

Klasifikasi Sentimen *Twitter* Berbasis Target Menggunakan *Long Short Term Memory* dan *Attention Network*

Muhamad Enrinal Zulhimar (14116101)

Pembimbing :

Pembimbing 1 : Dr. Masayu Leylia Khodra, S.T., M.T.

Pembimbing 2 : Rahman Indra Kesuma, S.Kom., M.Cs.

ABSTRAK

Klasifikasi sentimen berbasis target bertujuan untuk mengidentifikasi polaritas sentimen suatu target tertentu dalam konteksnya. Penelitian sebelumnya sudah menyadari pentingnya target dalam klasifikasi sentimen. Namun, penelitian sebelumnya mengabaikan pemodelan target yang terpisah. Dalam tugas akhir ini diteliti apakah dengan menggunakan arsitektur *Interactive Attention Network* (IAN) yang memodelkan target dan konteks secara terpisah dan secara interaktif mempelajari *attention* dari target dan konteks dapat memprediksi polaritas sentimen berbasis target pada dataset *twitter* dan membuktikan pentingnya koordinasi target dan konteks untuk meningkatkan kinerja klasifikasi sentimen berbasis target. Setiap kalimat pada data *twitter* diberikan indeks kata dan direpresentasikan embedding pada setiap kata menggunakan GloVe. Setelah itu dibangkitkan *hidden states* dari tiap target dan konteks untuk selanjutnya digunakan dalam pembangkitan *attention* masing-masing. Penelitian ini mendapatkan model terbaik dengan nilai akurasi 0.708 dan macro-f1 0.686 yang didapatkan dari arsitektur IAN menggunakan model LSTM dengan *pooling function average*, *batch size* 32, menggunakan *optimizer* adam, *dropout rate* diatur dengan nilai 0.5, dimensi dari *word embedding*, *attention vector* dan LSTM *hidden states* diatur dengan nilai 300 dan menggunakan *fix embedding*. Pentingnya koordinasi target dan konteks juga terlihat dengan membandingkan nilai akurasi dan *macro-f1* dengan arsitektur-arsitektur yang tidak memperhatikan koordinasi target dan konteks. Dalam penelitian selanjutnya dapat dilakukan modifikasi koordinasi target dengan menggunakan konteks dari kiri target dan konteks dari kanan target untuk meningkatkan performa model, menambahkan korelasi antara target dan konteks, mencoba *pre-trained word embedding* lain, eksplorasi tahapan *preprocessing* lain dan mencoba untuk menerapkan arsitektur IAN pada data *twitter* berbahasa Indonesia.

Kata kunci : Klasifikasi Sentimen; Target; Konteks; Interactive Attention Network; Twitter;

Target-Based Twitter Sentiment Classification Using Long Short Term Memory And Attention Network

Muhamad Enrinal Zulhimar (14116101)

Advisor :

Advisor 1 : Dr. Masayu Leylia Khodra, S.T., M.T.

Advisor 2 : Rahman Indra Kesuma, S.Kom., M.Cs.

ABSTRACT

Target-based sentiment classification aims to identify the polarity of sentiment of a particular target in its context. Previous research has recognized the importance of targets in sentiment classification. However, previous research ignored separate target modeling. In this final project investigated whether using Interactive Attention Network (IAN) architecture which models targets and context separately and interactively studying attention from target and context can predict the polarity of target-based sentiment on twitter dataset and prove the importance of target and context coordination to improve target-based sentiment classification performance. Each sentence in the Twitter data will be given a word index and represented by its word embedding using GloVe. After that, the hidden states of each target and the context will be generated for further use in generating each attention. This study obtained the best model with an accuracy value of 0.708 and macro-f1 0.686 obtained from the IAN architecture using the LSTM model with the pooling function average, batch size 32, using the Adam optimizer, the dropout rate is set with a value of 0.5, the dimensions of word embedding, attention vector and LSTM hidden states are set to 300 and use fix embedding. The importance of target and context coordination is also seen by comparing the accuracy and macro-f1 values with architectures that do not pay attention to the coordination of targets and contexts. In further research, modification of target coordination can be done by using the context from the left of the target and the context from the right of the target to improve model performance, add correlation between target and context, try other pre-trained word embedding, explore other preprocessing stages and try to implement IAN architecture on Twitter data in Bahasa.

Keywords: Sentiment Classification; Target; Context; Interactive Attention Network; Twitter;