

ABSTRAK

Media sosial twitter adalah salah satu media yang seringkali digunakan masyarakat untuk berbagi berbagai kisah dan opini yang mereka miliki, salah satunya ialah membicarakan mengenai tokoh-tokoh publik tertentu. Di dalam negara demokrasi seperti di indonesia, tokoh yang paling sering dinilai dan dibicarakan biasanya merupakan tokoh yang dianggap layak dan dipertimbangkan untuk dipilih sebagai pemimpin negara. Oleh karena itu, penelitian ini mencoba menganalisis twitter berbahasa indonesia yang memiliki opini tentang tokoh-tokoh publik di Indonesia. Penelitian ini sendiri terdiri dari beberapa tahap. Tahap pertama merupakan tahap *pre-processing* yang terbagi menjadi bagian *case folding*, *tokenisasi*, *stemming* dan *filtering stopword*. Selanjutnya pada tahap kedua adalah dengan melakukan *feature extraction*. Tahap ini dikerjakan dengan memanfaatkan penggunaan kamus *sentiment word indonesia* yang disebut sebagai *lexical resource*. Pada tahap terakhir ialah melakukan klasifikasi. Klasifikasi dilakukan dengan pendekatan *lexical based model*. Hasil yang diperoleh dari analisis terhadap tokoh publik menggunakan pendekatan *lexical based model* memberikan nilai akurasi sebesar 72,0% untuk pengujian datanya.

Kata kunci : *lexical based approach*, *sentiment analysis*, *sentiment word indonesia*

ABSTRACT

Twitter social media is one of the media that is often used by the public to share various stories and opinions they have, one of them is to talk about certain public figures. In a democratic country such as Indonesia, the figures who are most often assessed and discussed are usually those who are considered worthy and considered to be elected as the leader of the country. Therefore, this research attempts to analyze Indonesian-language Twitter which has opinions about public figures in Indonesia. This research consists of several stages. The first stage is the pre-processing, pre-processing process consisting of case folding, tokenization, stemming and stopword filtering sections. Furthermore, in the second stage is to perform feature extraction. This stage using the Indonesian sentiment word dictionary which is known as lexical resource. The last stage is doing the classification using tweet classification. The classification method used in this research is lexical based model approach. The results obtained from the analysis of public figures using the lexical-based model approach gave an accuracy value of 72.0% for data testing.

Keyword : *lexical based approach, sentiment analysis, sentiment word indonesia*