

Penerapan Algoritma K-Means untuk Mengelompokkan Perilaku Konsumen dalam Berbelanja (Studi Kasus : Platform Lazada)

Nurhidayati^{1*}, Almi Yulistia Alwanda², Fathurrahman³

^{1,2,3}Program Studi Sistem Informasi, Universitas Hamzanwadi

hidayati2188@gmail.com

Abstrak

Pesatnya perkembangan industri e-commerce telah meningkatkan kebutuhan untuk memahami lebih dalam perilaku konsumen. Pemahaman ini, menjadi aspek yang sangat penting karena dapat membantu perusahaan dalam menentukan strategi bisnis, khususnya dalam hal pengadaan barang dagangan, penetapan strategi promosi, pemasaran dan pelayanan pelanggan. Penelitian ini bertujuan untuk menerapkan algoritma K-Means dalam menganalisis dan mengelompokkan pola perilaku konsumen dalam bertransaksi pada platform Lazada berdasarkan kualitas produk, harga, promosi, layanan pelanggan, kecepatan pengiriman serta metode pembayaran yang digunakan. Metode penelitian yang digunakan adalah metode kuantitatif dengan teknik pengumpulan data Observasi dan studi pustaka. Observasi dilakukan dengan cara menyebar kuesioner secara offline dan online ke-Masyarakat Umum wilayah Pancor-Selong yang merupakan bagian dari sampel penelitian ini sebanyak 500 Responden setelah melalui tahapan pemrosesan menggunakan sistem *Knowledge Discovery in Databases* (KDD) untuk memastikan kualitas data siap untuk diolah menggunakan Algoritma K-Means. Adapun hasil yang didapatkan dalam pengelompokkan pola belanja konsumen pada platform Lazada pada penelitian ini menunjukkan bahwa perilaku konsumen tersebut dapat dikelompokkan menjadi tiga: (C1) kelompok tinggi dengan jumlah 166 responden, (C2) kelompok sedang 174 responden dan (C3) kelompok rendah sebanyak 160 responden. Setiap kelompok merepresentasikan perbedaan perilaku konsumen dalam berbelanja yang didasarkan pada kualitas produk, harga, promosi, pemasaran, layanan pelanggan, kecepatan pengiriman serta metode pembayaran yang digunakan.

Kata kunci : *Data Mining, E-Commerce, K-Means, Lazada, Perilaku Konsumen...*

Abstract

The rapid development of the e-commerce industry has increased the need for a deeper understanding of consumer behavior. This understanding is a very important aspect because it can help companies in determining business strategies, especially in terms of merchandise procurement, determining promotional strategies, marketing and customer service. This study aims to apply the K-Means algorithm in analyzing and grouping consumer behavior patterns in transactions on the Lazada platform based on product quality, price, promotion, customer service, delivery speed and payment methods used. The research method used is a quantitative method with data collection techniques of observation and literature study. Observations were conducted by distributing questionnaires offline and online to the General Public of the Pancor-Selong area which is part of the sample of this study as many as 500 Respondents after going through the processing stage using the Knowledge Discovery in Databases (KDD) system to ensure the quality of the data ready to be processed using the K-Means Algorithm. The results obtained in grouping consumer shopping patterns on the Lazada platform in this study indicate that consumer behavior can be grouped into three: (C1) high group with 166 respondents, (C2) medium group 174 respondents and (C3) low group as many as 160 respondents. Each group represents differences in consumer behavior in shopping based on product quality, price, promotion, marketing, customer service, delivery speed and payment methods used.

Keyword : *Data Mining, E-Commerce, K-Means, Lazada, Consumer Behavior*

1. Pendahuluan

Perkembangan teknologi informasi dan komunikasi yang semakin pesat telah membawa perubahan besar dalam kehidupan masyarakat modern, termasuk dalam aktivitas berbelanja online. Jika dahulu konsumen harus mendatangi toko fisik untuk memenuhi kebutuhan mereka, kini proses belanja dapat dilakukan dengan lebih praktis melalui platform e-commerce. Kemudahan akses, efisiensi waktu, serta beragam pilihan produk yang ditawarkan membuat belanja online semakin diminati. Salah satu platform e-commerce yang populer dan banyak digunakan di Indonesia adalah Lazada, yang berhasil menarik konsumen dari berbagai kalangan dengan menawarkan fitur menarik, promo besar-besaran, serta pelayanan yang relatif cepat. Perubahan tren belanja ini turut memengaruhi perilaku konsumen, baik dari segi preferensi produk, frekuensi pembelian, hingga cara mereka merespons promosi. Hal tersebut menandakan bahwa perilaku konsumen dalam berbelanja mengalami transformasi yang cukup signifikan seiring berkembangnya teknologi digital.

Meningkatnya jumlah pengguna e-commerce membawa dampak pada semakin beragamnya perilaku konsumen[1].

Bagi perusahaan seperti Lazada, memahami perilaku konsumen merupakan faktor penting dalam menjaga loyalitas pelanggan dan memenangkan persaingan pasar. Perilaku konsumen yang kompleks, seperti kecenderungan membeli produk tertentu, memilih metode pembayaran yang berbeda, hingga kepekaan terhadap diskon atau promosi, perlu dianalisis dengan baik agar perusahaan dapat menyusun strategi pemasaran yang sesuai. Namun, data konsumen yang dihasilkan dari aktivitas transaksi jumlahnya sangat besar dan tidak dapat dipahami secara sederhana tanpa adanya bantuan metode analisis data yang tepat. Oleh karena itu, diperlukan penerapan teknologi analisis data yang mampu menggali informasi dan menemukan pola pola tersembunyi dari perilaku konsumen tersebut.

Data mining hadir sebagai solusi yang efektif untuk menganalisis data dalam jumlah besar dan kompleks. Data mining merupakan proses menggali informasi yang bernilai dari sekumpulan data yang besar untuk menemukan pola atau hubungan tertentu yang sebelumnya tidak terlihat. Salah satu metode dalam data mining yang sering digunakan untuk analisis perilaku konsumen

adalah algoritma K-Means Clustering. K-Means merupakan algoritma pengelompokan (clustering) yang bekerja dengan cara membagi data ke dalam beberapa kelompok berdasarkan kesamaan karakteristik. Dengan menggunakan metode ini, data perilaku belanja konsumen dapat dipetakan ke dalam kelompok-kelompok tertentu, seperti kelompok konsumen yang sensitif terhadap harga, kelompok yang lebih tertarik pada kualitas produk, maupun kelompok yang responsif terhadap promosi. Informasi hasil pengelompokan ini sangat bermanfaat bagi perusahaan dalam menyusun strategi pemasaran, memberikan rekomendasi produk, serta meningkatkan kualitas layanan[2].

Penelitian ini bertujuan untuk mengidentifikasi pola perilaku konsumen di platform Lazada melalui pengelompokan menggunakan algoritma K-Means. Hasil dari penelitian ini diharapkan mampu memberikan gambaran yang lebih jelas mengenai segmentasi konsumen berdasarkan perilaku belanja mereka, sehingga Lazada dapat merancang strategi yang lebih efektif dan tepat sasaran. Selain itu, penelitian ini juga diharapkan dapat memberikan kontribusi dalam pengembangan ilmu pengetahuan, khususnya dalam penerapan

teknik data mining untuk analisis perilaku konsumen pada bidang e-commerce.

2. Tinjauan Pustaka

2.1 Penelitian terkait

Berikut beberapa referensi artikel yang digunakan peneliti sebagai acuan dalam penelitian ini adalah sebagai berikut:

- a. Penelitian yang dilakukan oleh Yahya , Mahpuz dalam jurnal yang berjudul “Penggunaan Algoritma K-Means Untuk Menganalisis Pelanggan Potensial Pada Dealer SPS Motor Honda Lombok Timur Nusa Tenggara Barat” Tujuan dari penelitian ini adalah untuk mengetahui kelompok umur dan profesi apa saja yang menjadi pelanggan potensial di SPS Motor Honda Lombok Timur. Metode penelitian yang digunakan adalah metode analitik deskriptif, dimana peneliti mengambil data-data penjualan sepeda motor dari SPS Motor Honda Lombok Timur, yang selanjutnya diolah dan dianalisis menggunakan software Rapidminer 5. Hasil analisis data tersebut selanjutnya dideskripsikan dengan detail mengenai kelompok pelanggan potensial. Berdasarkan analisis yang telah dilakukan, dapat ditarik kesimpulan bahwa kelompok yang termasuk dalam pelanggan potensial dari SPS Motor Honda Lombok Timur adalah kelompok dengan kriteria umur

berkisar 33-46 tahun yang berprofesi sebagai wiraswasta dan petani, Hal ini disebabkan karena, pelanggan tersebut memiliki penghasilan dengan usia produktif. Sedangkan kelompok pelanggan yang tidak potensial ada 2 irisan yaitu yang pertama apabila kategori usia produktif tapi tidak memiliki penghasilan dengan rentang umur 17-32 tahun dengan profesi pelajar dan wiraswasta. Kedua yaitu, apabila kategori usia tidak produktif dengan rentang usia 47-73 tahun dan tidak memiliki penghasilan dengan profesi petani dan wiraswasta[3].

- b. Penelitian yang dilakukan oleh Amri Muliawan nur, Muhammad Saiful, Hariman Bahtiar dkk dalam jurnal yang berjudul “Penerapan Algoritma K Means Clustering Dalam Mengelompokkan Smartphone Yang Rekomendasi Berdasarkan Spesifikasi” Berbagai jenis smartphone hadir dengan harga dan spesifikasi yang berbeda-beda, menyebabkan penjual kadang-kadang kesulitan dalam memberikan rekomendasi kepada konsumen yang ingin membeli smartphone yang sesuai dengan spesifikasi dan kisaran harga yang diinginkan. Tantangan ini muncul karena sulit bagi penjual untuk mengingat spesifikasi dari setiap smartphone yang dijual. K Means Clustering bertujuan untuk mengelompokkan data spesifikasi yang

ada ke dalam beberapa klaster, di mana data dalam satu klaster memiliki karakteristik yang serupa. Data yang digunakan adalah data private sebanyak 350 data, dengan melibatkan 4 atribut, antara lain Merek, Harga, Baterai, dan Penyimpanan Internal. Dengan membentuk kelompok-kelompok smartphone ini, penjual akan lebih mudah merekomendasikan smartphone yang sesuai kepada pelanggan. Hasil penelitian menunjukkan bahwa berbagai merek smartphone dikategorikan ke dalam tiga kelompok: Kelompok Rekomendasi, yang mencakup 225 item; Kelompok Sangat Direkomendasikan, yang mencakup 98 item; dan Kelompok Kurang Direkomendasikan, yang mencakup 27 item. Pengelompokan ini diharapkan dapat membantu penjual dengan mudah meningkatkan stok smartphone yang direkomendasikan sesuai dengan kebutuhan konsumen dalam hal harga dan spesifikasi[4].

- c. Penelitian yang dilakukan oleh Nurhidayati, Suhartini, Amri Muliawan Nur dkk dalam jurnal yang berjudul “Penerapan Sistem Informasi Penjualan Baju Berbasis Android Sebagai Peningkatan Layanan Konsumen” Dapat disimpulkan bahwa Perkembangan di dunia internet sangat diharapkan untuk masa sekarang ini karena semakin banyak perusahaan swasta, instansi dan institusi pendidikan sangatlah tinggi akan kebutuhannya untuk dijadikan layanan informasi. Dalam penelitian ini akan mencoba membuat Aplikasi Sistem Penjualan Baju

Jersey Berbasis Android yang beralamatkan di Toko Berugak Jersey Lombok. Toko Berugak Jersey Lombok hanya menggunakan sosial media dan manual dalam memasarkan produknya. Sehingga sebagian besar konsumen/pelanggan berasal dari orang-orang di sekitar dan sosial media yang digunakan. Saat ini suatu bidang usaha tentu kurang kompetitif jika tidak memiliki media pemasaran online seperti aplikasi/website e-commerce. Keberadaan suatu aplikasi/website akan dapat memperluas jangkauan dalam pemasaran produk-produk yang dijual pada toko ini. Metode yang digunakan dalam perancangan dan pembuatan perangkat lunak ini adalah metode waterfall. Dan menghasilkan rancangan aplikasi sistem penjualan yang dapat mempermudah konsumen untuk melakukan pemesanan/transaksi di toko Berugak Jersey Lombok dan sebagai sarana untuk mempromosikan produk agar lebih komunikatif dan informatif, sehingga dapat memperluas area pemasaran dan menambah jumlah konsumen tersebut[5].

- d. Penelitian yang dilakukan oleh Andrian Hidayat, Nurhidayati, Amri Muliawan Nur dalam jurnal yang berjudul “Implementasi Algoritma K Means Untuk Klasterisasi Peserta Keluarga Berencana Berdasarkan Tingkat Risiko Kehamilan di Desa Pringgasela Selatan” dalam Program Keluarga Berencana (KB) merupakan sebuah program yang bertujuan untuk mengendalikan angka

pertumbuhan populasi penduduk, meningkatkan kesejahteraan keluarga, dan upaya mencegah risiko kesehatan pada ibu dan anak. Salah satu tantangan dalam program KB adalah proses identifikasi dan pelayanan yang maksimal terhadap peserta KB yang berisiko mengalami kehamilan dengan risiko tinggi yang dapat diidentifikasi berdasarkan faktor 4T yaitu Terlalu Muda, Terlalu Tua, Terlalu Banyak, Terlalu Dekat. Penelitian ini bertujuan untuk melakukan klasterisasi pada peserta KB berdasarkan tingkat risiko kehamilan sehingga diperoleh informasi terkait kondisi peserta KB sebagai landasan penyuluhan terkait risiko kehamilan tingkat tinggi dan perancangan program yang sesuai dan tepat sasaran bagi peserta KB. Penelitian ini dilaksanakan di Desa Pringgasela Selatan memanfaatkan data mining menggunakan Algoritma K-Means Clustering pada 400 data peserta KB untuk membentuk kelompok atau klaster peserta KB berdasarkan tingkat risiko kehamilannya dengan atribut berupa Nama, Usia, Jumlah Anak, dan Usia Anak Terakhir. Hasil yang diperoleh yaitu terbentuk 2 (dua) cluster dengan Cluster 1 sebagai kelompok kehamilan berisiko tinggi dengan 170 item anggota dan Cluster 0 sebagai kelompok kehamilan berisiko rendah dengan 230 item anggota[6].

- e. Penelitian yang dilakukan oleh Amri Muliawan Nur, Hariman Bahtiar, Mila Agustiarini Jannah dalam jurnal yang berjudul “Implementasi Algoritma K-Means Clustering

Dalam Mengelompokkan Kepatuhan Wajib Pajak Bumi dan Bangunan Dengan Optimasi Elbow” Dapat di simpulkan bahwa Pajak Bumi dan Bangunan (PBB) merupakan salah satu sumber pendapatan asli daerah yang memiliki peran penting dalam mendukung pembangunan di berbagai daerah. Oleh karena itu tingkat kepatuhan pajak perlu ditingkatkan dengan berbagai strategi seperti sosialisasi maupun edukasi secara berkelanjutan untuk meningkatkan kesadaran atas pentingnya membayar pajak, selain itu pula peningkatan pelayanan yang lebih baik perlu dilakukan. Tujuan penelitian ini mengelompokkan tingkat kepatuhan wajib pajak PBB di Kecamatan Sakra menggunakan algoritma K-Means Clustering. Data yang digunakan dalam penelitian ini adalah data PBB Kecamatan Sakra tahun 2023, dengan jumlah dataset sebanyak 376 dengan melibatkan lima atribut utama, yaitu: luas bumi, luas bangunan, ketetapan PBB, status pembayaran, dan denda. Hasil yang diperoleh dari pengolahan menggunakan algoritma K-Means, memperoleh jumlah cluster yang sangat baik. Berikut adalah hasil cluster yang diperoleh : Cluster 1 tingkat kepatuhan tinggi, berjumlah 355 item. Cluster 2 tingkat kepatuhan sedang, berjumlah 18 item dan cluster 3 tingkat kepatuhan rendah, berjumlah 3 item. Hasil pengelompokan ini dapat menjadi acuan bagi pihak berwenang untuk merumuskan strategi peningkatan kepatuhan yang lebih terarah

melalui edukasi dan pelayanan yang lebih baik kedepannya[7].

2.2 Landasan Teori

1. Pengertian Data Mining

Data mining merupakan suatu pendekatan analitis yang digunakan untuk menemukan pola tersembunyi, hubungan, atau informasi bernilai dari sekumpulan data berukuran besar yang sebelumnya sulit diidentifikasi secara manual. Data mining terdiri dari dua kata yaitu data dan mining. Data adalah sekumpulan bahan, alat, teks yang belum mempunyai arti. Biasanya data bersifat tunggal dan kaku. Butuh proses pengolahan yang benar untuk menghasilkan data menjadi informasi. Sedangkan mining diartikan sebagai penambangan atau penggalian informasi. [8].

2. Metode data Mining[9].

a. Klasifikasi

Klasifikasi merupakan salah satu teknik utama dalam data mining yang digunakan untuk mengelompokkan data ke dalam kategori atau kelas tertentu berdasarkan pola yang ditemukan dari data sebelumnya.

b. Clustering

Clustering adalah metode dalam data mining yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok (cluster) berdasarkan tingkat kemiripan atau kedekatan antar data.

c. Estimasi

Metode estimasi dalam data mining adalah teknik yang digunakan untuk memperkirakan nilai atau informasi yang belum diketahui dari suatu data berdasarkan pola yang ditemukan dalam data yang ada. Asosiasi

Asosiasi merupakan salah satu metode dalam data mining yang digunakan untuk menemukan pola hubungan atau keterkaitan antar item dalam kumpulan data yang besar.

d. Prediksi

Prediksi adalah metode dalam data mining yang digunakan untuk memperkirakan nilai atau hasil yang belum terjadi, biasanya berkaitan dengan masa depan. Meskipun mirip dengan estimasi dan klasifikasi, prediksi lebih menekankan pada upaya meramalkan sesuatu berdasarkan pola atau data historis yang sudah tersedia.

3. Algoritma K-Means

Algoritma K-Means merupakan salah satu metode clustering non hierarkis yang berfungsi untuk mengelompokkan data ke dalam sejumlah klaster berdasarkan tingkat kesamaan atribut. Dalam konteks penelitian ini, K-Means digunakan untuk mengelompokkan konsumen Lazada berdasarkan perilaku belanja mereka, seperti frekuensi pembelian, jenis produk yang sering dibeli, nilai transaksi, dan waktu berbelanja. Proses K-Means diawali dengan menentukan titik pusat klaster (centroid) secara

acak, lalu setiap data konsumen akan ditempatkan ke dalam klaster terdekat berdasarkan jarak ke centroid tersebut. Selanjutnya, posisi centroid diperbarui berdasarkan rata-rata data dalam klaster, dan proses ini diulang hingga posisi centroid stabil. Tujuan penerapan metode ini adalah untuk meminimalkan perbedaan antar data dalam satu klaster dan memaksimalkan perbedaan antar klaster, sehingga menghasilkan segmentasi konsumen yang lebih jelas. Dengan hasil pengelompokan ini, Lazada dapat mengenali pola belanja tersembunyi dan merancang strategi pemasaran yang lebih tepat sasaran untuk setiap segmen konsumen[10].

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

- Tentukan nilai k sebagai jumlah klaster yang ingin dibentuk.
- Inisialisasi k pusat cluster ini bisa dilakukan dengan berbagai cara, namun yang paling sering dilakukan adalah dengan cara random yang diambil dari data yang ada.
- Menghitung jarak setiap data input terhadap masing – masing centroid menggunakan rumus jarak Euclidean (Euclidean Distance) hingga ditemukan jarak yang paling dekat dari setiap data dengan centroid
- Mengklasifikasikan setiap data berdasarkan kedekatannya dengan centroid (jarak terkecil).
- Memperbaharui nilai centroid

- Melakukan perulangan dari langkah 2 hingga 5, sampai anggota tiap cluster tidak ada yang berubah.

4. Metode Clustering

Clustering merupakan salah satu teknik dalam pembelajaran mesin tanpa supervisi (unsupervised learning) yang digunakan untuk mengelompokkan data berdasarkan kesamaan karakteristik tertentu. Tujuan utama dari clustering adalah membagi sekumpulan data ke dalam kelompok-kelompok kecil yang disebut dengan cluster, di mana setiap data dalam satu cluster memiliki kemiripan satu sama lain, dan berbeda secara signifikan dengan data yang berada di cluster lainnya. Teknik ini sangat berguna dalam proses eksplorasi data, terutama ketika belum tersedia label atau kategori yang jelas. Berbeda dengan klasifikasi yang membutuhkan data yang telah dilabeli sebelumnya, clustering bekerja secara mandiri tanpa memerlukan informasi kategori di awal. Oleh karena itu, metode ini sering digunakan untuk menemukan pola-pola tersembunyi atau struktur alami dalam data yang belum diketahui secara pasti. Dalam praktiknya, clustering banyak dimanfaatkan dalam berbagai bidang, seperti pemasaran untuk segmentasi pelanggan, pengelompokan dokumen dalam sistem informasi.

5. E-Commerce

E-commerce (electronic commerce) atau perdagangan elektronik merupakan suatu bentuk aktivitas jual beli barang atau jasa yang dilakukan melalui media elektronik, dengan internet sebagai

sarana utama. Proses ini memungkinkan konsumen dan penjual untuk berinteraksi, melakukan transaksi, serta memperoleh dan memberikan produk atau layanan tanpa harus bertemu secara fisik. Keberadaan e-commerce membawa dampak yang signifikan bagi perkembangan dunia bisnis. Salah satu keunggulan utama dari e-commerce adalah kemampuannya untuk menghilangkan batasan geografis dan waktu, memungkinkan transaksi dilakukan kapan saja dan di mana saja [11].

6. Platform Lazada

Lazada telah melihat pertumbuhan yang signifikan di pasar Asia Tenggara. Sejak akuisisi oleh Alibaba, Lazada telah memanfaatkan teknologi dan sumber daya yang lebih besar untuk memperluas pasar dan memperkenalkan lebih banyak layanan. Seiring berjalannya waktu, platform ini telah mencakup lebih banyak kategori produk dan melayani lebih banyak konsumen, dengan pengembangan pesat di negara-negara seperti Indonesia, Thailand, dan Vietnam. Lazada juga memanfaatkan data besar dan teknologi AI (artificial intelligence) untuk memahami kebiasaan konsumen, menciptakan pengalaman berbelanja yang lebih personal, serta mengoptimalkan logistik dalam pengiriman. Lazada menawarkan berbagai fitur dan layanan, termasuk sistem pembayaran terintegrasi (Lazada Wallet), layanan pengiriman yang efisien, serta fitur pengembalian barang yang memudahkan konsumen. Platform ini juga memiliki program promo besar besaran seperti

Lazada 11.11 dan 12.12 yang memberikan diskon besar bagi konsumen. Selain itu, Lazada menyediakan marketplace bagi penjual untuk membuka toko mereka di platform ini. Dalam konteks penelitian ini, Lazada dipilih sebagai objek studi kasus karena tingginya volume transaksi serta keragaman perilaku konsumennya. Selain itu, data dari pengguna Lazada cukup representatif untuk dianalisis menggunakan pendekatan data mining guna memprediksi kecenderungan perilaku belanja konsumen secara lebih mendalam[

7. Perilaku Konsumen

Perilaku konsumen merujuk pada serangkaian tindakan dan keputusan yang diambil oleh individu atau kelompok dalam memilih, membeli, menggunakan, dan mengonsumsi produk atau layanan. Perilaku ini dipengaruhi oleh berbagai faktor, termasuk kebutuhan pribadi, psikologis, sosial, serta pengaruh budaya dan lingkungan disekitar. Serta memahami perilaku konsumen sangat penting bagi perusahaan dalam merancang strategi pemasaran yang efektif, karena hal ini memungkinkan mereka untuk menyesuaikan produk, harga, promosi, dan distribusi dengan preferensi dan harapan konsumen. Proses pengambilan keputusan konsumen sendiri dapat dibagi menjadi beberapa tahap: pengenalan kebutuhan, pencarian informasi, evaluasi alternatif, keputusan pembelian, dan pasca pembelian. Dalam setiap tahap ini, konsumen mengumpulkan informasi, menilai opsi yang ada, dan akhirnya membuat keputusan berdasarkan

pertimbangan rasional dan emosional. Setelah pembelian evaluasi konsumen terhadap produk atau layanan yang diperoleh dapat mempengaruhi keputusan mereka untuk membeli lagi di masa depan, serta bagaimana mereka merekomendasikan produk itu sendiri kepada orang lain

8. Google Colab

Google Colaboratory, atau yang sering disebut Google Colab, adalah sebuah platform pemrograman berbasis cloud yang disediakan oleh Google yang memungkinkan pengguna untuk menulis dan mengeksekusi kode Python langsung di browser tanpa perlu menginstal perangkat lunak tambahan di komputer lokal[12].

9. Phyton

Python merupakan salah satu bahasa pemrograman yang saat ini sangat banyak digunakan oleh pengembang perangkat lunak dan peneliti data. Popularitas Python tidak terlepas dari struktur sintaksnya yang ringkas, mudah dibaca, dan relatif sederhana dibandingkan dengan bahasa pemrograman lainnya, sehingga cocok digunakan oleh pemula maupun profesional. Selain itu, Python bersifat open source, artinya pengguna dapat mengunduh, memodifikasi, dan mendistribusikan kembali perangkat lunak berbasis Python secara bebas[7].

3. Metodologi Penelitian

3.1 Lokasi Penelitian dan waktu penelitian

Penelitian ini menganalisis pola perilaku konsumen dalam berbelanja pada *platform e-commerce* Lazada. Penelitian ini dilakukan

secara daring dan online dengan objek penelitian yaitu pengguna platform e-commerce Lazada yang berada di wilayah Lombok Timur. Pengumpulan data dilakukan melalui penyebaran kuesioner secara online kepada responden yang pernah melakukan pembelian pada platform Lazada. Penelitian ini berlangsung dari bulan Maret hingga Juli tahun 2025.

3.2 Teknik pengumpulan data

Data yang diperlukan dalam penelitian ini diperoleh dari hasil penyebaran kuesioner kepada pengguna platform Lazada yang berada di wilayah Lombok Timur. Data tersebut berupa tanggapan responden terhadap beberapa indikator perilaku konsumen dalam berbelanja online. Selain berisi jawaban terhadap kebiasaan belanja, data yang dikumpulkan juga mencakup atribut-atribut yang menunjang proses analisis data menggunakan algoritma K-Means. Berikut beberapa metode pengumpulan data yang digunakan pada penelitian ini[13]:

- Observasi, dengan melakukan pengamatan terhadap aktivitas konsumen dalam menggunakan platform Lazada secara umum, serta memperhatikan faktor-faktor yang memengaruhi keputusan pembelian berdasarkan perilaku pengguna yang diamati melalui media digital.
- Kuesioner, dengan menyebarkan kuesioner kepada responden yang pernah berbelanja di

platform Lazada. Dengan jumlah data siap diolah sebanyak 500 Responden. Kuesioner disusun berdasarkan indikator seperti kualitas produk, harga, promosi, layanan pelanggan, kecepatan pengiriman, dan metode pembayaran yang merupakan atribut penelitian. Jawaban responden tersebut kemudian dikonversi ke dalam bentuk skala Likert agar dapat dianalisis menggunakan algoritma K-Means dengan skala 1 sampai 5 yang dimana nilai “1 = Sangat Tidak Setuju”, “2 = Tidak Setuju”, “3 Netral”, “4 = Setuju”, “5 = Sangat Setuju”.

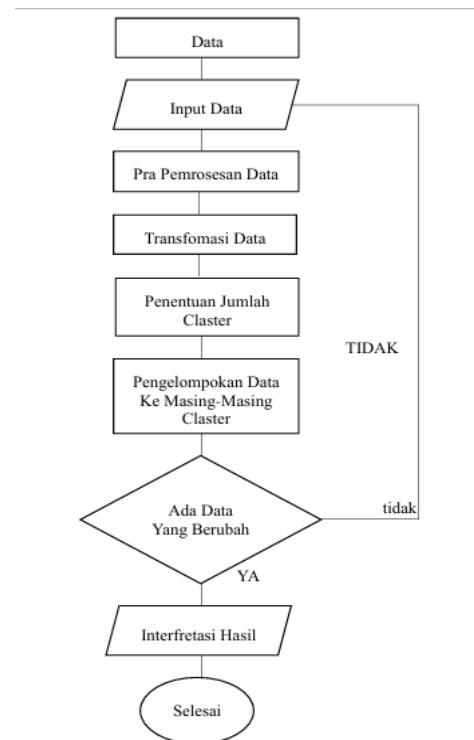
- Studi Pustaka, sebagai informasi tambahan dari berbagai referensi seperti jurnal ilmiah, artikel, buku, dan sumber daring lainnya yang berkaitan dengan perilaku konsumen, e-commerce, serta penerapan algoritma K-Means dalam proses data mining.

3.3 Tahapan penelitian

Berikut beberapa tahapan yang dilakukan pada penelitian ini adalah sebagai berikut:

- a) Tahapan sebelum pengolahan data
 - Pengumpulan Data, yang digunakan dalam penelitian ini adalah dataset private yang diperoleh melalui Kuesioner dalam kurun waktu 4 bulan yaitu mulai bulan Maret sampai dengan Juli 2025.
 - Penginputan Data, merupakan tahapan yang dilakukan setelah proses pengumpulan data.
 - Cleaning Data, berfungsi untuk memilih data mana saja yang akan digunakan, dan menghapus data yang tidak sesuai atau double.

- Analisis dan Seleksi Data, dilakukan untuk memilih atribut-atribut yang relevan dengan tujuan analisis.
- Pengolahan Data, Setelah melakukan seleksi data, tahap selanjutnya yaitu pengolahan data, untuk menghasilkan informasi yang sesuai dengan apa yang diinginkan.
- Penarikan Kesimpulan dari hasil pengolahan data, dari hasil ini diharapkan mampu memberikan informasi yang dapat digunakan dalam pengambilan keputusan.



Gambar 1 Flowchar pengolahan data

Pada tahapan ini, data konsumen dari platform Lazada diproses mulai dari pembersihan data hingga transformasi agar siap dianalisis. Selanjutnya dilakukan klasterisasi menggunakan algoritma K-Means dengan menentukan jumlah klaster optimal. Proses ini menghasilkan pengelompokan konsumen berdasarkan perilaku belanja, misalnya konsumen intensif, sedang, dan jarang berbelanja. Hasil klasterisasi kemudian divisualisasikan melalui scatter plot untuk melihat pola yang terbentuk. Dengan cara ini, K-Means dapat membantu memprediksi perilaku konsumen sehingga bermanfaat untuk strategi pemasaran yang lebih tepat sasaran[14].

4. Hasil dan Pembahasan

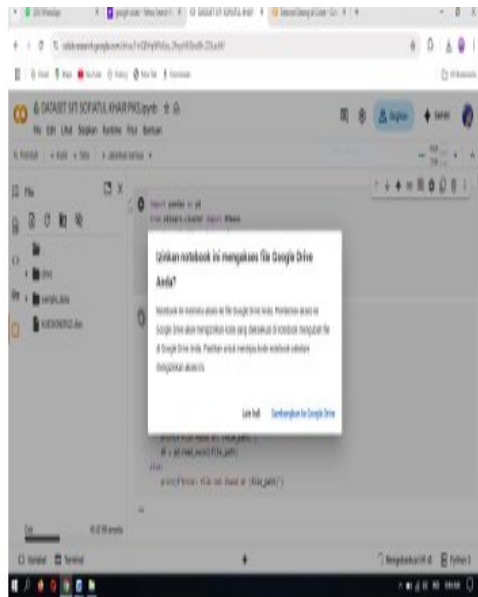
a. Pengolahan menggunakan *Google Colab*

Berikut tahapan yg dilakukan:

- Membuka lembar kerja Google colab

b) Tahapan pengolahan data

- Sinkronisasi ke google drive



Gambar 2. Proses sinkronisasi ke google drive

- Input data (*import library google colab, muat dataset, tampilkan data*)

```
[3]: df.head(10)
```

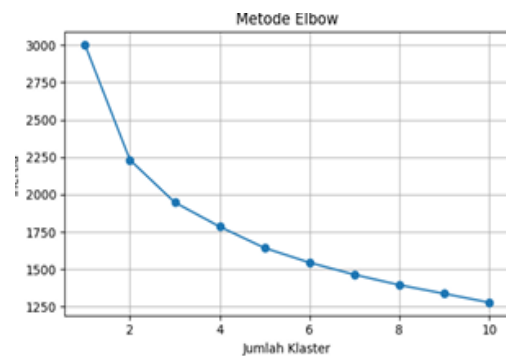
	KUALITAS PRODUK	HARGA PRODUK	PROMOSI DAN DISKON	LAYANAN PELANGGAN	KECEPATAN PENGIRIMAN	PETODE PEMBAYARAN
0	1	5	3	4	3	4
1	5	3	4	3	3	4
2	3	3	4	2	3	4
3	4	3	4	3	3	3
4	1	4	4	3	4	5
5	1	3	4	3	4	3
6	1	4	5	3	3	5
7	2	3	3	3	3	4
8	1	3	4	3	3	4
9	1	4	4	5	5	5

Gambar 3. Tampilan data setelah diinput

- Klasterisasi awal
- Analisis *Elbow*
- Plot SSE

```
plt.figure(figsize=(6, 4))
plt.plot(range(1, 11), inertia, marker='o')
plt.title('Metode Elbow')
plt.xlabel('Jumlah Kluster')
plt.ylabel('Inertia')
plt.grid(True)
plt.tight_layout()
plt.show()
```

Gambar 4. Plot SSE



Gambar 5. Titik Elbow

- Klusterisasi Optimal
- Visualisasi Klusterisasi

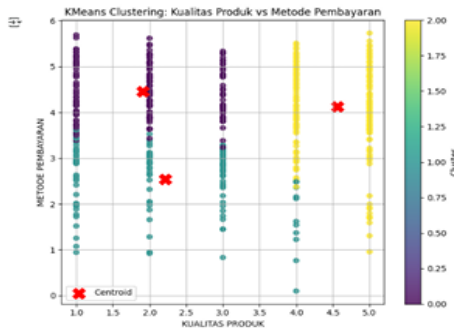
```
[67]: kualitas_produk = df['KUALITAS PRODUK'] + np.random.normal(0, 1.5, len(df))
      harga_produk = df['HARGA PRODUK'] + np.random.normal(0, 0.3, len(df))

# Clustering
X = df[['KUALITAS PRODUK', 'HARGA PRODUK']]
kmeans = KMeans(n_clusters=3, random_state=42, n_init=10) # Added n_init for reproducibility
labels = kmeans.fit_predict(X)

# Visualisasi
plt.figure(figsize=(6, 5))
scatter = plt.scatter(kualitas_produk, harga_produk, c=labels, cmap='viridis', alpha=0.8)
plt.title('KMeans Clustering')
plt.xlabel('KUALITAS PRODUK')
plt.ylabel('HARGA PRODUK')
plt.colorbar(scatter, label='Cluster')
plt.tight_layout()
plt.show()
```

Gambar 6 Visualisasi Klusterisasi

- Visualisasi Titik Centroid.



Gambar 6. Visualisasi Titik Centroid kualitas produk vs metode pembayaran

Berdasarkan hasil dari proses klustering dan analisis yang telah dilakukan menggunakan *Google Colab* dan *Ms Excel* dari data perilaku konsumen sebanyak 500 dataset, telah berhasil dikelompokkan menjadi tiga cluster.

Setiap cluster yang dihasilkan memiliki nilai persentase yang sama dengan perhitungan *Ms. Excel* dan pengolahan data menggunakan *Google Colab*.

1. Cluster 0 sebagai Cluster 1 (C1) memiliki cluster sebanyak 166 data, yang dimana kelompok ini mempresentasikan konsumen dengan tingkat perilaku konsumen tinggi dalam berbelanja di Lazada. Kelompok ini memiliki nilai rata-rata tertinggi di semua aspek.

Tabel 1. Hasil pengolahan C1

Kluster	anggota	Karakteristik cluster berdasarkan rata-rata perilaku konsumen
C1	166	Kualitas produk : 4,447 Harga produk : 4,189 Promosi dan diskon : 4, 145 Layanan pelanggan : 4,117

		Kecepatan pengiriman: 4,151 Metode pembayaran : 4,274
--	--	---

2. Untuk Cluster 1 sebagai Cluster 2 (C2) terdiri dari 174 data, yang dimana kelompok ini mempresentasikan konsumen dengan tingkat perilaku konsumen sedang.

Tabel 2 Hasil Pengolahan C2

Kluster	Anggota	Karakteristik cluster berdasarkan rata-rata perilaku konsumen
C2	174	Kualitas produk : 2,891 Harga produk : 2,724 Promosi dan diskon : 2,943 Layanan pelanggan : 2,667 Kecepatan pengiriman: 2,644 Metode pembayaran : 3,189

3. Untuk Cluster 2 sebagai Cluster 3 (C3) terdiri dari 160 data, yang dimana kelompok ini mempresentasikan konsumen dengan tingkat perilaku konsumen rendah.

Tabel 3. Hasil Pengolahan C3

Kluster	Anggota	Karakteristik cluster berdasarkan rata-rata perilaku konsumen
C3	160	Kualitas produk : 1,599 Harga produk : 3,653 Promosi dan diskon : 3,952 Layanan pelanggan : 4,007 Kecepatan pengiriman: 4,007 Metode pembayaran : 34,068

Hasil penelitian menunjukkan bahwa Cluster dengan perilaku konsumen tinggi memiliki karakteristik yang berbeda dibandingkan dengan Cluster perilaku konsumen sedang dan rendah. Berdasarkan hasil perhitungan data yang dilakukan menggunakan Google Colab dan Microsoft Excel untuk persentase kemiripan yang dihasilkan 100% sama. Berikut ini adalah persentase kemiripan nilai yang didapatkan dari hasil perhitungan Google Colab dan Excel sebagai berikut :

Tabel 4 Perbandingan Hasil Google Colab dan Ms Excel

5. Kesimpulan

Berdasarkan dari hasil penelitian yang telah dilakukan mengenai penerapan algoritma K-Means untuk pengelompokan perilaku konsumen dalam berbelanja pada platform Lazada, dapat ditarik kesimpulan, hasil analisis menunjukkan bahwa konsumen dapat

Cluster	Google Colab	Ms Excel	Persentase
C1	166	166	33,2%
C2	174	174	34,8%
C3	160	160	

dikelompokkan menjadi tiga klaster ($K=3$), yaitu kelompok konsumen dengan tingkat perilaku konsumen yang tinggi sebanyak 166 responden, kelompok konsumen dengan tingkat perilaku konsumen sedang sebanyak 174 responden, serta kelompok konsumen tingkat perilaku konsumen rendah sebanyak 160 responden. Berdasarkan hasil penelitian tersebut diharapkan bisa menjadi informasi yang berguna dan dapat membantu pihak Lazada dalam memahami pola perilaku konsumennya dalam berbelanja, sehingga perusahaan bisa mengambil keputusan dan dapat menyusun strategi pemasaran serta pengembangan layanan yang lebih tepat sasaran untuk meningkatkan daya tarik dan

loyalitas konsumen dalam berbelanja pada platform Lazada khususnya di Wilayah Lombok Timur Nusa Tenggara Barat.

DAFTAR PUSTAKA

- [1] B. Apriyanto And S. L. M. Sitio, "Penerapan K-Means Dalam Menganalisis Pola Pembelian Pelanggan Pada Data Transaksi E-Commerce," *Bit-Tech*, Vol. 7, No. 3, Pp. 790–797, 2025, Doi: 10.32877/Bt.V7i3.2195.
- [2] B. T. A. S. Ryanto And L. H. Tampubolon, "Pengaruh Promosi Sosial Media Terhadap Keputusan Pembelian E-Commerce Lazada Pada Pengguna Sosial Media Di Wilayah Dki Jakarta," *J. Transaksi*, Vol. 16, No. 1, Pp. 53–66, 2024.
- [3] R. Between, S. J. Impact, T. Related, E. Indexes, And Z. Lian-Feng, "邱殿明 1 , 张连峰 2 (1," Vol. 4, No. 1, Pp. 189–199, 2020.
- [4] A. Muliawan Nur, M. Saiful2, H. Bahtiar, And Muhammad Taufik Hidayat, "Penerapan Algoritma K-Means Clustering Dalam Mengelompokkan Smartphone Yang Rekomendasi Berdasarkan Spesifikasi," *Infotek J. Inform. Dan Teknol.*, Vol. 7, No. 2, Pp. 478–488, 2024, Doi: 10.29408/Jit.V7i2.26283.
- [5] Vizsgálatok, "濟無no Title No Title No Title," Vol. 7, No. 1, Pp. 1–23, 2016.
- [6] A. Hidayat And A. M. Nur, "Implementasi Algoritma K-Means Untuk Klasterisasi Peserta Keluarga Berencana Berdasarkan Tingkat Risiko Kehamilan Di Desa Pringgasela Selatan," Vol. 1, No. 2, Pp. 154–166, 2023.
- [7] V. No, A. M. Nur, H. Bahtiar, M. A. Jannah, And V. No, "Infotek : Jurnal Informatika Dan Teknologi Implementasi Algoritma K-Means Clustering Dalam Mengelompokkan Kepatuhan Wajib Pajak Bumi Dan Bangunan Dengan Optimasi Elbow," *Infotek J. Inform. Dan Teknol.*, Vol. 8, No. 1, 2025.
- [8] M. Sirotjudin And Meliza, "Journal Of Accounting And Management ' S Student (Jam ' S)," *J. Account. Manag. Student*, Vol. 1, Pp. 1–9, 2023.
- [9] V. No, Z. Amri, M. Rodi, M. N. Wathani, A. Bagja, And V. No, "Infotek : Jurnal Informatika Dan Teknologi Prediksi Diabetes Menggunakan Algoritma K-Nearest (Knn) Teknik Smote-Enn Infotek : Jurnal Informatika Dan Teknologi," Vol. 8, No. 1, Pp. 193–204, 2025.
- [10] F. Nuraeni, M. Hasby, And M. Firdaus, "Model Klasterisasi Data Pertanian Menggunakan Algoritma K-Means," Pp. 890–898, 2025, Doi: 10.33364/Algoritma/V.22-2.2054.
- [11] B. Com, M. Rafi, And H. Arsyad, "Klusterisasi Data Review Pengguna Aplikasi Marketplace," Vol. 09, Pp. 1–10, 2024.
- [12] M. Arifin, F. Helmi, And Iddrus, "Analisis Perbandingan Algoritma Asosiasi Data Mining Pada Minimarket Adi Poday Dengan Google Collab," *J. Algoritm.*, Vol. 22, No. 1, Pp. 103–114, 2025, Doi: 10.33364/Algoritma/V.22-1.2177.
- [13] A. A. Kari And V. Sofica, "Analisis Pelanggan Potensial Dengan Algoritma K-Means Untuk Meningkatkan Jumlah



Paket Pada Pt Cipta Teknindo Rekatama,” *Informatics Educ. Prof. J. Informatics*, Vol. 9, No. 2, P. 135, 2024, Doi: 10.51211/Itbi.V9i2.3037.

- [14] I. D. A. I. Saraswati, I. M. A. O. Gunawan, And I. M. A. W. Putra,

“Analisis Keranjang Belanja Pada Data Ritel Non-Toko Menggunakan Algoritma Fp-Growth,” *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 13, no. 3, pp. 1692–1704, 2024, doi: 10.35889/jutisi.v13i3.2238.