

BAB 5 SIMPULAN DAN SARAN

5.1 Simpulan

Implementasi algoritma *multinomial naive bayes* dan *logistic regression* untuk sentimen analisis opini yang diambil dari media sosial Twitter telah berhasil dibuat. Hasil ujicoba menunjukkan bahwa performa algoritma *logistic regression* lebih tinggi daripada performa *multinomial naive bayes* baik pada rasio perbandingan 70:30 maupun rasio perbandingan 80:20. Kedua algoritma memiliki performa yang lebih baik pada rasio perbandingan 80:20, namun algoritma *Logistic Regression* lebih unggul dengan nilai skor akurasi 0.864516129032258, presisi 0.8372911169744943, *recall* 0.99581589958159 dan *f1-score* 0.9096989966555185. Sedangkan *Multinomial Naive Bayes* mendapatkan nilai skor akurasi sebesar 0.7354838709677419, presisi 0.7225170583775588, *recall* 0.9968619246861925, dan *f1-score* sebesar 0.8378021978021979. Penerapan SMOTE memberikan peningkatan performa akurasi *Multinomial Naive Bayes* hingga 14% sedangkan *Logistic Regression* hanya meningkat hingga 5%. Namun, secara keseluruhan performa *Logistic Regression* tetap lebih baik daripada *Multinomial Naive Bayes* pada analisis sentimen penelitian ini.

5.2 Saran

Setelah selesai menjalani penelitian, terdapat beberapa saran yang dapat dipertimbangkan untuk penelitian selanjutnya, yaitu sebagai berikut.

1. Disarankan untuk menambahkan kata-kata slang terbaru pada kamus *slangword removal* yang sesuai dengan tren yang terjadi saat penelitian dilakukan sehingga kata-kata slang tersebut dapat digantikan dengan padanan kata Bahasa Indonesia yang sesuai dengan EYD V. Jika memungkinkan gunakan kamus *slangword* dengan ukuran yang besar sehingga mencakup banyak kata. Pengubahan *slangword* ke kata yang benar akan meningkatkan performa model.
2. Gunakan algoritma lain selain *multinomial naive bayes*, seperti *Support Vector Machine* atau *K-Nearest Neighbor* untuk membandingkan performanya dengan *logistic regression*.