

## ABSTRAK

### KONTEKSTUAL DAN NON-KONTEKSTUAL *WORD EMBEDDING* DALAM MENDETEKSI TWEET *MULTICLASS* BERNUANSA DEPRESIF DI TWITTER MENGGUNAKAN LONG SHORT-TERM MEMORY

Oleh

ARIF FARHAN BUKHORI  
18/430254/PA/18767

Salah satu bentuk emosi yang diekspresikan melalui media sosial adalah perasaan yang bernuansa depresif. Salah satu situs jejaring sosial yang paling banyak digunakan adalah Twitter, dengan 300 juta lebih pengguna aktif, twitter telah menjadi platform besar untuk mengekspresikan pandangan seseorang, sehingga banyak peneliti yang kini mulai menggunakannya sebagai sumber yang baik untuk memperoleh data yang berkaitan dengan Kesehatan mental. Analisis sentimen merupakan cara untuk mengidentifikasi perilaku depresif di media sosial. Beberapa jenis *word embedding* dalam melakukan analisis sentimen yang terkenal adalah *Word2Vec*, *FastText*, dan juga BERT. Penelitian ini bertujuan untuk melakukan perbandingan kontekstual dan non kontekstual *word embedding* *Word2Vec*, *FastText*, dan BERT dalam mendeteksi tweet bernuansa depresif di media sosial twitter menggunakan model LSTM. Berdasarkan hasil pengujian dan perbandingan, kombinasi model *word embedding* kontekstual BERT dan Long Short-Term Memory terbukti dapat memberikan nilai yang lebih baik dalam melakukan klasifikasi pada tweet bernuansa depresif, dengan skor *F1-Score* mencapai diatas 95% terhadap dataset bersifat *multiclass*.

Kata-kata kunci: Klasifikasi Teks, *Tweet Depressive*, BERT, Kontekstual *Embedding*, *Long Short-Term Memory*.

## ABSTRACT

### ***CONTEXTUAL AND NON-CONTEXTUAL WORD EMBEDDING IN DETECTING DEPRESSIVE MULTICLASS TWEETS ON TWITTER USING LONG SHORT-TERM MEMORY***

By

ARIF FARHAN BUKHORI  
18/430254/PA/18767

Feelings of depression are one type of emotion expressed on social media. With over 300 million active users, one of the most popular social networking sites, Twitter, has become a huge platform for people to express their feelings. So many researchers are now using it as a good source for obtaining data related to mental health. Sentiment analysis is a way to identify depressive behavior on social media. Word2Vec, FastText, and BERT are three well-known types of word embedding in sentiment analysis. This study aims to compare contextual and non-contextual word embedding Word2Vec, FastText, and BERT in detecting depressive tweets on Twitter social media using the LSTM model. Based on tests and comparisons, the combination of the BERT contextual word embedding model and Long Short-Term Memory proved to provide a better value in classifying tweets with depression nuances, with an F1-Score reaching 95% for a multiclass dataset.

Keywords: text classification, depressive tweets, BERT, contextual embedding, Long Short-Term Memory.