

INTISARI

PENGARUH PENDETEKSIAN SARKASME TERHADAP UKURAN KUALITAS ANALISIS SENTIMEN PADA TWITTER

Oleh

Fariz Prawira
12/334718/PA/14950

Di masa pemilihan umum, banyak opini-opini tentang tokoh politik yang disuarakan oleh masyarakat secara langsung atau melalui media sosial seperti Twitter. Dengan analisis sentimen, opini-opini tersebut dapat diklasifikasikan ke dalam opini positif yang berarti dukungan atau ke dalam opini negatif yang berarti ejekan atau cemoohan. Akan tetapi, terkadang ditemukan opini dengan sentimen sarkasme yang membuat suatu opini salah diklasifikasikan.

Penelitian ini menggabungkan analisis sentimen dengan pendeteksian sarkasme untuk pengklasifikasian opini-opini yang terdapat pada Twitter. Untuk analisis sentimen, terdapat 3 fitur yang diimplementasikan, yaitu *unigram*, *select k-best*, dan TF-IDF, sedangkan untuk pendeteksian sarkasme, terdapat 4 set fitur yang diimplementasikan, yaitu *sentiment-relate*, *punctuation-relate*, *lexical and syntactic*, dan *pattern-relate*. *Classifier* yang digunakan untuk analisis sentimen adalah *Multinomial Naïve Bayes* dan untuk pendeteksian sarkasme adalah *Random Forest Classifier*.

Pada penelitian ini, diperoleh hasil bahwa pendeteksian sarkasme dengan 1500 data latih dapat meningkatkan akurasi dari analisis sentimen sebesar 1.20%. Selain itu, pendeteksian sarkasme juga dapat meningkatkan presisi sebesar 2.43%. Akan tetapi, peningkatan akurasi dan presisi dari analisis sentimen berdampak pada menurunnya *recall* sebesar 0.27%.

Kata kunci: analisis sentimen, pendeteksian sarkasme, Multinomial Naïve Bayes, Random Forest Classifier.

ABSTRACT

THE EFFECT OF SARCASM DETECTION ON SENTIMENT ANALYSIS QUALITY MEASUREMENT IN TWITTER

By

Fariz Prawira
12/334718/PA/14950

In the period of a general election, there are much opinion about political figures that expressed by people directly or by social media such as Twitter. By sentiment analysis, opinions can be classified into positive opinion which mean support or negative opinion which mean mockery. However, sometimes there is sarcasm in the opinion that make opinion incorrectly classified.

This research combines sentiment analysis and sarcasm detection for opinions classification in Twitter. For sentiment analysis, there are 3 features, which are unigram, select k-best and TF-IDF. For sarcasm detection, there are 4 sets of features, which are sentiment-relate, punctuation-relate, lexical and syntactic, and pattern-relate. Classifier used for the sentiment analysis is Multinomial Naïve Bayes and for sarcasm detection is Random Forest Classifier.

In this research, the result of sarcasm detection combined with sentiment analysis are increased of accuracy by 1.20%. Beside that, sarcasm detection can increase precision by 2.43%. However, the impact on increasing of the accuracy and precision is decreasing of the recall by 0.27%.

Keyword: sentiment analysis, sarcasm detection, Multinomial Naïve Bayes, Random Forest Classifier