

INTISARI

PERBANDINGAN PERFORMA ANALISIS SENTIMEN DENGAN CNN-LSTM DAN CNN-BILSTM

Oleh

Reyzita Afrida

21/473666/PA/20415

Pada era digital saat ini, perkembangan data yang semakin pesat memungkinkan pengolahan informasi dalam jumlah besar dan kompleks. Salah satu perkembangan ilmu pengolahan data adalah kemampuan dalam menganalisis sentimen yang terdapat dalam suatu teks. Untuk mendapatkan performa model analisis sentimen yang lebih baik, arsitektur model klasifikasi dilakukan dengan menggabungkan model yang dapat mengekstrak fitur pola lokal seperti *Convolutional Neural Network* (CNN) dan juga model yang dapat mengakomodasi data dengan dependensi berurutan seperti *Long Short-Term Memory* (LSTM) serta *Bidirectional Long Short-Term Memory* (BiLSTM). Penelitian ini dilakukan untuk melihat performa model CNN, LSTM, BiLSTM, dan model *hybrid* CNN-LSTM dan CNN-BiLSTM dalam mengklasifikasikan sentimen ulasan aplikasi layanan transportasi, yaitu Access by KAI milik PT Kereta Api Indonesia (Persero) pada Google Play Store dan Apple App Store. Performa klasifikasi diukur menggunakan akurasi dan skor F1. Permasalahan kelas tidak seimbang pada data ditangani dengan melakukan *random over-sampling*. Hasil akhir menunjukkan performa klasifikasi sentimen terbaik diberikan oleh model CNN-LSTM pada data tanpa penanganan kelas tidak seimbang dengan perolehan akurasi sebesar 91,655%, skor F1 untuk kelas positif sebesar 63,817%, dan skor F1 untuk kelas negatif sebesar 95,284%. Begitu juga pada model dengan penanganan kelas tidak seimbang, performa klasifikasi terbaik diberikan oleh model CNN-LSTM dengan perolehan akurasi sebesar 90,755%, skor F1 untuk kelas positif sebesar 61,696%, dan skor F1 untuk kelas negatif sebesar 94,756%. Dengan demikian, pada studi kasus penelitian ini, menggabungkan CNN dan LSTM mampu meningkatkan performa klasifikasi lebih baik dibandingkan model CNN-BiLSTM, dan model dasar CNN, LSTM, dan BiLSTM.

Kata Kunci: analisis sentimen, *Convolutional Neural Network* (CNN), *Long Short-Term Memory* (LSTM), *Bidirectional Long Short-Term Memory* (BiLSTM), CNN-LSTM, CNN-BiLSTM.

ABSTRACT

COMPARISON OF SENTIMENT ANALYSIS PERFORMANCE WITH CNN-LSTM AND CNN-BILSTM

By

Reyzita Afrida

21/473666/PA/20415

In today's digital era, the rapid growth of data enables the processing of large and complex volumes of information. One significant development in data processing is the ability to analyze sentiments contained in text. To achieve better performance in sentiment analysis, this study's classification architecture combines models that can extract local pattern features through Convolutional Neural Networks (CNN), as well as models that accommodate sequential dependencies through Long Short-Term Memory (LSTM) and Bidirectional Long Short-Term Memory (BiLSTM). This research aims to examine the performance of CNN, LSTM, BiLSTM, and the hybrid models CNN-LSTM and CNN-BiLSTM in classifying user sentiment toward the "Access by KAI" application, owned by PT Kereta Api Indonesia (Persero), on Google Play Store and Apple App Store. Classification performance is measured using accuracy and F1-score. The issue of class imbalance is addressed through random over-sampling. The results indicate that the best classification performance is achieved by the CNN-LSTM model on data without class imbalance handling, attaining 91.655% accuracy, a 63.817% F1-score for the positive class, and a 95.284% F1-score for the negative class. For models with class imbalance handling, the CNN-LSTM model also achieves the highest performance, with 90.755% accuracy, a 61.696% F1-score for the positive class, and a 94.756% F1-score for the negative class. Therefore, in this study, combining CNN and LSTM proves to be more effective at improving classification performance compared to CNN-BiLSTM or the standalone CNN, LSTM, and BiLSTM models.

Keyword: sentiment analysis, *Convolutional Neural Network* (CNN), *Long Short-Term Memory* (LSTM), *Bidirectional Long Short-Term Memory* (BiLSTM), CNN-LSTM, CNN-BiLSTM.