

## BAB V PENUTUP

### 5.1 Kesimpulan

Penelitian ini menggunakan dua dataset yang diberi nama Dataset A dan Dataset B. Dataset A memiliki jumlah data sebanyak 600 baris data dan merepresentasikan dataset dengan jumlah data kecil. Dataset B memiliki jumlah data sebanyak 1.786 baris data dan merepresentasikan dataset dengan jumlah data besar. BERT menghasilkan performa terbaik pada jenis *truncation* yaitu *head-only truncation* pada Dataset A dan Dataset B. BERT dengan *head-only truncation* menghasilkan akurasi sebesar 67%, *precision* sebesar 53%, *recall* sebesar 48%, dan *f-score* sebesar 45% pada Dataset A serta menghasilkan akurasi sebesar 93%, *precision* sebesar 90%, *recall* sebesar 90%, dan *f-score* sebesar 89% pada Dataset B. Waktu pelatihan BERT pada Dataset A adalah sekitar 9 menit dibandingkan dengan waktu pelatihan pada Dataset B yang membutuhkan waktu sekitar 27 menit.

Perbandingan BERT dengan model *recurrent* menunjukkan bahwa LSTM dan GRU dapat mengungguli BERT dalam dataset yang kecil yaitu Dataset A. BERT hanya bisa mengungguli RNN pada dataset kecil. Hal tersebut karena model mengalami *overfit* pada *epoch* awal. Sedangkan pada Dataset B yang besar, performa BERT dapat mengungguli LSTM pada seluruh metrik evaluasi. BERT dapat meningkatkan akurasi sebesar 8%, *precision* 6%, dan *recall* sebesar 7% daripada LSTM.

### 5.2 Saran

Pada penelitian ini masih ada beberapa hal yang bisa dikembangkan untuk penelitian selanjutnya. Data masih dapat dilakukan pra-pemrosesan lebih lanjut seperti augmentasi data untuk meningkatkan performa BERT. Penggunaan model dapat dikembangkan dengan mencari model variasi dari BERT yang lebih efisien dengan memiliki parameter lebih sedikit dan membutuhkan waktu pelatihan yang lebih cepat namun masih dapat mempertahankan akurasi dari BERT<sub>BASE</sub>.