BAB 5 SIMPULAN DAN SARAN

5.1 Simpulan

Analisis sentimen mengenai program naturalisasi pemain timnas sepak bola Indonesia di Twitter menggunakan model *Naive Bayes* menunjukkan keberhasilan implementasi model *Multinomial Naive Bayes* dan *Complement Naive Bayes* dengan bobot TF-IDF. Awalnya, data berjumlah 11.105 baris, namun setelah proses *pre-processing* dan *labelling*, data terseleksi menjadi 7.501 baris pada label VADER, sedangkan pada label *pre-trained* data terseleksi menjadi 7.086. Model *Naive Bayes* terbaik dan stabil dalam hal *precision*, *recall*, dan *F1-score* sebelum penyeimbangan data dan sudah dilakukan *hyperparameter tuning* adalah model *Complement Naive Bayes* dengan akurasi 84%. Model ini memiliki *precision* tertinggi sebesar 97% untuk data negatif dan 65% untuk data positif. *Recall* model mencapai 81% untuk data negatif dan 94% untuk data positif. Sedangkan *F1-score* model mencapai 88% untuk data negatif dan 77% untuk data positif. Setelah penyeimbangan data dan dilakukan *hyperparameter tuning*, kedua model mengalami peningkatan performa dengan meningkatnya *precision*, *recall*, dan *F1-score* dengan hasil yang cukup seimbang. Akurasi kedua model yang mencapai sebesar 90%. Model ini memiliki *precision* tertinggi sebesar 98% untuk data negatif dan 89% untuk data positif. *Recall* model mencapai 89% untuk data negatif dan 94% untuk data positif. Sedangkan *F1-score* model mencapai 93% untuk data negatif dan 84% untuk data positif.

5.2 Saran

Saran yang dapat diberikan sebagai referensi untuk penelitian selanjutnya adalah:

1. Melakukan tahap *pre-processing* lebih detail terutama pada bagian normalisasi karena banyaknya kata yang tidak baku dari *Tweet* yang telah diambil.
2. Memperbanyak data dengan menggunakan kata kunci yang berbeda namun tetap dalam konteks penelitian sehingga sentimen kata lebih bervariasi.