

## JUDUL

# Analisis Perilaku Pelanggan E-commerce Menggunakan Model Klustering Dengan Algoritma K-Means

## BAB I PENDAHULUAN

### 1.1 Latar Belakang

E-commerce telah menjadi salah satu industri yang paling berkembang pesat dalam beberapa tahun terakhir. Hal ini didorong oleh meningkatnya penetrasi internet dan smartphone di seluruh dunia. Menurut laporan dari Statista, pasar e-commerce global diperkirakan akan mencapai nilai \$5,4 triliun pada tahun 2025. Dalam persaingan yang semakin ketat, pelaku bisnis e-commerce dituntut untuk dapat memahami perilaku pelanggan dengan lebih baik.

Hal ini diperlukan untuk dapat mengembangkan strategi pemasaran dan penjualan yang lebih efektif. Salah satu pendekatan yang dapat digunakan untuk memahami perilaku pelanggan adalah dengan menggunakan machine learning. Dengan menggunakan machine learning, pelaku bisnis e-commerce dapat menganalisis data perilaku pelanggan secara lebih mendalam. Hal ini memungkinkan mereka untuk mengidentifikasi pola dan tren yang dapat digunakan untuk memprediksi perilaku pelanggan di masa depan (Mauludin, M.S., 2022).

Urgensi penggunaan Machine Learning untuk menganalisis data pelanggan e-commerce diantaranya karena peningkatan volume data yang semakin besar, menjadi semakin sulit untuk menganalisis data secara manual. Machine learning dapat membantu pelaku bisnis e-commerce untuk menganalisis data secara lebih cepat dan akurat (Alfarizki, M.P et al., 2022). Pelaku bisnis e-commerce dituntut untuk dapat memberikan pengalaman belanja yang lebih baik bagi pelanggan. Machine learning dapat membantu pelaku bisnis e-commerce untuk memberikan rekomendasi produk yang lebih relevan, meningkatkan personalisasi, dan meningkatkan kepuasan pelanggan.

Perilaku konsumen adalah suatu kondisi ketika konsumen memilih, membeli, serta memanfaatkan produk dan jasa demi memenuhi kebutuhan atau keinginan mereka. Keputusan pembelian merupakan suatu sikap dimana seorang konsumen benar-benar membeli. Sifat konsumen yang selektif dalam memilih suatu produk sangat

mempengaruhi keputusan pembelian. Pada penelitian ini akan dilakukan analisis perilaku pelanggan berdasarkan metrik pengukuran yang ada di dalam analisis RFM (*Recency, Frequency, Monetary*) untuk mengidentifikasi dan memahami segmen pelanggan yang berbeda berdasarkan perilaku mereka dalam tiga dimensi tersebut.

Dengan memahami nilai RFM perusahaan dapat mengetahui pelanggan mulai dari yang loyal, berisiko tidak lagi menggunakan platform perusahaan, sampai pelanggan yang sudah hilang. Dengan informasi tersebut perusahaan dapat memperbaiki kualitas pelayanan dan menjaga pelanggan yang sudah loyal. Model yang digunakan pada penelitian ini adalah k-means untuk mengkluster pelanggan berdasarkan *recency, frequency, dan monetary*.

Selain menggunakan teknik analisis RFM, penelitian ini juga mengeksplorasi aktivitas pelanggan saat menggunakan platform e-commerce kemudian dibuatkan model untuk memprediksi faktor yang paling berpengaruh terhadap loyalitas pelanggan selama menggunakan platform dengan menggunakan metode ensemble learning untuk membandingkan kinerja masing-masing algoritma. Model prediksi ini bermanfaat untuk perusahaan mengetahui kesulitan pelanggan saat mengakses menu tertentu atau menu apa yang disukai, sehingga perusahaan dapat terus meningkatkan kualitas platform nya.

Penelitian terdahulu menerapkan Logistic Regression untuk menganalisa data yang ada pada E-Commerce untuk mengetahui tingkat loyalitas pelanggan mereka. Penelitian ini menunjukkan bahwa prediksi loyalitas pelanggan dapat di dapatkan dengan adanya kepercayaan yang berpengaruh positif dan signifikan terhadap Loyalitas Pelanggan e-commerce. Hal ini menunjukkan bahwa semakin tinggi tingkat kepercayaan dan kepuasan maka loyalitas pelanggan akan mengalami peningkatan. Dengan menggunakan Logistic Regression sebagai algoritma yang digunakan dalam penelitian ini mendapatkan tingkat keakuratan 0,82(Rahman, A., 2022). Kekurangan dari penelitian ini adalah metode yang sederhana menggunakan logistic regression yang mungkin tidak dapat menangkap kompleksitas dan hubungan non-linear antara variabel.

Penelitian lain membandingkan algoritme pembelajaran mesin konvensional dan jaringan syaraf tiruan untuk memprediksi niat pembelanja online berdasarkan berbagai fitur seperti perilaku penelusuran, informasi terkait produk, dan hari-hari khusus. Penelitian ini menyimpulkan bahwa jaringan syaraf tiruan dengan enam lapisan tersembunyi decoder-encoder, fungsi aktivasi sigmoid, pengoptimal adagrad,

menghasilkan tingkat akurasi 98.48%, skor F1 95.06%, presisi 97.36%, recall 96.81%, dan AUC 96.81% (Agustyaningrum et al., 2021). Namun kekurangan di penelitian tidak melakukan data *preprocessing* terlebih dahulu sehingga skor kinerja yang dihasilkan mungkin bias, skor yang terlalu tinggi juga bisa menyebabkan *overfitting*, dan belum melakukan evaluasi.

Penelitian lain menerapkan model deep learning berbasis Artificial Neural Network (ANN), tiga model ANN dibandingkan dengan jumlah layer dan node yang berbeda, dan menguji akurasi dan loss dari masing-masing model untuk memprediksi churn pelanggan berdasarkan data fitur seperti skor kredit, geografi, gender, usia, saldo, jumlah produk. Model terbaik adalah model 2 dengan 5 layer dan 12 node di setiap hidden layer, yang memiliki akurasi 86% dan loss 0.1233 (Satria, W.A et al., 2020). Penelitian ini memiliki kekurangan tidak mengevaluasi model untuk menguji apakah model terdapat bias.

Penelitian baru baru ini mencoba mengukur *customer lifetime value* dan mengelompokan pelanggan berdasarkan nilai tersebut menggunakan algoritma boosting, seperti AdaBoost dan gradient boosting decision tree, memiliki kinerja terbaik (Sun, Y et al., 2023). Penelitian ini mengihalmi penulisan ini karena memberikan analisis yang empiris dan teknik validasi silang yang mendukung kinerja model.

Berdasarkan latar belakang tersebut penulis membuat model pembelajaran mesin untuk memahami faktor yang menyebabkan seorang pelanggan memutuskan untuk melanjutkan pembelian atau meninggalkan, menggunakan metode ensemble learning untuk membandingkan model yang terbaik berdasarkan evaluasi dan validasi yang dilakukan. Data yang digunakan adalah data transaksi dan klik pengguna e-commerce, berbagai EDA dan fitur engineering akan dijelaskan pada tulisan ini.

## 1.2 Rumusan Masalah

1. Apakah terdapat pola perilaku pelanggan yang dapat diidentifikasi dengan menggunakan teknik machine learning dalam e-commerce, dan bagaimana pola tersebut memengaruhi keputusan pembelian?
2. Apakah terdapat faktor-faktor tertentu dalam perilaku pelanggan e-commerce yang dapat diidentifikasi melalui analisis machine learning yang memengaruhi retensi pelanggan dan loyalitas terhadap platform e-commerce?

3. Bagaimana machine learning dapat digunakan untuk mengidentifikasi pelanggan yang berpotensi berhenti berbelanja di platform e-commerce dan menerapkan strategi retensi yang tepat?
4. Apakah penggunaan algoritma machine learning dapat membantu e-commerce dalam memahami tren perubahan perilaku pelanggan seiring waktu dan mengadaptasi strategi pemasaran mereka?

### 1.3 Tujuan Penelitian

1. Mengembangkan model machine learning yang dapat memprediksi preferensi pembelian pelanggan e-commerce berdasarkan data historis mereka, dengan akurasi yang tinggi.
2. Mengeksplorasi faktor-faktor yang memengaruhi retensi pelanggan dan loyalitas terhadap platform e-commerce, dan mengidentifikasi tindakan yang dapat diambil berdasarkan analisis machine learning untuk meningkatkan retensi pelanggan.
3. Membantu perusahaan e-commerce untuk meningkatkan retensi pelanggan, mengembangkan strategi pemasaran yang lebih efektif, dan meningkatkan profitabilitas.

### 1.4 Batasan Masalah

1. Penelitian ini hanya menggunakan algoritma decision tree dan variasi nya, serta menggunakan satu algoritma kluster untuk membuat segmentasi pelanggan.
2. Pengujian model hanya menggunakan data testing yang dipecah dari dataset, tidak melakukan pengujian lagi pada data baru.
3. Pengujian kinerja model hanya berdasarkan evaluasi persisi, recall, skor F1, dan learning curve.

### 1.5 Kontribusi Penelitian

Penelitian ini dapat memberikan kontribusi kepada perusahaan e-commerce diantaranya:

1. Membantu perusahaan memahami bagaimana perilaku pelanggan berdasarkan kapan pembelian terakhir mereka (*Recency*), seberapa sering mereka melakukan pembelian (*Frequency*), dan berapa total pengeluaran mereka (*Monetary*). Kemudian apa hambatan pelanggan saat menggunakan platform.

2. Dengan menganalisis data perilaku pelanggan, perusahaan dapat memahami kebutuhan pelanggan dan mengeksplorasi aturan perilaku mereka.
3. Peningkatan pengalaman pelanggan, penelitian ini dapat digunakan untuk meningkatkan pengalaman pelanggan dengan meningkatkan kualitas pelayanan dan teknologi di platform e-commerce.
4. Analisis RFM dapat membantu perusahaan ecommerce dalam memahami perilaku pelanggan mereka, dan bagaimana informasi ini dapat digunakan untuk meningkatkan penjualan dan loyalitas pelanggan.

