

ABSTRAKSI

Gartenhutte merupakan destinasi wisata di Desa Selotapak, Trawas, yang menerima banyak ulasan pelanggan melalui media sosial. Banyaknya ulasan menyebabkan pengelola kesulitan melakukan analisis secara manual. Penelitian ini bertujuan membangun sistem analisis sentimen dan deteksi aspek ulasan (tempat, produk, dan pelayanan) menggunakan pendekatan multi-label berbasis Convolutional Neural Network–Long Short-Term Memory (CNN–LSTM). Data dikumpulkan dari Instagram, X, dan Google Review sebanyak 1.835 ulasan melalui proses crawling. Data diproses melalui tahapan preprocessing, tokenisasi, dan pembentukan representasi vektor menggunakan Word2Vec. Penelitian ini menguji 12 variasi konfigurasi model CNN–LSTM yang mencakup perbedaan jumlah unit, kombinasi kernel convolution, dropout, serta penerapan augmentasi data. Hasil pengujian menunjukkan bahwa pada klasifikasi sentimen, model dengan augmentasi data memberikan performa terbaik dengan nilai test accuracy sebesar 83,8%. Sebaliknya, pada klasifikasi aspek tempat, produk, dan pelayanan, model tanpa augmentasi menunjukkan performa yang lebih optimal dan konsisten dengan nilai test accuracy masing-masing sebesar 84,3%, 84,9%, dan 85,9%. Hasil penelitian ini menunjukkan bahwa penerapan augmentasi data efektif meningkatkan performa pada klasifikasi sentimen, namun tidak memberikan peningkatan yang konsisten pada deteksi aspek. Dengan konfigurasi arsitektur yang disesuaikan pada setiap label, model CNN–LSTM mampu menghasilkan performa klasifikasi yang optimal.

Kata Kunci: Analisis Sentimen, CNN–LSTM, Word2Vec, Multi-Label, Ulasan Pelanggan.

ABSTRAKSI

Gartenhutte is a tourist destination located in Selotapak Village, Trawas, which receives a large number of customer reviews through social media platforms. The high volume of reviews makes it difficult for management to conduct manual analysis. This study aims to develop a sentiment analysis and aspect detection system (place, product, and service) using a multi-label approach based on a Convolutional Neural Network–Long Short-Term Memory (CNN–LSTM) architecture. A total of 1,835 reviews were collected from Instagram, X, and Google Review through a crawling process. The data were processed through preprocessing, tokenization, and vector representation using Word2Vec. This research evaluated 12 variations of CNN–LSTM model configurations, including differences in the number of units, combinations of convolution kernel sizes, dropout rates, and the application of data augmentation. The experimental results show that for sentiment classification, the model with data augmentation achieved the best performance with a test accuracy of 83.8%. In contrast, for aspect classification (place, product, and service), models without augmentation demonstrated more optimal and consistent performance, achieving test accuracies of 84.3%, 84.9%, and 85.9%, respectively. These findings indicate that data augmentation is effective in improving sentiment classification performance but does not provide consistent improvements for aspect detection. With architecture configurations tailored to each label, the CNN–LSTM model is able to achieve optimal classification performance.

Keywords: Sentiment Analysis, CNN–LSTM, Word2Vec, Multi-Label, Customer Reviews.