

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan terdapat beberapa poin kesimpulan antara lain sebagai berikut :

- a. Penerapan seleksi fitur *Information Gain* dalam penelitian ini, mempengaruhi kinerja *Random Forest* dan *Logistic Regression* seperti *accuracy*, *precision*, *recall*, *F-measure*. Dari hasil ujicoba pada penelitian ini algoritma *Random Forest* ketika diterapkan seleksi fitur *Information Gain* mengalami peningkatan nilai akurasi dari 95.331 % menjadi 95.657 %. Sedangkan untuk algoritma *Logistic Regression* mengalami penurunan akurasi yang tidak signifikan dari 82.953 % menjadi 82.302 %.
- b. Setelah diterapkan seleksi fitur *Information Gain* algoritma *Random Forest* memiliki kinerja terbaik. Hal ini dapat dilihat dari *confusion matrix* dengan nilai 543 *True Positive (TP)*, 338 *True Negative (TN)*, 29 *False Positive (FP)*, dan 11 *False Negative (FN)*.

5.2 Saran

Berdasarkan hasil penelitian yang dilakukan peneliti, berikut adalah saran untuk penelitian lebih lanjut yang bisa dilakukan :

- a. Untuk peneliti selanjutnya dapat menambah algoritma selain *Random Forest* dan *Logistic Regression* untuk klasifikasi Email Spam.
- b. Penelitian selanjutnya bisa dengan menerapkan sistem deteksi email SPAM.

- c. Pada penelitian lebih lanjut bisa menggunakan metode seleksi fitur lainnya. Sehingga dapat dilakukan perbandingan dengan penelitian ini untuk mendapat kinerja terbaik dalam menyeleksi fitur untuk mendeteksi email SPAM.

