

INTISARI

ANALISIS SENTIMEN BERDASARKAN ASPEK PADA ULASAN RESTORAN BERBAHASA INDONESIA MENGGUNAKAN KOMBINASI *CONVOLUTIONAL NEURAL NETWORK (CNN)* DAN *CONTEXTUALIZED WORD EMBEDDING*

Oleh

Putri Rizki Amalia

17/409444/PA/17751

Opini seseorang terhadap suatu produk atau layanan yang dituangkan melalui sebuah ulasan merupakan suatu hal yang cukup penting bagi pemilik ataupun calon pelanggan produk atau layanan tersebut. Namun, banyaknya jumlah ulasan menyulitkan para calon pelanggan untuk menganalisis informasi yang terdapat pada ulasan-ulasan tersebut. Analisis sentimen berdasarkan aspek merupakan suatu cara untuk mengetahui polaritas sentimen suatu kalimat opini berdasarkan aspek yang telah ditentukan. Penelitian terkait sudah banyak dilakukan menggunakan kombinasi *deep learning* seperti CNN, GRU, dan LSTM dengan *traditional word embedding* seperti *Word2vec*, *GloVe*, dan *FastText*. Namun, penelitian analisis sentimen berdasarkan aspek dengan kombinasi *deep learning* dengan *contextualized word embedding* masih sangat terbatas pada ulasan berbahasa Indonesia.

Model yang digunakan pada penelitian ini adalah kombinasi CNN dengan *contextualized word embedding*, BERT dan ELMo, lalu akan dibandingkan dengan model kombinasi CNN dengan *traditional word embedding*, *Word2vec*. Pengujian klasifikasi aspek terhadap tiga model; BERT-CNN, ELMo-CNN, dan *Word2vec*-CNN, memberikan hasil terbaik pada model ELMo-CNN dengan nilai *micro-average precision* sebesar 0.86, *micro-average recall* sebesar 0.90, dan *micro-average f1-score* sebesar 0.88. Sedangkan untuk klasifikasi sentimen memberikan hasil terbaik pada model BERT-CNN dengan rata-rata nilai *precision* sebesar 0.88, *recall* sebesar 0.89, dan *f1-score* sebesar 0.91. Klasifikasi yang dilakukan dengan data tanpa *stemming* pada ketiga model memiliki hasil yang hampir mirip bahkan lebih baik dibandingkan menggunakan data dengan *stemming*.

Kata kunci: analisis sentimen berdasarkan aspek, ulasan restoran, *convolutional neural network*, *contextualized word embedding*, ELMo, BERT, *Word2vec*.

ABSTRACT

ASPECT-BASED SENTIMENT ANALYSIS ON INDONESIAN RESTAURANT REVIEW USING A COMBINATION OF CONVOLUTIONAL NEURAL NETWORK (CNN) AND CONTEXTUALIZED WORD EMBEDDING

By

Putri Rizki Amalia

17/409444/PA/17751

Someone's opinion on a product or service that is poured through a review is something that is quite important for the owner or potential customer of the product or service. However, the large number of reviews makes it difficult for them to analyze the information contained in the reviews. Aspect-based sentiment analysis is the process of determining the sentiment polarity of a sentence based on predetermined aspects. Many related studies have been using a combination of deep learning such as CNN, GRU, and LSTM with traditional word embedding such as Word2vec, GloVe, and FastText. However, aspect-based sentiment analysis with a combination of deep learning with contextualized word embedding is still very limited to reviews of Indonesian-language.

The model used in this research is a combination of CNN with contextualized word embedding; BERT and ELMo, then it will be compared with the CNN combined with traditional word embedding; Word2vec. The result of aspect-classification on three models; BERT-CNN, ELMo-CNN, and Word2vec-CNN give the best results on the ELMo-CNN model with micro-average precision of 0.86, micro-average recall of 0.90, and micro-average f1-score of 0.88. Meanwhile, the sentiment-classification gives the best results on the BERT-CNN model with a precision value of 0.88, a recall of 0.89, and an f1-score of 0.91. Classification using data without stemming on the three models have almost similar results, even better than using data with stemming.

Keywords: aspect-based sentiment analysis, indonesian restaurant review, convolutional neural network, contextualized word embedding, ELMo, BERT, Word2vec