

PERTAMINA OIL SALES PREDICTION AT PT PRESOL INDO PRIMA USING K-NEAREST NEIGHBOR

ABSTRACT

PT. Presol Indo Prima is Pertamina's leading oil distribution company located in South Sumatra, which offers a variety of oil products. Based on sales data from the previous year, oil sales predictions are needed to identify the best-selling products. To achieve this, this research uses a classification technique using the k-Nearest Neighbor (k-NN) algorithm, which was chosen because of its effectiveness in classifying data based on similarity to previous data. This research carries out a data mining process on sales data from PT. Presol Indo Prima from 2021-2023. Algorithm model evaluation was carried out by measuring accuracy, precision, recall, and f1-score for the value $k = 3$. The test results showed that the value $k = 3$ provided good performance with Class 1 or less popular produces 96% precision, 98% recall, and 97% f1-score, class 2 or popular produces 50% precision, 60% recall, and 55% f1-score, class 3 or very popular produces 0% precision, 0% recall, and 0% f1-score. These results are the basis for calculating accuracy using the confusion matrix method which obtained an accuracy of 0.9456 or 94.56% with an error rate of 5.44%.

Keywords : *K-Nearest Neighbor (KNN), Classification, Prediction*

PREDIKSI PENJUALAN OLI PERTAMINA PADA PT PRESOL INDO PRIMA MENGGUNAKAN K-NEAREST NEIGHBOR

ABSTRAK

PT. Presol Indo Prima merupakan perusahaan distribusi minyak terkemuka Pertamina yang berlokasi di Sumatera Selatan, yang menawarkan beragam produk minyak. Berdasarkan data penjualan dari tahun sebelumnya, diperlukan prediksi penjualan minyak untuk mengidentifikasi produk yang paling laku. Untuk mencapai hal ini, penelitian ini menggunakan teknik klasifikasi menggunakan algoritma k-Nearest Neighbor (k-NN), yang dipilih karena efektivitasnya dalam mengklasifikasikan data berdasarkan kesamaan dengan data sebelumnya. Penelitian ini melakukan proses data mining pada data penjualan dari PT. Presol Indo Prima dari tahun 2021-2023. Evaluasi model algoritma dilakukan dengan mengukur akurasi, presisi, recall, dan f1-score untuk nilai $k = 3$. Hasil pengujian menunjukkan bahwa nilai $K = 3$ memberikan performa yang baik dengan Kelas 1 atau kurang laris menghasilkan precision 96%, recall 98%, dan f1-score 97%, kelas 2 atau laris menghasilkan precision 50%, recall 60%, dan f1-score 55%, kelas 3 atau sangat laris menghasilkan precision 0%, recall 0%, dan f1-score 0%. Hasil tersebut menjadi dasar perhitungan akurasi menggunakan metode confusion matriks yang diperoleh akurasi sebesar 0,9456 atau 94,56% dengan tingkat kesalahan sebesar 5,44% .

Kata Kunci : *K-Nearest Neighbor* (KNN), Klasifikasi, Prediksi