

ABSTRAK

Penelitian ini mengkaji pentingnya analisis data review hotel dan penggunaan teknologi Natural Language Processing (NLP) dalam memprediksi sentimen review hotel. Dalam penelitian ini, model deep learning seperti Long Short-Term Memory (LSTM) dan Efficiently Learning an Encoder that Classifies Token Replacements Accurately (ELECTRA) digunakan untuk memprediksi sentimen review hotel dalam bahasa Indonesia. Data review hotel diperoleh melalui proses data scraping dengan webscraper.io dari situs Tripadvisor dan diperoleh sebanyak 977 data review hotel dari Grand Mercure Maha Cipta Medan Angkasa. Sebelum dilakukan proses prediksi sentimen, data review hotel harus melalui tahap text preprocessing untuk menghilangkan punctuations, uppercase, stopwords, dan dilakukan proses lemmatizer untuk memudahkan pemrosesan data selanjutnya. Selain itu, sentimen yang tadinya tidak seimbang perlu diseimbangkan dengan proses undersampling. Data yang sudah dibersihkan dan diseimbangkan kemudian dilabeli sebagai sentimen negatif (0), netral (1), dan positif (2). Hasil pengujian menunjukkan bahwa model ELECTRA menghasilkan performa yang lebih baik dibandingkan dengan LSTM dengan akurasi sebesar 47% oleh ELECTRA dan 30% oleh LSTM.

Kata Kunci : Analisis Sentimen; *Natural Language Processing (NLP); ELECTRA; LSTM; Data Scraping*