

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian dan analisis yang telah dilakukan, dapat disimpulkan sebagai berikut:

1. Algoritma *Decision Tree* sebagai model dasar berhasil melakukan klasifikasi *spam email* dengan performa yang cukup baik, mencapai akurasi 91,24%, *precision* 91,37%, *recall* 91,24%, dan *F1-Score* 91,21%. Namun, model ini masih menunjukkan tanda-tanda *overfitting* ringan hingga sedang, terlihat dari gap antara *training accuracy* (sangat tinggi dan stabil ~0,91–0,92) dengan *validation accuracy* yang lebih rendah.
2. Penerapan metode *ensemble Adaptive Boosting* (AdaBoost) pada *Decision Tree* memberikan peningkatan performa yang sangat signifikan, mencapai akurasi 98,08%, *precision* 98,09%, *recall* 98,08%, dan *F1-Score* 98,08%. Model ini juga berhasil mengurangi *overfitting* (gap *training-validation* sangat kecil, *validation accuracy* mendekati 0,98, dan AUC-ROC tertinggi 0,9969), menunjukkan generalisasi yang baik.
3. Penerapan metode *ensemble Bootstrap Aggregating* (Bagging) pada *Decision Tree* juga meningkatkan performa dibandingkan model tunggal, dengan akurasi 91,66%, *precision* 92,01%, *recall* 91,66%, dan *F1-Score* 91,59%. Bagging berhasil mengurangi *varians* dan meningkatkan stabilitas model (*learning curve* lebih stabil daripada *Decision Tree* tunggal), tetapi peningkatannya tidak sebesar AdaBoost. Bagging lebih cocok untuk dataset teks seimbang dengan fokus pada ketahanan terhadap variasi data baru, daripada peningkatan akurasi.
4. Dari perbandingan ketiga model, *Decision Tree* + AdaBoost merupakan metode yang memberikan performa terbaik secara keseluruhan dalam hal akurasi, keseimbangan *precision-recall*, *F1-Score*, serta kemampuan meminimalkan risiko *overfitting* pada tugas klasifikasi *spam email*.
5. Secara keseluruhan, penggunaan metode *ensemble* terbukti efektif untuk

meningkatkan ketahanan dan akurasi klasifikasi *email* dibandingkan menggunakan algoritma *Decision Tree* secara tunggal.

5.2 Saran

Berdasarkan hasil penelitian dan keterbatasan yang ditemui, berikut beberapa saran untuk pengembangan penelitian selanjutnya:

1. Mencoba algoritma *base classifier* lain selain *Decision Tree* (misalnya *Random Forest*, *Gradient Boosting/XGBoost*, *LightGBM*, dll)
2. Menggunakan dataset yang lebih besar dan lebih beragam (misalnya Enron, CSDMC2010, atau kombinasi beberapa dataset publik *spam email* terbaru) serta dataset yang mengandung *email* dalam bahasa Indonesia atau multibahasa untuk menguji generalisasi model di konteks lokal Indonesia.
3. Melakukan *hyperparameter tuning* yang lebih mendalam (menggunakan *Grid Search*, *Random Search*, atau *Bayesian Optimization*) pada parameter *Decision Tree* (*max_depth*, *min_samples_split*, dll.) maupun parameter *ensemble* (jumlah estimator, learning rate pada AdaBoost) untuk mencari kombinasi optimal yang mungkin meningkatkan performa lebih lanjut.