

ABSTRAK

Evelyn Angelica (01112210020)

ANALISIS SEGMENTASI PELANGGAN MENGGUNAKAN METODE *MACHINE LEARNING* BERDASARKAN DATA PENJUALAN *ONLINE*

Skripsi, Fakultas Sains dan Teknologi (2025)

(xiv + 71 halaman, 25 gambar, 26 tabel)

Strategi pemasaran yang sesuai diperlukan dalam meningkat penjualan. Oleh karena itu, segmentasi pelanggan dilakukan guna mengetahui karakteristik dari setiap pelanggan. Penelitian ini bertujuan untuk mengklasifikasi pelanggan ke dalam kelompok guna melihat karakteristik dari setiap pelanggan berdasarkan data penjualan sebuah toko *online*. Penelitian ini menggunakan 4 metode klasifikasi, yaitu *Decision Tree*, *Random Forest*, *XGBoost*, dan *Support Vector Machine*, serta menggunakan 1 metode klastering, yaitu *Hierarchical Clustering*. Metode terbaik yaitu *XGBoost* dengan nilai akurasi sebesar 70,02% dengan *Hour Difference* sebagai variabel yang memiliki hubungan tinggi terhadap menentukan pelanggan loyal atau tidak. Metode *Hierarchical Clustering* menunjukkan bahwa mayoritas pelanggan melakukan transaksi dengan kategori *sparepart*, tidak menggunakan promo gratis biaya pengiriman, pada hari kerja (Senin - Jumat), dan pada siang hari dengan mayoritas pelanggan berdomisili di Jakarta dan sekitarnya.

Kata Kunci: segmentasi pelanggan, *Decision Tree*, *Random Forest*, *XGBoost*,
Support Vector Machine, *Hierarchical Clustering*.

Referensi : 32 (1986-2024)

ABSTRACT

Evelyn Angelica (01112210020)

CUSTOMER SEGMENTATION ANALYSIS USING MACHINE LEARNING METHOD BASED ON ONLINE SALES DATA

Thesis, Faculty of Science and Technology (2025)

(xiv + 71 pages, 25 figures, 26 tables)

An appropriate marketing strategy is needed to increase sales. Therefore, customer segmentation is carried out to determine the characteristics of each customer. This study aims to classify customers into groups to see the characteristics of each customer based on sales data from an online store. This study uses 4 classifications methods, namely Decision Tree, Random Forest, XGBoost, and Support Vector Machine, and uses 1 clustering method, namely Hierarchical Clustering. The best method is XGBoost with an accuracy value of 70.02% and Hour Difference as the variable that has a high relationship in determining whether a customer is loyal or not. The Hierarchical Clustering method shows that most customers make transactions in the spare part category, do not use free shipping promotions, shop on weekdays (Monday to Friday), during the afternoon, and are predominantly located in Jakarta and its surrounding areas.

Keywords: customer segmentation, Decision Tree, Random Forest, XGBoost,
Support Vector Machine, Hierarchical Clustering.

Reference: 32 (1986-2024)