

## ABSTRACT

### **FINE-GRAINED SENTIMENT CLASSIFICATION ON COVID-19 VACCINATION IN INDONESIA USING BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS (BERT) AND CONVOLUTIONAL NEURAL NETWORK (CNN)**

Fedora Ramadhanty Widiyanto Putri

18/423106/PA/18189

In order to improve qualities, opinions are essential as feedback. Sentiment analysis is one of natural language processing field that studies about opinions, sentiments, and emotions that are available on texts. Sometimes, opinion can only not be good or bad, it might be somewhere in between.

Fine-grained sentiment classification can be used to classify opinion into many classes to understand the opinions better. There are not a lot of sentiment classification research using classes more than three, but some have implemented it in English with the combination of deep learning model, however, to find about this topic in Bahasa is not easy.

This research uses the combination model of BERT as the contextualized word embedding and CNN, which is compared to the conventional word embedding Word2Vec and CNN. The result of fine-grained sentiment classification on both model gives the best result from the combination of BERT-CNN model with macro-average precision score of 61.08, macro-averaged recall score of 59.49, macro-averaged f1-score of 59.57, and accuracy of 61.32.

*Keywords: Sentiment classification, fine-grained sentiment classification, BERT, Word2Vec, convolutional neural network, deep learning, tweets*

## ABSTRAK

### ***FINE-GRAINED SENTIMENT CLASSIFICATION PADA VAKSIN COVID-19 DI INDONESIA MENGGUNAKAN BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS (BERT) DAN CONVOLUTIONAL NEURAL NETWORK (CNN)***

Fedora Ramadhanty Widiyanto Putri

18/423106/PA/18189

Untuk meningkatkan kualitas, pendapat merupakan hal penting sebagai masukan. Analisis sentimen adalah salah satu bidang pemrosesan bahasa alami yang mempelajari tentang opini, sentimen, dan emosi yang tersedia pada teks. Terkadang, opini tidak bisa digolongkan menjadi hanya baik atau buruk, bisa jadi di antara keduanya.

Klasifikasi sentiment secara *fine-grained* dapat digunakan untuk mengklasifikasikan opini ke dalam banyak kelas dengan tujuan untuk memahami opini lebih baik. Tidak banyak penelitian klasifikasi sentimen yang menggunakan kelas lebih dari tiga, tetapi beberapa telah menerapkannya dalam bahasa Inggris dengan kombinasi model *deep learning*, namun untuk menemukan topik ini dalam Bahasa Indonesia tidaklah mudah.

Penelitian ini menggunakan model kombinasi BERT sebagai contextualized word embedding dan CNN, yang dibandingkan dengan *conventional word embedding* Word2Vec dan CNN. Hasil klasifikasi sentimen pada kedua model memberikan hasil terbaik dari kombinasi model BERT-CNN dengan skor *macro-average precision* sebesar 61.08, *macro-averaged recall* sebesar 59.49, *macro-averaged f1-score* sebesar 59.57, dan akurasi sebesar 61.32.

*Kata kunci: klasifikasi sentimen, klasifikasi sentimen secara fine-grained, BERT, Word2Vec, convolutional neural network, deep learning, tweets*