

ABSTRAK

Perkembangan teknologi informasi dan media digital telah mempercepat penyebaran berita hoaks di berbagai *platform digital* di Indonesia, dengan jumlah konten hoaks yang terus meningkat setiap tahunnya. Kondisi ini menuntut adanya sistem deteksi hoaks otomatis yang mampu membantu masyarakat dalam memilah informasi secara lebih efektif. Penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja algoritma *Naive Bayes*, *Logistic Regression*, dan *Support Vector Machine* (SVM) dalam mendeteksi berita hoaks berbasis teks berbahasa Indonesia menggunakan dua metode representasi fitur, yaitu TF-IDF dan BERT *embedding*. Dataset yang digunakan terdiri dari 23.632 data berita seimbang, meliputi 11.816 berita hoaks yang bersumber dari TurnBackHoax.id dan 11.816 berita non-hoaks yang diperoleh dari portal berita Kompas, Tempo, dan Detik. Tahapan penelitian mencakup *preprocessing* teks berupa *text cleaning*, *case folding*, *tokenisasi*, dan *stopword removal* dilanjutkan dengan *ekstraksi fitur* menggunakan TF-IDF dan IndoBERT, serta klasifikasi dalam enam skenario eksperimen yang merupakan kombinasi dari ketiga algoritma dan kedua representasi fitur. Evaluasi model dilakukan menggunakan *confusion matrix* dengan *metrix accuracy*, *precision*, *recall*, dan *F1-score*. Hasil penelitian menunjukkan bahwa kombinasi *Logistic Regression* dengan BERT (S5) menghasilkan performa terbaik dengan *accuracy* 98,24%, *precision* 98,47%, *recall* 98,01%, dan *F1-score* 98,24%, diikuti oleh SVM dengan BERT (S6) sebesar 98,18%. *Naive Bayes* secara konsisten menghasilkan performa paling rendah pada kedua representasi fitur, dengan *F1-score* 93,71% pada TF-IDF dan 91,78% pada BERT. Temuan ini mengindikasikan bahwa *Logistic Regression* dengan representasi BERT merupakan kombinasi paling optimal dan direkomendasikan untuk pengembangan sistem deteksi berita hoaks berbahasa Indonesia.

Kata kunci: Berita Hoaks, Metode Klasifikasi, Representasi Fitur

ABSTRACT

Advances in information technology and digital media have accelerated the spread of fake news across various digital platforms in Indonesia, with the volume of fake news content continuing to rise each year. This situation calls for an automated fake news detection system capable of helping the public filter information more effectively. This study aims to analyze and compare the performance of the Naive Bayes, Logistic Regression, and Support Vector Machine (SVM) algorithms in detecting Indonesian-language text-based fake news using two feature representation methods: TF-IDF and BERT embedding. The dataset consists of 23,632 balanced news items, including 11,816 hoax news items sourced from TurnBackHoax.id and 11,816 non-hoax news items obtained from the Kompas, Tempo, and Detik news portals. The research process included text preprocessing—such as text cleaning, case folding, tokenization, and stopword removal—followed by feature extraction using TF-IDF and IndoBERT, as well as classification across six experimental scenarios combining the three algorithms and the two feature representation methods. Model evaluation was conducted using a confusion matrix with the metrics accuracy, precision, recall, and F1-score. The results of the study show that the combination of Logistic Regression and BERT (S5) yielded the best performance, with an accuracy of 98.24%, precision of 98.47%, recall of 98.01%, and an F1-score of 98.24%, followed by SVM with BERT (S6) at 98.18%. Naive Bayes consistently produced the lowest performance on both feature representations, with an F1-score of 93.71% on TF-IDF and 91.78% on BERT. These findings indicate that Logistic Regression with BERT representation is the most optimal combination and is recommended for the development of an Indonesian-language fake news detection system.

Keywords: Fake News, Clasification Methods, Feature Representation