
12	Aspirin 0.001	cluster_0	0	2020	10	100000
13	Aspirin 0.001	cluster_0	0	2020	10	100000
14	Aspirin 0.001	cluster_0	0	2020	10	100000
15	Aspirin 0.001	cluster_0	0	2020	10	100000
16	Aspirin 0.001	cluster_0	0	2020	10	100000
17	Aspirin 0.001	cluster_0	0	2020	10	100000
18	Aspirin 0.001	cluster_0	0	2020	10	100000
19	Aspirin 0.001	cluster_0	0	2020	10	100000
20	Aspirin 0.001	cluster_0	0	2020	10	100000
21	Aspirin 0.001	cluster_0	0	2020	10	100000
22	Aspirin 0.001	cluster_0	0	2020	10	100000
23	Aspirin 0.001	cluster_0	0	2020	10	100000
24	Aspirin 0.001	cluster_0	0	2020	10	100000
25	Aspirin 0.001	cluster_0	0	2020	10	100000
26	Aspirin 0.001	cluster_0	0	2020	10	100000
27	Aspirin 0.001	cluster_0	0	2020	10	100000
28	Aspirin 0.001	cluster_0	0	2020	10	100000
29	Aspirin 0.001	cluster_0	0	2020	10	100000
30	Aspirin 0.001	cluster_0	0	2020	10	100000
31	Aspirin 0.001	cluster_0	0	2020	10	100000
32	Aspirin 0.001	cluster_0	0	2020	10	100000
33	Aspirin 0.001	cluster_0	0	2020	10	100000
34	Aspirin 0.001	cluster_0	0	2020	10	100000
35	Aspirin 0.001	cluster_0	0	2020	10	100000
36	Aspirin 0.001	cluster_0	0	2020	10	100000
37	Aspirin 0.001	cluster_0	0	2020	10	100000
38	Aspirin 0.001	cluster_0	0	2020	10	100000
39	Aspirin 0.001	cluster_0	0	2020	10	100000
40	Aspirin 0.001	cluster_0	0	2020	10	100000
41	Aspirin 0.001	cluster_0	0	2020	10	100000
42	Aspirin 0.001	cluster_0	0	2020	10	100000
43	Aspirin 0.001	cluster_0	0	2020	10	100000
44	Aspirin 0.001	cluster_0	0	2020	10	100000
45	Aspirin 0.001	cluster_0	0	2020	10	100000
46	Aspirin 0.001	cluster_0	0	2020	10	100000
47	Aspirin 0.001	cluster_0	0	2020	10	100000
48	Aspirin 0.001	cluster_0	0	2020	10	100000
49	Aspirin 0.001	cluster_0	0	2020	10	100000
50	Aspirin 0.001	cluster_0	0	2020	10	100000
51	Aspirin 0.001	cluster_0	0	2020	10	100000
52	Aspirin 0.001	cluster_0	0	2020	10	100000
53	Aspirin 0.001	cluster_0	0	2020	10	100000
54	Aspirin 0.001	cluster_0	0	2020	10	100000
55	Aspirin 0.001	cluster_0	0	2020	10	100000
56	Aspirin 0.001	cluster_0	0	2020	10	100000
57	Aspirin 0.001	cluster_0	0	2020	10	100000
58	Aspirin 0.001	cluster_0	0	2020	10	100000
59	Aspirin 0.001	cluster_0	0	2020	10	100000
60	Aspirin 0.001	cluster_0	0	2020	10	100000
61	Aspirin 0.001	cluster_0	0	2020	10	100000
62	Aspirin 0.001	cluster_0	0	2020	10	100000
63	Aspirin 0.001	cluster_0	0	2020	10	100000
64	Aspirin 0.001	cluster_0	0	2020	10	100000
65	Aspirin 0.001	cluster_0	0	2020	10	100000
66	Aspirin 0.001	cluster_0	0	2020	10	100000
67	Aspirin 0.001	cluster_0	0	2020	10	100000
68	Aspirin 0.001	cluster_0	0	2020	10	100000
69	Aspirin 0.001	cluster_0	0	2020	10	100000
70	Aspirin 0.001	cluster_0	0	2020	10	100000
71	Aspirin 0.001	cluster_0	0	2020	10	100000
72	Aspirin 0.001	cluster_0	0	2020	10	100000
73	Aspirin 0.001	cluster_0	0	2020	10	100000
74	Aspirin 0.001	cluster_0	0	2020	10	100000
75	Aspirin 0.001	cluster_0	0	2020	10	100000

Gambar 7 Hasil Seleksi Attribute

4.2. Data Transformation

Pada tahap data *transformation* ini dilakukan untuk merubah format *type attribut*, dikarenakan pada penelitian tidak ada data yang diperlu dirubah maka pada tahap ini tidak digunakan, proses dapat dilanjutkan ke tahap selanjutnya.

4.3. Data Mining



Gambar 8. Operator K-Means Clustering

Algoritma *K-Means Clustering* digunakan pada tahap data mining pada proyek tugas akhir ini. Algoritma *K-Means Clustering* digunakan untuk melakukan pengelompokan pada dataset penjualan produk kesehatan ke dalam jumlah *cluster* terbaik dengan nilai DBI (*Davies Bouldin Index*) teroptimal.



Gambar 9. Parameters K-Means Clustering

Berdasarkan Gambar 9. penerapan parameter jumlah sebanyak 7 ($k=7$) dengan 10 kali iterasi mengubah jumlah kluster (k). Pada *parameter measure type* diterapkan *Numerical Measure* dan menggunakan *Eucludin Distance*.

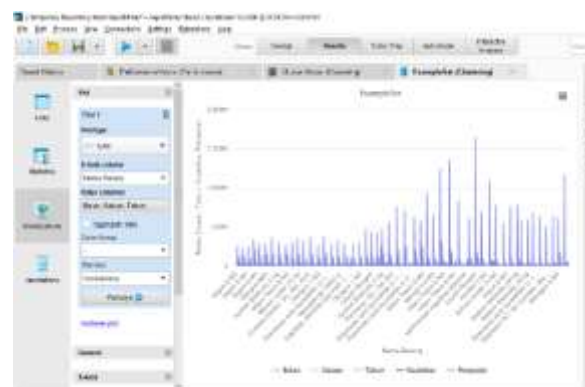
4.4. Interpretation/Evaluation



Gambar 10. Performance

Proses evaluasi pada penelitian ini digunakan operator *Performance* (*Cluster Distance Performance*). Operator ini memiliki 2 jenis kriteria pada parameternya, yaitu *Average Within Centroid Distance* dan juga *Davies Bouldin Index* (DBI) dan juga menerapkan *parameter Maximize* agar dapat dihasilkan nilai DBI non-negatif (≥ 0).

Setelah proses pemilihan informasi, penentuan informasi, dan perubahan informasi pada dataset, informasi tersebut dapat ditangani dalam proses penggalan informasi dengan menggunakan perhitungan *K-Means Clustering*.



Gambar 11. Visualisasi Hasil Klasterisasi

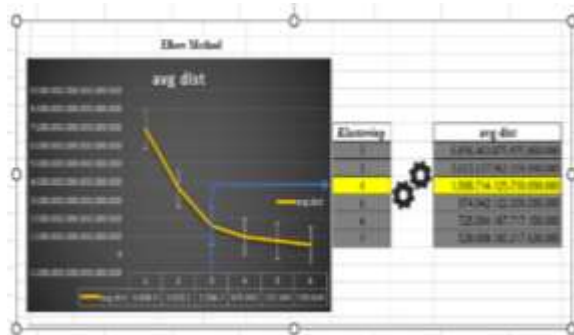
Pada parameter operator *K-Means* telah diterapkan jumlah kluster (K) sebanyak 7. Penerapan ini menghasilkan klasterisasi dataset penjualan produk kesehatan di PT. Citra Teknik Medica terbagi kedalam 7 kluster dengan jumlah items yang berbeda.

Tabel 2. Jumlah Item setiap Cluster

No.	Cluster	Jumlah Items
1.	Klaster 0	630
2.	Klaster 1	4
3.	Klaster 2	10
4.	Klaster 3	15
5.	Klaster 4	49
6.	Klaster 5	1
7.	Klaster 6	26
Total Number Of Items		735

Berdasarkan tabel 2. di jelaskan items yang masuk kedalam klaster 0 sebanyak 630 items, kemudian items yang masuk kedalam klaster 1 sebanyak 4, items yang masuk kedalam klaster 2 sebanyak 10, items yang masuk kedalam klaster 3 sebanyak 15, item yang masuk kedalam klaster 4 sebanyak 49, items yang masuk kedalam klaster 5 sebanyak 1, items yang masuk kedalam klaster 6 sebanyak 26, jadi jumlah total *number of items* adalah 735 items.

Penerapan operator *performance* (*Cluster Distance Performance*) untuk tahap evaluasi pada *RapidMiner* memiliki salah satu output berupa *Davies Bouldin Index*. Dilakukan 7 kali percobaan untuk menemukan nilai *Davies Bouldin Index* (DBI) teroptimal, dimulai dari penerapan ke dalam klaster berjumlah 2 ($K = 2$) sampai penerapan ke dalam klaster berjumlah 7 ($K = 7$). Percobaan ini dilakukan untuk mencari jumlah klaster terbaik berdasarkan nilai *Davies Bouldin Index* (DBI) nya.



Gambar 12. Avg Dist

Berdasarkan informasi pada gambar 12. setelah dilakukan percobaan untuk menentukan jumlah klaster dari $K = 2$ sampai $K = 7$, gambar avg dist pada siku line ditentukan bahwa nilai *Davies Bouldin Index* (DBI) teroptimal dimiliki oleh klaster 4 ($K = 4$) dengan nilai *Davies Bouldin Index* (DBI) yang paling mendekati 0 sebesar 1.566.714.125.739.650.000.

5. KESIMPULAN DAN SARAN

Kesimpulan yang diperoleh dari penelitian data penjualan produk kesehatan yaitu pemilihan data-data yang digunakan ini data tahun 2018-2022 dengan jumlah 735 record. Visualisasi Hasil Klasterisasi Pada parameter operator K-Means telah diterapkan jumlah klaster (k) sebanyak 7. Peneliti mendalami 7 klaster dengan jumlah item yang berbeda. Percobaan ini

dilakukan untuk mencari jumlah klaster terbaik berdasarkan nilai Davies Bouldin Index (DBI) nya. Setelah dilakukan percobaan untuk menentukan jumlah klaster dari $K = 2$ sampai $K = 7$, gambar rata-rata dist pada garis siku ditentukan bahwa nilai Davies Bouldin Index (DBI) teroptimal yang dimiliki oleh klaster 4 ($K = 4$) dengan nilai Davies Bouldin Index (DBI) yang paling mendekati 0 sebesar 1.566.714.125.739.650.000. Adapun saran dari penulis untuk penelitian selanjutnya adalah analisis dapat lebih dikembangkan menggunakan algoritma yang lainnya seperti algoritma Apriori, X-Means. Atau menggunakan metode analisis software yang lainnya.

DAFTAR PUSTAKA

- [1] F. Indriyani and E. Irfiani, "Clustering Data Penjualan pada Toko Perlengkapan Outdoor Menggunakan Metode K-Means," *JUITA J. Inform.*, vol. 7, no. 2, p. 109, 2019, doi: 10.30595/juita.v7i2.5529.
- [2] I. S. Mangku Negara, P. Purwono, and I. A. Ashari, "Analisa Cluster Data Transaksi Penjualan Minimarket Selama Pandemi Covid-19 dengan Algoritma K-means," *JOINTECS (Journal Inf. Technol. Comput. Sci.)*, vol. 6, no. 3, p. 153, 2021, doi: 10.31328/jointecs.v6i3.2693.
- [3] Y. Dharma Putra, M. Sudarma, and I. B. A. Swamardika, "Clustering History Data Penjualan Menggunakan Algoritma K-Means," *Maj. Ilm. Teknol. Elektro*, vol. 20, no. 2, p. 195, 2021, doi: 10.24843/mite.2021.v20i02.p03.
- [4] C. Purnama, W. Witanti, and P. Nurul Sabrina, "Klasterisasi Penjualan Pakaian untuk Meningkatkan Strategi Penjualan Barang Menggunakan K-Means," *J. Inf. Technol.*, vol. 4, no. 1, pp. 35–38, 2022, doi: 10.47292/joint.v4i1.79.
- [5] F. Amin, D. S. Anggraeni, and Q. Aini, "Penerapan Metode K-Means dalam Penjualan Produk Souq.Com," *Appl. Inf. Syst. Manag.*, vol. 5, no. 1, pp. 7–14, 2022, doi: 10.15408/aism.v5i1.22534.
- [6] M. R. Fahlevi, D. Ridha, D. Putri, and E. Syahrin, "Analisis Pengelompokan Data Pelelangan Barang Dengan Metode K-Means Clustering," *Jurasik (Jurnal Ris. Sist. Inf. dan Tek. Inform.)*, vol. 8, no. 1, pp. 53–61, 2023, [Online]. Available: <http://tunasbangsa.ac.id/ejurnal/index.php/jurasik/article/view/541>
- [7] A. S. Normah, Siti Nurajizah, "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.
- [8] N. Novianti, M. Mulyawan, D. A. Kurnia, and A. R. Rinaldi, "Clustering Data Penjualan Produk Makanan pada Toko Toserba Yogya Siliwangi dengan Menggunakan Metode K-Means," *MEANS (Media Inf. Anal. dan Sist.)*, vol. 7, no. 1,

- pp. 77–84, 2022, doi: 10.54367/means.v7i1.1850.
- [9] D. Ramdhan, G. Dwilestari, R. D. Dana, A. Ajiz, and K. Kaslani, “Clustering Data Persediaan Barang Dengan Menggunakan Metode K-Means,” *MEANS (Media Inf. Anal. dan Sist.*, vol. 7, no. 1, pp. 1–9, 2022, doi: 10.54367/means.v7i1.1826.
- [10] N. Afiasari, N. Suarna, and N. Rahaningsi, “Implementasi Data Mining Transaksi Penjualan Menggunakan Algoritma Clustering dengan Metode K-Means,” *J. SAINTEKOM*, vol. 13, no. 1, pp. 100–110, 2023, doi: 10.33020/saintekom.v13i1.402.