BAB V
PENUTUP

5.1 Kesimpulan

Berdasarkan seluruh tahap penelitian yang telah dilakukan, mulai dari pengumpulan data, *preprocessing*, hingga pembangunan dan evaluasi model *Long Short-Term Memory (LSTM)* untuk klasifikasi emosi pada teks berbahasa Indonesia, dapat ditarik beberapa kesimpulan sebagai berikut:

1. Berdasarkan hasil pengujian pada data uji, model menghasilkan tingkat akurasi sebesar 81%. Analisis metrik menunjukkan performa terbaik justru diraih oleh kelas *Love* dengan *F1-Score* 0.83 dan *Recall* 0.87, sedangkan nilai *Precision* tertinggi dicapai oleh kelas *Fear* sebesar 0.85. Secara keseluruhan, rata-rata tertimbang (*weighte avg*) untuk *Precision, Recall,* dan *F1-Score* yang seragam berada di angka 0.81, membuktikan bahwa model memiliki stabilitas yang baik dan seimbang dalam mengklasifikasikan kelima kategori emosi.
2. Model memiliki karakteristik paling kuat dalam mengenali kelas *Love*, disusul oleh kelas *Anger, Fear,* dan *Joy* yang memiliki performa setara *F1-Score* 0.82. Di sisi lain, kelas *Sad* teridentifikasi sebagai kategori yang paling sulit dikenali dengan *F1-Score* terendah yaitu, 0.77. Analisis *Confusion Matrix* mengungkap pola kesalahan signifikan berupa ambiguitas antara emosi negatif. Terdapat saling tukar prediksi yang tinggi di mana 30 data Sad salah diprediksi sebagai *Anger*, sementara 11 data *Fear* dan 27 data *Anger* justru keliru diprediksi masuk ke kategori *Sad*. Temuan ini menunjukkan bahwa dalam data teks media sosial di Indonesia, ekspresi kesedihan seringkali mengandung nuansa frustasi, marah atau ketakutan.

5.2 Saran

Berdasarkan kendala yang dihadapi selama proses penelitian dan analisis terhadap hasil yang diperoleh, terdapat beberapa saran yang dapat diajukan untuk pengembangan penelitian selanjutnya agar hasil yang didapatkan lebih maksimal:

1. Penelitian selanjutnya disarankan untuk memperbanyak jumlah dataset secara signifikan, baik dengan memperluas rentang waktu pengambilan data maupun menambah sumber platform lain. Semakin besar volume data yang digunakan, semakin kaya variasi kosakata dan pola bahasa yang dapat dipelajari oleh model, sehingga mampu meningkatkan akurasi dan generalisasi model dalam mengenali emosi.
2. Penelitian selanjutnya disarankan untuk mencoba variasi arsitektur yang lebih kompleks, seperti *Bidirectional LSTM (Bi-LSTM)* atau model berbasis *Transformer* seperti *IndoBERT*. Penggunaan model tersebut diharapkan dapat menangkap konteks kalimat dua arah dengan lebih baik, sehingga mampu mengurangi kesalahan prediksi antara kelas emosi yang memiliki kemiripan konteks.
3. Karena cepatnya perkembangan bahasa gaul di media sosial, disarankan untuk menggunakan atau mengembangkan kamus normalisasi teks yang lebih luas dan terkini, atau menerapkan teknik *spell checking* otomatis guna menangani variasi penulisan kata *(typo)* yang belum terdaftar di kamus manual.
4. Penelitian selanjutnya, bisa meneliti secara linguistik kenapa ambiguitas antara takut dan sedih itu sangat tinggi dalam konteks opini publik pada media sosial di Indonesia. Perlu dikaji apakah ini merupakan cerminan budaya berekspresi, di mana rasa ketakutan sering diungkapkan dengan narasi keluh kesah (sehingga terbaca sedih), atau memang karena keterbatasan leksikon bahasa slang yang digunakan oleh pengguna saat ini.