

ABSTRAK

Pelanggan yang setia merupakan aset berharga dari sebuah perusahaan telekomunikasi. Oleh karena itu, perusahaan akan terus berusaha untuk meningkatkan kepuasan pelanggan guna menghindari kerugian akibat *churn*. Dari permasalahan ini diperlukan sebuah model klasifikasi pelanggan yang akurat untuk membantu perusahaan mengurangi *churn*. Pelanggan yang *churn* biasanya tidak banyak sehingga disebut sebagai kelas minor, 3.75% dari keseluruhan *dataset*, yang mengakibatkan ketidakseimbangan data. Ketidakseimbangan data mengakibatkan sulitnya membangun model prediksi. Salah satu solusi untuk menyelesaikan *imbalance data* yaitu dengan menerapkan teknik *undersampling*.

Data yang digunakan dalam Tugas Akhir ini berasal dari data pelanggan PT. Telekomunikasi Indonesia Regional 7. Pada penelitian Tugas Akhir ini dilakukan penanganan *imbalance data* menggunakan *Neighborhood Cleaning Rule* (NCL) dan *Random Undersampling* (RUS). Kemudian setelah itu, hasilnya dibuat model klasifikasi menggunakan metode *Weighted Random Forest* (WRF) untuk mengklasifikasikan kelas *churn* atau *not churn*. Penelitian yang dilakukan terbukti dapat meningkatkan performansi model prediksi yang dibangun menggunakan WRF dengan menerapkan kombinasi NCL dan RUS, yakni menghasilkan tingkat akurasi model klasifikasi sebesar 93.799% dengan nilai F1 terbaik, yaitu 90.651% dan *oob error rate* 7.36%.

Kata kunci: *churn prediction, imbalance data, undersampling, weighted random forest.*