

BAB V PENUTUP

5.1 Kesimpulan

Dalam penelitian ini, analisis sentimen untuk film 'Gadis Kretek' di Twitter melibatkan proses pengumpulan data, pre-processing, dan penerapan algoritma K-Nearest Neighbors (KNN). Data dikumpulkan dari Twitter dengan total 13.180 tweet, dan proses pre-processing memastikan data bersih dan konsisten. Data komentar yang sudah dipisahkan berdasarkan bahasa menunjukkan distribusi 2.227 komentar negatif dan 2.105 komentar positif dalam bahasa Indonesia, yang telah diproses lebih lanjut untuk analisis sentimen.

Model KNN diterapkan pada data yang telah diolah dengan menggunakan TF-IDF untuk transformasi teks menjadi fitur numerik. Proses ini memberikan bobot pada kata-kata dalam dokumen berdasarkan frekuensi dan kepentingannya dalam konteks tertentu. Evaluasi model dilakukan dengan menggunakan teknik cross-validation untuk menentukan nilai K yang optimal. Hasil cross-validation menunjukkan bahwa nilai K terbaik adalah 48, dengan akurasi mencapai 69,58%.

Analisis lebih lanjut menunjukkan bahwa ukuran data uji yang lebih kecil (10%) memberikan hasil yang lebih baik dalam hal recall untuk kelas positif, dengan akurasi keseluruhan yang meningkat menjadi 71,20%. Precision untuk kelas negatif tetap pada 0,78, tetapi recall untuk kelas positif meningkat menjadi 0,86. Hasil ini menunjukkan bahwa model KNN lebih efektif dalam mendeteksi komentar positif ketika ukuran data uji lebih kecil.

Secara keseluruhan, penelitian ini menunjukkan bahwa model KNN dapat dioptimalkan dengan memilih ukuran data uji yang sesuai dan melakukan fine-tuning parameter model untuk meningkatkan akurasi dan efektivitas dalam klasifikasi komentar. Dengan pendekatan yang seimbang antara ukuran data pelatihan dan uji, serta penerapan pre-processing yang menyeluruh, analisis sentimen dapat dilakukan dengan lebih akurat dan relevan dalam konteks aplikasi praktis.

5.2 Saran

Berdasarkan hasil penelitian ini, beberapa saran untuk penelitian selanjutnya yang dapat meningkatkan kualitas dan kedalaman analisis adalah sebagai berikut:

Eksplorasi Algoritma Lain:

- 1) Model Machine Learning Lain: Selain KNN, eksperimen dengan algoritma lain seperti *Support Vector Machines* (SVM), Random Forest, atau XGBoost dapat memberikan perbandingan kinerja yang lebih luas.
- 2) Model Deep Learning: Menggunakan model deep learning seperti LSTM (*Long Short-Term Memory*) atau Transformer-based models seperti BERT untuk analisis sentimen bisa menghasilkan hasil yang lebih akurat dengan memahami konteks yang lebih dalam dalam teks.

Pengolahan Data Lebih Lanjut:

- 1) Penggabungan Data Multisumber: Gabungkan data dari berbagai platform media sosial (seperti Instagram, Facebook) untuk mendapatkan gambaran yang lebih komprehensif tentang sentimen publik.
- 2) Ekspansi Kamus Normalisasi: Menambahkan lebih banyak variasi bahasa gaul dan jargon yang mungkin relevan dengan komentar terkait film untuk meningkatkan akurasi normalisasi.

Pengembangan dan Penerapan Model:

- 1) Implementasi Model di Aplikasi Nyata: Pertimbangkan untuk mengimplementasikan model dalam aplikasi nyata seperti analisis sentimen otomatis di platform media sosial untuk feedback langsung.

Dengan menerapkan saran-saran ini, penelitian selanjutnya dapat mencapai hasil yang lebih mendalam dan akurat, memberikan wawasan yang lebih kaya dan aplikatif dalam analisis sentimen, serta meningkatkan relevansi model yang digunakan.