

ABSTRAK

Twitter merupakan sosial media yang sampai saat ini sangat digemari dan menjadi penyebar informasi yang sangat cepat. Informasi yang beredar juga sangat banyak mulai dari berita, opini, komentar, kritik dan semuanya ada yang bersifat positif, negatif, dan netral. Menurut data yang dilansir Statista berdasarkan hasil penelitian PeerReach, Indonesia termasuk pengguna *twitter* yang paling aktif ke 3 di dunia dibawah Amerika Serikat dan Jepang. *Twitter* menjadi salah satu media sosial yang banyak digunakan di Bandung dimana kota Bandung sempat menjadi kota peringkat 6 dunia jumlah pengguna *twitter*. Dari kumpulan data tersebut kita dapat melakukan analisis sentimen terhadap suatu keadaan untuk melihat respon masyarakat terhadap objek tersebut. Penelitian analisis sentimen ini dilakukan untuk mengklasifikasi sebuah *tweet* ke dalam kelas sentimen positif, netral, dan negatif. Dimulai dengan proses pengumpulan data, pemilihan fitur, dan proses klasifikasi itu sendiri. Metode yang digunakan dalam penelitian analisis sentimen ini adalah *lexicon based* dengan *opinion lexicon* milik Hu dan Liu yang telah diterjemahkan kedalam Bahasa Indonesia dan *naïve bayes classifier* berdasarkan James Hennessey. Penelitian ini dilakukan dengan beberapa skenario yang berhubungan dengan *text preprocessing* dan *feature selection*. Kata kunci yang digunakan pada penelitian ini adalah “@ridwankamil” dengan jumlah sebanyak 300 *tweet* yang telah dilabeli secara manual ke dalam tiga kelas yaitu positif, netral, dan negatif. Hasil dari penelitian ini adalah perolehan nilai akurasi tertinggi yang dihasilkan adalah sebesar 68,67% dengan penerapan skenario pengujian melalui tahapan *text preprocessing* dan *feature selection*.

Kata kunci : analisis sentimen, *twitter*, *lexicon based*, *naïve bayes*

ABSTRACT

Twitter is a social media that until now very popular and became disseminators of information very quickly. Information circulating is also very much starting from news, opinion, comment, criticism and everything there are positive, negative, and neutral. According to data reported by the research statista PeerReach, Indonesia including the most active Twitter users to third in the world following the United States and Japan. Twitter became one of social media are widely used in the city of Bandung, where it became world number six ranked city of twitter users. From the collection of these data we can perform sentiment analysis on a state to see the public response to the object. Research sentiment analysis is performed to classify a tweet into a class sentiment positive, neutral, and negative. Starting with collecting data, feature selection and classification process itself. The method used in this research sentiment analysis is based with the lexicon lexicon opinion belongs to Hu and Liu have been translated into Indonesian and naïve Bayes classifier based on James Hennessey. This study will test with several scenarios related to text preprocessing and feature selection. Keywords used in this study is "@ridwankamil" with a total of 300 tweets that have been labeled manually into three classes: positive, neutral, and negative. Results from this study is the acquisition value generated the highest accuracy is equal to 68.67% with the implementation of test scenarios through the stages of text preprocessing and feature selection.

Keywords: sentiment analysis, twitter, lexicon based, naïve Bayes