

ABSTRAK

Judul

PERBANDINGAN *WORD EMBEDDING WORD2VEC, GLOVE, DAN FASTTEXT* MENGGUNAKAN *DEEP LEARNING* PADA ULASAN KONDISI PENGGUNA OBAT KESEHATAN

Oleh

Fiqih Aulia Pradana

Produk kesehatan seperti obat merupakan salah satu aspek penting dalam kesehatan. Sebagai masyarakat awam perlu rekomendasi obat dari penyintas penyakit. Rekomendasi tersebut berupa ulasan yang dapat dijadikan bahan pengetahuan untuk masyarakat dalam memilih obat sesuai kondisi yang dialami. Dalam statistika khususnya dalam metode klasifikasi, data berbentuk ulasan diproses melalui NLP. *Text mining* merupakan salah satu bagian dari NLP. Dalam penerapannya *text mining* harus berdampingan dengan *machine learning* untuk mengolah algoritma untuk mendapat hasil yang baik. *Word embedding* merupakan teknik analisis kebahasaan terbaru, dimana metode ini memberikan informasi struktur, urutan, semantik dan konteks di sekitar kata. *Word Embedding* yang dipakai adalah *Word2Vec*, *Glove*, dan *FastText*. Metode yang digunakan untuk mengklasifikasikan yaitu *deep learning* LSTM. LSTM bisa digunakan untuk menyimpan memori jangka panjang. Perpaduan *word embedding* dan LSTM dapat melihat hasil kinerja masing-masing *word embedding* dan keunggulan dari *word embedding* yang bisa dimaksimalkan. Hasil dari perbandingan word embedding *Word2Vec*, *Glove*, dan *FastText* digabungkan dengan metode LSTM mampu mendapatkan hasil nilai akurasi berturut-turut 85.20%, 84.19%, 86.22%. Sedangkan nilai *F1-Score word embedding Word2Vec, Glove, dan FastText* berturut-turut 85%, 84%, 86%.

Kata Kunci : Ulasan, Kondisi, Klasifikasi, *Word2Vec, Glove, FastText, LSTM*

ABSTRACT

Judul

COMPARISON OF WORD EMBEDDING WORD2VEC, GLOVE, AND FASTTEXT USING DEEP LEARNING IN REVIEWS ON CONDITIONS OF HEALTH MEDICINE USERS

By

Fiqih Aulia Pradana

Health products such as medicines are one of the important aspects of health. As ordinary people, we need drug recommendations from disease survivors. These recommendations are in the form of reviews that can be used as material for knowledge for the public in choosing drugs according to the conditions experienced. In statistics, especially in classification methods, data in the form of reviews are processed through NLP. Text mining is a part of NLP. In its application, text mining must coexist with machine learning to process algorithms to get good results. Word embedding is the latest linguistic analysis technique, where this method provides structure, sequence, semantic and context information around words. The Word Embedding used are Word2Vec, Glove, and FasText. The method used to classify is deep learning LSTM. LSTM can be used to store long-term memory. The combination of word embedding and LSTM can see the performance results of each word embedding and the advantages of word embedding that can be maximized. The results of the word embedding comparison Word2Vec, Glove, and FasText combined with the LSTM method were able to obtain accuracy values of 85.20%, 84.19%, 86.22%, respectively. While the F1-Score value of word embedding Word2Vec, Glove, and FasText respectively 85%, 84%, 86%.

Key Words : Reviews, Conditions, Classification, Word2Vec, Glove, FastText, LSTM