

**ANALISIS SENTIMEN TERHADAP RESESI 2023 PADA MEDIA SOSIAL
TWITTER MENGGUNAKAN METODE NAÏVE BAYES DENGAN
SELEKSI FITUR *INFORMATION GAIN***

Hafidz Ashabi Muhammad

ABSTRAK

Dalam menghadapi dinamika ekonomi global, resesi menjadi kritikal. Media sosial, khususnya Twitter, menjadi saluran utama untuk berekspresi dan berbagi pandangan terkait kondisi ekonomi. Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap resesi tahun 2023 berdasarkan data yang diambil dari percakapan di media sosial Twitter. Metode analisis sentimen yang digunakan adalah Naïve Bayes, yang dikenal karena kesejukannya dalam penanganan data teks. Selain itu, seleksi fitur *Information Gain* diaplikasikan untuk memilih kata-kata atau fitur yang paling informatif dalam memprediksi sentimen. *Information Gain* memungkinkan kita untuk mengevaluasi kontribusi relatif dari setiap kata terhadap penentuan sentimen. Data yang digunakan dalam penelitian ini terdiri dari sejumlah besar tweet yang diambil pada periode yang relevan dengan konteks resesi 2023. Proses pengolahan data melibatkan tahap-tahap seperti tokenisasi, penghilangan stop words, dan normalisasi teks. Penelitian ini juga akan melakukan perbandingan penggunaan *Information Gain* dan tidak menggunakan *Information Gain* sebagai seleksi fitur. Pengumpulan data akan dilakukan *crawling* menggunakan bahasa pemrograman R dan terintegrasi pada API yang telah disediakan oleh *twitter*. 2. Dari hasil penelitian menunjukkan adanya peningkatan model Naïve Bayes apabila menggunakan seleksi fitur *Information Gain* dengan nilai pengambilan *top ranking* '>0.01' yaitu akurasi 0.96, *recall* 1, *precision* 0.93, *f1 score* 0.96 dan *specificity* 0.93 dibandingkan sebelumnya tanpa seleksi fitur *information gain* yaitu akurasi 0.92, *recall* 1, *precision* 0.85, *f1 score* 0.91 dan *specificity* 1.

Kata kunci: Analisis sentimen, *Twitter*, *Naïve Bayes*, *Information Gain*

ABSTRACT

In facing global economic dynamics, recession becomes critical. Social media, especially Twitter, has become the main channel for expression and sharing views regarding economic conditions. This research aims to conduct sentiment analysis regarding the 2023 recession based on data taken from conversations on Twitter social media. The sentiment analysis method used is Naïve Bayes, which is known for its coolness in handling text data. In addition, Information Gain feature selection is applied to select the most informative words or features in predicting sentiment. Information Gain allows us to evaluate the relative contribution of each word to sentiment determination. The data used in this research consists of a large number of tweets taken in a period relevant to the context of the 2023 recession. The data processing process involves stages such as tokenization, removal of stop words, and text normalization. This research will also compare the use of Information Gain and not using Information Gain as feature selection. Data collection will be carried out by crawling using the R programming language and integrated with the API provided by Twitter. 2. The research results show that there is an improvement in the Naïve Bayes model when using the Information Gain feature selection with a top ranking value of '>0.01', namely accuracy 0.96, recall 1, precision 0.93, f1 score 0.96 and specificity 0.93 compared to before without information gain feature selection, namely accuracy 0.92, recall 1, precision 0.85, f1 score 0.91 and specificity 1.

Keywords: *Sentiment Analysis, Twitter, Naïve Bayes, Information Gain*