

ABSTRAK

Banyaknya ulasan pengunjung pada *Google Maps* Lokawisata Hutan Pinus Limpakuwus yang sulit dianalisis secara manual karena data berupa teks tidak terstruktur dan jumlahnya terus bertambah. Penelitian ini bertujuan untuk menganalisis sentimen ulasan pengunjung serta mengetahui efektivitas algoritma *Naïve Bayes* dalam mengklasifikasikan sentimen positif dan negatif. Metode yang digunakan adalah analisis sentimen dengan tahapan pengumpulan data menggunakan *Apify* dan *Google Maps Reviews Scraper*. Dari 7.000 ulasan yang diperoleh, proses pembersihan data menghasilkan 3.558 data valid dan 2.378 data siap dianalisis. Tahap *preprocessing* meliputi *cleaning*, *case folding*, normalisasi kata, *tokenizing*, dan *stemming*. Selanjutnya dilakukan ekstraksi fitur menggunakan *TF-IDF*, kemudian pelabelan data menggunakan metode *lexicon-based* dengan kamus sentimen positif dan negatif untuk menentukan label sentimen berdasarkan nilai *polarity score*. Setelah itu dataset dibagi menjadi 80% data latih dan 20% data uji, kemudian dilakukan proses klasifikasi menggunakan algoritma *Naïve Bayes*. Hasil penelitian menunjukkan bahwa sebagian besar ulasan pengunjung memiliki sentimen positif terhadap keindahan alam dan kenyamanan Lokawisata Hutan Pinus Limpakuwus. Evaluasi model menghasilkan tingkat akurasi sebesar 84%, sehingga menunjukkan bahwa algoritma *Naïve Bayes* cukup efektif dalam melakukan klasifikasi sentimen ulasan wisatawan pada *Google Maps*. Penelitian selanjutnya diharapkan dapat menggunakan metode klasifikasi lain agar hasil analisis menjadi lebih akurat.

Kata kunci: Analisis Sentimen, *Naïve Bayes*, *Google Maps*, Ulasan Pengunjung, Hutan Pinus Limpakuwus

ABSTRACT

The large volume of visitor reviews for the Limpakuwus Pine Forest tourist site on Google Maps makes manual analysis difficult, as the data consists of unstructured text that continues to grow. This study aims to analyze visitor sentiment and evaluate the effectiveness of the Naïve Bayes algorithm in classifying positive and negative sentiments. The methodology employed is sentiment analysis, with data collection carried out using Apify and the Google Maps Reviews Scraper. Out of 7,000 collected reviews, data cleaning yielded 3,558 valid entries, with 2,378 entries ultimately ready for analysis. The preprocessing stage involved cleaning, case folding, word normalization, tokenization, and stemming. Subsequently, feature extraction was performed using TF-IDF, followed by data labeling via a lexicon-based method—utilizing positive and negative sentiment dictionaries—to assign sentiment labels based on polarity scores. The dataset was then split into 80% training data and 20% testing data, and the classification process was executed using the Naïve Bayes algorithm. The results indicate that the majority of visitor reviews express positive sentiment regarding the natural beauty and comfort of the Limpakuwus Pine Forest. Model evaluation yielded an accuracy rate of 84%, demonstrating that the Naïve Bayes algorithm is effective for classifying tourist sentiment in Google Maps reviews. Future research is encouraged to explore other classification methods to further enhance analysis accuracy.

Keywords: Sentiment Analysis, Naïve Bayes, Google Maps, Visitor Reviews, Limpakuwus Pine Forest