

ABSTRAK

PT Astra International Tbk - Honda Sales Operation (Astra Motor) bergerak dalam bidang distribusi dan penjualan retail kendaraan roda dua bermerek Honda terbesar di Indonesia. Namun hanya terdapat sekitar 30% orang yang melakukan pembelian sepeda motor dari sejumlah orang yang telah menjadi prospek calon pelanggan. Strategi untuk menawarkan atau mempromosikan produk kepada konsumen harus dilakukan dengan tepat. Dari permasalahan yang ada, maka diperlukan sebuah segmentasi pelanggan serta pemodelan untuk memprediksi jenis motor yang akan dibeli agar pemasaran dapat dilakukan secara efektif dalam aspek biaya, waktu, dan tenaga. Sebelum melakukan segmentasi pelanggan, peneliti menggunakan metode Random Under-Sampling untuk menangani data imbalanced agar dapat menghasilkan model yang optimal. Kemudian peneliti melakukan segmentasi pelanggan dengan memanfaatkan metode K-Modes. Setelah melakukan segmentasi, maka peneliti menyusun model prediktif untuk memprediksi jenis motor yang akan dibeli. Metode yang digunakan adalah Multinomial Logistik dikarenakan variabel dependen berkategori lebih dari dua. Variabel independen yang digunakan pada model tersebut mencakup faktor eksternal, faktor internal, dan hasil segmentasi pelanggan menggunakan *clustering*. Pemodelan data yang dilakukan menghasilkan nilai akurasi dan F1-score yang hampir sama antara data *training* dan *testing*. Maka, dapat disimpulkan bahwa model yang dihasilkan memiliki performa dan kemampuan beradaptasi yang sangat baik dengan 9 variabel yang signifikan terhadap model.

Kata kunci: jenis motor, k-modes, *random under-sampling*, regresi logistik multinomial, segmentasi pelanggan

ABSTRACT

PT Astra International Tbk - Honda Sales Operation (Astra Motor) is engaged in distribution and retail sales of the largest Honda branded two-wheeled vehicles in Indonesia. However, there are only about 30% of people who buy motorbikes from a number of people who have become prospects for potential customers. The strategy for offering or promoting products to consumers must be done appropriately. From the existing problems, a customer segmentation and modeling is needed to predict the type of motorbike to be purchased so that marketing can be carried out effectively in the aspects of cost, time and effort. Before conducting customer segmentation, researchers used the Random Under-Sampling method to handle imbalanced data in order to produce an optimal model. Then researchers perform customer segmentation by utilizing the K-Modes method. After segmenting, the researchers developed a predictive model to predict the type of motorbike to be purchased. The method used is Multinomial Logistics because the dependent variable has more than two categories. The independent variables used in this model include external factors, internal factors, and the results of customer segmentation using clustering. The data modeling carried out produces almost the same accuracy and F1-score values between the training and testing data. So, it can be concluded that the resulting model has very good performance and adaptability with 9 variables that are significant to the model.

Keywords: *customer segmentation, k-modes, motor type, multinomial logistic regression, random under-sampling*