

PENGELOMPOKAN TRANSAKSI PENJUALAN AKSESORIS HP DAN PULSA DENGAN METODE K-MEANS UNTUK MENINGKATKAN STRATEGI PEMASARAN DI TOKO BAGUS CELLULER

Muhammad Syam Al ghifari ¹, Martanto ², Umi Hayati ³

^{1,3} Teknik Informatika, ² Manajemen Informatika

Sekolah Tinggi Manajemen Informatika dan Komputer IKMI Cirebon

Jl. Perjuangan No.10B, Karyamulya, Kec. Kesambi, Kota Cirebon, Jawa Barat 45135, (0231) 490480.

syamalghifari7@gmail.com

ABSTRAK

Dalam bidang Informatika, menghadapi tantangan penting dalam pengklasifikasian transaksi bisnis menjadi sebuah isu yang mendesak, khususnya bagi sektor ritel seperti Bagus Celluler. Isu utama yang muncul adalah kerumitan dan jumlah data yang bertambah seiring kemajuan teknologi. Studi ini diarahkan untuk mengeksplorasi dan mengimplementasikan metode klusterisasi data dengan menggunakan algoritma K-Means, khususnya pada transaksi penjualan aksesoris telepon seluler dan voucher pulsa di Bagus Celluler. Tujuan klusterisasi ini adalah untuk menemukan tren pembelian yang seragam di antara berbagai segmen konsumen. Data yang terkumpul dari Bagus Celluler akan ditelaah dengan teliti melalui algoritma K-Means untuk mengkategorikan konsumen berdasarkan pola pembelian mereka. Studi kasus ini dilaksanakan dengan mengambil contoh Bagus Celluler, yang mencerminkan dinamika pasar aksesoris telepon seluler dan voucher pulsa di wilayah spesifik. Dengan penerapan algoritma K-Means, diharapkan temuan dari studi ini akan menyediakan wawasan strategis yang bermanfaat bagi Bagus Celluler untuk meningkatkan strategi pemasaran mereka, dengan menyesuaikan penawaran produk dan jasa sesuai dengan kecenderungan konsumen yang beragam. Studi ini bertujuan untuk memberikan pemahaman yang lebih dalam mengenai klusterisasi dalam sektor ritel, terutama pada pasar aksesoris telepon seluler dan voucher pulsa. Selain itu, studi ini juga menunjukkan bagaimana analisis data dapat dimanfaatkan untuk memperbaiki strategi pemasaran dan meningkatkan kepuasan konsumen. Berdasarkan hasil klusterisasi, jumlah kelompok optimal untuk data transaksi penjualan aksesoris telepon seluler dan voucher pulsa dengan menggunakan algoritma K-Means dan indeks Davies Bouldin adalah tiga, dengan nilai Davies Bouldin sebesar 0.205. Nilai maksimum iterasi yang menghasilkan kluster terbaik untuk data transaksi ini, yang berkisar dari satu hingga sepuluh dan dihitung secara numerik, menunjukkan konsistensi nilai DBI sebesar -0.205 untuk kluster K=3. Hasil studi ini diharapkan dapat menjadi fondasi dalam merancang strategi pemasaran yang lebih efisien dan sesuai dengan setiap segmen konsumen yang telah diidentifikasi.

Kata kunci : Pengelompokan data, K-Means, analisis transaksi penjualan, preferensi konsumen, strategi pemasaran, aksesoris HP, pulsa, studi kasus

1. PENDAHULUAN

Di zaman yang ditandai dengan kemajuan teknologi yang cepat, Informatika telah menjadi pilar penting dalam evolusi berbagai bidang kehidupan. Inovasi berkelanjutan dalam teknologi informasi telah merevolusi cara kita berkomunikasi, bekerja, dan mengelola bisnis. Dampak dari inovasi ini terasa luas, mulai dari evolusi model bisnis hingga perubahan dalam pendidikan dan layanan publik. Dalam ranah teknologi, kemajuan di bidang kecerdasan buatan, cloud computing, dan Internet of Things (IoT) telah memodifikasi cara kita memanfaatkan dan menyatukan teknologi dalam kehidupan kita sehari-hari. Sementara itu, dunia bisnis mengalami transformasi digital, dengan e-commerce sebagai salah satu contoh paling menonjol dari pengaruh teknologi informasi terhadap operasional perusahaan dan interaksi dengan konsumen. Dalam konteks ini, toko-toko seperti Bagus Celluler dihadapkan pada tantangan dan kesempatan baru dalam menyesuaikan strategi pemasaran mereka agar tetap relevan dan kompetitif di pasar yang terus berubah. Karena itu, penelitian ini

dirancang untuk meneliti penggunaan metode K-Means dalam klasifikasi transaksi penjualan aksesoris telepon seluler dan voucher pulsa di Bagus Celluler, dengan tujuan untuk memberikan wawasan yang dapat meningkatkan strategi pemasaran dalam industri yang dinamis ini.

Dalam lingkup Informatika, tantangan signifikan yang dihadapi adalah pengelompokan transaksi bisnis, yang sangat relevan bagi sektor ritel, termasuk toko Bagus Celluler. Toko tersebut menghadapi beberapa isu yang memerlukan solusi. Salah satu isu utama adalah peningkatan kompleksitas dan volume data yang berkembang pesat sejalan dengan kemajuan teknologi. Dalam operasionalnya, toko mungkin mengalami kesulitan dalam menganalisis dan memahami pola pembelian pelanggan dengan efektif. Di samping itu, persaingan pasar yang meningkat menuntut pemanfaatan data transaksi yang optimal untuk memperkuat strategi pemasaran dan mempertahankan daya saing. Kendala praktis seringkali muncul dari kurangnya pemahaman tentang pengelolaan dan penerapan metode klusterisasi data

yang sesuai. Ini menuntut pengembangan metode yang cerdas dan efisien dalam pengelompokan transaksi penjualan, yang akan memberikan wawasan berharga untuk manajemen toko. Mengatasi masalah ini menjadi krusial, sebab dapat membantu toko Bagus Celluler dan bisnis serupa untuk meningkatkan efisiensi operasional. Penyelesaian masalah ini diharapkan akan membuka peluang baru dalam memaksimalkan penggunaan data transaksi untuk keuntungan bisnis. Dengan demikian, toko Bagus Celluler dapat mengidentifikasi peluang untuk inovasi dan pertumbuhan dalam pasar yang kompetitif. Akhirnya, solusi yang ditemukan akan memungkinkan toko untuk menyediakan layanan yang lebih baik dan memenuhi kebutuhan pelanggan secara lebih efektif.

Studi-studi yang telah dilakukan sebelumnya telah mengeksplorasi dimensi pengelompokan dalam transaksi penjualan dan preferensi konsumen di berbagai sektor. Misalnya, penelitian oleh [1], menunjukkan bagaimana metode K-Means dapat digunakan untuk mengklasifikasikan konsumen di sektor ritel. Hasil mereka menunjukkan bahwa pengelompokan tersebut dapat membantu dalam mengidentifikasi tren pembelian yang dapat dimanfaatkan untuk memperkuat strategi pemasaran. Akan tetapi, penelitian tersebut tidak menyoroti sektor aksesoris telepon seluler dan pulsa, yang memiliki ciri khas tersendiri baik dari sisi konsumen maupun produk. Penelitian lain yang tidak disebutkan namanya, meneliti tentang preferensi konsumen dalam industri aksesoris teknologi, termasuk aksesoris telepon seluler, namun tidak memanfaatkan metode pengelompokan. Hasil penelitian ini menekankan bahwa memahami preferensi konsumen merupakan kunci dari strategi pemasaran yang efektif. Namun, masih terdapat ruang untuk penelitian lebih mendalam di area spesifik ini, seperti di pasar aksesoris telepon seluler dan pulsa di toko Bagus Celluler. Penelitian oleh [2] menerapkan algoritma K-Means untuk mengelompokkan produk unggulan PT. Koko Pelli. Penelitian ini menyoroti pentingnya bagi perusahaan untuk mengenali pola pembelian konsumen dan mengelompokkan produk sesuai dengan preferensi dan kebutuhan mereka. Hasil penelitian ini mungkin termasuk penilaian terhadap efektivitas algoritma K-Means dalam mengelompokkan produk berdasarkan kriteria seperti kejelasan kelompok, homogenitas, dan relevansi dengan kebutuhan bisnis. Meskipun penelitian ini telah menggunakan algoritma K-Means, masih terbuka peluang untuk penelitian lebih lanjut, seperti mengintegrasikan data tambahan seperti demografi atau perilaku pembelian untuk meningkatkan presisi pengelompokan. Penelitian mendatang juga bisa mempertimbangkan penggunaan algoritma pengelompokan lain atau pengembangan algoritma baru yang disesuaikan dengan kebutuhan bisnis tertentu. Terakhir penelitian oleh [3], meneliti pengaruh analisis data transaksi terhadap peningkatan strategi pemasaran. Studi ini menegaskan pentingnya pemasaran berbasis data dalam dunia bisnis yang

modern. Walaupun literatur yang ada telah membahas aspek-aspek penting dari pengelompokan pelanggan dan analisis preferensi, penelitian yang berfokus pada transaksi penjualan aksesoris telepon seluler dan pulsa, terutama yang menggunakan metode K-Means, masih jarang. Oleh karena itu, penelitian ini bertujuan untuk melengkapi literatur yang ada dan memberikan wawasan lebih dalam mengenai cara-cara mengoptimalkan strategi pemasaran dalam konteks yang spesifik ini.

Inti dari penelitian ini adalah penerapan metode K-Means untuk menganalisis data transaksi penjualan aksesoris telepon dan voucher di Bagus Celluler, dengan penekanan pada segmentasi konsumen. Tujuan utamanya adalah menemukan segmen konsumen dengan preferensi pembelian yang serupa untuk produk aksesoris telepon dan voucher. Pentingnya penelitian ini terletak pada upayanya untuk menutup celah dalam literatur yang ada, yang belum sepenuhnya menjelajahi pemanfaatan metode K-Means dalam konteks bisnis tertentu ini. Penelitian ini juga bertujuan untuk memberikan wawasan yang lebih dalam tentang bagaimana strategi pemasaran dapat dioptimalkan dalam bisnis aksesoris telepon dan voucher, dengan pendekatan yang lebih terarah dan disesuaikan dengan preferensi pelanggan. Hasil penelitian ini diharapkan dapat memberikan keuntungan praktis bagi Bagus Celluler dan bisnis serupa, dengan menyediakan panduan strategis yang lebih efektif untuk bersaing di pasar ritel teknologi yang sangat kompetitif. Kesimpulannya, penelitian ini diharapkan dapat memberikan sumbangan signifikan dalam bidang Informatika dan pemasaran, dengan mengisi kekosongan dalam pemahaman tentang analisis data konsumen dalam konteks yang sangat spesifik ini.

Penelitian ini akan menerapkan teknik K-Means untuk mengevaluasi data transaksi dari penjualan aksesoris telepon dan pulsa di Bagus Celluler. Teknik ini akan digunakan untuk mengkategorikan pelanggan menurut pola pembelian mereka. Data yang dikumpulkan akan diteliti dengan teliti menggunakan algoritma K-Means. Selanjutnya, penelitian ini akan menggunakan metode analisis data yang maju untuk menemukan tren pembelian yang konsisten di antara kelompok-kelompok pelanggan yang terbentuk. Pendekatan eksperimental ini akan membantu dalam memahami kecenderungan pelanggan dan merancang strategi pemasaran yang lebih personal. Metode ini diharapkan dapat memberikan wawasan yang lebih luas mengenai perilaku konsumen dalam pasar aksesoris telepon dan pulsa, serta strategi untuk meningkatkan efektivitas pemasaran.

Penelitian ini diharapkan memberikan dampak yang berarti dalam bidang Informatika dan industri aksesoris telepon serta voucher. Dengan berhasilnya penelitian ini, akan tersedia wawasan yang lebih dalam mengenai penerapan metode K-Means dalam menganalisis data transaksi penjualan dan segmentasi konsumen untuk kasus tertentu. Hal ini akan

memperkaya pemahaman mengenai cara praktis menggunakan metode segmentasi dalam industri ritel teknologi, serta memberikan kejelasan lebih lanjut tentang preferensi konsumen, tren pembelian, dan perilaku konsumen dalam pasar aksesoris telepon dan voucher. Praktisi di bidang ini akan mendapat keuntungan dari panduan strategis yang lebih efektif untuk pengambilan keputusan pemasaran dan penawaran produk yang lebih sesuai dengan keinginan pelanggan. Temuan dari penelitian ini juga diharapkan dapat memberikan perspektif baru bagi peneliti lain dalam Informatika dan mungkin akan mendorong inovasi dalam teknologi pintar untuk analisis data konsumen, sehingga memungkinkan pemahaman dan respons pasar yang lebih efisien—suatu aspek krusial di era digital. Secara keseluruhan, penelitian ini berpotensi memberikan kontribusi penting dalam memahami analisis data konsumen dan penerapannya dalam bisnis aksesoris telepon dan voucher, serta memberikan manfaat praktis bagi stakeholder di sektor ini

2. TINJAUAN PUSTAKA

2.1. Penelitian terdahulu

Hasil tinjauan literatur yang telah dilakukan pada topik “Pengelompokan Transaksi Penjualan Aksesoris Hp & Pulsa Dengan K-Means Untuk Strategi Pemasaran Toko Bagus Celuler” dapat dirangkum sebagai berikut:

Jurnal [1] yang berjudul " Penerapan K-Means pada Pengelompokan Penjualan Produk Smartphone" oleh [4], membahas pentingnya para pengembang untuk menemukan strategi atau suatu pola yang dapat meningkatkan penjualan dan pemasaran produk, seperti halnya pada PT. Vivo Communication Indonesia cabang Pematangsiantar. Salah satu strateginya adalah memanfaatkan data transaksi dengan melakukan pengelompokan untuk melihat produk mana yang lebih laku dipasaran dan mana yang tidak, agar dapat dilakukan evaluasi dalam perencanaan promosi produk Vivo selanjutnya. Metode pengelompokan pada penelitian ini menggunakan metode Clustering K-Means. Metode K-Means adalah metode data mining yang mampu mengelompokkan beberapa data menjadi bagian-bagian tertentu. Pada makalah ini data penjualan yang telah diperoleh akan dibagi menjadi 3 kelompok yaitu penjualan rendah, penjualan sedang dan penjualan tinggi. Berdasarkan hasil Pengujian menggunakan metode K-Means terhadap data penjualan smartphone Vivo, diperoleh bahwa penjualan kelompok tertinggi hanya ada 1 data, yakni Vivo Y12 3+32GB. Sehingga dapat disimpulkan bahwa metode K-Means dapat diterapkan untuk melakukan pengelompokan penjualan smartphone Vivo, karena sesuai dengan hasil penjualan yang sebenarnya.

Jurnal [2] yang berjudul “Penerapan Algoritma K-Means Pada Penjualan Saldo Transportasi Online Studi Kasus Konter XYZ” oleh [5]. Penelitian ini bertujuan untuk mengelompokkan penjualan saldo

transportasi online agar dapat memberikan masukan khususnya dalam strategi promosi konter XYZ. Metode yang diterapkan adalah klasterisasi menggunakan algoritma K-Means untuk menghasilkan kelompok berdasarkan nominal harga pembelian. Awalnya, dataset terdiri dari 4.867 entri, tetapi setelah dilakukan pembersihan data, tersisa 2.642 entri. Data ini diperoleh dari penjualan konter XYZ selama bulan Oktober 2022. Hasil penelitian menunjukkan bahwa terdapat dua kelompok utama, yaitu Cluster C1 dengan jumlah 2.499 data atau sekitar 94,59% dari total, dan Cluster C2 dengan jumlah 143 data atau sekitar 5,41%. Cluster C1 memiliki nominal pembelian yang rendah, sementara Cluster C2 memiliki nominal pembelian yang tinggi. Temuan ini mengindikasikan bahwa algoritma K-Means efektif digunakan untuk mengelompokkan penjualan saldo transportasi online di konter XYZ, dan juga menunjukkan bahwa penjualan saldo transportasi online cukup menggembirakan karena mayoritas data berada pada kelompok dengan nominal pembelian rendah.

Jurnal [3] berjudul “Penerapan Metode K-Means Clustering Pada Data Penjualan Optik Chantika” oleh [6]. Tujuan dari penelitian ini adalah untuk melakukan analisis data mining menggunakan metode clustering K-Means pada data penjualan merek kacamata di Optik Chantika. Evaluasi performa metode K-Means dilakukan menggunakan nilai Davies Bouldin Index, dan hasil optimal menunjukkan jumlah klaster 10. Namun, penelitian ini memilih untuk menggunakan tiga klaster untuk klasterisasi: penjualan tertinggi, penjualan sedang, dan penjualan terendah. Melalui metode K-Means, merek kacamata dengan penjualan tertinggi selama tiga bulan tergabung dalam Cluster_0, sedangkan Cluster_1 berisi merek kacamata dengan penjualan sedang, dan Cluster_2 berisi merek kacamata dengan penjualan terendah. Hasil penelitian ini menunjukkan bahwa 50% merek kacamata termasuk dalam kategori penjualan tertinggi, 27,27% termasuk dalam kategori penjualan sedang, dan 22,72% termasuk dalam kategori penjualan terendah. Dengan adanya pengelompokan berdasarkan pola penjualan ini, Optik Chantika dapat mengoptimalkan pengelolaan data penjualan dan memenuhi kebutuhan konsumen dengan lebih baik. Dengan pemahaman yang lebih baik tentang pola penjualan, perusahaan dapat mengambil keputusan yang lebih efektif dalam hal pengadaan stok dan strategi pemasaran. Diharapkan hal ini dapat meningkatkan pendapatan perusahaan dan memberikan manfaat yang signifikan bagi Optik Chantika.

Jurnal [4] berjudul “Penerapan Data Mining Pada Penjualan Produk Digital Konter Leppangeng Cell Menggunakan Metode K-Means Clustering” oleh [7]. Metode K-Means dapat diterapkan pada Konter Leppangeng Cell untuk menganalisis penjualan produk digital yang laris dan tidak laris. Penerapan metode K-Means di Konter Leppangeng Cell melibatkan pengelompokan data stok produk digital

dan memilih tiga kelompok acak sebagai centroid awal. Setelah data tiap kelompok stabil, hasil pengolahan data menunjukkan bahwa dari total 133 produk yang dianalisis, 114 di antaranya terjual dengan baik, 5 produk kurang laku, dan 14 produk tidak laku. Selanjutnya, penerapan metode K-Means dilakukan menggunakan perangkat lunak RapidMiner dengan memasukkan data stok produk awal, stok terjual, dan stok akhir yang akan diubah menjadi kelompok. Setelah proses ini selesai, RapidMiner akan menghasilkan informasi mengenai produk yang diminati secara tinggi dan rendah berdasarkan analisis K-Means. Dengan demikian, melalui penerapan metode K-Means ini, Konter Leppangeng Cell dapat mendapatkan wawasan yang lebih baik mengenai penjualan produk digital mereka, memungkinkan mereka untuk mengoptimalkan strategi pemasaran dan manajemen stok guna meningkatkan penjualan dan mengurangi produk yang tidak laku.

Jurnal [5] berjudul “Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Rotan” oleh [8]. Analisis diperlukan untuk memahami pola penjualan produk agar dapat menghasilkan prediksi penjualan yang akurat. Penggalan informasi dari data berskala besar dapat dilakukan menggunakan teknik data mining, dengan metode clustering sebagai dasar pengambilan keputusan dalam proses prediksi. Hasil pengujian data mining ini dapat dibagi menjadi tiga kelompok utama: Tidak Laku, Sedang Laku, dan Laris. Setelah pengujian, hasil kluster menunjukkan bahwa: Cluster 0 memiliki nilai sebesar 4,636,996,106,454.166, Cluster 1 memiliki nilai sebesar 1,088,684,561,455.892, Cluster 2 memiliki nilai sebesar 4,624,429,728,751.111. Dari hasil kluster tersebut, diketahui bahwa: Kelompok Tidak Laku terdiri dari 6 data, Kelompok Sedang Laku memiliki 114 data, Kelompok Laris memiliki 9 data. Dengan informasi ini, perusahaan dapat mengidentifikasi pola penjualan yang berbeda-beda dan membuat keputusan strategis yang lebih tepat dalam mengelola produknya, seperti mengoptimalkan stok, menyesuaikan harga, atau meningkatkan promosi untuk meningkatkan penjualan produk yang kurang laku dan memperkuat strategi untuk produk yang laris.

Jurnal [6] berjudul “Mengidentifikasi Strategi Promosi pada Jasa Penjualan Saldo Digital menggunakan Pendekatan Clustering” oleh [9]. Tujuan dari penelitian ini adalah untuk mengidentifikasi pola penjualan saldo digital dan menyusun strategi promosi yang efektif untuk konter ZAR. Penelitian ini menggunakan metode clustering dengan Algoritma K-Means untuk mengelompokkan transaksi berdasarkan besaran nominal harga penjualan. Silhouette digunakan untuk menentukan jumlah cluster yang optimal. Dataset yang digunakan adalah data penjualan saldo digital konter ZAR dari bulan November 2021 hingga Oktober 2022. Hasil penelitian menunjukkan bahwa terdapat dua cluster utama yang dihasilkan. Cluster pertama terdiri dari penjualan dengan nominal harga rendah, sebanyak

49.213 data, sementara cluster kedua terdiri dari penjualan dengan nominal harga tinggi, sebanyak 3.076 data. Meskipun demikian, ditemukan bahwa ada beberapa data yang termasuk dalam cluster nominal harga tinggi. Dari hasil penelitian ini, dapat disimpulkan bahwa algoritma K-Means efektif digunakan untuk mengelompokkan penjualan saldo digital konter ZAR. Meskipun penjualan saldo digital konter ZAR tergolong baik, ada beberapa data yang masuk ke dalam cluster dengan nominal harga tinggi. Oleh karena itu, pemilik konter ZAR dapat merancang strategi promosi dengan memberikan potongan harga untuk produk dengan nominal harga di bawah 142.000. Hal ini diharapkan dapat mendorong pelanggan yang biasanya membeli dengan nominal harga tinggi untuk beralih ke produk dengan nominal harga rendah, sehingga meningkatkan penjualan secara keseluruhan.

Jurnal [7] berjudul “Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko OJ CELL” oleh [10]. Penelitian ini bertujuan untuk menghasilkan data yang dapat digunakan untuk keperluan pemasaran, yang diharapkan akan membantu pertumbuhan toko. Metode yang digunakan adalah Clustering K-Means untuk mengelompokkan jenis transaksi menjadi sangat laris, laris, dan kurang laris. Penelitian ini bertujuan untuk mengetahui provider atau paket data yang diminati oleh masyarakat sekitar Oj Cell. Dalam penelitian ini, juga digunakan aplikasi RapidMiner untuk memudahkan analisis data dan menghasilkan hasil yang lebih akurat. Dengan demikian, diharapkan penelitian ini akan memberikan wawasan yang berharga bagi Oj Cell dalam merancang strategi pemasaran yang lebih efektif, yang pada gilirannya akan membantu pertumbuhan dan kesuksesan toko mereka.

Jurnal [8] berjudul “Analisis Clustering Pengelompokan Penjualan Paket Data Menggunakan Metode K-Means” oleh [11]. Penelitian ini dilakukan untuk menganalisis pengelompokan data penjualan paket yang sering dibeli oleh pelanggan di counter menggunakan metode K-Means. Metode K-Means dipilih karena algoritma ini tidak terpengaruh oleh tatanan awal dari objek-objek yang digunakan. Hal ini terbukti ketika penulis mencoba menentukan cluster awal yang berpusat secara acak dari salah satu objek dalam perhitungan pertama penjualan paket data di counter. Variabel yang digunakan dalam penelitian ini antara lain Harga, Masa aktif, dan beberapa data lainnya mengenai paket. Algoritma K-Means Cluster Analysis pada dasarnya diterapkan pada masalah pemahaman kebutuhan konsumen, dengan tujuan mengidentifikasi jenis paket data yang sering dibeli oleh pelanggan. Algoritma K-Means dapat digunakan untuk menggambarkan karakteristik dari masing-masing kelompok dengan merangkum sejumlah besar objek, sehingga mempermudah pemahaman dan analisis lebih lanjut.

Jurnal [9] berjudul “Penentuan Segmentasi Konsumen Pada Marketing Data IFood Menggunakan Metode K – Means Clustering” oleh [12]. Berdasarkan informasi tersebut, dapat disimpulkan bahwa perusahaan IFood perlu melakukan segmentasi konsumen. Segmentasi ini akan dilakukan menggunakan metode K-Means Clustering. Pertama, variabel jumlah pembelian produk per kategori akan dikelompokkan menjadi 2 kluster. Kedua, variabel metode pembelian akan dikelompokkan menjadi 2 kluster. Terakhir, tingkatan pendidikan konsumen akan dikelompokkan menjadi 5 tingkatan. Dengan menerapkan metode K-Means Clustering, diharapkan perusahaan dapat lebih mudah menentukan rencana promosi atau pemasaran untuk kampanye berikutnya dengan tingkat keberhasilan mencapai 100%.

Jurnal [10] berjudul “Analisis Clustering Pada Pengguna Brand Hp Menggunakan Metode K-Means” oleh [13]. Tujuan dari penelitian ini adalah menerapkan proses k-means clustering pada data merek ponsel dan menerapkan algoritma k-means pada data penggunaan telepon seluler siswa di Perguruan Tinggi Bandar Lampung agar dapat dikluster menjadi 3 kelompok, yaitu kelompok "sangat baik", "baik", dan "cukup" dalam hal kebutuhan teknologi. Metode yang digunakan adalah algoritma k-means clustering. Hasil penelitian ini menunjukkan bahwa dari total 322 siswa yang disurvei, 120 siswa memiliki telepon seluler dengan kategori "sangat baik", 148 siswa memiliki telepon seluler dengan kategori "baik", dan 54 siswa memiliki telepon seluler dengan kategori "cukup". Hal ini mengindikasikan bahwa mayoritas siswa di Daerah Bandar Lampung menggunakan merek HP dengan kategori "baik" dalam memenuhi kebutuhan teknologi.

2.2. Machine Learning

Machine Learning, salah satu cabang ilmu komputer yang berkaitan dengan kecerdasan buatan, berfokus pada pembuatan sistem yang belajar dari data. Data adalah aspek kunci dari algoritma Machine Learning. Data pelatihan yang dibagi menjadi data pelatihan dan data pengujian digunakan untuk mengembangkan model, sedangkan data pengujian digunakan untuk mengevaluasi kinerja algoritma yang dilatih pada data baru. Dengan memanfaatkan data, Machine Learning meningkatkan kemampuannya untuk memahami pola, membuat prediksi, dan belajar dari pengalaman.[10]

2.3. Data Mining

Data mining merupakan suatu proses yang kompleks dengan menggunakan statistik, matematika, kecerdasan buatan, dan machine learning guna mengekstraksi informasi yang berharga dari database berskala besar. Informasi yang terhimpun diolah dalam database untuk mendukung pengambilan keputusan.[8] menguraikan tahapan-tahapan data mining, mencakup pembersihan data, integrasi data dari beragam sumber, pemilihan data yang relevan, transformasi data, proses mining, evaluasi pola, dan

presentasi pengetahuan. Tahap terakhir mencakup perumusan keputusan berdasarkan hasil analisis yang didapatkan.[14]

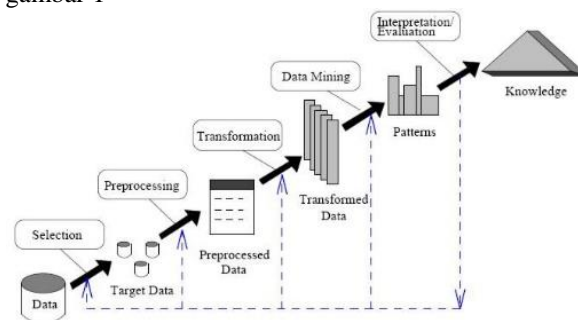
2.4. Clustering K-Means

Clustering K-Means adalah salah satu metode analisis data yang digunakan untuk mengelompokkan data menjadi beberapa kelompok yang saling berhubungan berdasarkan kesamaan atribut. Metode ini bekerja dengan cara mendefinisikan pusat kluster (centroid) secara acak, kemudian menghitung jarak antara setiap data dengan centroid tersebut. Setiap data akan dikelompokkan ke dalam kluster yang memiliki centroid terdekat. Setelah semua data terkelompok, akan dihitung kembali centroid baru berdasarkan rata-rata dari data dalam satu kluster.[4]

3. METODE PENELITIAN

3.1. Metode penelitian

Penelitian ini menerapkan metode Knowledge Discovery in Database (KDD) dengan pendekatan kuantitatif. Proses KDD merupakan serangkaian tahapan global dalam penggalian data yang melibatkan pengolahan informasi dari data yang besar dengan tujuan untuk mengeksplorasi dan menganalisis data tersebut sehingga menghasilkan informasi yang bermanfaat untuk pembentukan pengetahuan baru. Adapun tahapan metode penelitian yang dilakukan untuk menyusun penelitian ini dapat dilihat pada gambar 1



Gambar 1 Tahapan Metode Penelitian

3.2. Tahapan Metode Penelitian

Tabel di bawah ini memberikan penjelasan mengenai aktivitas dan deskripsi aktivitas dari setiap tahapan-tahapan yang telah divisualisasikan pada gambar 3. 1 diatas. Analisis lebih lanjut terkait tabel akan dijelaskan dalam bagian table 3.1.

Tabel 1 Tahapan Metode Penelitian

No	Tahapan	Aktivitas	Deskripsi Aktivitas
1	Selection	Memilih dan mengumpulkan data	Memilih dan mengumpulkan data yang relevan dengan tujuan penelitian. Proses ini melibatkan pemilihan sumber data yang sesuai dan pembatasan data yang diperlukan.

No	Tahapan	Aktivitas	Deskripsi Aktivitas
2	Preprocessing	Persiapan data sebelum proses analisis	Melibatkan langkah-langkah persiapan data sebelum proses analisis, seperti membersihkan data dari noise, mengisi nilai yang hilang, dan standarisasi data.
3	Transformation	Mengubah struktur atau format data agar sesuai dengan kebutuhan analisis	Mengubah struktur atau format data agar sesuai dengan kebutuhan analisis, termasuk pengkodean variabel, pembuatan variabel baru, atau transformasi lainnya.
4	Data Mining	Proses utama penggunaan algoritma K-Means Clustering	Proses utama penggunaan algoritma K-Means Clustering untuk menggali pola dan hubungan tersembunyi dalam data penjualan aksesoris hp dan pulsa.
5	Interpretation/Evaluation	Interpretation/Evaluation	Menganalisis hasil clustering dengan menggunakan Davies Bouldin Index (DBI) sebagai metode evaluasi untuk mengukur keberhasilan dan kualitas clustering.

3.3. Sumber Data

Sumber data yang digunakan dalam penelitian ini adalah data primer. Data primer merupakan data yang diperoleh secara langsung dari objek yang diteliti atau dari sumber lainnya. Dalam konteks penelitian ini, sumber data primer berasal dari seluruh transaksi penjualan aksesoris HP dan pulsa di toko Bagus Celluler selama periode dari bulan Juli hingga Oktober, dengan total 1021 record. Data ini diperoleh langsung melalui wawancara dengan staf penjualan yang bertujuan untuk memperoleh informasi terkait dengan seluruh transaksi penjualan aksesoris HP dan pulsa di toko Bagus Celluler

3.4. Teknik Pengumpulan Data

Teknik pengumpulan data merupakan langkah yang sangat penting dalam. Penelitian. Pengumpulan data adalah prosedur yang sistematis dan standar untuk memperoleh data yang diperlukan. Tehnik pengumpulan data dalam penelitian ini dilaksanakan dalam dua cara yaitu:

a. Observasi

Observasi penelitian dilakukan dengan melakukan kunjungan langsung toko dan mengajukan izin kepada peilik toko Bagus Celluler. Izin tersebut diperlukan untuk mendapatkan data terkait transaksi penjualan aksesoris hp & pulsa. Data tersebut akan digunakan sebagai dataset untuk

proses analisis menggunakan algoritma K-Means Clustering dalam aplikasi RapidMiner versi 10.2.

b. Wawancara

Wawancara merupakan proses interaksi langsung antara peneliti dan responden dengan tujuan mendapatkan informasi secara langsung dari narasumber. Melalui pertanyaan-pertanyaan yang telah disiapkan sebelumnya. Dalam konteks penelitian ini, wawancara dilakukan dengan pemilik toko yaitu Bapak Bagus., yang bertujuan untuk memperoleh informasi terkait data terkait transaksi penjualan aksesoris hp & pulsa toko Bagus Celluler.

c. Dokumentasi

Dokumentasi adalah metode pengumpulan data yang melibatkan pengambilan foto, dokumen, dan tulisan. Tujuan dari dokumentasi ini adalah untuk memperoleh data atau informasi yang sesuai dengan variabel penelitian yang telah direncanakan sebelumnya

3.5. Teknik Analisis Data

Teknik analisa data K-Means merupakan salah satu metode data clustering non hirarki yang berusaha mempartisi data yang ada ke dalam bentuk satu atau lebih cluster/kelompok. Metode ini mempartisi data ke dalam cluster/kelompok sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu cluster yang sama dan data yang mempunyai karakteristik yang berbeda dikelompokkan ke dalam kelompok yang lain.

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

a. Tentukan Jumlah Kluster k:

(k) adalah jumlah kluster yang diinginkan.

b. Menentukan Titik Pusat Awal dari Setiap Kluster:

(C = { c₁, c₂, ..., c_k }) adalah himpunan titik pusat awal dari setiap kluster.

c. Alokasikan Semua Data/Objek ke Kluster Terdekat:

Jarak antara titik data (x_i) dan titik pusat kluster (c_j) dapat dihitung menggunakan fungsi jarak, misalnya dengan menggunakan jarak Euclidean:

$$[D(x_i, c_j) = \sqrt{\sum_{d=1}^D (x_{id} - c_{jd})^2}]$$

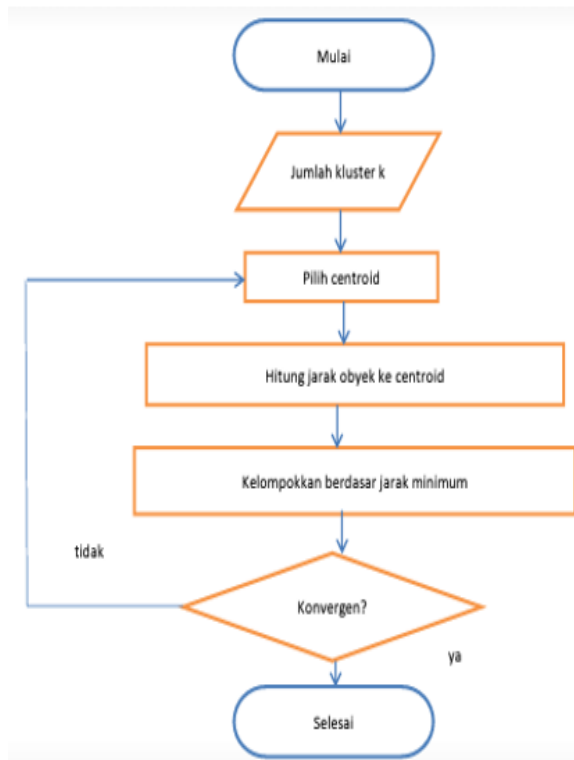
di mana (D) adalah dimensi dari data, (x_{id}) adalah nilai dimensi (d) dari data (x_i), dan (c_{jd}) adalah nilai dimensi (d) dari pusat kluster (c_j).

d. Hitung Kembali Pusat Kluster dengan Keanggotaan Cluster yang Sekarang:

Pusat kluster (c_j) dihitung ulang sebagai rata-rata dari semua data/objek dalam kluster (j): [c_j = $\frac{1}{n_j} \sum_{x_i \in S_j} x_i$] di mana (S_j) adalah himpunan data yang termasuk dalam kluster (j), dan (n_j) adalah jumlah data dalam kluster (j).

e. Tugaskan Lagi Setiap Objek Menggunakan Pusat Kluster yang Baru:

Ulangi langkah 3 dan 4 hingga pusat kluster tidak berubah lagi atau iterasi konvergen.



Gambar 1. tahapan K-means

4. HASIL DAN PEMBAHASAN

4.1. Hasil Penelitian

Berdasarkan hasil penelitian penerapan *Knowledge Discovery in Database (KDD)* dalam penelitian ini yaitu sebagai berikut:

4.1.1. Dataset

Dataset yang akan di olah pada penelitian ini yaitu berdasarkan data transaksi penjualan aksesoris hp dan pulsa di toko Bagus Celluler dengan jumlah data sebanyak 1021 record. Data ini diperoleh langsung dengan melalui tahap observasi, wawancara dan dokumentasi yang bertujuan untuk memperoleh informasi yang berkaitan dengan transaksi penjualan aksesoris hp dan pulsa di toko Bagus Celluler. Yang mencakup variabel id pelanggan, jenis/nama item, jumlah barang, harga, total, dan bulan. Data tersebut akan digunakan sebagai dataset untuk proses analisis menggunakan algoritma K-Means Clustering dalam aplikasi RapidMiner versi 10.2. Berikut sampel data transaksi penjualan aksesoris hp dan pulsa di toko Bagus Celluler yang akan digunakan pada penelitian dapat dilihat pada gambar 2.

Tabel 1. Sampel Data transaksi

No	Id pelanggan	Nama item	Qty	Harga	Total	Bulan
1.	JL230701001	Vcr XL Flex 32GB	1	82000	82000	Juli
2.	JL230701002	Vcr XL Flex 3,5GB	1	18000	18000	Juli
3.	JL230701003	Vcr Axis 5.5GB 15H	1	23000	23000	Juli
4.	JL230701004	Kartu XL Flex 3,5GB	1	22000	22000	Juli
5.	JL230701004	Kartu Telkom 22GB	1	40000	40000	Juli
6.	JL230701004	Vcr Tsel 2GB 30h	1	12000	12000	Juli
...
1019	JL230704087	MMC Vivan 64GB V64U10	1	145000	145000	oktober
1020	JL230704087	Macaron Gambar List 3D	1	25000	25000	oktober
1021	JL230704088	Kartu XL Flex 19GB	1	58000	58000	oktober

4.1.2. Selection

Tahapan pertama adalah melibatkan pencarian file lokasi yang telah dibuat sebelumnya dalam format xlsx menggunakan Operator Read Excel. Fungsi dari Operator Read Excel ini adalah untuk membaca data yang tersimpan dalam format xlsx. Operatore Read Excel ini digunakan untuk mengimport atau memasukan data excel yang berada di komputer ke dalam proses pada rapidminer. Gambar 2 memvisualisasikan data yang telah diimport, mengindikasikan bahwa tombol pada operator Read Excel tidak lagi menunjukkan tanda seru berwarna kuning, menandakan bahwa operator tersebut telah berhasil mengandung data dan siap untuk diolah



Gambar 2. Read Excel

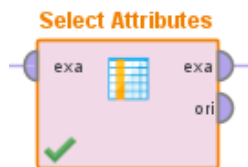
Parameter pada operator Read Excel menggunakan parameter default. Dari hasil pembacaan operator *Read Excel* didapat informasi yang bisa dilihat pada tabel 2 berikut.

Tabel 2. Statistik Dataset transaksi penjualan

No	Uraian	Keterangan
1	Record	1021
2	Special Attributes	2
3	Regular Attributes	4
4	Attributes:	
	Id pelanggan	Polynomial, Missing 0
	Nama item	Polynomial, Missing 0
	Qty	Integer, Missing 0
	Harga	Integer, Missing 0
	Total	Integer, Missing 0
	Bulan	Polynomial, Missing 0

Dalam Tabel 2 di atas terdapat dua jenis data yang digunakan, yaitu integer dan polynomial. Bagian yang menggunakan tipe data integer mencakup tiga atribut, yakni Qty, Harga, dan total. Sementara itu, atribut yang menggunakan tipe data polinomial melibatkan tiga variabel, id pelanggan, nama item, dan bulan.

Select Attributes, operator ini digunakan untuk memisahkan data apa saja yang akan dipergunakan dalam memproses data.



Gambar 3. Select Attributes

Parameter pada operator *Select Attributes* yang digunakan tampak pada tabel 3.

Tabel 3. Parameter dan atribut yang dipilih pada operator Select Attributes

No	Parameter	Isi
1	Type	Include attributes
2	Attribute filter type	a subset
3	Select attribute	ID pelanggan, nama item, bulan juli, bulan agustus, bulan september, bulan oktober, Qty, harga, total.

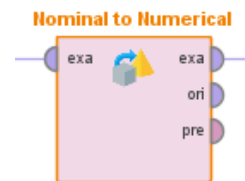
4.1.3. Preprocessing

Pada tahapan preprocessing ini bertujuan untuk melakukan persiapan data sebelum masuk ke proses analisis lebih lanjut. Proses ini melibatkan serangkaian langkah untuk membersihkan, memformat, dan menyesuaikan data agar sesuai dengan kebutuhan analisis atau pemodelan yang akan dilakukan. Tujuan utama dari preprocessing data adalah meningkatkan kualitas data, menghilangkan noise atau gangguan, serta memastikan data siap untuk diolah dengan tepat. Dikarenakan pada dataset transaksi penjualan tidak ditemukan *missing value* atau data yang tidak memiliki nilai maka *preprocessing* tidak dilakukan. Adapun hasil statistik dari data transaksi penjualan dapat dilihat pada gambar 5.

Gambar 4. hasil preprocessing

4.1.4. Transformation

Dilakukannya tahap transformation ini bertujuan untuk mengubah format atau struktur data menjadi bentuk yang lebih sesuai dan dapat diolah dengan efisien dalam proses pemodelan atau analisis data selanjutnya. *Nominal to Numerical*, digunakan untuk mengubah tipe atribut yang bersifat non-numerik menjadi tipe numerik. Dalam dataset transaksi penjualan atribut id nama dan bulan.



Gambar 5. Nominal to Numerical

Tabel 4. Parameter pada Operator Nominal to Numerical

No	Parameter	Isi
1	Attribute Filter Type	Subset
2	Attributes	Id pelanggan, nama item, bulan.
3	Coding Type	Unique Integers

4.1.5. Data Mining

Berikut adalah desain proses untuk membuat cluster K-Means data penjualan aksesoris hp & Pula menggunakan tool Rapid Miner versi 10.2. :



Gambar 6. Model Clustering Menggunakan Algoritma K-Means

Operator utama dalam proses pemodelan ini adalah Cluster K-Means, yang bertindak sebagai alat untuk mendapatkan hasil dari pengelompokan data. Algoritma K-Means digunakan untuk menentukan jumlah k cluster dan menetapkan setiap contoh ke dalam satu cluster yang sesuai. Pada gambar 8 adalah operator Clustering dengan menggunakan algoritma K-Means.



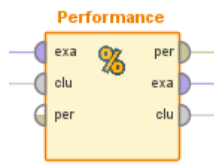
Gambar 7. K-Means Clustering

Penggunaan parameter pada operator K-Means Clustering ini mengacu pada penggunaan parameter default. Dengan mengubah nilai k dimulai dari 2 sampai dengan 10, mengubah nilai *maxruns 1-10*, *measure type* diubah menjadi *Numerical Measure*.

Tabel 5. Parameter pada Operator K-Means Clustering

No	Parameter	Isi
1	K	2-6
2	Max_Runs	1-10
3	Measure Type	Numerical Measure

Proses selanjutnya yaitu menambahkan Operator Performance dengan *Main Criterion* Davies-Bouldin dengan tujuan untuk mengevaluasi performa model klaster yang dihasilkan.



Gambar 8. Performance

Operator Performance digunakan untuk mengevaluasi kinerja model atau proses dalam analisis data. Dalam konteks K-Means clustering, operator ini memberikan informasi tentang seberapa baik model K-Means membentuk kelompok atau klaster yang sesuai dengan pola data sebenarnya. Evaluasi ini membantu memahami sejauh mana model K-Means berhasil dalam mengelompokkan data, penting untuk menentukan keandalan model dan sesuai dengan tujuan analisis. Parameter pada operator *Performance* yang digunakan tampak pada tabel 6.

Tabel 6. Parameter pada operator Performance

No	Parameter	Isi
1	Main Criterion	Devies Bouldin

Dari hasil pembacaan operator *Performance* didapat informasi mengenai pencarian nilai K terbaik yang dapat dilihat pada tabel 7.

Tabel 7. Hasil Operator Performance

No	K	DBI
1	K2	-0.293
2	K3	-0.205
3	K4	-0.456
4	K5	-0.535
5	K6	-0.423
6	K7	-0.541
7	K8	-0.436
8	K9	-0.373
9	K10	-0.347

Berikut merupakan hasil pembacaan operator K-Means Clustering dengan parameter default seperti yang ada pada tabel 7, didapat informasi pada table 8 sebagai berikut.

PerformanceVector

```
PerformanceVector:
Avg. within centroid distance: -4586715842.812
Avg. within centroid distance_cluster_0: -3345741158.013
Avg. within centroid distance_cluster_1: -320104529453.000
Davies Bouldin: -0.293
```

Gambar 9. Hasil Performance Vector

Tabel 8. Hasil Eksperimen Cluster (*Numerical Measure - Chebychev Distance*)

K	Main Criterion	Max Runs	DBI
2	Davies Bouldin	1	-0.293
		2	-0.293
		3	-0.293
		4	-0.293
		5	-0.293
		6	-0.293
		7	-0.293
		8	-0.293
		9	-0.293
		10	-0.293
3	Davies Bouldin	1	-0.205
		2	-0.205
		3	-0.205
		4	-0.205
		5	-0.205
		6	-0.205
		7	-0.205
		8	-0.205
		9	-0.205
		10	-0.205
4	Davies Bouldin	1	-0.456
		2	-0.456
		3	-0.456
		4	-0.456
		5	-0.456
		6	-0.456
		7	-0.456
		8	-0.456
		9	-0.456
		10	-0.456
5	Davies Bouldin	1	-0.535
		2	-0.535
		3	-0.535
		4	-0.535
		5	-0.535
		6	-0.535
		7	-0.535
		8	-0.535
		9	-0.535
		10	-0.535
6	Davies Bouldin	1	-0.423
		2	-0.423
		3	-0.423
		4	-0.423
		5	-0.423
		6	-0.423
		7	-0.423
		8	-0.423
		9	-0.423
		10	-0.423
7	Davies Bouldin	1	-0.541
		2	-0.541

K	Main Criterion	Max Runs	DBI
		3	-0.541
		4	-0.541
		5	-0.541
		6	-0.541
		7	-0.541
		8	-0.541
		9	-0.541
8	Davies Bouldin	10	-0.541
		1	-0.436
		2	-0.436
		3	-0.436
		4	-0.436
		5	-0.436
		6	-0.436
		7	-0.436
		8	-0.436
		9	-0.436
9	Davies Bouldin	10	-0.436
		1	-0.373
		2	-0.373
		3	-0.373
		4	-0.373
		5	-0.373
		6	-0.373
		7	-0.373
		8	-0.373
		9	-0.373
10	Davies Bouldin	10	-0.373
		1	-0.347
		2	-0.347
		3	-0.347
		4	-0.347
		5	-0.347
		6	-0.347
		7	-0.347
		8	-0.347
		9	-0.347

4.1.6. Interpretation/Evaluation

Dari hasil penelitian dengan *Cluster Distance Performance* berdasarkan *Davies Bouldin* pada *numerical measure*, ditemukan bahwa K=3 merupakan nilai K terbaik. Hal ini disebabkan karena cluster K=3 memiliki nilai DBI yang paling rendah atau mendekati nol. Hasil pengukuran nilai *max runs* dari 1 sampai 10, dengan *numerical measure* menunjukkan nilai DBI yang sama, yaitu -0.205. Dengan jumlah *cluster* 0 sebanyak 1017 *items*, *cluster* 1 sebanyak 2 *items*, dan *cluster* 2 sebanyak 2 *items*.

Tabel 9. Hasil Clustering K=3 yang merupakan nilai DBI terbaik

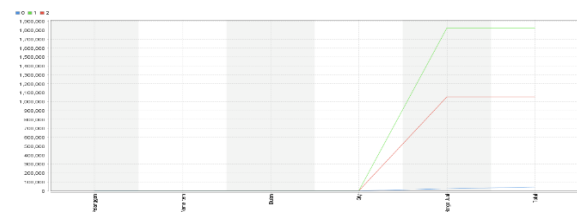
No	Cluster	Keterangan
1	Cluster 0	1017 <i>Items</i>
2	Cluster 1	2 <i>Items</i>
3	Cluster 2	2 <i>Items</i>
	Total Number of Items	1021 <i>Items</i>

Titik centroid dari tiap cluster yang digunakan untuk identifikasi karakteristik dari cluster ditunjukkan pada tabel 10.

Tabel 10. Titik-titik centroid dari hasil clustering

No	Atribute	Cluster_0	Cluster_1	Cluster_2
1	ID Pelanggan	472.61356932153393	95.5	229.5
2	Nama Item	49.47295968534907	61.5	70.5
3	Bulan	1.6293018682399214	0.5	0.5
4	Qty	2.831858407079646	1.0	1.0
5	Harga Jual	24036.22320550639	1824500.0	1052500.0
6	Total	43067.02654867257	1824500.0	1052500.0

Berikut adalah hasil grafik dari setiap atribut yang divisualisasikan dalam plot menggunakan RapidMiner 10.2.



Gambar 10. Hasil grafik pada atribut data

Dari gambar 11, dapat disimpulkan bahwa hasil pengelompokan data penjualan aksesoris hp dan pulsa menggunakan Algoritma K-Means berdasarkan pada atribut ID pelanggan, nama item, bulan juli, bulan agustus, bulan september, bulan oktober, dan Qty. pada ketiga klaster memberikan hasil yang serupa, serta memiliki Evaluasi dan nilai Support terendah dengan akurasi sebesar 0. Sementara itu, atribut harga jual dan total menghasilkan nilai akurasi yang berbeda, di mana klaster 0 harga jual memiliki akurasi sebesar 10 dan total memiliki akurasi sebesar 40 untuk klaster 1 harga dan total memiliki akurasi sebesar 1035 sedangkan klaster 2 harga dan total memiliki akurasi 1810.

4.2. Pembahasan

4.2.1. Mencari nilai K terbaik

Melalui proses analisis dengan menggunakan algoritma K-Means pada Rapid Miner versi 10.2, penelitian ini mencari nilai K terbaik untuk dataset yang relevan dengan rentang K2 hingga K10. Evaluasi hasil menunjukkan bahwa K3 memberikan kinerja optimal dengan nilai mendekati 0, yaitu -0.205. Hasil tersebut menunjukkan keberhasilan algoritma K-Means dalam mengelompokkan data pada K3, dengan *Cluster Distance Performance* berdasarkan *Davies Bouldin* sebesar -0.205. Signifikansi nilai DBI mendekati 0 menandakan bahwa pembagian kelompok yang dihasilkan oleh algoritma K-Means pada K3 sangat seragam dan memiliki sedikit perbedaan internal. Oleh karena itu, dapat disimpulkan bahwa K-

Means memberikan hasil yang memuaskan untuk tugas pengelompokan data penjualan ini. Berikut adalah hasil eksperimen dari K2 hingga K10:

Tabel 11. Hasil K terbaik

No	K	DBI
1	K2	-0.293
2	K3	-0.205
3	K4	-0.456
4	K5	-0.535
5	K6	-0.423
6	K7	-0.541
7	K8	-0.436
8	K9	-0.373
9	K10	-0.347

Hasil ini memberikan gambaran menyeluruh tentang seberapa baik kinerja algoritma K-Means untuk berbagai nilai K yang digunakan, dengan K3 sebagai pilihan terbaik berdasarkan evaluasi DBI.

4.2.2. Pengukuran nilai max runs

Berdasarkan hasil pengukuran nilai max runs dari 1 sampai 10, menggunakan numerical measure, menunjukkan nilai DBI yang konsisten, yakni -0.205, pada kelompok K=3. Pada konfigurasi ini, terdapat 948 items dalam cluster 0, 2 items dalam cluster 1, dan 2 items dalam cluster 2. Perlu dicatat bahwa nilai DBI yang stabil pada K=3 menunjukkan kualitas pengelompokan yang baik. Dengan demikian, hasil eksperimen memberikan indikasi bahwa numerical measure efektif dalam membentuk tiga kelompok yang signifikan. Kesimpulan ini memperkuat keandalan metode pengukuran numerik yang digunakan dalam analisis clustering ini, yang dapat menjadi dasar bagi langkah-langkah selanjutnya dalam pemrosesan dan interpretasi data lebih lanjut.

5. KESIMPULAN DAN SARAN

Berikut adalah kesimpulan berdasarkan penelitian yang telah dilakukan pada dataset penjualan aksesoris hp dan pulsa menggunakan algoritma K-Means. Jumlah kelompok terbaik pada data penerimaan peserta didik baru menggunakan algoritma k-means berdasarkan *Davies Bouldin* terdapat pada K=3 dengan nilai *Davies Bouldin* -0.205. Nilai max runs yang mengashilkan kelompok terbaik pada data penerimaan peserta didik baru menggunakan algoritma K-Means berdasarkan *Davies Bouldin*, yaitu dari 1 sampai 10, menggunakan numerical, menunjukkan nilai DBI yang konsisten, yakni -0.205, pada kelompok K=3.

Berdasarkan hasil dan pembahasan dari penelitian ini, maka terdapat beberapa saran untuk pengembangan penelitian selanjutnya yaitu penelitian ini hanya melakukan eksperimen pada nilai k2 hingga k10, parameter numerical measures, dan jumlah Max Runs 1 hingga 10. Sehingga dapat dikembangkan untuk lebih lanjut yaitu. Mempertimbangkan variasi nilai K yang lebih luas dalam rentang pengukuran numerik guna mengeksplorasi kemungkinan

pengelompokan yang lebih optimal. Sebagai langkah selanjutnya, perlu dilakukan analisis lebih mendalam terhadap item dalam setiap cluster untuk memahami karakteristik dan pola yang mendasari pembentukan kelompok tersebut. Penting juga mengevaluasi kecocokan metode pengukuran numerik terhadap sifat khusus dataset, dengan mempertimbangkan efektivitas metode pada jenis data tertentu. Visualisasi data, seperti plot distribusi cluster, dapat digunakan untuk memberikan gambaran yang lebih jelas tentang sebaran dan karakteristik setiap kelompok. Untuk meningkatkan keamanan interpretasi hasil, disarankan melakukan validasi silang atau metode evaluasi tambahan yang memperkuat reliabilitas temuan dari analisis clustering. Diharapkan, penelitian selanjutnya dapat mengeksplorasi lebih lanjut dalam pemilihan metode pengukuran numerik dan nilai K yang optimal, serta menguji efektivitasnya pada berbagai jenis dataset. Penelitian juga dapat menelusuri potensi penggunaan algoritma clustering dalam konteks aplikasi praktis atau industri tertentu. Fokus pada interpretabilitas hasil clustering dan pemahaman lebih mendalam terhadap konteks domain spesifik diharapkan dapat memperkaya kontribusi penelitian. Terakhir, penelitian berikutnya bisa mempertimbangkan pengembangan model prediktif atau integrasi dengan teknik pembelajaran mesin lainnya untuk memberikan solusi yang lebih menyeluruh dan relevan terhadap tantangan praktis di bidang pengelompokan data.

DAFTAR PUSTAKA

- [1] I. Nawangsih, R. Puspita, and Suherman, "Implementasi Algoritma K-Means Dalam Mengkategorikan Produk Terlaris Dan Kurang Laris Pada Toko Alfamart Cikarang," *Pelita Teknol.*, vol. 16, no. 1, pp. 79–87, 2021, [Online]. Available: https://jurnal.pelitabangsa.ac.id/index.php/pelita_tekno/article/view/674
- [2] M. Dahria, R. Gunawan, and Z. Lubis, "Implementasi K-Means Untuk Pengelompokan Produk Terbaik PT. Koko Pelli," *Semin. Nas. Sains dan ...*, pp. 495–498, 2019, [Online]. Available: <http://seminar-id.com/prosiding/index.php/sensasi/article/view/352%0Ahttp://seminar-id.com/prosiding/index.php/sensasi/article/download/352/345>
- [3] E. Febrianty, L. Awalina, and W. I. Rahayu, "Optimalisasi Strategi Pemasaran dengan Segmentasi Pelanggan Menggunakan Penerapan K-Means Clustering pada Transaksi Online Retail Optimizing Marketing Strategies with Customer Segmentation Using K-Means Clustering on Online Retail Transactions," *J. Teknol. dan Inf.*, vol. 13, no. September, pp. 122–137, 2023, doi: 10.34010/jati.v13i2.
- [4] F. P. A. Hasibuan, S. Sumarno, and I. Parlina, "Penerapan K-Means pada Pengelompokan

- Penjualan Produk Smartphone,” *SATESI J. Sains Teknol. dan Sist. Inf.*, vol. 1, no. 1, pp. 15–20, 2021, doi: 10.54259/satesi.v1i1.3.
- [5] H. Mutaqin, “Penerapan Algoritma K-Means Pada Penjualan Saldo Transportasi Online Studi Kasus Konter XYZ,” *J. Algoritm.*, vol. 20, no. 1, pp. 57–64, 2023, doi: 10.33364/algoritma/v.20-1.1244.
- [6] M. Rizki and M. Mulyawan, “Penerapan Metode K-Means Clustering Pada Data Penjualan Optik Chantika,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 2, pp. 1303–1307, 2023, doi: 10.36040/jati.v7i2.6562.
- [7] N. Astuti, J. N. Utamajaya, and A. Pratama, “Penerapan Data Mining Pada Penjualan Produk Digital Konter Leppangeng Cell Menggunakan Metode K-Means Clustering,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 3, p. 754, 2022, doi: 10.30865/jurikom.v9i3.4351.
- [8] Puji Rahayu, Ika Anikah, Dias Bayu Saputra, Tri Anelia, and Martanto, “Penerapan Data Mining Metode K-Means Clustering Untuk Analisa Penjualan Rotan,” *KOPERTIP J. Ilm. Manaj. Inform. dan Komput.*, vol. 4, no. 2, pp. 42–50, 2020, doi: 10.32485/kopertip.v4i2.118.
- [9] H. S. Nugraha, H. Mutaqin, A. Fathah, and C. Juliane, “Mengidentifikasi Strategi Promosi pada Jasa Penjualan Saldo Digital menggunakan Pendekatan Clustering,” *Edumatic J. Pendidik. Inform.*, vol. 7, no. 1, pp. 11–19, 2023, doi: 10.29408/edumatic.v7i1.7385.
- [10] M. Rafi Nahjan, Nono Heryana, and Apriade Voutama, “Implementasi Rapidminer Dengan Metode Clustering K-Means Untuk Analisa Penjualan Pada Toko Oj Cell,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 1, pp. 101–104, 2023, doi: 10.36040/jati.v7i1.6094.
- [11] D. G. Ramadhan, I. Prihatini, and F. Liantoni, “Analisis Clustering Pengelompokan Penjualan Paket Data Menggunakan Metode K-Means,” *Ultim. J. Tek. Inform.*, vol. 13, no. 1, pp. 33–38, 2021, doi: 10.31937/ti.v13i1.1981.
- [12] M. I. N. Rais, “Penentuan Segmentasi Konsumen Pada Marketing Data Ifood Menggunakan Metode K-Means Clustering,” 2022, [Online]. Available: [http://repository.unpas.ac.id/61693/%0Ahttp://repository.unpas.ac.id/61693/1/Muhammad Irsyad Nur Rais_203010158_Teknik Industri.pdf](http://repository.unpas.ac.id/61693/%0Ahttp://repository.unpas.ac.id/61693/1/Muhammad%20Irsyad%20Nur%20Rais_203010158_Teknik%20Industri.pdf)
- [13] I. Nuryani and D. Darwis, “Analisis Clustering Pada Pengguna Brand Hp Menggunakan Metode K-Means,” *Pros. Semin. Nas. Ilmu Komput.*, vol. 1, no. 1, p. 2021, 2021.
- [14] R. Maoulana, B. Irawan, and A. Bahtiar, “DATA MINING DALAM KONTEKS TRANSAKSI PENJUALAN HIJAB DENGAN MENGGUNAKAN ALGORITMA CLUSTERING K-MEANS (STUDI KASUS : HIJABER TRENDY),” vol. 8, no. 1, pp. 515–521, 2024.