

BAB 5

SIMPULAN DAN SARAN

5.1 Simpulan

Berdasarkan hasil penelitian yang dilakukan mengenai analisis sentimen terhadap *rebranding* twitter menjadi X dengan pendekatan *machine learning* dapat disimpulkan bahwa:

- Pengumpulan data *tweet* berhasil dilakukan dengan memanfaatkan *library tweet-harvest* menggunakan beberapa kata kunci. Data yang berhasil diperoleh sebanyak 4375 data dengan distribusi kelasnya setelah dilakukan *labeling* adalah 2147 data netral, 1532 data positif, dan 696 data negatif.
- Proses analisis sentimen dilakukan dengan mengimplementasikan algoritma Naive Bayes dan Support Vector Machine (SVM) sebagai *base learner*. Seluruh data diproses melalui tahapan *data pre-processing*. Data yang telah diproses akan diberi label menggunakan TextBlob. Kemudian data akan dibagi menjadi data latih dan data uji dengan rasio 90:10. Setelah itu data diubah menjadi representasi vektor menggunakan metode ekstraksi yang berbeda, yaitu CountVectorizer dan TF-IDF. Kemudian model akan dilatih terlebih dahulu dan diuji untuk melakukan prediksi. Model akan dievaluasi menggunakan *confusion matrix* dan nilai dari *classification report* yang meliputi nilai *accuracy*, *precision*, *recall*, dan *f1-score*.
- Hasil dari penelitian yang dilakukan ini menghasilkan model terbaik pada model SVM dengan TF-IDF yang dilakukan *hyperparameter tuning*. Model ini menghasilkan nilai akurasi tertinggi dibandingkan model dengan algoritma dan *feature extraction* lainnya. Model ini memperoleh tingkat akurasi sebesar 92%. Model ini mendapatkan nilai *precision* untuk kelas negatif, netral, dan positif secara berturut-turut sebesar 89%, 94%, dan 91%. Kemudian diperoleh juga nilai *recall* sebesar 77%, 98%, dan 91%. Nilai *recall* yang tinggi pada kelas netral mengindikasikan bahwa sentimen *rebranding* yang dilakukan cenderung ke arah netral.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran untuk peneliti selanjutnya yang hendak melakukan penelitian serupa.

- Menerapkan sistem yang telah dibangun dengan Web atau Aplikasi untuk melakukan analisis sentimen secara *real-time*.
- Menerapkan metode *feature selection* untuk memilih fitur yang relevan sehingga dapat meningkatkan performa model.
- Melibatkan pakar bahasa untuk melakukan *labeling* manual sebagai upaya validasi terhadap hasil pelabelan otomatis. Hasil dari *labeling* yang dilakukan oleh pakar bahasa dan otomatis dapat dibandingkan keakuratan dan konsistensi dari teknik *labeling* yang digunakan.
- Menggunakan algoritma lain untuk melakukan analisis sentimen dan melakukan perbandingan dengan algoritma yang sudah digunakan pada penelitian ini.

