

ABSTRAK

Penelitian ini didasari dari kondisi banyaknya pelanggan data internet PT Telkomsel yang berpindah menjadi pelanggan *non data*. Hal ini salah satunya disebabkan karena adanya keinginan pelanggan untuk terus mencari produk atau layanan yang benar-benar mereka butuhkan dan memberikan *benefit* yang lebih bagus. Selain itu, persaingan antar operator telekomunikasi juga menjadi salah satu penyebabnya terkait dengan penawaran harga yang lebih kompetitif dan *benefit* yang lebih menguntungkan.

Dengan penelitian ini, penulis ingin memberikan *insight* kepada Telkomsel dalam memprediksi pelanggan data yang berubah menjadi *non data* user secara akurat menggunakan *predictive analytics*. Kemudian dibuatkan segmentasi terhadap hasil dari prediksi berdasarkan *customer behavior*, guna mengelompokkan pelanggan tersebut yang berpotensi berubah menjadi pelanggan *non data* dengan penawaran layanan yang sesuai.

Sumber data yang dipergunakan berasal dari data Telkomsel yang berkaitan dengan *behavior* pelanggan meliputi total tagihan, penggunaan data, SMS, percakapan, dan informasi *lapse*, dalam periode bulan Januari sampai dengan bulan Maret 2019. Dilakukan pengolahan data melalui *data mining* dengan model *logistic regression* menggunakan SPSS data modeler.

Logistik regression ini digunakan untuk memprediksi pelanggan yang akan *lapse* yang sebelumnya *data user* menjadi *non data user* dengan menggunakan 14 input variabel dengan tipe numerik. *Analytics based table* diambil dari data pelanggan yang *lapse* di bulan April 2019 sebesar 40.991 dan mengambil *sample* pelanggan yang tetap menjadi pelanggan data dibulan tersebut sebesar 12.297 pelanggan dari total 390.567 pelanggan. *Analytics based table* tersebut dibagi menjadi 2 *data set* yaitu 70% *training data set* & 30% *testing data set* baik untuk pelanggan yang *lapse* maupun yang tidak *lapse* dari *data user*. Maka dari total *training* dan *testing data set* sebesar 53.288 pelanggan, diprediksi 44.818 pelanggan akan *lapse* menjadi *non data user* dengan akurasi model mencapai 86% dan menghasilkan top 6 *significant variable* yaitu *Day of Broadband Average*, *Total Recharge Average*, *Length of Stay Average*, *Volume quota chat Average*, *Volume quota Video Average*, dan *Payload Average*.

Adapun hasil segmentasi menggunakan *k-means*, diperoleh jumlah *cluster* sebanyak 5 dengan *silhouette index* 0.43 yang berarti validitas segmentasi masuk dalam kategori *fair*.

Dari hasil penelitian ini diharapkan dapat memberikan gambaran model dan hasil prediksi migrasi dan juga karakteristik dari segmentasi pelanggan prabayar yang akan *lapse* menjadi *non data user*, dan juga dapat digunakan oleh operator telekomunikasi dalam memberikan penawaran khusus dan program pemasaran yang tepat sesuai dengan karakteristik segmen.

Kata Kunci: Segmentasi pasar, perilaku pelanggan, strategi pemasaran, *logistic regression*, *K-Means*.