

BAB 5 SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan hasil penelitian dan implementasi sistem yang telah dilakukan, diperoleh kesimpulan sebagai berikut.

1. Penelitian ini berhasil menganalisis sentimen terhadap pelatih baru Tim Nasional Indonesia dengan menggunakan komentar dari Instagram sebagai objek analisis. Melalui tahapan preprocessing teks yang meliputi pembersihan data, *ase folding*, tokenisasi, normalisasi, *stopword removal*, dan *stemming*, data komentar berhasil dipersiapkan untuk dianalisis secara sistematis. Proses klasifikasi dilakukan menggunakan algoritma *Naïve Bayes* dengan dukungan metode TF-IDF untuk ekstraksi fitur, dan penyeimbangan data menggunakan SMOTE. Hasilnya menunjukkan bahwa pendekatan ini mampu mengidentifikasi opini publik terkait pelatih baru secara otomatis.
2. Model klasifikasi sentimen yang dibangun menggunakan algoritma *Multinomial Naïve Bayes* menunjukkan performa yang baik dan konsisten pada berbagai rasio pembagian data. Pada rasio 80:20, model menghasilkan akurasi sebesar 85,88%. Pengujian lanjutan dilakukan dengan dua rasio tambahan, yaitu 70:30 dengan akurasi 85,77%, di mana nilai *precision* tertinggi berada pada label netral (0,9734), dan nilai *recall* tertinggi terdapat pada label negatif (1,0000). Sementara itu, pada rasio 90:10, model memperoleh akurasi tertinggi sebesar 88,39%, serta performa sangat tinggi pada label positif dan negatif. Hasil tersebut menunjukkan bahwa algoritma *Naïve Bayes* mampu mempertahankan stabilitas dan akurasi klasifikasi dalam berbagai skenario pembagian data.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran yang dapat diberikan untuk pengembangan dan penelitian selanjutnya adalah sebagai berikut.

1. Pengumpulan data komentar dilakukan secara manual, sehingga jumlah data masih terbatas. Disarankan agar pengambilan data dilakukan secara otomatis

menggunakan API atau teknik web scraping agar jumlah data lebih besar dan variatif, serta lebih representatif terhadap opini publik.

2. Proses pelabelan data dalam penelitian ini menggunakan metode otomatis berbasis *TextBlob*, yang meskipun praktis, memiliki keterbatasan dalam akurasi pada teks berbahasa Indonesia. Penelitian selanjutnya disarankan untuk melakukan pelabelan secara manual atau semi-otomatis dengan melibatkan anotator manusia agar hasil pelatihan model lebih optimal.
3. Algoritma *Naïve Bayes* memang sederhana dan cepat, namun memiliki keterbatasan dalam menangani data yang kompleks. Oleh karena itu, disarankan untuk mengeksplorasi metode lain seperti *Support Vector Machine* (SVM), *Random Forest*, atau bahkan pendekatan *deep learning* (seperti LSTM atau BERT) untuk memperoleh performa klasifikasi yang lebih baik.
4. Penelitian ini hanya fokus pada sentimen dalam tiga kategori: positif, negatif, dan netral. Penelitian selanjutnya dapat memperluas cakupan analisis, misalnya dengan klasifikasi berbasis emosi (seperti marah, senang, kecewa), atau dengan analisis temporal untuk melihat bagaimana opini publik berubah dari waktu ke waktu.

