

ABSTRAK

Perkembangan teknologi informasi meningkatkan interaksi pengguna pada kolom komentar berita, namun turut memunculkan komentar buatan akun bot yang menurunkan kualitas diskusi, menyebarkan misinformasi, dan memengaruhi opini publik. Metode deteksi berbasis aturan maupun *machine learning* konvensional masih terbatas dalam memahami konteks bahasa secara mendalam. Penelitian ini menerapkan model *IndoBERT*, varian BERT yang dioptimalkan untuk bahasa Indonesia, untuk mendeteksi komentar *bot* berdasarkan pola linguistik dan konteks semantik menggunakan pendekatan *Natural Language Processing* (NLP). Data penelitian berupa 1.000 komentar yang terbagi seimbang menjadi 500 komentar *bot* dan 500 komentar manusia dikumpulkan melalui teknik web scraping dari enam portal berita dan forum Indonesia, yaitu Detik.com, Kompas.com, Kompasiana.com, Kaskus.co.id, Quora.com, dan CNN Indonesia. Tahapan penelitian meliputi pengumpulan data, pelabelan, *preprocessing*, pembagian data menjadi 70% data latih dan 30% data uji, pelatihan dan *fine-tuning* model *IndoBERT*, serta evaluasi menggunakan *Accuracy*, *Precision*, *Recall*, *F1-Score*, dan *Matthews Correlation Coefficient* (MCC). Hasil pengujian menunjukkan performa yang sangat baik dengan *Accuracy* 95,67%, *Precision* 91,56%, *Recall* 100%, *F1-Score* 95,59%, dan MCC 0,9164, mengungguli model pembandingan *Random Forest* yang memperoleh *F1-Score* 91,84%. Model *IndoBERT* selanjutnya diimplementasikan dalam aplikasi web sebagai solusi praktis untuk mendukung proses deteksi komentar *bot* secara otomatis. Hasil penelitian menunjukkan bahwa *IndoBERT* efektif mengidentifikasi komentar *bot* dan berpotensi dikembangkan lebih lanjut sebagai sistem moderasi komentar pada platform berita di masa mendatang.

Kata Kunci: Deteksi Komentar Bot, BERT, *IndoBERT*, *Natural Language Processing*, Klasifikasi Teks.

ABSTRACT

Advances in information technology have increased user interaction in news comment sections, but this has also given rise to bot-generated comments that degrade discussion quality, spread misinformation, and influence public opinion. Rule-based and conventional machine learning detection methods remain limited in understanding language context in depth. This study applies IndoBERT, a BERT variant optimized for Indonesian, to detect bot comments based on linguistic patterns and semantic context using a Natural Language Processing (NLP) approach. The data, comprising 1,000 comments evenly divided into 500 bot and 500 human comments, were collected through web scraping from six Indonesian news portals and forums: Detik.com, Kompas.com, Kompasiana.com, Kaskus.co.id, Quora.com, and CNN Indonesia. The research stages included data collection, labeling, preprocessing, splitting the data into 70% training and 30% testing data, training and fine-tuning IndoBERT, and evaluation using Accuracy, Precision, Recall, F1-Score, and Matthews Correlation Coefficient (MCC). Test results showed excellent performance, with an Accuracy of 95.67%, Precision of 91.56%, Recall of 100%, F1-Score of 95.59%, and MCC of 0.9164, outperforming the Random Forest baseline, which achieved an F1-Score of 91.84%. IndoBERT was further implemented in a web application as a practical solution to support automatic bot comment detection. These findings indicate that IndoBERT effectively identifies bot comments and has potential for further development as a comment moderation system for news platforms.

Keywords: *Bot Comment Detection, BERT, IndoBERT, Natural Language Processing, Text Classification.*