

BAB 5 SIMPULAN DAN SARAN

5.1 Simpulan

Penelitian ini menghasilkan beberapa kesimpulan sebagai berikut:

1. Implementasi model IndoBERT dalam mendeteksi berita hoaks berbahasa Indonesia berhasil dilakukan melalui proses fine-tuning dengan mengoptimalkan dua hyperparameter. Konfigurasi optimal dicapai pada epoch kedua dengan learning rate 3×10^{-5} , dimana model menunjukkan keseimbangan terbaik antara kemampuan pembelajaran dan generalisasi. Peningkatan epoch ke-3 atau learning rate yang lebih tinggi (4×10^{-5}) justru mengakibatkan penurunan performa dan memicu overfitting, yang ditandai dengan divergensi antara training loss dan validation loss.
2. Pada data training, model ini mencapai akurasi 99.35%, presisi 98.97%, recall 99.67%, dan F_1 -score 99.32%. Performa ini berhasil dipertahankan pada data validasi dengan akurasi 99.09%, presisi 98.87%, recall 99.22%, dan F_1 -score 99.04%. Pada data testing, model ini tetap mempertahankan performa pada partisi data sebelumnya dengan akurasi 98.71%, presisi 98.79%, recall 98.51%, dan F_1 -score 98.65%. Model ini juga menunjukkan distribusi error yang seimbang antara false positive dan false negative.

5.2 Saran

Meskipun model IndoBERT ini sudah membawa hasil yang cukup memuaskan, namun beberapa keterbatasan perlu ditangani dalam penelitian mendatang. Pertama, model ini hanya dapat mendeteksi berita hoaks berbasis teks saja. Dan dengan berkembangnya platform sosial media yang mengandalkan video sebagai media utamanya, diperlukan suatu model yang dapat mengklasifikasikan hoaks selain teks. Selain itu, model ini juga memiliki keterbatasan dalam mengandalkan konteks bahasa tanpa mempertimbangkan akurasi faktual konten. Lebih lanjut, dataset yang digunakan dalam studi ini tidak seimbang, dengan artikel hoaks yang secara signifikan lebih pendek dibandingkan artikel berita faktual. Keterbatasan ini menunjukkan area-area yang perlu diperbaiki, terutama dalam hal

mengintegrasikan mekanisme pemeriksaan fakta dan mengatasi ketidakseimbangan dataset.

Penelitian selanjutnya dapat mengeksplorasi integrasi kemampuan *multimodal* untuk mendeteksi hoaks dalam gambar dan video, memperluas kemampuan model pada berbagai jenis konten; tidak terbatas hanya pada konten berbasis teks saja. Dan dikarenakan dataset yang digunakan hanya berfokus pada konteks berita saja, penggunaan dataset yang lebih beragam serta eksplorasi teknik seperti *data augmentation* atau *synthetic hoax generation* dapat meningkatkan ketangguhan dan generalisasi model ini di berbagai media dan konteks. Dengan mengatasi keterbatasan-keterbatasan ini dan memperluas cakupannya, penelitian selanjutnya memiliki potensi untuk secara signifikan meningkatkan kemampuan sistem deteksi hoaks otomatis, memberikan kontribusi pada pertahanan keamanan siber yang lebih efektif menghadapi lanskap ancaman siber yang terus berkembang.

