

ABSTRAK

Kebun Binatang Surabaya(KBS) merupakan bagian penting dari kehidupan kota Surabaya. KBS yang telah berdiri dari tahun 1916 memiliki berbagai ulasan terhadapnya sehingga membuat para wisatawan seringkali menimbang kembali ketika ingin mengunjunginya. Ulasan tersebut berisi tanggapan positif dan negatif dari berbagai kalangan masyarakat di Tripadvisor dan *Google maps*. Penelitian ini dilakukan untuk analisis sentimen pada sentimen ulasan Wisata KBS dengan metode *Support Vector Machine* (SVM). Kelebihan metode SVM adalah dapat mengolah data secara efektif untuk dimensi tinggi dan SVM relatif menghabiskan sedikit memori. Namun metode SVM memiliki kelemahan terhadap data yang memiliki fitur yang besar. Adapun tujuan dari penelitian ini menerapkan seleksi fitur *Recursive Feature Elimination* (RFE) untuk mengeliminasi fitur terendah agar dapat meningkatkan performa SVM. Data yang digunakan adalah data tanggapan dari ulasan KBS yang berasal dari proses *scraping* data melalui Tripadvisor dan *Google maps* yang telah diformat menjadi data *excel*. Data yang digunakan berjumlah 760 data. Selanjutnya dilakukan proses *text preprocessing*, meliputi *case folding*, *cleaning*, *normalization*, *tokenizing*, *filtering*, dan *stemming*. Setelah itu dilakukan pembobotan kata menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF). Data setelah dibobotkan akan dilakukan penyeimbangan data dengan metode *Synthetic Minority Oversampling Technique* (SMOTE). Hasil dari pengujian model klasifikasi SVM terjadi peningkatan performa apabila data telah diseimbangkan dengan *accuracy* sebesar 95,78%, *precision* 96%, *recall* 97%, *f1-score* 97% dan waktu pemrosesan 0,17 detik.

Kata kunci : *Analisis sentimen, Kebun Binatang Surabaya, TF-IDF, SVM, SMOTE, RFE*

ABSTRACT

Surabaya Zoo is an important part of Surabaya city life. Surabaya Zoo, which has been established since 1916, has various reviews about it that make tourists often reconsider when they want to visit it. The reviews contain positive and negative responses from various groups of people on Tripadvisor and Google maps. This research was conducted for sentiment analysis on the sentiment of Surabaya Zoo Tourism reviews with the Support Vector Machine (SVM) method. The advantages of the SVM method are that it can process data effectively for high dimensions and SVM consumes relatively little memory. However, the SVM method has a weakness against data that has large features. The purpose of this research is to apply Recursive Feature Elimination (RFE) feature selection to eliminate the lowest features in order to improve SVM performance. The data used is response data from Surabaya Zoo reviews derived from the process of scraping data through Tripadvisor and Google maps that have been formatted into excel data. The data used amounted to 760 data. Next, the text preprocessing process is carried out, including case folding, cleaning, normalization, tokenizing, filtering, and stemming. After that, word weighting is done using Term Frequency-Inverse Document Frequency (TF-IDF). Data after being weighted will be balanced using the Synthetic Minority Oversampling Technique (SMOTE) method. The results of testing the SVM classification model show an increase in performance when the data has been balanced with an accuracy of 95.78%, precision 96%, recall 97%, f1-score 97% and processing time 0.17 seconds.

Keywords: Sentiment analysis, Surabaya Zoo, TF-IDF, SVM, SMOTE, RFE

