

ABSTRAK

Soraya Bedsheet memasarkan produknya di berbagai marketplace salah satunya Shopee. Seiring dengan banyaknya penjualan pada marketplace Shopee terdapat beberapa persoalan yang dialami pada toko Soraya Bedsheet. Persoalan yang pertama yaitu tidak dapat memperkirakan stok produk, kedua ketika produk sudah soldout Soraya Bedsheet tidak dapat mengupdate stok, dan ketiga nomor pembeli tidak bisa dilihat. Untuk menangani hal tersebut, maka dilakukan prediksi penjualan pada marketplace Shopee di toko Soraya Bedsheet. Prediksi yang dilakukan bertujuan untuk mendapatkan informasi dan pengetahuan dengan menggunakan Hybrid Data Mining dengan teknik klasterisasi menggunakan algoritma K-Means dan klasifikasi menggunakan algoritma K-Nearest Neighbor. Pada algoritma K-Means menghasilkan 2 cluster yaitu “Sedikit Peminat” berjumlah 64 data dan “Banyak Peminat” berjumlah 25 data. Selanjutnya data cluster Banyak Peminat dijadikan sebagai data training pada algoritma K-Nearest Neighbor dengan menghasilkan 2 klasifikasi yaitu Penjualan Rendah dan Penjualan Tinggi. Berdasarkan pengujian menggunakan algoritma K-Nearest Neighbor dengan menggunakan jumlah $k=3$, $k=5$, dan $k=7$ menghasilkan prediksi “Penjualan Tinggi”. Hal ini membuktikan produk “Soraya Bedsheet Katun Gold Motif Dallas Ask Grey Tua” jika dilakukan pengujian menggunakan $k=3$, $k=5$, dan $k=7$ maka produk tersebut dapat diprediksi termasuk dalam “Penjualan Tinggi”. Berdasarkan pengujian tingkat akurasi yang dilakukan dengan menggunakan 3 kali pengujian nilai k yaitu $k=3$ akurasi sebesar 88%, $k=5$ akurasi sebesar 72%, dan $k=7$ akurasi sebesar 64% dapat disimpulkan bahwa semakin tinggi nilai k maka semakin turun nilai akurasinya.

Kata Kunci: *Hybrid* Data Mining, K-Means, K-Nearest Neighbor, Prediksi Penjualan, Shopee

ABSTRACT

Soraya Bedsheet markets its products in various marketplaces, one of which is Shopee. Along with the large number of sales on the Shopee marketplace, there are several problems experienced at the Soraya Bedsheet shop. The first problem is not being able to estimate product stock, secondly, when the product is sold out, Soraya Bedsheet cannot update stock, and thirdly, the buyer number cannot be seen. To handle this, sales predictions were made on the Shopee marketplace at the Soraya Bedsheet shop. The predictions carried out aim to obtain information and knowledge using Hybrid Data Mining with clustering techniques using the K-Means algorithm and classification using the K-Nearest Neighbor algorithm. The K-Means algorithm produces 2 clusters, namely "A Few Interests" totaling 64 data and "A Lot of Interest" totaling 25 data. Next, the Many Interested Cluster data is used as training data for the K-Nearest Neighbor algorithm by producing 2 classifications, namely Low Sales and High Sales. Based on testing using the K-Nearest Neighbor algorithm using the numbers $k=3$, $k=5$, and $k=7$, it produces a prediction of "High Sales". This proves that the product "Soraya Bedsheet Cotton Gold Motif Dallas Ask Gray Tua" if tested using $k=3$, $k=5$, and $k=7$ then the product can be predicted to be included in "High Sales". Based on accuracy level testing carried out using 3 k value tests, namely $k=3$ accuracy of 88%, $k=5$ accuracy of 72%, and $k=7$ accuracy of 64%, it can be concluded that the higher the k value, the lower the accuracy value.

Keywords: Hybrid Data Mining, K-Means, K-Nearest Neighbor, Sales Prediction, Shopee