

## BAB 5 SIMPULAN DAN SARAN

### 5.1 Simpulan

Berdasarkan hasil pengujian yang telah dilakukan, dapat disimpulkan bahwa implementasi algoritma *Recurrent Neural Network* dengan model *Long Short-Term Memory* pada proses klasifikasi ulasan film dari situs IMDb telah berhasil. Pengujian dilakukan menggunakan data ulasan film dari situs IMDb dengan total 320.747 ulasan. Pengujian dilakukan dengan menggunakan beberapa skenario yang memiliki pembagian rasio berbeda, yaitu 60:40, 70:30, dan 80:20. Hasil pengujian menunjukkan bahwa skenario dengan rasio 80% *training* dan 20% *testing* memberikan performa lebih tinggi, dengan akurasi sebesar 96%, presisi sebesar 97%, *recall* sebesar 98%, dan *f1-score* sebesar 97%.

Selain itu, penelitian ini juga menunjukkan bahwa pembagian data *training* dan data *testing* memiliki pengaruh yang signifikan terhadap hasil akurasi. Data yang semakin banyak digunakan dalam proses *training*, maka hasil akurasi yang diperoleh cenderung lebih baik. Dari total 320.747 data yang digunakan, sebanyak 80,36% memiliki sentimen positif, sementara 19,63% memiliki sentimen negatif. Analisis sentimen pada ulasan film IMDb menunjukkan bahwa sentimen positif memiliki persentase yang lebih tinggi dibandingkan sentimen negatif.

### 5.2 Saran

Berdasarkan hasil penelitian yang dilakukan, terdapat beberapa rekomendasi yang dapat dijadikan sebagai pedoman untuk pengembangan penelitian lanjut, antara lain:

1. Pada penelitian selanjutnya, diharapkan bisa mencoba *library stemming* lain seperti *spaCy* karena pada *library NTLK* masih ada beberapa kata-kata yang tidak baku yang dapat mempengaruhi performa model.
2. Dapat menggunakan algoritma *Support Vector Machine* pada dataset IMDb 320.000 *reviews* ini untuk dijadikan perbandingan hasil dengan algoritma *Recurrent Neural Network*.
3. Coba menyamaratakan jumlah label positif, negatif sebelum pelabelan dengan *library* dan sesudah menggunakan *library*.