

ABSTRAKSI

Pelanggan cabut telah menjadi masalah yang signifikan dan juga menjadi tantangan bagi Perusahaan Telekomunikasi. Hal ini perlu untuk dievaluasi apa permasalahan besar dari Pelanggan cabut sehingga pihak manajemen Perusahaan dapat membuat kebijakan yang tepat terkait permasalahan ini dan dapat mempertahankan Pelanggan yang ada. Data Pelanggan cabut yang termasuk kategori cabut Atas Permintaan Sendiri (APS) di Perusahaan ini sangat kecil. Kekurangan data ini mengakibatkan ketidakseimbangan data dan ini menyebabkan kesulitan-kesulitan dalam pengembangan model prediksi yang baik. Isu ini kemudian menjadi permasalahan yang menantang dalam pembelajaran mesin.

Penelitian ini menyelidiki bagaimana menangani ketidakseimbangan kelas dalam prediksi cabutan menggunakan *data Preprocessing*, kombinasi metode *sampling* antara *Synthetic Minority Over Sampling* (SMOTE) dan *Random Under Sampling* (RUS). Usulan kedua Kombinasi teknik tersebut mencoba untuk menangani kedua masalah diatasehingga dapat menghasilkan performansi model prediksi cabut yang lebih baik. SMOTE akan memperbanyak data *minor* dan menghindari dari masalah *overfitting*. RUS akan mengurangi jumlah data mayor dan mempertahankan dari pengurangan yang berlebihan sehingga menghindari kehilangan informasi. Teknik pertama adalah *data preprocessing* dan teknik kedua adalah *classification*. Kumpulan data yang digunakan dalam penelitian ini adalah data Internet *Broadband* PT. Telkom Regional 6 Kalimantan.

Hasil penelitian ini mencoba membangun model prediksi cabut menggunakan salah satu pendekatan *ensemble*, metode *Bagging* dikombinasikan dengan C4.5 sebagai classifier tunggal. Karena kesederhanaannya, metode *Bagging* dapat meningkatkan hasil kinerja dengan penggunaan data *preprocessing* dan melatih *classifier* tunggal-nya. Hasil penelitian ini menunjukkan kombinasi metode *sampling* SMOTE dan RUS dapat membuat data seimbang dan menaikkan kinerja prediksi cabut menjadi 571% dibandingkan tanpa *sampling* dan penggunaan metode *Bagging* memberikan sinergi yang positif untuk meningkatkan kinerja sampai 56%.

Kata kunci: *Random Under Sampling*, SMOTE, Metode *Bagging*, Masalah ketidakseimbangan, Prediksi Cabut