

ABSTRACT

OPINION MINING USING NAIVE BAYES ALGORITHM ON ONLINE ACCOMMODATION CUSTOMER REVIEWS

BINTANG ARIGO KAUTSAR URUMSAH
15/376918/PA/16438

An online customer review contains an opinion made by a user posted on an online accommodation booking website. Within an accommodation booking website such as airbnb.com, there is an overload of information from the amount of reviews within various accommodation listings. An opinion mining system can be used to analyze and summarize the online customer reviews in order to provide the users with classifications of the reviews accordingly the polarity being either positive, negative or neutral.

Within the purpose of classifying text, a term weighting method such as TF-IDF can be utilized to further assist in the process of classifying the text. In this research, the goal is to implement an opinion mining system that utilizes TF-IDF weighting and Naive Bayes algorithm to produce a high performance for each evaluation parameter in the purpose of classifying the polarity of the online customer reviews acquired from airbnb.com. The dataset used in this research consists of 750 online customer reviews, in which consists of 250 positive reviews, 250 negative reviews and 250 neutral reviews.

The results obtained in this research gave an accuracy of 69%, precision of 70%, recall of 70% and f-measure of 70%. However, upon implementing stop word reduction within the stop word removal procedure, the results obtained gave an overall improvement in accuracy of 72%, precision of 75%, recall of 72% and f-measure of 73%.

Key Words: *Naive Bayes, Term Frequency-Inverse Document Frequency, TF-IDF, Opinion Mining, Online, Customer Reviews.*

ABSTRAK

PENAMBANGANG OPINI MENGGUNAKAN ALGORITMA NAIVE BAYES PADA ULASAN ONLINE PELANGGAN AKOMODASI

BINTANG ARIGO KAUTSAR URUMSAH
15/376918/PA/16438

Ulasan pelanggan online berisi pendapat yang dibuat oleh pengguna yang diposting di situs web pemesanan akomodasi online. Dalam situs web pemesanan akomodasi seperti *airbnb.com*, terdapat informasi yang berlebihan dari jumlah ulasan dalam berbagai daftar akomodasi. Sistem penambangan opini dapat digunakan untuk menganalisis dan meringkas ulasan pelanggan online untuk memberikan pengguna klasifikasi sesuai dengan polaritas yang positif, negatif atau netral.

Dalam tujuan mengklasifikasikan teks, metode pembobotan istilah seperti *TF-IDF* dapat digunakan untuk membantu lebih lanjut dalam proses mengklasifikasikan teks. Dalam penelitian ini, tujuannya adalah untuk menerapkan sistem penambangan opini yang menggunakan pembobotan *TF-IDF* dan algoritma *Naive Bayes* untuk menghasilkan performa tinggi untuk setiap parameter evaluasi dengan tujuan mengklasifikasikan polaritas ulasan pelanggan online yang diperoleh dari *airbnb.com*. Dataset yang digunakan dalam penelitian ini terdiri dari 750 ulasan pelanggan online, yang terdiri dari 250 ulasan positif, 250 ulasan negatif dan 250 ulasan netral. Hasil yang diperoleh dalam penelitian ini memberikan akurasi 69%, ketepatan 70%, recall 70% dan f-ukur 70%. Namun, setelah menerapkan pengurangan *stop word* dalam prosedur *stop word removal*, hasil yang diperoleh memberikan peningkatan keseluruhan dalam akurasi 72%, ketepatan 75%, penarikan kembali 72% dan pengukuran-f 73%.

Kata Kunci: *Naive Bayes, Term Frequency-Inverse Document Frequency, TF-IDF, Opinion Mining, Online, Ulasan Pelanggan.*