

ABSTRACT

The development of the internet and social media can affect the source of data dissemination, including textual data. In recent years, the sentiment analysis research has been increased to obtain opinion patterns on the textual data. Sentiment analysis can be done using several approaches such as machine learning and deep learning. However, machine learning has limitation to process a large amount of data, therefore deep learning was created to solve this problem.

In this research, the writer used a deep learning approach namely the LSTM method. The LSTM method was chosen because it can process the large amount of data and long-term dependency. In addition, the LSTM method was also combined with a bidirectional layer so that data can be processed both forward and backward. While the attention mechanism was used to sorting out the important parts of words in response to a particular aspect. In some tests of applying these methods, the writer compared the affect of using dissimilar types of word embedding such as Word2Vec, FastText, and GloVe. The dataset in this research was taken from Indonesian novel reviews in Goodreads.com website.

The result of this research shows that the implementation of Word2Vec CBOW model on Attention Bidirectional LSTM produces the best accuracy of 82.67%. The development of deep learning architecture by including bidirectional and attention to LSTM layer can improve a better accuracy result and can produce varying accuracy values on every single use of word embedding variations.

Keywords : sentiment analysis, deep learning, word embedding, attention, bidirectional, LSTM.

INTISARI

Seiring laju perkembangan internet dan media sosial memberikan pengaruh terhadap sumber penyebaran data, termasuk data dalam bentuk tekstual. Beberapa tahun terakhir, penelitian analisis sentimen mengalami peningkatan untuk mendapatkan pola opini pada suatu data teks. Analisis sentimen dapat dilakukan dengan menggunakan beberapa pendekatan seperti *machine learning* maupun *deep learning*. Namun pendekatan *machine learning* memiliki keterbatasan dalam menangani jumlah data yang besar. Sehingga *deep learning* hadir untuk menangani masalah tersebut.

Pada penelitian ini, penulis menggunakan pendekatan *deep learning* yaitu metode LSTM. Metode LSTM dipilih karena dapat menangani jumlah data yang besar sekaligus data yang berukuran panjang (*long-term dependency*). Selain itu metode LSTM juga dikombinasikan dengan lapisan *bidirectional* agar data dapat diproses baik secara *forward* maupun *backward*. Sedangkan mekanisme *attention* digunakan untuk memilah bagian penting kata sebagai tanggapan terhadap aspek tertentu. Pada beberapa pengujian metode tersebut, penulis membandingkan pengaruh terhadap penggunaan jenis *word embedding* yang berbeda seperti Word2Vec, FastText, dan GloVe. Adapun dataset yang digunakan merupakan kumpulan ulasan novel berbahasa Indonesia yang diambil melalui situs Goodreads.com.

Hasil pengujian menunjukkan bahwa implementasi model Word2Vec CBOW pada *Attention Bidirectional LSTM* dapat menghasilkan akurasi terbaik sebesar 82.67%. Pengembangan arsitektur *deep learning* dengan menambahkan *bidirectional* dan *attention* pada lapisan LSTM juga mampu meningkatkan hasil akurasi yang lebih baik dan menghasilkan nilai akurasi yang bervariasi pada masing-masing penggunaan variasi *word embedding*.

Kata kunci – analisis sentimen, deep learning, word embedding, attention, bidirectional, LSTM.