

BAB 5 SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan hasil dari implementasi dan uji coba yang sudah dilakukan, dapat disimpulkan bahwa model BLOOM yang telah di-*fine-tuning* untuk mendeteksi *hate speech* bilingual bahasa Indonesia-Inggris telah selesai dibuat. Beberapa data yang digunakan dalam penelitian ini adalah data dari Twitter bahasa Indonesia dengan total 13.168 data yang terdiri dari 5.561 baris data *hate speech* dan 7.607 baris data *non-hate speech*, data dengan Bahasa Inggris dengan total 998 data yang terdiri dari 433 baris data *hate speech* dan 565 baris data *non-hate speech*, dan data bilingual yang berjumlah 700 baris data yang terdiri dari 232 baris data *hate speech* dan 468 baris data *non-hate speech*. Hasil dari model BLOOM Filter yang sudah di-*fine-tuning* didapatkan performa terbaik pada 6 epoch dengan nilai metrik evaluasi accuracy 94,28% , precision 88,63%, recall 92,85%, dan f1-score 90,69%. Model BLOOM Filter dapat mendeteksi 6 teks yang dimasukkan manual dengan baik, model berhasil mendeteksi 5 dari 6 teks yang dimasukkan.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan mengenai IMPLEMENTASI BLOOM FILTER UNTUK PENDETEKSI HATE SPEECH BILINGUAL BAHASA INDONESIA DAN INGGRIS, terdapat beberapa saran yang dapat dilakukan untuk peneliti selanjutnya dikarenakan adanya keterbatasan waktu, keterbatasan *dataset*, dan keterbatasan perangkat lunak, diantaranya adalah sebagai berikut.

1. Penggunaan *dataset* dengan kelas yang seimbang, dan *dataset* bilingual yang lebih dominan.
2. Membangun model BLOOM Filter untuk klasifikasi jenis *hate speech* bilingual.
3. Model belum bisa mendeteksi kalimat bilingual yang mengandung sarkasme, sehingga untuk penelitian selanjutnya dapat membangun model yang bisa mendeteksi sarkasme didalam kalimat bilingual.