

ABSTRACT

Riena Pribadi Gronloh (01804230020)

PREDICTION MODEL USING MACHINE LEARNING : ANALYSIS OF DETERMINANTS OF CUSTOMER CHURN AT PT XYZ

(xv + 77 pages ; 8 figures ; 12 tables ; 3 appendices)

This study aims to identify the factors influencing customer churn at PT XYZ, a B2B application-based company selling essential goods. Machine learning algorithms such as Random Forest and Logistic Regression were used to predict churn based on demographic and behavioral variables, including age, membership duration, monthly transaction averages, spending value, and product variety. Transaction data from January 2023 to August 2024 was analyzed to understand partner behavior patterns. This study demonstrated that the Random Forest algorithm achieved superior predictive performance in identifying potential customer churn at PT XYZ, with an accuracy of 87.17%, precision of 74.40%, recall of 71.10%, and an ROC-AUC score of 0.908. In comparison, Logistic Regression yielded an accuracy of 78.07%, precision of 57.70%, recall of 33.30%, and an ROC-AUC score of 0.733. This study provides strategic insights for PT XYZ to reduce churn and maintain customer purchase retention through a data-driven approach.

Reference : 60 (2002 – 2024)

Keywords : Customer Churn, Machine Learning, Random Forest, Logistic Regression, PT XYZ, Purchase Retention.

ABSTRAK

Riena Pribadi Gronloh (01804230020)

MODEL PREDIKSI MENGGUNAKAN *MACHINE LEARNING*: ANALISIS FAKTOR-FAKTOR PENENTU *CUSTOMER CHURN* DI PT XYZ

(xv + 77 halaman ; 8 gambar ; 12 tabel ; 3 lampiran)

Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor yang memengaruhi churn pelanggan di PT XYZ, sebuah perusahaan B2B berbasis aplikasi yang menjual kebutuhan pokok. Algoritma machine learning seperti Random Forest dan Regresi Logistik digunakan untuk memprediksi churn berdasarkan variabel demografis dan perilaku pelanggan, seperti usia, lama keanggotaan, rata-rata transaksi bulanan, nilai belanja, dan variasi produk. Data transaksi periode Januari 2023 hingga Agustus 2024 dianalisis untuk memahami pola perilaku mitra. Penelitian ini menunjukkan bahwa algoritma Random Forest memiliki kinerja prediktif yang unggul dalam mengidentifikasi potensi churn pelanggan di PT XYZ, dengan akurasi sebesar 87,17%, precision 74,40%, recall 71,10%, dan skor ROC-AUC sebesar 0,908. Sebagai perbandingan, Regresi Logistik menghasilkan akurasi sebesar 78,07%, precision 57,70%, recall 33,30%, dan skor ROC-AUC sebesar 0,733. Penelitian ini memberikan wawasan strategis bagi PT XYZ untuk mengurangi churn dan mempertahankan retensi pembelian pelanggan melalui pendekatan berbasis data.

Referensi : 60 (2002 – 2024)

Kata kunci : *Customer Churn*, *Machine Learning*, Random Forest, Regresi Logistik, Prediksi, PT XYZ, Retensi Pembelian.