

## ABSTRAK

*Twitter* merupakan wadah bagi para penggunanya untuk mengutarakan perasaannya atau memberikan tanggapan terhadap suatu hal atau barang ke dalam sebuah status (*tweet*). Tanggapan tersebut dapat berupa tanggapan positif atau negatif yang juga dapat mempengaruhi banyak orang. Dengan adanya media sosial *twitter*, telah memicu peluang besar dalam melakukan *analysis sentiment/opinion mining* untuk mengetahui respon atau tanggapan orang lain terhadap suatu hal yang dapat dimanfaatkan dalam proses pengambilan keputusan. Penelitian analisis sentimen yang dilakukan adalah mengklasifikasi sebuah *tweet* ke dalam kelas sentimen positif dan kelas negatif dengan melalui tahap *text preprocessing*, *feature selection* dan dengan proses klasifikasi itu sendiri. Metode yang digunakan dalam analisis sentimen ini adalah metode *lexicon-based* dengan *opinion lexicon* milik Hu dan Liu dan metode *k-nearest neighbor*. Penelitian ini dilakukan dengan beberapa skenario yang berhubungan dengan penggunaan proses *text preprocessing* dan *feature selection*. *Corpus* yang digunakan merupakan kumpulan *tweet* yang berasal dari *Twitter* mengenai dua jenis hotel yaitu *Hyatt* dan *Marriot* dengan jumlah data sebanyak 192 *tweet* dan telah dilabeli secara manual ke dalam dua kelas yaitu positif dan negatif. Hasil pengujian dari setiap skenario adalah diperolehnya nilai akurasi keseluruhan dari metode *lexicon based* untuk kasus hotel *Hyatt* sebesar 81,25% dan hotel *Marriot* sebesar 77%. Sedangkan nilai akurasi keseluruhan dari metode *k-nearest neighbor* sebesar 57,5%.

**Kata Kunci:** *sentiment analysis, opinion mining, lexicon-based, k-nearest neighbor, tweet*

## **ABSTRACT**

*Twitter is a place for users to express their feelings or to respond to a case or an item into a status (tweet). The response can be either positive or negative feedback which can also affect many people. With there are social media twitter,its has trigger a great opportunity for sentiment analysis / opinion mining to determine the responses or reactions of other people to a case that can be used in the decision making process. The sentiment analysis research that conducted are classifying a tweets into a positive sentiment class and negative class via text preprocessing phase, feature selection and classification process itself. The method used in this sentiment analysis is lexicon-based with opinion lexicon owned by Hu and Liu and k-nearest neighbor method. This research are conducted with several scenarios related to the use of the text preprocessing and feature selection.The corpus that is used are a collection of tweets from Twitter about two types of hotel that is Hyatt and Marriot with the amount of data as much as 192 tweets and has been labeled manually into two classes, which is positive and negative. The test results of each scenario are obtaining an overall accuracy rate of lexicon-based method for the case of the Hyatt hotel at 81, 25% and the Marriot hotel at 77% . While the value of the overall accuracy of the k-nearest neighbor method at 57.5%.*

**Keywords:** *sentiment analysis, opinion mining, lexicon- based, k-nearest neighbor, tweet*