

KLASIFIKASI PENDAPAT MENGENAI E-TILANG DI SOSIAL MEDIA MENGUNAKAN METODE NAIVE BAYES DAN SELEKSI FITUR QUERY EXPANSION RANKING

Migita Maulina Jannah
(190441100107)

Pembimbing 1 : Dr. Bain Khusnul Khotimah, ST., M.Kom.

Pembimbing 2 : Dr. Budi Dwi Satoto, ST., M.Kom

ABSTRAK

Penerapan e-tilang efektif dalam menindak pelanggaran peraturan lalu lintas, meskipun belum semua masyarakat Indonesia menguasai teknologi. Peraturan e-tilang menimbulkan opini positif dan negatif di masyarakat. Analisis sentimen digunakan untuk mengetahui pernyataan positif dan negatif di *Twitter*. Analisis sentimen dilakukan dengan menggunakan klasifikasi Multinomial *Naive bayes*. Metode klasifikasi yang sederhana, tetapi sensitif terhadap pemilihan fitur sehingga jumlah data ditentukan oleh hasil pemilihan seleksi fitur *Query Expansion Ranking*. Data di ambil dengan melakukan *crawling* data. Pada penelitian ini metode yang digunakan adalah *Naive bayes* dengan seleksi fitur *Query Expansion Ranking* untuk mengurangi jumlah fitur pada proses klasifikasi. Pengaruh seleksi fitur *Query Expansion Ranking* untuk metode *naive bayes* bertujuan untuk mengurangi dimensi fitur sehingga metode multinomial *naive bayes* bekerja secara efektif dan menghasilkan komputasi yang maksimal. Pada penelitian ini mencapai tingkat akurasi tertinggi sebesar 91.91% dengan menggunakan seluruh fitur yang tersedia atau 100% dalam proses seleksi fitur. Kecepatan atau waktu yang dibutuhkan oleh metode Naive Bayes dalam proses komputasi adalah $6,25 \times 10^7$ nanodetik[ns].

Kata Kunci: Analisis Sentimen , *Twitter*, *Naive bayes*, Seleksi Fitur *Query expansion ranking*

ABSTRACT

The implementation of e-ticketing is effective in enforcing traffic regulations, although not all Indonesian society has mastered the technology. E-ticketing regulations generate both positive and negative opinions in the community. Sentiment analysis is used to determine positive and negative statements on Twitter. Sentiment analysis is performed using the Multinomial Naive Bayes classification. The classification method is simple but sensitive to feature selection, so the amount of data is determined by the feature selection Query Expansion Ranking. Data is collected by crawling data. In this study, the method used is Naive Bayes with Query Expansion Ranking feature selection to reduce the number of features in the classification process. The influence of Query Expansion Ranking feature selection for the Naive Bayes method aims to reduce feature dimensions, allowing the multinomial Naive Bayes method to work effectively and produce optimal computations. This study achieved the highest accuracy rate of 91.91% by using all available features or 100% in the feature selection process. The computational speed or time required by the Naive Bayes method in the computation process is 62500000.000000 nanoseconds.

Keywords: Sentiment Analysis, *Twitter*, *Naive bayes*, Feature Selection *Query expansion ranking*