

BAB II

TINJAUAN PUSTAKA

2.1. Studi Sebelumnya

Penelitian yang dilakukan oleh Hanum dkk. bertujuan untuk mengelompokkan pelanggan ke dalam klaster berdasarkan perilaku pelanggan untuk mengenali kelompok pelanggan yang serupa dan merencanakan strategi pemasaran yang sesuai pada PT Ekspedisi Indonesia. Penelitian ini dilakukan dengan menggunakan metode *K-Means clustering*. Pengenalan pola yang bermanfaat dan segmentasi pelanggan lebih mudah dengan metode ini. Teknik *clustering* dalam *data mining* membantu bisnis memahami perilaku pelanggan, menemukan kelompok pelanggan yang sebanding, dan membuat strategi pemasaran yang tepat. Hasil penelitian menunjukkan bahwa *cluster* terbaik dibuat dengan nilai $k = 4$, dan data dinormalisasi dengan metode Normalisasi Min-Max. Pengelompokan pelanggan dapat memungkinkan identifikasi profil pelanggan untuk membantu perusahaan membuat keputusan [9].

Kemudian berdasarkan perancangan *Business Intelligence* dan segmentasi pelanggan yang dilakukan oleh Karami dan Widharto pada PT XYZ yang merupakan perusahaan penyedia solusi optimasi, informasi, dan pengukuran untuk berbagai sektor, seperti manufaktur, energi, infrastruktur, pertambangan, atau industri lainnya. Penelitian ini bertujuan untuk menggunakan *Business Intelligence* sebagai kerangka acuan untuk mengembangkan sistem informasi yang memberikan pemahaman lebih baik mengenai kontribusi pelanggan terhadap bisnis perusahaan. Dengan menggunakan teknik *Data Mining K-Means Clustering*, perusahaan dapat mengelompokkan pelanggan berdasarkan kriteria tertentu. Model RFM yang terkenal, yang dapat menjelaskan nilai pelanggan, adalah kriteria yang dipilih. Penelitian ini menghasilkan *XYZ dashboard* sebagai salah satu bentuk *Business Intelligence* memuat *profiling* pelanggan yang memungkinkan perusahaan untuk melacak kontribusi pelanggan serta mengoptimalkan sumber daya dan upaya pemasaran mereka, meningkatkan efisiensi, dan memaksimalkan keuntungan dari setiap segmen pelanggan [10].

Penelitian berikutnya dilakukan oleh Awalina dan Rahayu dengan tujuan untuk mengelompokkan pelanggan berdasarkan perilaku dan karakteristik melalui segmentasi pelanggan. Dalam penelitian ini, metode *clustering* dengan algoritma *K-Means* menjadi salah satu pilihan yang efektif dalam melakukan segmentasi pelanggan. Penelitian ini menggunakan pemrograman *python* untuk proses pemrosesan dan *pre-processing* data. Proses ini dilakukan melalui analisis data *exploratory* untuk memperoleh pemahaman tentang informasi yang digunakan dalam data sebelum melakukan klusterisasi. Hasil dari metode *elbow* menunjukkan bahwa penggunaan empat kluster adalah pilihan yang tepat untuk situasi ini. Selanjutnya, dengan menggunakan variabel *quantity*, *unit price*, dan *customer id*, pemodelan *K-Means* dengan 4 kluster dihasilkan, yang menghasilkan empat kluster yang berbeda dengan masing-masing karakteristik unik. Ada bukti bahwa perilaku konsumen sangat dipengaruhi oleh harga satuan dan kuantitas[11].

Penelitian berikutnya dilakukan oleh Rinjani dan Hamida dengan tujuan untuk memetakan pembagian jalur untuk sebaran titik navigasi pengiriman barang jadi yang lebih optimal pada salah satu perusahaan manufaktur otomotif di Jakarta. Metode yang digunakan pada penelitian ini adalah metode *K-Means clustering*. Dengan penerapan metode tersebut penelitian ini menghasilkan standar deviasi jarak supir dengan rata-rata jarak yang ditempuh adalah 95,067 km dibandingkan dengan sebelum penggunaan metode, yaitu 182,27. Dengan hasil tersebut didapatkan pembagian rute jalur yang lebih optimal dengan rata – rata jarak tempuh 46,767 Km dibandingkan sebelum menggunakan metode, yaitu 61,57 Km[12].

Penelitian lainnya dilakukan oleh Marsono yang bertujuan untuk mengelompokkan item produk berdasarkan tingkat penjualan pada PT Aquasolve yang merupakan salah satu perusahaan manufaktur di Indonesia. Penelitian ini menggunakan metode *K-Means clustering* yang menghasilkan 3 *cluster* atau 3 pengelompokan item produk berdasarkan tingkat *supply* dan produksi item tersebut. Dalam penelitian ini digunakan Metode *hard K-Means* digunakan untuk membagi semua data dalam kumpulan yang paling dekat dengan titik pusat. *Cluster 1* merupakan kelompok item produk yang memiliki tingkat permintaan tinggi, *cluster 2* merupakan kelompok item produk yang memiliki tingkat permintaan sedang, dan *cluster 3* merupakan kelompok item produk yang memiliki tingkat permintaan yang rendah [13].

Tabel 2.1. Penelitian Sebelumnya

No	Penulis	Tahun	Tujuan	Metode	Hasil
1	Hanum, dkk.	2023	Pengelompokan pelanggan dalam bentuk <i>cluster</i> untuk memahami perilaku pelanggan yang serupa, dan merencanakan strategi pemasaran yang sesuai	<i>K-Means Clustering</i>	<i>Cluster</i> terbaik dihasilkan dengan nilai k 4 atau pelanggan dikelompokkan menjadi 4 <i>cluster</i> , serta data dinormalisasi menggunakan metode <i>Min-Max Normalization</i> .
2	Karami dan Widharto	2023	Menggunakan <i>Business Intelligence</i> sebagai kerangka acuan untuk mengembangkan sistem informasi yang memberikan pemahaman lebih baik mengenai kontribusi pelanggan terhadap bisnis perusahaan.	<i>K-Means Clustering</i> Berdasarkan RFM Model	Menghasilkan <i>XYZ dashboard</i> sebagai salah satu bentuk <i>Business Intelligence</i> yang memuat <i>profiling</i> pelanggan untuk melihat kontribusi pelanggan dan pemasaran dari setiap segmen pelanggan.
3	Awalina dkk.	2023	Mengelompokkan pelanggan berdasarkan perilaku dan karakteristik khusus dengan menggunakan 3 variabel	<i>K-Means Clustering</i>	<ul style="list-style-type: none"> • Kluster 0 membeli barang dengan harga satuan yang rendah dan kode pelanggan yang tinggi.

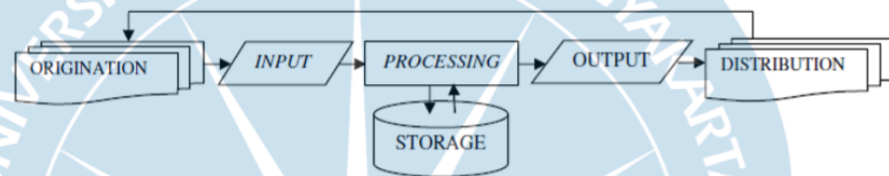
No	Penulis	Tahun	Tujuan	Metode	Hasil
			data yaitu <i>quantity</i> , <i>unit price</i> dan <i>customer id</i> .		<ul style="list-style-type: none"> • Kluster 1 membeli barang dengan harga satuan yang menengah dan kode pelanggan yang tinggi, • Kluster 2 membeli barang dengan harga satuan yang tinggi dan kode pelanggan yang menengah. • Kluster 3 membeli barang dengan harga satuan yang rendah dan kode pelanggan yang menengah.
4	Rijani dan Hamida	2019	Optimasi pembagian jalur pengiriman barang jadi	<i>K-Means Clustering</i>	<ul style="list-style-type: none"> • <i>Cluster 1</i> : rata – rata rute sopir 23,9 km • <i>Cluster 2</i> : rata – rata rute sopir 22,1 km • <i>Cluster 3</i> : rata – rata rute sopir 94,3 km
5	Marsono	2019	Mengelompokkan item produk berdasarkan tingkat penjualan	<i>K-Means Clustering</i>	<ul style="list-style-type: none"> • Kelompok 1 adalah kelompok dengan tingkat permintaan

No	Penulis	Tahun	Tujuan	Metode	Hasil
					<p>tertinggi berdasarkan jumlah <i>supply</i> dan produksi produk.</p> <ul style="list-style-type: none"> • Kelompok 2 adalah kelompok dengan tingkat permintaan sedang berdasarkan jumlah <i>supply</i> dan produksi pada produk. • Kelompok 3 adalah kelompok dengan tingkat permintaan rendah berdasarkan jumlah <i>supply</i> dan produksi pada produk.

2.2. Dasar Teori

2.2.1 Pengelolaan Data

Pengelolaan data dikenal sebagai proses perubahan data mentah menjadi bentuk yang lebih relevan, berguna, dan bermakna yang berupa informasi. Pengelolaan data terdiri dari enam tahap dasar, yaitu *orgination*, *input*, *prosesing*, *output*, *distribution* dan *storage*. Dalam pengelolaan data, siklus pengelolaan data berkembang sambil komponennya bertambah[14]. Prinsip-prinsip pengelolaan data di sektor ini mencakup pengumpulan, pemasukan, pengolahan, keluaran, penyimpanan dan penyampaian data [15].



Gambar 2.1. Siklus pengolahan Data[15]

a. *Orgination*

Orgination merupakan tahap awal dalam siklus pengelolaan data yang melibatkan pengumpulan data dari berbagai sumber, seperti survei, transaksi, sensor, atau basis data.

b. *Input*

Input merupakan proses di mana data yang telah dikumpulkan dimasukkan ke dalam sistem pengolahan data. Proses ini melibatkan konversi data ke format yang dapat diterima oleh komputer.

c. *Processing*

Processing merupakan pemrosesan data mentah yang diolah menggunakan teknik dan metode tertentu termasuk perhitungan, transformasi, dan analisis data untuk menghasilkan informasi berguna.

d. *Output*

Output adalah hasil pengolahan dan analisis data yang disajikan berupa informasi yang jelas dan mudah dipahami untuk mempermudah pengambilan keputusan.

e. *Storage*

Storage merupakan proses dan teknologi yang digunakan untuk menyimpan data dalam berbagai format secara aman dan terorganisir sehingga dapat diakses dan digunakan kapan saja saat dibutuhkan.

f. *Distribution*

Distribution merupakan proses penyampaian informasi atau hasil pengolahan dan analisis data yang tersimpan kepada pihak yang membutuhkan.

Selain itu, optimasi pengelolaan data juga melibatkan upaya untuk memahami dan memanfaatkan data sebagai aset berharga. Strategi pengelolaan data yang baik membantu perusahaan untuk mengidentifikasi tren, meramalkan kebutuhan produksi, dan membuat keputusan yang informasional. Oleh karena itu, implementasi teknologi terkini, seperti sistem manajemen basis data canggih, memainkan peran penting dalam membentuk fondasi pengelolaan data yang efisien dan adaptif di lingkungan industri remanufaktur yang dinamis [16].

2.2.2 *K-Means Clustering*

Clustering adalah metode penelitian formal tentang pengelompokan atau partisi data tanpa label kategori. Berbeda dengan klasifikasi, *clustering* tidak memiliki tahap pelatihan data yang diawasi. *Clustering* merupakan sebuah model dalam *Machine learning* yang menggabungkan data secara alami berdasarkan kemiripan objek data, sehingga dapat mengurangi kemiripan dengan kluster lain[17].

K-Means merupakan salah satu algoritma *clustering* yang paling umum digunakan dalam analisis data dan *machine learning*. Tujuan utama dari algoritma *K-Means* adalah mengelompokkan data menjadi beberapa kluster berdasarkan kesamaan antar-observasi. *K-Means* dalam *clustering* termasuk dalam metode *unsupervised learning* yang digunakan untuk mengelompokkan data berdasarkan variabel atau fitur. *Unsupervised Learning* merupakan algoritma mempelajari pola dari data tanpa label[18].

2.2.3 Overhaul

Overhaul adalah sebuah perbaikan yang umumnya terdiri dari serangkaian tindakan pemeliharaan preventif seperti pergantian oli, pembersihan, pelumasan, dan penggantian beberapa komponen yang aus pada sistem. Periode *overhaul* termasuk jadwal perawatan berkala, yang merupakan waktu penting di mana setiap komponen produk diperbarui secara menyeluruh untuk meningkatkan kinerja dan memperpanjang umur pakainya [19] [20].

2.2.4 Machine learning

Machine learning adalah sub bidang kecerdasan buatan (AI) di mana algoritma yang dilatih pada kumpulan data digunakan untuk membuat model pembelajaran mandiri[21]. Model ini memiliki kemampuan untuk mengklasifikasikan dan memprediksi hasil tanpa intervensi manusia. Saat ini, *machine learning* digunakan untuk berbagai tujuan komersial, termasuk memberikan rekomendasi produk kepada konsumen berdasarkan pembelian mereka sebelumnya, memprediksi fluktuasi pergerakan pasar maupun saham, dan menerjemahkan teks dari satu bahasa ke bahasa lainnya [22].

2.2.5 Segmentasi Pelanggan

Proses segmentasi pelanggan menjadi strategi perusahaan dalam usaha peningkatan kualitas *marketing*. Melalui segmentasi ini perusahaan mendapatkan data pendukung untuk menentukan strategi pemasaran yang efektif dan efisien dalam kegiatan promosi penjualan produk berdasarkan karakteristik dari sebaran pelanggan[23]. Dalam pemasaran, segmentasi masih

merupakan konsep penting. Segmentasi pelanggan adalah salah satu langkah awal dalam membuat model bisnis bagi perusahaan yang membagi pelanggan menjadi beberapa kelompok berdasarkan tingkat loyalitas pelanggan untuk membangun strategi pemasaran dan akan membuat hubungan dengan pelanggan lebih menarik[24]. Segmentasi pelanggan merupakan bentuk usaha untuk meningkatkan ketepatan pemasaran perusahaan yang terdiri dari kelompok besar yang dapat diidentifikasi dalam sebuah pasar dengan keinginan, daya beli, perilaku pembelian dan kebiasaan pembelian yang serupa [23].

2.2.6 Market Overview Dashboard

Dashboard adalah *platform* yang populer untuk visualisasi data dan analisis data untuk menilai persentase dan nilai pencapaian bisnis [25]. Dalam analisis pembelajaran, *dashboard* adalah panel kontrol yang dapat disesuaikan yang menampilkan fitur analitik pembelajaran yang dapat disesuaikan dan disesuaikan dengan proses pembelajaran secara *real-time*. Fitur ini dapat mencakup penilaian diri, peluang pembelajaran, saran, interaksi sosial, perbandingan dengan teman sekelas, dan tautan tambahan [26].

Market overview adalah gambaran umum tentang kondisi pasar pada suatu waktu tertentu. Ini mencakup berbagai aspek yang mempengaruhi pasar, termasuk tren, permintaan, penawaran, perilaku konsumen, dan faktor-faktor eksternal. *Market overview* menjadi alat penting bagi manajer pemasaran dan pengambil kebijakan untuk memahami pasar secara menyeluruh dan mengambil langkah-langkah yang tepat untuk mencapai tujuan bisnis.