

ABSTRACT

APPLICATION RATING CLASSIFICATION BASED ON TEXT REVIEWS USING BERT ANNOTATION AND HYBRIDIZATION OF RNN WITH BIDIRECTIONAL LSTM

Oleh

Faris Zaidan Nafis

21/482156/PA/21028

Ratings are a crucial factor in users' decisions to download an application. Users express their experiences and provide insights into app quality through reviews and ratings. However, there is often a mismatch between the actual star rating and what can be predicted from the collected review data due to inherent bias.

This study aims to improve rating classification performance by applying a hybrid approach that combines Recurrent Neural Network (RNN) and Bidirectional Long Short-Term Memory (BiLSTM). The BiLSTM architecture enables the model to capture contextual information from both preceding and succeeding words in a sentence, leading to a more comprehensive understanding of the text.

The model is evaluated using performance metrics such as accuracy, precision, recall, and F1-score. The dataset used consists of user reviews from the Google Play Store, which have been preprocessed and annotated to minimize bias. Evaluation results show that the hybrid RNN-BiLSTM model, with parameter tuning on `num_words`, achieved better performance compared to previous methods in classifying ratings based on text reviews, reaching 80.97% accuracy and 53.45% F1-score.

Kata Kunci: Rating classification, Sentiment annotation, Deep learning, BERT, RNN, BiLSTM.

INTISARI

KLASIFIKASI *RATING* APLIKASI BERDASARKAN ULASAN TEKS DENGAN ANOTASI BERT DAN HIBRIDISASI DARI RNN DENGAN BIDIRECTIONAL LSTM

Oleh

Faris Zaidan Nafis

21/482156/PA/21028

Rating menjadi aspek paling penting dalam keputusan pengguna dalam mengunduh aplikasi. Pengguna mengutarakan pengalaman dan informasi tentang kualitas aplikasi melalui ulasan dan *rating*. Namun, seringkali terdapat ketidakakuratan dalam prediksi *rating* bintang dengan dataset tidak bias dari ulasan yang terkumpul.

Penelitian ini bertujuan untuk meningkatkan performa dalam klasifikasi *rating* dengan menerapkan pendekatan hibridisasi antara Recurrent Neural Network (RNN) dan Bidirectional Long Short-Term Memory (BiLSTM). Pendekatan BiLSTM memungkinkan pemodelan informasi dari dua arah, sehingga mampu memahami konteks secara lebih menyeluruh baik dari kata-kata sebelumnya maupun sesudahnya dalam suatu kalimat.

Model dievaluasi menggunakan metrik evaluasi seperti *accuracy*, *precision*, *recall*, dan *F1-score* untuk mengukur performa klasifikasi. Dataset yang digunakan merupakan kumpulan ulasan aplikasi dari *Google Play Store* yang telah melalui proses anotasi dan pembersihan guna meminimalkan bias. Hasil evaluasi menunjukkan bahwa model hibridisasi RNN-BiLSTM dengan penyesuaian parameter *num_words* memberikan performa yang lebih baik dibandingkan metode sebelumnya dalam mengklasifikasikan *rating* berdasarkan ulasan teks yaitu sebesar 80.97% untuk akurasi dan 53.45% untuk *F1-score*.

Kata Kunci: Klasifikasi *rating*, Anotasi sentimen, *Deep learning*, BERT, RNN, BiLSTM.