

INTISARI

PERBANDINGAN METODE LONG SHORT-TERM MEMORY (LSTM) DAN BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS (BERT) DALAM PENGKLASIFIKASIAN BERITA PALSU

Oleh

SURYO MUQSITHO

20/456589/PA/19776

Penyebaran berita palsu atau hoaks yang merupakan berita tanpa dasar faktual semakin marak terjadi terutama menjelang dilaksanakannya pemilihan umum. Berita palsu digunakan sebagai alat politik yang dapat memanipulasi opini publik, merusak reputasi, dan menjatuhkan lawan politik. Untuk hal itu, penelitian ini dilakukan untuk mempelajari penggunaan metode *Long Short-Term Memory* (LSTM) dan *Bidirectional Encoder Representations from Transformers* (BERT), serta menentukan metode yang lebih optimal dalam melakukan analisis klasifikasi teks dengan menggunakan studi kasus pengklasifikasian berita palsu. Pengklasifikasian tersebut menggunakan kumpulan data yang diperoleh melalui CNN, Kompas, dan Tempo, serta portal berita yang berisi hasil pengecekan fakta dari berita palsu, turnbackhoax.id. Performa dari kedua metode diukur menggunakan metrik evaluasi akurasi, *recall*, *F1-score*, dan *Matthews Correlation Coefficient* (MCC). Perbandingan performa juga dilakukan pada kasus kelas data tidak seimbang dan kelas data yang telah diseimbangkan dengan menggunakan metode *random over sampling*. Hasil penelitian ini menunjukkan bahwa secara keseluruhan, metode BERT memberikan performa yang lebih optimal dalam melakukan analisis klasifikasi teks berbahasa Indonesia pada dua kasus data yang diberikan.

Kata kunci: *Long Short-Term Memory* (LSTM), *Bidirectional Encoder Representations from Transformers* (BERT), analisis klasifikasi teks.

ABSTRACT

COMPARISON OF LONG SHORT-TERM MEMORY (LSTM) AND BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS (BERT) METHOD IN CLASSIFYING FAKE NEWS

By

SURYO MUQSITHO

20/456589/PA/19776

The spread of fake news, or hoaxes, has become increasingly frequent, especially during election periods. These fabricated news articles are often used as political tools to manipulate public opinion, discredit opponents, and damage their reputations. To combat this issue, this study investigates the application of both Long Short-Term Memory (LSTM) and Bidirectional Encoder Representations from Transformers (BERT) for text classification analysis in identifying fake news. The study utilizes a dataset compiled from various news sources, including CNN, Kompas, Tempo, and turnbackhoax.id, a fact-checking news portal. The performance of both methods is evaluated using metrics such as accuracy, recall, F1-score, and Matthews Correlation Coefficient (MCC). Additionally, a comparison is conducted between balanced and imbalanced datasets, employing the random over-sampling method. The findings demonstrate that BERT outperforms LSTM overall in text classification analysis for the Indonesian language, effectively handling both balanced and imbalanced datasets.

Keywords: Long Short-Term Memory (LSTM), Bidirectional Encoder Representations from Transformers (BERT), text classification.