

BAB V PENUTUP

5.1 Kesimpulan

Pembahasan pada bab IV di atas membawa penelitian ini memasuki bab V dimana peneliti akan menarik kesimpulan. Adapun beberapa kesimpulan yang didapat pada penelitian ini, yaitu:

1. Model *BiLSTM* terbaik pada penelitian ini adalah model P2 dengan menggunakan *stemming* NDETC yang memiliki parameter *hidden size* 128, *dropout* 0,4, dan *learning rate* 0,001.
2. Hasil klasifikasi pada data test lirik musik yang *distemming* dengan NDETC menggunakan model P2 menunjukkan nilai *precision* 88,89%, *recall* 88,28%, *F1* 88,58%, dan akurasi 88,93%. Model tersebut memprediksi data test musik dengan hasil *confusion matrix* True Positive 128, True Negative 137, False Positive 16 dan False Negative 17 dengan actual value positive (emosi senang) sebanyak 145, actual value negative 153 (emosi sedih).
3. Pada penelitian ini, model *BiLSTM* dengan parameter *dropout* 0.4 memiliki rata-rata akurasi model yang lebih baik dari pada *dropout* 0,3 baik itu menggunakan data *stemming* NDETC dan sastrawi. Pada parameter *hidden size* 64 dan 128 pada data *stemming* NDETC rata-rata akurasinya adalah 80% dan 79%, sedangkan pada data *stemming* sastrawi rata-rata akurasinya 81% dan 82%. *Hyperparameter learning rate* terbaik berada pada nilai 0,001 baik menggunakan *stemming* NDETC dan sastrawi yang memiliki nilai rata-rata akurasi sebesar 85%.
4. Efektifitas dan kecocokan algoritma *BiLSTM* dalam hal klasifikasi teks terkhusus lirik musik berbahasa Indonesia cukup baik dengan nilai akurasi tertinggi 88,93% dengan penyesuaian terhadap beberapa parameter seperti *hidden size*, *dropout* dan *learning parameter*.
5. Penggunaan *stemming* NDETC dan sastrawi dapat mempengaruhi kinerja model, dimana hasil klasifikasi menggunakan model P2 mendapat nilai akurasi sebesar 88,93% pada data *stemming* NDETC dan 88,59% pada data *stemming* sastrawi. Hasil *stemming* pada data lirik musik juga mempunyai

perbedaan pada *vocab size*, dimana *vocab size* pada data *stemming* NDETC berjumlah 5408 sedangkan *vocab size* data *stemming* sastrawi berjumlah 5458.

6. Hasil pengujian pada aplikasi web dengan 10 data lirik musik yang sudah dilabeli emosi dan diklasifikasikan dengan model *BiLSTM* terbaik yaitu model P2, mendapatkan nilai akurasi 90%, *precision* 100%, *recall* 83,33% dan *F1-score* 90,91%.

5.2 Saran

Berdasarkan dari penelitian yang telah dilakukan, penulis menyadari bahwa aplikasi klasifikasi emosi, analisis dan implementasinya sendiri memiliki beberapa kekurangan. Dalam hal ini penulis berharap agar dikemudian hari penelitian ini dapat dikembangkan oleh peneliti-peneliti mendatang menjadi lebih baik dengan beberapa saran dari penulis seperti:

1. Menambahkan jumlah data lirik musik dengan parameter emosi yang lebih kompleks.
2. Melakukan teknik penyeimbangan dataset jika salah satu kelas memiliki ketimpangan.
3. Menggunakan teknik *negation handling* pada tahap *preprocessing*.
4. Menggunakan *pretrained word embedding*, seperti *Word2Vec*, *Fasttext* atau *Glove*.
5. Menambahkan skenario model dengan penyesuaian parameter dan *hyperparameter* lebih mendalam serta membuat model algoritma dengan lebih banyak *layer* lagi.
6. Menambahkan beberapa algoritma *deep learning* lain sebagai pembanding untuk mendapatkan model terbaik.
7. Membuat tampilan dan fungsi pada aplikasi web klasifikasi emosi lirik musik lebih baik dan rapi.