

## INTISARI

Perkembangan teknologi terutama internet telah membawa kita ke era digital yang memberi dampak pada berbagai sektor. Salah satu dampak positif dari perkembangan teknologi adalah adanya situs belanja *online* atau *e-commerce* yang memudahhi interaksi jual beli antara penjual dan pembeli. Sebagai kontributor 61,07% PDP Indonesia, usaha mikro, kecil dan menengah (UMKM) harus memanfaatkan potensi dari *e-commerce*. Apalagi dengan meningkatnya jangkauan sinyal internet di Indonesia setiap tahun seperti peningkatan sebesar 8,06% pada 2018.

Salah satu cara peningkatan pemanfaatan *e-commerce* sebagai penjual adalah dengan mengetahui faktor apa yang mempengaruhi jumlah penjualan. Penelitian ini akan meneliti faktor tersebut pada dua jenis produk yaitu *daily need (facewash)* dan elektronik (*handphone*). Data diambil dari halaman produk dan toko di Tokopedia meliputi jumlah gambar, jumlah kata, jumlah stok, jenis *merchant*, terakhir *online*, *badge point*, durasi balasan, *rating* toko, asal toko, waktu pemrosesan, opsi pengiriman, persentase diskon, harga akhir, jumlah diskusi, dan *rating* produk.

Model dipilih dari 6 metode yaitu Regresi Linier, Regresi Ridge, Regresi Lasso, *Random Forest*, dan *Artificial Neural Network* yang mewakili metode linier dan non-linier. Model *Random Forest* terpilih untuk kedua jenis produk dan menghasilkan model prediksi penjualan *facewash* dengan  $R^2=0,959$ ; RMSE=17,914; MAE=10,779; MAPE=24,8% serta model prediksi penjualan *handphone* dengan  $R^2=,954$ ; RMSE=10,487; MAE=5,648; MAPE=18,2%. Setelah dilakukan analisis, faktor yang paling berpengaruh dalam jumlah penjualan adalah jumlah diskusi. Hal itu dilihat dari nilai *importances* sebesar 0,791 (produk *daily need*) dan 0,770 (produk elektronik). Faktor lainnya tidak signifikan, hanya memiliki nilai *importances* di bawah 0,1.

**Kata kunci:** *e-commerce*, *machine learning*, *random forest*, faktor signifikan, analisis regresi

## ***ABSTRACT***

*The development of technology, especially the internet, has brought us to the digital era which has had many impacts. One of the positive impacts of technological developments is the existence of online shopping or e-commerce sites that accommodate buying and selling interactions between sellers and buyers. As a contributor to 61.07% of Indonesia's PDP, micro, small and medium enterprises (MSMEs) must take advantage of the potential of e-commerce. Especially with the increase in internet signal coverage in Indonesia every year, such as an increase of 8.06% in 2018.*

*A way to increase the use of e-commerce as a seller is to find out what factors are significantly affecting the number of sales. In this study, those factors will be examined using regression analysis on machine learning with two types of products, namely daily need and electronic. Data taken from product and store pages on Tokopedia includes number of images, number of words, number of stocks, type of merchant, last online, badge points, reply duration, store rating, store origin, processing time, shipping options, discount percentage, final price, quantity discussion, and product ratings.*

*The model was chosen from 6 methods, namely Linear Regression, Ridge Regression, Lasso Regression, Random Forest and Artificial Neural Network. The Random Forest model was selected for both product types and resulted in a model with  $R^2=0,959$ ;  $RMSE=17,914$ ;  $MAE=10,779$ ;  $MAPE=24,8\%$  (daily need product) dan  $R^2=,954$ ;  $RMSE=10,487$ ;  $MAE=5,648$ ;  $MAPE=18,2\%$  (electronic product). After conducting the further analysis, the most influential factor in the number of sales is the number of discussions. This can be seen from the importance value of 0.770 (electronic products). Other factors are not significant, only have an important value below 0.1.*

**Keywords:** *e-commerce, machine learning, random forest, significant factor, regression analysis*