

ABSTRAK

Strategi pemasaran saat ini memiliki peran yang sangat penting dalam persaingan bisnis antar perusahaan. Tidak hanya mengutamakan strategi pemasaran berorientasi produk, perusahaan juga harus mampu melakukan strategi pemasaran berorientasi pelanggan. Segmentasi pelanggan dapat dimanfaatkan untuk mengidentifikasi karakteristik pelanggan dan mengelompokkan pelanggan berdasarkan perilaku atau kebiasaan yang sama. Untuk mengetahui kategori dari segmentasi pelanggan dapat dilakukan dengan menggunakan klasifikasi. Dua dari algoritma klasifikasi adalah *Random Forest* dan *Artificial Neural Network*. Penelitian ini menggunakan data *Customer Segmentation Classification* dari Kaggle menghasilkan akurasi yang berbeda. Hasil akurasi menggunakan algoritma *Random Forest* tertinggi diperoleh dari data dengan *missing value* dihapus dan data dinormalisasi menggunakan *Cross Validation* = 7, *nodesize* = 6, *maxnode* = 10, dan *ntree* = 100 sebesar 35,76%. Sementara percobaan menggunakan algoritma *Artificial Neural Network* yang *missing value* dihapus dan data dinormalisasi dengan nilai *hidden layer*, *stepmax*, *threshold* masing-masing sebesar 6, 100.000, dan 1,0 mendapatkan hasil akurasi sebesar 34,12%. Hasil akurasi dari percobaan kedua algoritma tersebut diperoleh dengan menggunakan atribut *Ever_Marriage*, *Age*, *Graduated*, *Profession*, *Spending_Score*, dan *Family_Size* sebagai atribut biasa, dan *Segmentation* sebagai atribut tujuan atau label.

Kata kunci : Segmentasi Pelanggan, Klasifikasi, *Random Forest*, *Artificial Neural Network*, *Confusion Matrix*

ABSTRACT

Today's marketing strategy has a very important role in business competition between companies. Not only prioritizing product-oriented marketing strategies, companies must also be able to carry out customer-oriented marketing strategies. Customer segmentation can be used to identify customer characteristics and group customers based on similar behaviour or habits. To find out the category of customer segmentation can be done by using classification. Two of the classification algorithms are Random Forest and Artificial Neural Network. This research uses Customer Segmentation Classification data from Kaggle to produce different accuracy. The highest accuracy results using the Random Forest algorithm were obtained from data with missing values removed and data normalized using Cross Validation = 7, nodesize = 6, maxnode = 10, and ntree = 100 of 35,76%. While the experiment using the Artificial Neural Network algorithm where the missing value was deleted and the data normalized with the hidden layer, stepmax, threshold valued of 6, 100.000, and 1,0, obtained an accuracy of 34,12%. Accuracy results from the experiments of the two algorithms were obtained using the Ever_Married, Age, Graduated, Profession, Spending_Score, and Family_Size attributes as regular attributes, and Segmentation as destination attribute or label.

Keywords : Customer Segmentation, Classification, *Random Forest*, Artificial Neural Network, Confusion Matrix