

INTISARI

Analisis Sentimen Berdasarkan Aspek Terhadap Review Toko Online Berbahasa Indonesia Menggunakan *Machine Learning*

Oleh

Zulfa Dzakiyyatun Nisa

16/398532/PA/17493

Analisis sentimen pada level dokumen dan kalimat sering dilakukan namun kurang memberikan informasi yang mendetail. Oleh karena itu, dibutuhkan analisis sentimen berdasarkan aspek. Analisis sentimen berdasarkan aspek dibagi menjadi dua *subtask* yaitu klasifikasi aspek dan klasifikasi sentimen. Data yang digunakan adalah data *review* produk kategori kosmetik bahasa Indonesia pada shopee. Ekstraksi fitur TFIDF dan *n*-gram digunakan pada penelitian ini. *Support Vector Machine* dengan kernel rbf, linear dan poly serta *Multinomial Naive Bayes* diimplementasikan pada kedua *subtask*.

Fitur unigram-bigram dapat meningkatkan nilai performa klasifikasi pada kedua *subtask* dengan selisih akurasi pada klasifikasi aspek sebesar 1% untuk SVM dan 4% untuk MNB sedangkan pada klasifikasi sentimen sebesar 5% pada kedua metode. Hasil penelitian SVM dengan fitur unigram-bigram memperoleh akurasi, presisi, *recall* dan *f1-score* sebesar 77%, 96%, 78% dan 86% sedangkan MNB sebesar 75%, 93%, 77% dan 84%. Klasifikasi sentimen setiap aspek, SVM memiliki nilai rata - rata akurasi sebesar 88% sedangkan MNB sebesar 85%. SVM mendapatkan akurasi tertinggi pada 5 dari 8 aspek yaitu aspek harga, tekstur, ketahanan, pelayanan dan keaslian sedangkan MNB memiliki akurasi tertinggi pada aspek pengemasan dan warna. Metode SVM memperoleh nilai performa klasifikasi yang lebih baik pada kedua *subtask* dibandingkan MNB.

Kata kunci : *analisis sentimen berdasarkan aspek, Support Vector Machine, Multinomial Naive Bayes, n-gram, TFIDF*

ABSTRACT

Aspect Based Sentiment Analysis Of Indonesian Online Shop Review Using *Machine Learning*

By

Zulfa Dzakiyyatun Nisa

16/398532/PA/17493

Sentiment analysis at the document level is often carried out but does not provide detail information. Therefore, aspect based sentiment analysis is needed. Aspect based sentiment analysis is divided into two subtasks, aspect categorization and sentiment each aspect. The data used is the Indonesian cosmetic product review data on shopee. TFIDF and n-gram are used in this study. Support Vector Machine and Multinomial Naive Bayes are implemented in both subtasks.

Unigram-bigram feature can increase accuracy in both subtask with difference of 1% for SVM and 4% for MNB on aspect classification, while on sentiment classification the difference is 5% in both methods. Result for SVM with unigram-bigram feature has accuracy, precision, recall and f1-score of 77%, 96%, 78% and 86% while MNB has 75%, 93%, 77% and 84%. In sentiment classification each aspect, average accuracy of SVM is 88% while MNB is 85%. SVM has the highest accuracy in 5 of 8 aspects which are price, texture, durability, originality and service while MNB has better on packaging and color. SVM has better accuracy in both subtask than MNB.

Keywords : *aspect based sentiment analysis, Support Vector Machine, Multinomial Naive Bayes, n-gram, TFIDF*