

Abstrak

Churn prediction adalah salah satu aplikasi dari task *data mining* yang bertujuan untuk memprediksi para pelanggan yang berpotensi untuk *churn*. *Churn prediction* merupakan salah satu dari kasus *imbalance class*. Permasalahan *imbalance class* pada Clementine 9.0 di atasi dengan *node balancing* (menggunakan teknik *undersampling*, *oversampling* atau gabungan keduanya). Permasalahan yang ditimbulkan adalah model yang dihasilkan terlalu *overfitting*. *DataboostIM* adalah sebuah metoda yang dikembangkan dari metoda *boosting* yang dipadukan dengan konsep *data generation*. *Boosting* merupakan salah satu *ensemble method* yang biasa digunakan untuk memperbaiki akurasi secara umum pada permasalahan *imbalance class*. Kelemahan dari *boosting* adalah terlalu menekankan pada *hard examples* di kelas *minor*, hal ini menyebabkan akurasi dari kelas *major* cenderung turun walaupun akurasi kelas *minor* meningkat. *DataboostIM* tidak hanya memperbaiki akurasi dari kelas minor tetapi juga mempertahankan akurasi dari kelas mayor. Didalam Tugas Akhir ini, dibuat perangkat lunak yang mengimplementasikan metoda *DataboostIM* dengan *base classifier*-nya terintegrasi dengan model klasifikasi pada Clementine 9.0. Implementasi dilakukan dengan mengambil study kasus *churn prediction* pada sebuah perusahaan Telekomunikasi di Indonesia. Hasil yang didapat menunjukkan bahwa *DataboostIM* mampu memperbaiki kelas minor dan tetap mempertahankan akurasi pada kelas mayor, selain itu *DataboostIM* mampu memberikan hasil *lift curve* yang lebih baik daripada *node balancing* Clementine 9.0.

Kata kunci: *databoostIM*, *boosting*, *imbalance class*, *data generation*, *churn prediction*.