

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan seluruh tahapan penelitian yang telah dilaksanakan, mulai dari pengumpulan data komentar YouTube, pra-pemrosesan teks, hingga evaluasi model klasifikasi, dapat ditarik beberapa kesimpulan utama sebagai berikut:

1. Berdasarkan hasil pengujian pada data uji, model SVM mencatatkan kinerja yang cukup baik secara global dengan nilai akurasi (*Accuracy*) sebesar 86%. Hal ini menunjukkan bahwa metode SVM efektif dalam mengklasifikasikan mayoritas komentar dengan benar.
2. Analisis mendalam melalui *Confusion Matrix* dan *Classification Report* mengungkapkan adanya tantangan ketidakseimbangan kelas (*imbalanced dataset*) pada data komentar, di mana sentimen positif jauh lebih mendominasi dibandingkan sentimen negatif. Akibatnya, model sangat sensitif dalam mendeteksi sentimen positif (*Recall* 0.98), namun masih lemah dalam mengenali sentimen negatif (*Recall* 0.20). Hal ini mengindikasikan bahwa meskipun akurasi tinggi, model cenderung bias ke arah kelas mayoritas.

5.2 Saran

Penelitian ini menyadari bahwa sistem yang dibangun memiliki keterbatasan, sehingga **topik ini masih bisa dikembangkan untuk penelitian selanjutnya** guna mendapatkan hasil yang lebih optimal. Adapun beberapa saran pengembangan yang dapat dilakukan peneliti mendatang adalah:

1. **Penanganan Ketidakseimbangan Data:** Mengingat rendahnya kemampuan model dalam mendeteksi kelas minoritas (negatif), penelitian selanjutnya sangat disarankan untuk menerapkan teknik penanganan data tidak seimbang, seperti *Oversampling* (misalnya metode SMOTE) pada data minoritas atau *Undersampling* pada data mayoritas agar proporsi data latih menjadi lebih seimbang.

2. **Eksplorasi Algoritma Lanjutan:** Peneliti selanjutnya dapat mencoba membandingkan kinerja SVM dengan algoritma *Deep Learning* seperti *Long Short-Term Memory* (LSTM) atau model berbasis *Transformer* (seperti IndoBERT) yang sering kali memiliki kemampuan lebih baik dalam menangkap konteks kalimat yang kompleks.
3. **Perluasan Dataset:** Menambah jumlah dataset komentar secara signifikan atau menggabungkan sumber data dari berbagai video/kanal YouTube yang relevan dapat membantu memperkaya variasi kata dan meningkatkan kemampuan generalisasi model.
4. **Preprocessing Lanjutan:** Mengingat ragam bahasa di YouTube yang sangat informal, penerapan kamus normalisasi bahasa gaul (*slang*) yang lebih komprehensif atau teknik *Lemmatization* khusus Bahasa Indonesia dapat diterapkan untuk membersihkan data dengan lebih optimal.

