ABSTRAK

IMPLEMENTASI ALGORITMA *K-NEAREST NEIGHBOR (KNN)* UNTUK ANALISIS SENTIMEN PENGGUNA APLIKASI TOKOPEDIA

Oleh :

M. Rival Ridautal Lillah

1197050063

Penelitian ini bertujuan untuk menganalisis sentimen pengguna aplikasi Tokopedia di platform *Play Store* dan *App Store* menggunakan algoritma *K-Nearest Neighbor (KNN).* Metodologi penelitian mengadopsi pendekatan *CRISP-DM,* terdiri dari 5 tahap yaitu *Business Understanding, Data Understanding, Data Preparation, Modeling,* dan *Evaluation.* Tahap *Business Understanding* mengidentifikasi fokus pada peningkatan kualitas layanan dan pemahaman pelanggan terhadap Tokopedia. Data dari 1000 komentar *Play Store* dan 1000 komentar *App Store* dipilih sebagai sampel utama, dengan dua label sentimen yaitu positif dan negatif. Pada tahap *Data Preparation*, data diolah dengan *casefolding, cleansing, tokenization*, normalisasi, penghapusan *stopwords*, *stemming*, dan perhitungan nilai *TF-IDF*. Tahap *Modeling* mengeksplorasi hasil prediksi sentimen menggunakan *KNN*. Eksperimen dengan dataset *Play Store* menghasilkan variasi performa tergantung pada pembagian data. Skema 80:20 memiliki akurasi tertinggi 0,905. Evaluasi pada dataset *App Store* menunjukkan konsistensi performa model, dengan skema 90:10 dan 80:20 memiliki akurasi tertinggi masing-masing 0,9 dan 0,88. Perubahan nilai K dan proposisi data dari algoritma *KNN* ini mempengaruhi nilai akurasi dari tiap eksperimen. Hasil prediksi menunjukkan bahwa terdapat 89,2% prediksi label positif dan 10,8% prediksi label negatif untuk dataset *play store*. Adapun hasil prediksi pada dataset *app store* menunjukkan terdapat 97,0% prediksi label positif dan hanya 3,0% prediksi label negatif.

Kata Kunci: *K-Nearest Neighbor (KNN),* analisis sentimen, aplikasi Tokopedia, *CRISP-DM,* akurasi prediksi.

ABSTRACT

IMPLEMENTATION OF THE K-NEAREST NEIGHBOR ALGORITHM (KNN) FOR USER SENTIMENT ANALYSIS TOKOPEDIA APPLICATION

By :

M. Rival Ridautal Lillah

1197050063

This research aims to analyze user sentiment on the Tokopedia application on the Play Store and App Store platforms using the K-Nearest Neighbor (KNN) algorithm. The research methodology adopts the CRISP-DM approach, consisting of 5 stages: Business Understanding, Data Understanding, Data Preparation, Modeling, and Evaluation. The Business Understanding stage identifies a focus on improving service quality and customer understanding of Tokopedia. Data from 1000 Play Store comments and 1000 App Store comments were selected as the main samples, with two sentiment labels: positive and negative. In the Data Preparation stage, the data was processed through case folding, cleansing, tokenization, normalization, removal of stopwords, stemming, and TF-IDF value calculation. The Modeling stage explores the results of sentiment prediction using KNN. Experiments with the Play Store dataset resulted in performance variations depending on the data split. The 80:20 scheme had the highest accuracy of 0,905. Evaluation on the App Store dataset showed consistent model performance, with the 90:10 and 80:20 schemes achieving the highest accuracies of 0,9 and 0,88, respectively. Changes in the value of K and data splitting in the KNN algorithm affected the accuracy values in each experiment. The prediction results indicate that there are 89.2% predictions of positive labels and 10.8% predictions of negative labels for the Play Store dataset. As for the App Store dataset, the prediction results show 97.0% predictions of positive labels and only 3.0% predictions of negative labels.

Keywords: *K-Nearest Neighbor (KNN), sentiment analysis, Tokopedia application, CRISP-DM, prediction accuracy.*