

INTISARI

Klasifikasi Berita Hoaks dengan Metode BERT dan IndoBERT

Oleh

TARA DWIPA ARDHATU TIMUR

20/459374/PA/20035

Penyebaran berita hoaks, khususnya yang bertopik politik menjadi permasalahan yang semakin marak di Indonesia, terutama selama periode pemilihan umum. Penelitian ini bertujuan untuk mengklasifikasikan berita hoaks dengan topik politik menggunakan teknik *Natural Language Processing* (NLP) dengan metode *Bidirectional Encoder Representations from Transformers* (BERT) dan Indonesia *Bidirectional Encoder Representations from Transformers* (IndoBERT). Selain itu, *Long Short-Term Memory* (LSTM) dan *Bidirectional Long Short-Term Memory* (BiLSTM) digunakan sebagai pembanding untuk mengevaluasi performa klasifikasi. Data yang digunakan dalam penelitian ini terdiri dari berita valid dan berita hoaks. Berita valid diperoleh dari portal berita terpercaya seperti CNN Indonesia, Tempo, dan Antara News, sementara pengecekan berita hoaks dilakukan melalui *platform* seperti Turnbackhoax.id, Kompas Cek Fakta, dan Jaringan Antara Cegah Hoaks. Hasil penelitian menunjukkan bahwa metode berbasis *transformer*, khususnya IndoBERT, memiliki performa yang lebih baik dibandingkan metode LSTM dan BiLSTM dalam mendeteksi berita hoaks.

Kata Kunci: berita hoaks, *natural language processing*, *Bidirectional Encoder Representations from Transformers* (BERT), IndoBERT

ABSTRACT

Hoax News Classification Using BERT and IndoBERT

By

TARA DWIPA ARDHATU TIMUR

20/459374/PA/20035

The spread of hoax news, particularly in the political domain, has become a growing issue in Indonesia, especially during critical events like elections. This research focuses on classifying political hoax news using Natural Language Processing (NLP) techniques, specifically Bidirectional Encoder Representations from Transformers (BERT) and Indonesian Bidirectional Encoder Representations from Transformers (IndoBERT). To provide a comprehensive analysis, Long Short-Term Memory (LSTM) and Bidirectional Long Short-Term Memory (BiLSTM) methods are also employed as benchmarks for performance comparison. The dataset consists of valid and hoax news articles. Valid news articles were collected from trusted Indonesian news portals such as CNN Indonesia, Tempo, and Antara News, while hoax news verification was sourced from platforms such as Turnbackhoax.id, Kompas Cek Fakta, and Jaringan Antara Cegah Hoaks. Results show that Transformer-based methods, especially IndoBERT, outperform LSTM and BiLSTM methods in hoax news detection.

Keywords: hoax news, natural language processing, Bidirectional Encoder Representations from Transformers (BERT), IndoBERT