

ABSTRACT

SENTIMENT CLASSIFICATION OF DRUG ASSESSMENT REVIEWS USING THE BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMER (BERT) METHOD

By

YAZID ZINEDINE HDIANA

In this era, highly accessible health products such as medicines sold freely through pharmacies and e-commerce platforms have caused individuals who may have limited experience in the health sector to rely on user reviews to understand the types of drugs to treat certain conditions. These reviews are not only beneficial for individuals but also contribute to improving the quality of medicines produced by healthcare professionals. The problem of unstructured data that exists in reviews in the form of text data can be overcome with an innovative approach called BERT. BERT is a transformer-based model that uses pre-training to improve contextual understanding of text. This research evaluates BERT's performance in sentiment analysis using the BERT BASE Uncased pre-training model. The dataset used consists of user reviews of medicines available on the Kaggle platform. This research involves hyperparameter tuning to identify optimal classification models and uses two data labeling methods, namely rating-based labeling and textblob-based labeling. Experimental results show that the BERT BASE Uncased model achieves extraordinary accuracy, reaching a peak of 98% with certain hyperparameters in textblob-based data labeling. In addition, BERT BASE fine-tuning results for rating-based data classification achieved 97% accuracy.

Keywords : Sentiment Classification, Deep Learning, Transformers, BERT

ABSTRAK

KLASIFIKASI SENTIMEN ULASAN PENILAIAN OBAT DENGAN METODE *BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMER (BERT)*

Oleh

YAZID ZINEDINE HDIANA

Pada era ini, produk kesehatan sangat mudah diakses seperti obat-obatan yang dijual secara bebas melalui apotek dan platform *e-commerce* telah menyebabkan individu yang mungkin memiliki pengalaman dalam bidang kesehatan yang terbatas mengandalkan ulasan pengguna untuk memahami jenis obat dalam mengatasi kondisi tertentu. Ulasan ini tidak hanya bermanfaat bagi individu tetapi juga berkontribusi terhadap peningkatan kualitas obat yang diproduksi oleh para profesional kesehatan. Permasalahan data tidak terstruktur yang ada pada ulasan dalam bentuk data teks dapat diatasi dengan pendekatan inovatif yang disebut BERT. BERT adalah model berbasis *transformer* yang menggunakan *pre-training* untuk meningkatkan pemahaman teks kontekstual. Penelitian ini mengevaluasi kinerja BERT dalam analisis sentimen menggunakan model *pre-training* BERT BASE *Uncased*. *Dataset* yang digunakan terdiri dari ulasan pengguna terhadap obat-obatan yang tersedia di platform Kaggle. Penelitian ini melibatkan penyetelan *hyperparameter* untuk mengidentifikasi model klasifikasi optimal dan menggunakan dua metode pelabelan data, yaitu pelabelan berbasis *rating* dan pelabelan berbasis *textblob*. Hasil eksperimen menunjukkan bahwa model BERT BASE *Uncased* mencapai akurasi yang luar biasa, mencapai puncak 98% dengan *hyperparameter* tertentu dalam pelabelan data berbasis *textblob*. Selain itu, hasil *fine-tuning* BERT BASE untuk klasifikasi data berbasis *rating* mencapai akurasi 97%.

Kata Kunci : Klasifikasi Sentimen, *Deep Learning*, *Transformers*, BERT