

## ABSTRACT

The issue of government policy in passing the omnibus law was once trending on Twitter. The policy evoked pros and cons in the community. Public opinions on Twitter can be used as a decision support system in making appropriate policies to evaluate government performance. The sentiment classification method is fast in analyzing public opinion. Machine Learning-based sentiment classification is the method that has the best performance. However, the method has several problems, such as noisy features and high dimensionality of features, which significantly affect the performance of the sentiment classification. Thus, it requires feature selection. Therefore, to overcome the problems, this study presents a novel feature selection method using an integration of Query Expansion Ranking (QER) and metaheuristic algorithms (GA and PSO). QER is a filter-based feature selection method that has low-computational time, while metaheuristic algorithm is a wrapper-based feature selection method that has better performance than filter-based methods but has high-computational time. Thus, the combination of them aims to overcome the wrapper method's drawbacks by taking advantage of the filter method. In this study, Optimized-SVM is employed as classifiers. Based on the experimental results, the integration of QER and metaheuristic algorithm, both QG and QP can significantly improve sentiment classification performance and decrease computational time. The integration of QER and GA (QG) achieved the best performance in average precision, recall, and f-measure with the value of 0.968, respectively. The proposed method also outperformed all feature selections on the four additional datasets. It proves that the proposed method has robust performances.

**Keywords:** Feature Selection, Query Expansion Ranking, Metaheuristic Algorithm, Sentiment Classification, Machine Learning

## INTISARI

Keputusan pemerintah untuk mengesahkan *omnibus law* banyak dibahas di Twitter saat ini. Kebijakan tersebut menimbulkan pro dan kontra di tengah masyarakat. Opini publik yang ada di Twitter dapat digunakan sebagai sebuah sistem pendukung keputusan untuk membuat kebijakan yang tepat dalam mengevaluasi kinerja pemerintah. Metode analisis atau klasifikasi sentimen merupakan metode yang sangat cepat dalam melihat opini publik. Klasifikasi sentimen berbasis *machine learning* merupakan metode yang memiliki performa terbaik tetapi memiliki masalah fitur *noise* dan dimensi fitur yang tinggi. Hal tersebut sangat mempengaruhi kinerja klasifikasi sentimen sehingga dibutuhkan seleksi fitur. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan metode seleksi fitur dengan mengintegrasikan *Query Expansion Ranking* (QER) dan algoritme metaheuristik (GA dan PSO). QER merupakan metode seleksi fitur berbasis *filter* yang memiliki biaya komputasi yang rendah sedangkan algoritme metaheuristik merupakan metode berbasis *wrapper* yang memiliki performa lebih baik dari metode berbasis *filter* tetapi memiliki waktu komputasi yang tinggi. Sehingga, gabungan keduanya bertujuan untuk mengatasi kekurangan dari metode *wrapper* dengan mengambil keuntungan dari metode *filter*. *Optimized-SVM* digunakan sebagai *classifier*. Berdasarkan hasil eksperimen, integrasi QER dan algoritme metaheuristik, baik QG maupun QP, mampu meningkatkan performa klasifikasi sentimen secara signifikan dan dapat menurunkan waktu komputasi. Integrasi QER dan GA (QG) memiliki performa terbaik dan mampu mengalahkan *state-of-the-art* seleksi fitur yang diuji dengan mendapatkan nilai *precision*, *recall*, dan *f-measure* masing-masing sebesar 0,968. Metode seleksi fitur yang dikembangkan juga mampu mengungguli semua seleksi fitur pada keempat *dataset* tambahan. Hal ini membuktikan bahwa metode yang diusulkan memiliki performa yang sangat baik.

**Kata kunci** – *Seleksi Fitur, Query Expansion Ranking, Algoritme Metaheuristik, Klasifikasi Sentimen, Machine Learning*