

INTISARI

Keberhasilan suatu produk di pasar sangat dipengaruhi oleh sentimen konsumen. Sentimen konsumen terhadap suatu produk mudah ditemukan secara masif di internet dalam bentuk ulasan *online*. Penggunaan ulasan *online* konsumen sebagai objek analisis sentimen memiliki permasalahan pada jumlah data yang dibutuhkan dan orisinalitas data. *Vlog* merupakan salah satu bentuk alternatif dari ulasan *online* yang bisa digunakan sebagai objek analisis sentimen karena dapat mewakili jumlah data dari konsumen secara masif dan tingkat orisinalitas data yang terjaga. Pada penelitian sebelumnya, analisis sentimen dengan metode VADER telah dilakukan dengan berbagai macam ulasan *online* konsumen dalam bentuk teks sebagai objeknya. Akan tetapi, analisis sentimen dengan metode VADER dengan objek isi konten *vlog* yang merupakan data berbentuk video belum pernah dilakukan. Tujuan dari penelitian ini adalah untuk merancang kerangka kerja yang dapat melakukan analisis sentimen dengan metode VADER pada data berbentuk *vlog* serta menguji kemampuan VADER pada data *vlog* dengan bahasa terjemahan. Hasil analisis sentimen *vlog* akan divisualisasikan dalam bentuk kumpulan kata perwakilan sentimen positif dan negatif. Kata – kata tersebut dapat digunakan oleh pengembang produk sebagai bentuk sentimen konsumen yang diwakili oleh *vlogger* terhadap produk yang dikembangkan.

Fokus penelitian ini adalah proses analisis sentimen pada data *vlog* dengan menggunakan metode VADER dan bahasa pemrograman Python. Untuk mengetahui kemampuan VADER dalam menganalisis sentimen pada suatu topik, VADER digunakan pada data ulasan *online* terlebih dahulu dan dilanjutkan dengan data *vlog native* Inggris. Lebih lanjut, VADER juga digunakan pada data *vlog* terjemahan Inggris guna mengetahui kemampuan VADER dalam analisis sentimen pada data dengan bahasa terjemahan. Proses analisis sentimen ini memiliki empat tahapan utama, yaitu: 1) Pengambilan Data, 2) Preprocessing Data, 3) Tahap Analisis Sentimen, dan 4) Tahap Visualisasi Hasil.

Parameter evaluasi hasil yang digunakan pada penelitian ini adalah *F1-Score* dikarenakan data yang digunakan tidak tersebar secara merata. *F1-Score* bertujuan untuk mengetahui kemampuan analisis sentimen pada data dengan mempertimbangkan nilai presisi dan *recall*. Analisis sentimen yang memiliki nilai *F1-Score* melebihi 80% dianggap berkemampuan baik dalam memahami sentimen data. Hasil penelitian ini menunjukkan bahwa kerangka kerja analisis sentimen dengan metode VADER dapat digunakan pada data *vlog native* Inggris secara baik ditandai dengan nilai *F1-Score* yang diperoleh mencapai 86.4%. Di sisi lain, kerangka kerja analisis sentimen dengan metode VADER memiliki kemampuan kurang baik pada data *vlog* yang menggunakan bahasa terjemahan Inggris, dimana nilai *F1-Score* yang diperoleh berada di bawah 80%, yaitu 72.7%. Untuk menyampaikan hasil analisis secara lebih informatif, kumpulan kata perwakilan sentimen berhasil diidentifikasi dan divisualisasikan.

Kata Kunci: *Online Review, Vlog, VADER, Python, Sentiment Analysis, Word Identification, Word Visualization*

ABSTRACT

The success of a product on the market is strongly influenced by consumer sentiment. Consumer sentiment towards a product is easily found massively on the internet in the form of online reviews. The use of consumer online reviews as sentiment analysis objects has problems with the amount of data needed and the originality of the data. Vlog is an alternative form of online review that can be used as an object of sentiment analysis because it can represent a massive amount of data from consumers and the level of data originality is maintained. In previous studies, sentiment analysis using the VADER method has been carried out with a variety of online consumer reviews in the form of text as the object. However, sentiment analysis using the VADER method with vlog content as an object has never been done. The purpose of this study is to design a sentiment analysis framework with the VADER method on vlog data and to test its performance on data vlog with translation language. The results of the vlog sentiment analysis will be visualized in the form of a collection of positive and negative sentiment words. These words can be used by product developers as a consumer sentiment represented by vloggers on their products.

The focus of this research is the process of sentiment analysis on vlog data using the VADER method and the Python programming language. To find out VADER's ability to analyze sentiments on a topic, VADER is used in online review data first and followed by native English vlog data. Furthermore, VADER is also used in English translation vlog data to determine the ability of VADER in sentiment analysis on data with the language of translation. This sentiment analysis process has four main stages, namely: 1) Data Collection, 2) Data Preprocessing, 3) Sentiment Analysis Phase, and 4) Visualization Stage.

The evaluation parameter used in this study is F1-Score because the data used are not evenly distributed. F1-Score aims to determine the ability of sentiment analysis on data by considering the value of precision and recall. Sentiment analysis which has an F1-Score value exceeding 80% is considered capable of understanding data sentiment. The results of this study indicate that the sentiment analysis framework using the VADER method can be used on native English vlog data well characterized by the F1-Score obtained reaching 86.4%. On the other hand, the sentiment analysis framework using the VADER method has a poor ability in vlog data using English translation languages, where the F1-Score value obtained is below 80%, which is 72.7%. To convey the results of the analysis in a more informative manner, a collection of sentiments from sentiment representatives was identified as a collection of words and visualized.

Kata Kunci: *Online Review, Vlog, VADER, Python, Sentiment Analysis, Word Identification, Word Visualization*