

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian dan analisis yang sudah dilakukan maka dapat disimpulkan beberapa hal sebagai berikut:

1. Algoritma *Naive Bayes Multinomial*, *Gaussian*, dan *Bernoulli* memiliki performa berbeda dalam mengklasifikasikan sentimen masyarakat terhadap dampak *AI*. Seluruh proses dilakukan, mulai dari *Crawling* komentar *YouTube*, *Preprocessing text*, pelabelan sentimen, vektorisasi TF-IDF, pembagian data, penyeimbangan dengan SMOTE, hingga pelatihan dan evaluasi model. Hasilnya menunjukkan perbandingan performa antar model yang kemudian model terbaik diimplementasikan ke dalam aplikasi berbasis *Streamlit*.
2. Penggunaan nilai parameter *Alpha* pada *Multinomial Naive Bayes* menunjukkan bahwa nilai *Alpha* dapat mempengaruhi kinerja model. Saat diuji menggunakan parameter *Alpha* 0.1 mendapatkan akurasi sebesar 87,04%, diikuti dengan *Alpha* 0.3 dan 0.5 dengan akurasi sebesar 87,50%.
3. Dari hasil pengujian yang ada menunjukkan tingkat akurasi terbaik dicapai oleh *Multinomial Naive Bayes* tanpa parameter *Alpha* dengan nilai akurasi sebesar 90,28% dalam mengklasifikasi opini masyarakat terhadap dampak *Artificial Intelligence*. Sementara itu, *Gaussian* dan *Bernoulli Naive Bayes* hanya mencapai akurasi sebesar 68,06% dan 68,52% yang menunjukkan performa yang lebih rendah karena kurang cocok dengan data teks seperti komentar *YouTube*.

5.2 Saran

1. Untuk penelitian selanjutnya adalah memperluas cakupan data dengan memasukkan sentimen netral, sehingga analisis sentimen dapat mencerminkan keseluruhan spektrum opini publik secara lebih utuh.

Selain itu, karena dalam penelitian ini digunakan teknik SMOTE untuk menyeimbangkan data latih, penelitian selanjutnya disarankan menggunakan data yang seimbang secara alami atau mempertimbangkan teknik balancing lain agar lebih mendekati distribusi sesungguhnya.

2. Penggunaan algoritma *Naive Bayes* juga dapat diperluas dengan membandingkan hasilnya dengan model yang lebih kompleks seperti *Support Vector Machine (SVM)*, *Random Forest*, atau model *Deep Learning* seperti BERT dan LSTM. Penelitian lanjutan juga dapat mengevaluasi pengaruh berbagai representasi fitur teks, seperti penggunaan bigram, trigram, atau embedding berbasis konteks, untuk meningkatkan performa klasifikasi sentimen.

