

ABSTRAK

Perkembangan teknologi dan internet telah mengubah cara masyarakat Indonesia mengakses informasi, dengan lebih dari 132 juta pengguna internet kini mengandalkan media digital sebagai sumber berita utama. Transformasi ini menimbulkan risiko penyebaran berita hoaks yang semakin mengkhawatirkan, terutama melalui *platform* media sosial seperti Facebook, YouTube, dan Instagram yang digunakan oleh sekitar 50% masyarakat untuk memperoleh informasi. Fenomena ini berdampak negatif terhadap opini publik, kepercayaan masyarakat, dan stabilitas sosial, sehingga diperlukan metode yang efektif untuk mengklasifikasi berita (hoaks dan asli) berbahasa Indonesia yang beredar di media digital. Penelitian ini mengusulkan model deteksi berita hoaks dengan mengintegrasikan metode *Indonesian Bidirectional Encoder Representations from Transformers* (IndoBERT) sebagai *feature extractor* dan *Support Vector Machine* (SVM) sebagai *classifier* untuk mengoptimalkan akurasi klasifikasi. Proses implementasi dimulai dengan *fine-tuning model pre-trained* IndoBERT menggunakan *AdamW Optimizer* dengan mengombinasikan *weight decay*, *learning rate*, dan *batch size* paling optimal sehingga menghasilkan *validation accuracy* sebesar 96,25%. Token [CLS] berukuran 1×768 dari hasil *training* IndoBERT kemudian diekstraksi sebagai *input* untuk model SVM. Optimasi parameter SVM dilakukan menggunakan *Grid Search* dan *K-Fold Cross Validation* dengan 20 kombinasi parameter berbeda, di mana kombinasi terbaik adalah $C = 10$ dan $\gamma = 0,001$ dengan *validation accuracy* sempurna (100%). Model yang dihasilkan kemudian dievaluasi dan mencapai performa yang sangat baik pada *data testing* dengan *accuracy* 93,75%, *precision* 97,30%, *recall* 90%, dan *F1-score* 93,51%. Hasil ini menunjukkan bahwa IndoBERT dan SVM mampu mengklasifikasikan berita berbahasa Indonesia secara efektif dan seimbang, memberikan solusi praktis untuk meningkatkan keakuratan penyaringan informasi di media digital dan mengurangi penyebaran disinformasi.

Kata kunci : Klasifikasi berita, *natural language processing*, IndoBERT, SVM