

INTISARI

ANALISIS SENTIMEN MENGGUNAKAN SELEKSI FITUR INFORMATION GAIN DIINTEGRASIKAN DENGAN BINARY PARTICLE SWARM OPTIMIZATION WITH ADAPTIVE INERTIA WEIGHT

Oleh

Prima Dermawan

21/484214/PPA/06182

ChatGPT saat ini telah menjadi alat bagi masyarakat untuk mendapatkan informasi dengan cepat. Keunggulan ChatGPT dari NLP lainnya adalah interaksinya yang sangat natural sehingga pengguna merasa seperti melakukan percakapan dengan manusia. Keunggulan ini menimbulkan sentimen yang positif dan juga negatif terhadap ChatGPT. Menganalisis sentimen ChatGPT diperlukan untuk dapat memberikan umpan balik terhadap model NLP sehingga pengembangan NLP selanjutnya dapat menghasilkan produk yang lebih baik.

Metode analisis sentimen dengan menggunakan data dari Twitter adalah metode yang tepat untuk dapat mengetahui bagaimana masyarakat menilai ChatGPT. Namun, untuk dapat melakukan *training* sentimen diperlukan *tweet* yang banyak agar akurasi yang didapatkan sangat baik. *Tweet* yang banyak akan menghasilkan fitur yang besar. Fitur yang besar itu jika dimasukkan ke dalam metode *Machine Learning*, tentu akan menurunkan performa *Machine Learning* tersebut. Maka diperlukan metode pengurangan fitur agar fitur-fitur yang tidak relevan dan bahkan membuat bias dapat dibuang. Pada kasus ini terdapat 17292 fitur yang harus dikurangi, metode seleksi fitur adalah metode yang tepat untuk kasus ini.

Penelitian ini mengusulkan penggabungan metode *information gain* dan *binary particle swarm optimization with adaptive inertia weight* sebagai seleksi fitur. *Information gain* yang merupakan *filter based selection* akan memilih *subset* fitur dengan nilai *gain* yang tinggi. Fitur yang telah berkurang kemudian akan diseleksi lagi dengan *wrapper based selection* yaitu *AIW-BPSO* yang akan mencari subset fitur yang optimum. Berdasarkan hasil eksperimen, metode seleksi fitur yang diusulkan mendapatkan hasil yang memuaskan dengan nilai akurasi 0,938, lebih baik dari beberapa seleksi fitur yang diujikan. Hal ini membuktikan bahwa metode yang diusulkan memiliki performa yang sangat baik.

Kata Kunci: Teks, *Information Gain*, *Binary Particle Swarm Optimization*, Analisis Sentimen, Klasifikasi, *SVM*, *Adaptive Inertia Weight*

ABSTRACT

Integrating Information Gain and Binary Particle Swarm Optimization with Adaptive Inertia Weight as Feature Selection Method on Sentiment Analysis

by

Prima Dermawan

21/484214/PPA/06182

ChatGPT has now become a tool for people to obtain information quickly. The advantage of ChatGPT over other NLP (Natural Language Processing) systems is its highly natural interaction, making users feel like they're having a conversation with a human. This advantage gives rise to both positive and negative sentiments towards ChatGPT. Analyzing the sentiment of ChatGPT is necessary to provide feedback to the NLP model, enabling further development of NLP that can produce better products.

Using Twitter data for sentiment analysis is an appropriate method to determine how the public evaluates ChatGPT. However, in order to train sentiment, a large number of tweets are needed to achieve high accuracy. A large number of tweets will result in a large feature set. If this large feature set is used in a Machine Learning method, it will decrease the performance of the Machine Learning model. Therefore, a feature reduction method is necessary to discard irrelevant features and even eliminate biases. In this case, there are 17292 features that need to be reduced, and feature selection is the appropriate method for this case.

This research proposes the combination of the information gain method and binary particle swarm optimization with adaptive inertia weight as a feature selection technique. Information gain, which is a filter-based selection, selects a subset of features with high gain values. The reduced features are then further selected using a wrapper-based selection, AIW-BPSO, which searches for the optimum subset of features. Based on the experimental results, the proposed feature selection method achieves satisfactory results with an accuracy value of 0.938, outperforming several other feature selection methods that were tested. This proves that the proposed method performs exceptionally well.

Keywords: Text, Information Gain, Binary Particle Swarm Optimization, Sentiment Analysis, Classification, SVM, Adaptive Inertia Weight