

PENGEMBANGAN SISTEM IDENTIFIKASI BERITA HOAX BAHASA INDONESIA MENGGUNAKAN MULTINOMIAL NAIVE BAYES

DEVELOPMENT OF AN INDONESIAN HOAX NEWS IDENTIFICATION SYSTEM USING MULTINOMIAL NAIVE BAYES

Muhamad Rizal Rifaldi^{1,*}, Supriyanto¹, Mohammad Hilman¹

Informatika, Universitas Sultan Ageng Tirtayasa, Jl. Jendral Sudirman Km 3, Kotabumi,
Kec. Purwakarta, Cilegon, 42121
*muhamadrizalrifaldi@gmail.com

disubmit: DD Month YYYY, direvisi: DD Month YYYY, diterima: DD Month YYYY

ABSTRAK

Penelitian ini mengembangkan sistem deteksi berita hoax berbahasa Indonesia menggunakan algoritma Multinomial Naive Bayes dengan pembobotan TF-IDF. Dataset yang digunakan terdiri dari berita daring dan media sosial yang telah melalui tahap pra-pemrosesan teks. Model dilatih secara terpisah untuk masing-masing domain dan dievaluasi menggunakan metrik akurasi. Hasil pengujian menunjukkan akurasi sebesar 98,1% pada data berita daring dan 81,4% pada data media sosial. Selain itu, penggunaan teks penuh menghasilkan tingkat keyakinan prediksi yang lebih tinggi dibandingkan headline. Sistem diimplementasikan dalam bentuk aplikasi web berbasis Streamlit sebagai alat bantu verifikasi awal terhadap informasi hoax.

Kata Kunci: *Multinomial Naive Bayes, TF-IDF, Deteksi Hoax, Machine Learning*

ABSTRACT

This study develops an Indonesian hoax news detection system using the Multinomial Naive Bayes algorithm with TF-IDF weighting. The dataset consists of online news articles and social media texts that undergo text preprocessing. Separate models are trained for each domain and evaluated using accuracy metrics. Experimental results show an accuracy of 98.1% for online news data and 81.4% for social media data. In addition, full-text inputs produce higher prediction confidence than headline-based inputs. The system is implemented as a Streamlit-based web application to support early verification of potential hoax information.

Keywords: *Multinomial Naive Bayes, TF-IDF, Hoax Detection, Machine Learning*