

## ABSTRAK

Penelitian ini bertujuan untuk meningkatkan akurasi prediksi dalam analisis penjualan dan kepuasan dengan menangani data yang hilang menggunakan imputasi rata-rata dan model machine learning. Nilai numerik yang hilang ditangani dengan imputasi rata-rata, sementara nilai kategori yang hilang dikecualikan. Dataset dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian. Model prediktif dibangun menggunakan algoritma Decision Tree, k-Nearest Neighbors (KNN), dan Random Forest. Kinerja model dievaluasi menggunakan metrik seperti AUC, akurasi, F1-score, presisi, recall, dan MCC. Hasil penelitian menunjukkan bahwa proses imputasi secara signifikan meningkatkan kinerja model, dengan Random Forest mencapai AUC tertinggi (0,999) dan akurasi klasifikasi tertinggi (0,982). Hal ini menekankan peran penting imputasi dalam meningkatkan kualitas data dan keandalan prediksi. Selain itu, penelitian ini menetapkan Random Forest sebagai metode yang kuat dalam menangani data yang hilang serta menghasilkan hasil prediksi yang unggul pada dataset serupa.

Kata Kunci: Desicion Tree, Imputation, Missing Data, K-Nearest Neighbors, Random Forest