

BAB V

PENUTUP

5.1 Kesimpulan

1. Penelitian ini membuktikan bahwa kombinasi algoritma Naive Bayes dan teknik undersampling efektif dalam menangani ketidakseimbangan kelas pada analisis sentimen Twitter terkait Pilpres Indonesia 2024. Teknik undersampling berhasil menyeimbangkan distribusi kelas, memungkinkan model mempelajari pola sentimen secara lebih adil.
2. Naive Bayes menunjukkan performa yang baik dengan akurasi 90% dan rata-rata weighted F1-score 90%. Model mampu mendeteksi sentimen positif dengan baik, meskipun performa pada kelas netral sedikit lebih rendah karena sifat teks yang ambigu. Validasi dengan K-Fold Cross Validation mengonfirmasi konsistensi model dalam skenario serupa.

Kesimpulan ini menunjukkan bahwa pendekatan yang digunakan dapat menjadi solusi efektif untuk analisis sentimen pada data sosial media dengan ketidakseimbangan kelas yang signifikan.

5.2 Saran

Untuk pengembangan lebih lanjut, disarankan:

1. Peningkatan Kualitas Data – Menggunakan dataset lebih besar dan teknik data augmentation untuk mengurangi bias.
2. Eksperimen dengan Algoritma Lain – Mencoba algoritma seperti Logistic Regression, Random Forest, LSTM, atau BERT untuk perbandingan performa.
3. Optimisasi Fitur – Mengeksplorasi n-grams, word embeddings, atau tf-idf weighting guna meningkatkan akurasi model.
4. Implementasi Real-Time – Mengintegrasikan model dengan API Twitter untuk analisis sentimen secara langsung.
5. Teknik Ensemble – Menggunakan kombinasi algoritma seperti voting atau stacking untuk meningkatkan ketahanan model.

Dengan pengembangan lebih lanjut, penelitian ini dapat memberikan kontribusi signifikan dalam analisis sentimen, baik untuk penelitian akademik maupun implementasi nyata.