

BAB 5 SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan hasil pengujian yang telah dilakukan, Berikut hal yang dapat disimpulkan.

1. Implementasi algoritma BERT pada proses klasifikasi sentimen pengguna Twitter terhadap ChatGPT telah berhasil. Pengujian dilakukan menggunakan data tweet berbahasa Indonesia mengenai ChatGPT dari situs Twitter dari rentang tanggal 1 Desember 2024 sampai dengan 31 Desember 2024 dengan total 4.174 tweet. Model yang digunakan pada penelitian ini adalah BERT based uncased untuk digunakan pada tahap pretraining dan fine-tuning.
2. Pengujian dilakukan dengan menggunakan 3 skenario pengujian yang berbeda, yaitu menggunakan dataset yang tidak seimbang, data yang seimbang dengan metode *undersampling*, dan data yang seimbang dengan metode *oversampling*. Hasil pengujian menunjukkan bahwa skenario pengujian pada data yang seimbang dengan metode *oversampling* memberikan performa yang lebih tinggi dibandingkan dengan skenario yang lain, dengan akurasi sebesar 95%, presisi sebesar 95%, recall sebesar 95%, dan f1-score sebesar 95%.

5.2 Saran

Berikut adalah beberapa saran yang dapat diberikan untuk peneliti selanjutnya.

1. Melakukan proses *text cleaning* secara minimal, dikarenakan tokenizer pada model BERT dapat melakukan sebagian besar *text preprocessing*. Beberapa tahap seperti konversi ke lowercase, menghapus hyperlink, dan menghapus teks yang tidak valid seperti baris kosong masih diperlukan untuk meningkatkan akurasi pada hasil training model.
2. Menggunakan model Natural Language Processing yang lain seperti GPT-3, GPT-4, atau RoBERTa untuk dijadikan perbandingan hasil dengan model BERT.