

ABSTRAK

Penelitian ini bertujuan untuk mengimplementasikan model berbasis *Deep Learning*, khususnya arsitektur Transformer IndoBERT, dalam mendeteksi *clickbait* pada judul berita berbahasa Indonesia dengan mengintegrasikan fitur identitas penerbit. Dengan berkembangnya media daring dalam ekosistem *Attention Economy*, perhatian pengguna menjadi komoditas utama yang diperebutkan. Kondisi ini mendorong banyak penerbit berita untuk memprioritaskan trafik di atas akurasi dengan menggunakan judul *clickbait* yang memicu respons emosional serta memanfaatkan kesenjangan informasi pembaca. Meskipun berbagai model deteksi *clickbait* telah dikembangkan sebelumnya, sebagian besar penelitian terdahulu hanya berfokus pada analisis fitur teks (judul) dan mengabaikan fakta bahwa setiap media penerbit memiliki tingkat sensasionalisme serta gaya redaksional yang berbeda-beda.

Penelitian ini menggunakan dataset sekunder CLICK-ID, khususnya subset *all_agree*, yang memuat data judul berita dari 12 penerbit berita daring utama di Indonesia. Data tersebut melalui tahap prapemrosesan sistematis berupa pembersihan data, *case folding*, tokenisasi menggunakan *tokenizer* bawaan IndoBERT, serta rekayasa fitur untuk mengubah identitas penerbit menjadi representasi vektor numerik (*embedding*). Dua skenario model dikembangkan untuk evaluasi komparatif: Model *Baseline* yang hanya mengekstraksi 768 dimensi input dari teks judul, dan Model Usulan yang menggabungkan (melalui operasi *concatenation*) vektor teks judul tersebut dengan 8 dimensi vektor *embedding* identitas penerbit. Proses *fine-tuning* pada model dilakukan menggunakan *optimizer* AdamW, fungsi *Cross Entropy Loss*, aktivasi *Softmax* untuk prediksi kelas, serta penerapan strategi *early stopping* guna mencegah terjadinya *overfitting* selama pelatihan.

Hasil pengujian menunjukkan bahwa pengintegrasian fitur identitas penerbit mampu meningkatkan kinerja deteksi secara drastis. Model Usulan yang menggunakan fitur penerbit berhasil mencapai tingkat akurasi sebesar 95,93%, presisi 95,95%, *recall* 93,33%, dan *F1-Score* 94,62% dengan nilai *Area Under the Curve* (AUC) mencapai 0,99. Sebagai perbandingan, Model *Baseline* yang hanya menggunakan teks judul berita mencapai akurasi 93,26% dan *F1-Score* 91,21% dengan nilai AUC 0,98. Signifikansi peningkatan performa ini divalidasi lebih lanjut melalui uji statistik *McNemar's Test* yang menghasilkan *p-value* sebesar 0,0000662, jauh di bawah tingkat signifikansi 0,05. Penelitian ini mengonfirmasi secara empiris bahwa reputasi dan gaya bahasa spesifik dari media penerbit merupakan indikator yang sangat kuat dalam klasifikasi *clickbait*. Hasil dari pengembangan model klasifikasi ini diharapkan dapat diimplementasikan lebih lanjut sebagai sistem penyaringan informasi otomatis, seperti ekstensi peramban (*browser extension*), untuk membantu pengguna internet menghindari informasi yang sensasional atau menyesatkan.

Kata Kunci: deteksi clickbait, IndoBERT, identitas penerbit, Natural Language Processing, deep learning.

ABSTRACT

This research aims to implement a Deep Learning-based model, specifically the IndoBERT Transformer architecture, in detecting clickbait in Indonesian news headlines by integrating the publisher identity feature. With the development of online media within the Attention Economy ecosystem, user attention has become a primary contested commodity. This condition drives many news publishers to prioritize traffic over accuracy by using clickbait headlines that trigger emotional responses and exploit the readers' information gap. Although various clickbait detection models have been developed previously, most prior studies focused only on analyzing textual features (headlines) and ignored the fact that each news publisher has a different level of sensationalism and editorial style.

This study utilizes the CLICK-ID secondary dataset, specifically the all_agree subset, which contains news headline data from 12 major online news publishers in Indonesia. The data went through a systematic preprocessing stage consisting of data cleansing, case folding, tokenization using IndoBERT's built-in tokenizer, and feature engineering to convert the publisher's identity into a numerical vector representation (embedding). Two model scenarios were developed for comparative evaluation: a Baseline Model that only extracts 768 input dimensions from the headline text, and a Proposed Model that combines (through a concatenation operation) the headline text vector with an 8-dimensional publisher identity embedding vector. The fine-tuning process on the model was carried out using the AdamW optimizer, Cross Entropy Loss function, Softmax activation for class prediction, and the implementation of an early stopping strategy to prevent overfitting during training.

The test results show that the integration of the publisher identity feature significantly improves detection performance. The Proposed Model, which uses the publisher feature, successfully achieved an accuracy rate of 95.93%, precision of 95.95%, recall of 93.33%, and an F1-Score of 94.62%, with an Area Under the Curve (AUC) value reaching 0.99. For comparison, the Baseline Model, which only uses the news headline text, achieved an accuracy of 93.26% and an F1-Score of 91.21%, with an AUC value of 0.98. The significance of this performance improvement was further validated through McNemar's Test, which yielded a p-value of 0.0000662, well below the 0.05 significance level. This study empirically confirms that the reputation and specific linguistic style of the publishing media are very strong indicators in clickbait classification. The results of this classification model development are expected to be further implemented as an automatic information filtering system, such as a browser extension, to help internet users avoid sensational or misleading information.

Keywords: clickbait detection, IndoBERT, publisher identity, Natural Language Processing, deep learning.