BAB 5
SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan implementasi dalam penelitian yang telah dilakukan, dapat disimpulkan bahwa algoritma *convolutional neural network* dengan teknik *word embedding* menggunakan *pre-trained model* FastText dapat diimplementasikan untuk melakukan analisis sentimen pada twit (*tweet*) *cyberbullying*. Hasil model berdasarkan penggunaaan FastText sebagai metode *word embedding*, *callback function* `ModelCheckpoint()` dan `EarlyStopping()` menunjukkan bahwa akurasi serta *loss terbaik* adalah pada 117 dari 420 *epochs* menggunakan *learning rate* sebesar 0,01, serta `batch_size` sebesar 2048. Adapun akurasi yang diperoleh adalah sebesar 89,18%, dengan *loss* sebesar 29% karena ketika proses pelatihan hampir menuju epoch ke-90, nilai akurasi dan loss sama-sama tidak mengalami peningkatan sehingga berhenti di nilai tersebut karena adanya `EarlyStopping()` yang menghentikan proses pelatihan di saat sudah optimal. Berarti, model ini walaupun model sudah memiliki performa baik sebesar 89,18%, model ini masih membuat *error* dalam sistem prediksi sebesar 29% dalam data.

Pada sisi lain, jika dibandingkan dengan penelitian-penelitian sebelumnya, misalkan, menggunakan algoritma Naıve Bayes di tahun 2016, hasil penelitian ini memiliki akurasi yang lebih tinggi, yaitu 89,18%. Hasil penelitian ini jelas lebih tinggi dibandingkan penelitian sebelumnya, yang memiliki hasil akurasi 87,67%. Akan tetapi, ada faktor lain yang mempengaruhi perbedaan hasil tersebut, salah satunya, jumlah data. Penelitian ini menggunakan 47.692 data, sedangkan penelitian sebelumnya menggunakan 339.698 data. Hal lain yang bisa dibandingkan dengan penelitian selanjutnya adalah, penelitian ini menggunakan *callback function* `EarlyStopping()` dari Keras yang dapat menghentikan proses pelatihan jika sudah mendapatkan *validation loss* terkecil, sedangkan penelitian menggunakan Naive Bayes di tahun 2016 tidak menggunakannya.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang bisa dikembangkan untuk penelitian di masa depan antara lain.

1. Menggunakan teknik-teknik gabungan dengan *convolutional neural network*,

supaya bisa mendapat hasil akurasi yang lebih baik dan *loss* yang lebih kecil, seperti menggunakan gabungan *pre-trained model* untuk *word embedding* lain dengan algoritma yang telah diimplementasikan.

2. Menggunakan kombinasi *epoch* dan variasi *learning rate* untuk mencari akurasi yang lebih baik.

3. Menggunakan *dataset* dengan tingkat keseimbangan yang relatif lebih tinggi yang diharapkan dapat memberikan hasil dan akurasi yang lebih baik serta *loss* yang lebih kecil dalam melakukan klasifikasi *multi-class*.

4. Menggunakan *synonym extraction* untuk mengelompokkan kata-kata yang memiliki banyak singkatan(contoh: ”WDYM?” ditulis ”What do you mean?”)

5. Dalam menggunakan *callback function*, gunakan parameter waktu seperti `patience` yang lebih besar nilainya, sehingga lebih lama berhenti waktu pelatihan algoritma dan berpotensi mendapatkan akurasi yang lebih baik.

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA