

INTISARI

ANALISIS PENGARUH *EMOTICON* DAN SARKASME PADA ANALISIS SENTIMEN BERBAHASA INDONESIA DI *TWITTER*

Oleh

Debby Alita
16/403670/PPA/05184

Indonesia merupakan pengguna aktif twitter yang menempati peringkat kelima terbesar di dunia. *Tweet* yang ditulis oleh pengguna twitter bermacam-macam, mulai dari tweet berupa tanggapan positif hingga negatif. Komentar tersebut jika ditambahkan dengan *emoticon* dianggap dapat memudahkan seseorang dalam mengungkapkan perasaannya tetapi tidak sedikit juga peneliti yang berpendapat lain dengan mengabaikan *emoticon* dikarenakan dapat mengganggu proses analisis sentiment, tidak hanya permasalahan dalam penggunaan *emoticon* saja akan tetapi dalam proses analisis sentimen juga sering sekali terjadi kesalahan dalam proses klasifikasi yaitu terdapat *tweet* yang bernilai positif namun memiliki makna negatif yang menyebabkan nilai positif tersebut berubah menjadi nilai negatif yang disebut juga dengan sarkasme.

Penelitian ini menganalisis data komentar dengan mempertimbangkan emoticon sebagai fitur tambahan selain fitur text dan mengkombinasikan analisis sentimen dengan pendeteksian sarkasme. Metode klasifikasi yang digunakan untuk proses analisis sentimen adalah metode *naïve bayes classifier* dan *support vector machine* sedangkan untuk pendeteksian sarkasme adalah metode *random forest classifier*, *naïve bayes classifier* dan *support vector machine* dengan menggunakan empat fitur untuk proses ekstraksi data yaitu *sentiment relate feature*, *punctuation relate feature*, *lexical* dan *syntactic feature*, dan *pattern relate feature*.

Pada penelitian ini, diperoleh hasil bahwa penggunaan *emoticon* tanpa pendeteksian sarkasme dapat menaikkan nilai akurasi pada proses analisis sentimen sebesar 25,30% dan penggunaan *emoticon* dengan pendeteksian sarkasme dapat menaikkan nilai akurasi pada proses analisis sentimen sebesar 14,80% .Metode terbaik dalam proses analisis sentimen dengan penggunaan *emoticon* adalah metode *support vector machine* dengan nilai sebesar 69,74% dan tanpa penggunaan *emoticon* adalah metode *naïve bayes classifier* dengan nilai sebesar 49,51%.

Kata kunci: analisis sentimen, *Emoticon*, pendeteksian sarkasme, *Naïve Bayes Classifier*, *Support vector machine*, *Random Forest Classifier*.

ABSTRACT

ANALYSIS OF EMOTICON AND SARCASM EFFECT ON SENTIMENT ANALYSIS OF INDONESIA LANGUAGE IN TWITTER

By

Debby Alita
16/403670/PPA/05184

Indonesia is an active user of twitter which is ranked fifth largest in the world. Tweets written by twitter users vary, ranging from tweet in the form of positive to negative responses. These comments when added with emoticons are considered to facilitate a person in expressing his feelings but not a few researchers who argue otherwise by ignoring emoticons because it can interfere with the process of analysis of sentiment, not only the problem in the use of emoticons alone but in the process of sentiment analysis is also often a mistake in the process of classification is there is a tweet that has a positive value but has a negative meaning that causes the positive value is turned into a negative value that is also called sarcasm.

This reasearch analyze comment data by considering emoticons as additional features in addition to text features and combine sentiment analysis with sarcasm detection. The classification method used for sentimental analysis is the naïve bayes classifier and support vector machine, while for the detection of sarcasm is the random forest classifier, naïve bayes classifier and support vector machine method using four features for data extraction process: sentiment relate feature, punctuation relate feature, lexical and syntactic features, and pattern relate features.

In this research, it was found out that the use of emoticon without sarcasm detection can increase the accuracy value in the sentiment analysis process by 25,30% and the use of emoticon with sarcasm detection can increase the accuracy value in the sentiment analysis process by 14,80%. The best method in the analysis process sentiment with emoticon use is the support vector machine method with a value of 69,74% and without the use of emoticons is the naïve bayes classifier method with a value of 49,51%.

Keyword: sentiment analysis, *Emoticon*, sarcasm detection, *Naïve Bayes Classifier*, *Support vector machine*, *Random Forest Classifier*.