

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Penelitian ini berhasil membangun sebuah model klasifikasi emosi berdasarkan teks bahasa Indonesia menggunakan algoritma CNN dan layer LSTM dengan tingkat akurasi sebesar 58%. Dari ke-5 jenis emosi yang ada pada penelitian ini, emosi *sadness* adalah jenis emosi yang memiliki performa paling rendah. Hal ini bisa disebabkan karena adanya *overfitting* pada dataset yang digunakan ataupun karena adanya varian yang tinggi. Secara keseluruhan, performa model yang dihasilkan dalam penelitian ini masih berada di bawah *baseline* yang telah ditetapkan yaitu akurasi sebesar 69.73%, sehingga dapat disimpulkan bahwa dataset pada penelitian ini akan lebih baik jika menggunakan model yang dibangun dengan kombinasi *feature* dan algoritma *linear regression* dari pada model yang dibangun pada penelitian ini.

5.2. Saran

Saran untuk penelitian ini adalah:

- a. Mencoba menggunakan algoritma yang berbeda dengan dataset yang sama.
- b. Mengurangi jumlah data label *anger*, *happy*, dan *sadness* dengan mengambil sebagian dari jumlah data sebenarnya secara acak atau menyamakan jumlah data dengan jumlah rata-rata label *fear* dan *love*.

- c. Gunakan lebih banyak sumber data khususnya data dengan emosi *sadness* untuk melakukan training pada model untuk mengurangi varian.
- d. Membangun kamus bahasa Indonesia yang lebih besar dengan menggunakan sumber diluar dataset yang digunakan.
- e. Menggunakan lebih banyak Kamus *Typography* Indonesia atau membangun Kamus *Typography* Indonesia sendiri untuk mengubah kata tidak baku yang memiliki makna yang sama menjadi satu kata yang sama.
- f. Menggunakan *pre-trained* model yang sudah ada seperti BERT.