

Abstrak—Menghadapi persaingan yang semakin ketat dan meningkatnya permintaan akan layanan digital, perusahaan telekomunikasi menghadapi tantangan yang lebih besar dalam mengelola retensi pelanggan dan mengurangi *churn* (tingkat pelanggan yang berhenti berlangganan). Prediksi *churn* pelanggan menjadi sangat penting, karena mempertahankan pelanggan yang sudah ada lebih hemat biaya dibandingkan dengan mendapatkan pelanggan baru. Penelitian ini menyelidiki efektivitas algoritma *machine learning*, dengan fokus khusus pada model *Random Forest*, dalam memprediksi *churn* pelanggan di sektor telekomunikasi. Menggunakan dataset *Telco Customer Churn* dari Kaggle, penelitian ini melakukan pra-pemrosesan data yang ekstensif, termasuk menangani nilai yang hilang, mengkodekan variabel kategorikal, melakukan penskalaan fitur numerik, dan menyeimbangkan data melalui *Synthetic Minority Oversampling Technique* (SMOTE) untuk meningkatkan akurasi dan interpretabilitas model. Analisis komparatif terhadap beberapa algoritma, termasuk *Random Forest*, *Logistic Regression*, *K-Nearest Neighbors*, *Support Vector Classifier* (SVC), dan *Gradient Boosting*, dilakukan berdasarkan skor ROC-AUC dan metrik klasifikasi tambahan untuk mengevaluasi kinerja prediktif. Hasil penelitian menunjukkan bahwa *Random Forest* mencapai akurasi tertinggi dengan skor ROC-AUC sebesar 0.85, yang menggarisbawahi ketangguhan dan kesesuaiannya dalam menangani dataset kompleks dan berdimensi tinggi yang sering ditemui dalam prediksi *churn*. Analisis pentingnya fitur menyoroti prediktor kritis, seperti *Total Charges*, masa berlaku pelanggan, dan *Monthly Charges*, yang memberikan wawasan yang dapat ditindaklanjuti untuk mengembangkan strategi retensi pelanggan yang terarah. Selain itu, dengan memasukkan analisis biaya-manfaat, penelitian ini menunjukkan keunggulan finansial dari mengurangi tingkat *churn*, menekankan implikasi praktis bagi perusahaan telekomunikasi dalam merancang strategi retensi yang efektif. Penelitian ini menyoroti keefektifan algoritma *Random Forest* dalam prediksi *churn* di sektor telekomunikasi, menawarkan wawasan berbasis model yang mendukung praktik berbasis data untuk meningkatkan loyalitas pelanggan dan mengoptimalkan strategi retensi.

Kata Kunci—Prediksi *churn* pelanggan, *Random Forest*, *Telco Customer Churn*, pentingnya fitur, SMOTE, *machine learning*, pertumbuhan ekonomi. .