

## **BAB V PENUTUP**

### **5.1 Kesimpulan**

Berdasarkan tahapan analisis yang dilakukan dan hasil yang didapatkan, maka diperoleh beberapa kesimpulan berikut :

1. Penelitian ini menggunakan data komentar dari YouTube terkait pembangunan Ibu Kota Nusantara sebanyak 4.733 komentar awal, yang telah melalui proses filtering dan pelabelan manual. Data akhir yang dianalisis berjumlah 3.000 komentar, dengan distribusi 1.492 sentimen positif (49,7%) dan 1.508 sentimen negatif (50,3%).
2. Penelitian ini menggunakan dua metode validasi, yaitu K-Fold Cross Validation (5-fold) dan split data 80:20, untuk membandingkan hasil akurasi, precision, recall, dan F1-score pada data latih dan uji. SVM RBF mencapai akurasi 74% pada K-Fold dan 73,17% pada split data 80:20, sementara Random Forest mencatat 75% pada K-Fold dan 72,22% pada split data 80:20. Kedua metode menunjukkan keseimbangan precision, recall, dan F1-score, menggambarkan kemampuan klasifikasi sentimen yang konsisten.
3. Hasil model dibandingkan dengan metode lain seperti SVM Linear, Polynomial, Adaboost, dan Bagging. Analisis ini menunjukkan bahwa kombinasi SVM Kernel RBF dan Random Forest lebih unggul dalam menangani data teks yang kompleks dibandingkan metode pembanding lainnya.

## 5.2 Saran

Berdasarkan hasil dan kesimpulan yang telah dicapai dalam penelitian ini, peneliti memberikan saran untuk meningkatkan kualitas penelitian sebagai berikut:

1. Disarankan untuk memperluas jumlah data dengan mengumpulkan komentar dari lebih banyak video YouTube atau platform lain yang relevan. Ini akan memperkaya model dan memberikan hasil yang lebih representatif tentang opini publik terkait pembangunan Ibu Kota Nusantara.
2. Untuk meningkatkan akurasi, disarankan untuk mencoba algoritma berbasis deep learning seperti BERT atau LSTM, yang dapat menangani konteks dan bahasa informal dengan lebih baik dibandingkan model tradisional seperti *SVM* dan *Random forest*.
3. Walaupun penelitian ini sudah menggunakan teknik word normalization dan penanganan bahasa informal, pengembangan lebih lanjut pada preprocessing, seperti penerapan entity recognition atau contextual spelling correction, dapat meningkatkan kualitas data teks dan lebih mengatasi variasi kata yang kompleks, terutama dalam komentar media sosial.
4. Mencoba pendekatan lainnya, selain *SVM* dan *Random forest*, metode stacking dan boosting (seperti XGBoost) dapat dicoba untuk mengevaluasi apakah model-model ini dapat memberikan akurasi yang lebih tinggi dalam analisis sentimen.