

INTISARI

PENDEKATAN EKSTRAKSI FITUR *SENT2VEC* DAN METODE *BIDIRECTIONAL LONG SHORT TERM MEMORY* UNTUK DETEKSI *HATE SPEECH* BAHASA INDONESIA

Oleh

Ari Wahyono

17/418625/PPA/05408

Fasttext telah digunakan oleh beberapa peneliti untuk mendeteksi ujaran kebencian. *Fasttext* memiliki keunggulan dalam merepresentasikan kata meski kata tersebut tidak muncul saat proses *training* karena mempelajari representasi kata dengan mempertimbangkan morfologi dan melakukan representasi kata dengan jumlah karakter n-gram. Meskipun *Fasttext* memiliki keunggulan dalam sintatik namun memiliki kelemahan dalam hal semantik yaitu kata yang memiliki struktural *character* yang sama akan dianggap memiliki kedekatan makna karena representasi kata pada *fasttext* yang dibangun dengan struktur morfologi.

Pendekatan ekstraksi fitur *sent2vec* yang dibangun berdasarkan *library fasttext* dilakukan untuk meningkatkan hasil dari penelitian sebelumnya. *Sent2vec* mampu merepresentasikan kata, frase atau kalimat. Model *sent2vec* menggunakan konteks yang jauh lebih besar dan karenanya dapat menggunakan lebih banyak informasi *n-gram* untuk mendapatkan representasi konteks semantik yang lebih baik. Untuk menangkap informasi semantik secara efektif, BiLSTM digunakan untuk metode klasifikasi yang mempertimbangkan informasi sebelumnya dan selanjutnya. BiLSTM merupakan pengembangan dari metode LSTM yang menangkap informasi hanya satu arah ke depan.

Penelitian ini mengusulkan kombinasi baru dari *sent2vec* dan BiLSTM untuk meningkatkan hasil penelitian sebelumnya. Dataset ujaran kebencian dari penelitian sebelumnya digunakan untuk mengetahui pengaruh ekstraksi fitur *sent2vec* dengan parameter kata n-gram yang berbeda. Hasil penelitian menunjukkan kinerja menghasilkan 85,94% dari F-Measure dan akurasi 84,61% dan mengungguli hasil sebelumnya.

Kata Kunci: *Bidirectional LSTM, Fasttext, Feature Extraction, Hate speech, Sent2vec*.

ABSTRACT

SENT2VEC FEATURE EXTRACTION APPROACH AND BIDIRECTIONAL LONG SHORT TERM MEMORY TO DETECT HATE SPEECH IN INDONESIAN

By

Ari Wahyono
17/418625 / PPA / 05408

Many researchers have used Fasttext to detect Indonesian hate speech. Fasttext has the advantage of representing the word even if the word does not appear during the training process because it learns word representation by considering morphology. However, fasttext has a disadvantage in terms of semantics. Words with the same structural character have a closeness in meaning because of a fasttext made by morphological structure.

The sent2vec feature extraction approach built based on the fasttext library is carried out to improve the results of previous studies. Sent2vec is capable of representing words, phrases or sentences. The sent2vec model uses a much larger context and can therefore use more n-gram information to represent the semantic context better. BiLSTM is used for a classification method that considers previous and subsequent information to capture semantic information effectively. BiLSTM is the development of the LSTM method, which captures information in only one way forward.

This paper proposed a new combination of sent2vec and BiLSTM to improve the previous research results. Dataset of hate speech from previous research is used to determine the effect of sent2vec feature extraction with the different parameters of the word n-grams. The results showed performance yields 85,94% of F-Measure and 84,61% accuracy and outperformed previous results.

Kata Kunci: *Bidirectional LSTM, Fasttext, Feature Extraction, Hate speech, Sent2vec .*