

INTISARI

Penggunaan Twitter sekarang ini tidak hanya sebatas pada media sosial untuk pertemanan, akan tetapi Twitter juga digunakan sebagai alat promosi dan kampanye. Pengguna Twitter bebas untuk mengeluarkan pendapat dan opininya, termasuk opini tentang calon Presiden Indonesia 2014. Penelitian ini menampung pendapat dan opini masyarakat tersebut dengan membaginya menjadi lima *class attribute*, yaitu sangat positif, positif, netral, negative dan sangat negatif. Proses klasifikasi pada penelitian ini menggunakan metode klasifikasi *Naïve Bayes Classifier* (NBC) dengan *preprocessing* data menggunakan tokenisasi, *cleansing* dan *filtering*. Data yang digunakan adalah *tweet* dalam bahasa Indonesia tentang calon Presiden Indonesia 2014, dengan jumlah *dataset* sebanyak 900 *tweet* yang didistribusikan secara merata kedalam lima *class attribute*. Hasil akurasi tertinggi didapatkan saat menggunakan tokenisasi *n-gram*, *stopword list WEKA* dan *emoticons*, yaitu dengan nilai akurasi 71,9%, nilai presisi 71,6%, nilai *recall* 71,9% nilai *TP rate* 66,1% dan nilai *TN rate* 65%.

Kata kunci – analisis sentimen, calon presiden indonesia 2014, lima *class attribute*.

ABSTRACT

Nowadays, Twitter is not only used for social media to maintain friendship, but also Twitter is used to promote and campaign. Twitter users are free to express their opinions, including opinions about candidates of Indonesian President 2014. This research accommodates the public opinions by classifying it into five class attributes : very positive, positive, neutral, negative and very negative. The classification process using Naïve Bayes Classifier (NBC) with data preprocessing using tokenization, cleansing and filtering. The data used in this research are Indonesian tweets about candidates of Indonesian President 2014, with 900 tweets of dataset and distributed to five class attributes equally. As result, highest accuracy obtained when the experiment using combination of tokenization n-gram, stopword list WEKA and emoticons, which is the values consisting 71,88% accuracy, 71,6% precision, 71,9% recall, 6,1% TP rate and 65% TN rate.

Keywords - sentiment analysis, candidate of Indonesian president 2014, five class attribute.