

ABSTRAK

Pemusatan penjualan yang terfokus kepada produk yang sedang banyak peminat dan populer adalah salah satu teknik penjualan supermarket. Teknik penjualan *seasonal* seperti ini terkadang memberikan dampak yang dapat terlihat sangat jelas dengan adanya ketidakseimbangan penjualan pada produk yang telah ada di supermarket. Ketidakseimbangan penjualan dapat menjadi penyebab awal suatu produk kehilangan peminat dan menjadi produk yang akhirnya tersingkirkan dari penjualan ditoko. Dengan adanya model klasifikasi yang dibuat untuk memprediksi produk yang akan terseleksi atau churn, dapat membantu staff dalam melakukan pemerataan penjualan setiap produk. Semakin banyak produk yang churn akibat kekurangan peminat dapat mempengaruhi penjualan keseluruhan dari supermarket. Tujuan dari penelitian ini adalah untuk membantu staff dalam mengklasifikasi produk yang berpotensi churn. Model klasifikasi terdiri dari 3 model dengan algoritma berbeda dan didapatkan hasil bahwa penerapan algoritma Random Forest lebih efektif untuk memprediksi data dengan 96% keakuratan dibandingkan 81% pada algoritma Logistic Regression dan 46% pada algoritma Support Vector Machine.

ABSTRACT

Concentration of sales that are focused on products that are in great demand and are popular is one of the supermarket sales techniques. Seasonal sales techniques like this sometimes have an impact that can be seen obviously by the imbalance in sales of existing products in supermarkets. Sales imbalance can be the initial cause for a product to lose interest and become a product that is eventually removed from store. With a classification model made to predict which products will be eliminated or churn, it can assist staff in distributing the sales of each product. The more products are churn due to lack of enthusiasts which can affect the overall sales of the supermarket. The purpose of this study is to assist staff in classifying potentially churn products. The classification model consists of 3 models with different algorithms and the results show that the application of the Random Forest algorithm is more effective for predicting data with 96% accuracy compared to 81% for the Logistic Regression algorithm and 46% for the Support Vector Machine algorithm.