

INTISARI

BIDIRECTIONAL LONG SHORT TERM MEMORY DAN EKSTRAKSI FITUR DOC2VEC UNTUK DETEKSI HATE SPEECH BAHASA INDONESIA

Oleh

Edo Cahaya Putra
20/466403/PPA/05969

Ujaran kebencian merupakan komunikasi yang meremehkan seseorang atau kelompok berdasarkan karakteristik seperti ras, etnis, jenis kelamin, kewarganegaraan, agama dan organisasi. Deteksi ujaran kebencian sangat berperan dalam mencegah konflik yang mungkin dapat terjadi. Dalam mendeteksi ujaran kebencian penggunaan terdapat banyak metode yang digunakan namun harus memilih metode yang terbaik agar mendapatkan hasil yang optimal.

Penelitian ini bertujuan untuk mendeteksi ujaran kebencian dengan menggunakan metode *Bidirectional Long Short Term Memory* dan *doc2vec embedding* serta membandingkan dengan *word2vec embedding*. Untuk pengujian metode *Bidirectional LSTM* dengan perhitungan nilai akurasi, presisi, *recall*, dan *F-measure*.

Pada pengujian perbandingan *doc2vec* dan *word2vec*, *doc2vec* lebih baik dibandingkan dengan *word2vec*. Pada model *doc2vec* dengan arsitektur DBOW mendapatkan nilai akurasi sebesar 0.7767 dan *F1-score* sebesar 0.7721. Untuk arsitektur DM mendapatkan nilai akurasi sebesar 0.7957 dan *F1-score* sebesar 0.7907. Sedangkan model *word2vec* dengan arsitektur *skip-gram* mendapatkan nilai akurasi sebesar 0.7732 dan *F1-score* sebesar 0.7685. Untuk arsitektur CBOW mendapatkan nilai akurasi sebesar 0.7751 dan *F1-score* sebesar 0.7686.

**Kata Kunci : Ujaran Kebencian, *Bidirectional LSTM*, *Doc2vec*, *Word2vec*,
Klasifikasi**

ABSTRACT

BIDIRECTIONAL LONG SHORT TERM MEMORY AND DOC2VEC FEATURE EXTRACTION FOR INDONESIAN HATE SPEECH DETECTION

By

Edo Cahaya Putra
20/466403/PPA/05969

Hate speech is communication that disparages a person or group based on characteristics such as race, ethnicity, gender, nationality, religion and organization. Hate speech detection plays an important role in preventing conflicts that might occur. In detecting hate speech, there are many methods used, but you have to choose the best method in order to get optimal results.

This study aims to detect hate speech using the *Bidirectional Long Short Term Memory* method and *doc2vec embedding* and compare it with *word2vec embedding*. For testing the *Bidirectional LSTM* method with the calculation of accuracy, precision, *recall*, and *F-measure* values.

In testing the comparison of *doc2vec* and *word2vec*, *doc2vec* is better than *word2vec*. The *doc2vec* model with DBOW architecture obtains an accuracy value of 0.7767 and an *F1-score* of 0.7721. For the DM architecture, an accuracy value of 0.7957 and an *F1-score* of 0.7907 are obtained. While the *word2vec* model with a *skip-gram* architecture gets an accuracy value of 0.7732 and an *F1-score* of 0.7685. For the CBOW architecture, the accuracy value is 0.7751 and the *F1-score* is 0.7686.

Keywords: Hate speech, *Bidirectional LSTM*, *Doc2vec*, *Word2vec*,
Classification