

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dalam penelitian ini dilakukan pengujian model dengan menggunakan Algoritma klasifikasi yaitu algoritma CART, C4.5 dan C5.0 dengan menggunakan data text email spam dan ham yang di olah dengan tools *rapidminer* dengan tahapan text preprocessing. Berdasarkan hasil penelitian, dapat ditarik kesimpulan bahwa algoritma yang memberikan nilai *accuracy* tertinggi adalah algoritma CART (*Classification and Regression Tree*) dengan nilai *accuracy* 92.29% diikuti oleh algoritma C5.0 mendapatkan nilai *accuracy* sebesar 91.97% dan algoritma C4.5 nilai *accuracy* yang didapatkan adalah 91.53%. Algoritma CART adalah algoritma yang sederhana tapi bisa mendapatkan nilai *accuracy* yang paling tinggi dibandingkan dengan model yang lainnya. Dari penelitian ini juga dapat disimpulkan algoritma CART adalah algoritma yang paling tepat untuk klasifikasi email spam dan ham menggunakan data text email. Meskipun demikian, semua algoritma memiliki performa yang baik dengan *accuracy* di atas 90%, sehingga dapat dianggap kompetitif dalam proses klasifikasi.

5.2 Saran

Saran-saran dari penulis untuk penelitian selanjutnya adalah sebagai berikut:

- a. Gunakan dataset yang lebih besar dan beragam untuk meningkatkan kemampuan model.
- b. Eksperimen dengan teknik text preprocessing lain, seperti stemming atau Word2Vec.
- c. Uji algoritma machine learning yang berbeda supaya mendapatkan hasil yang lebih luas
- d. Uji model dalam aplikasi nyata untuk mengukur performa di lingkungan produksi.