

## BAB V PENUTUP

### 5.1 Kesimpulan

Pada penelitian ini menganalisis sentimen terhadap film “Agak Laen” di *Twitter* melalui beberapa tahap, yaitu pengumpulan data, pemrosesan awal (*pre-processing*), dan penerapan algoritma *K-Nearest Neighbors* (KNN). Data yang dianalisis terdiri dari 4378 *tweet* yang dikumpulkan dari *Twitter*, dengan tahap *pre-processing* dilakukan untuk memastikan data dalam kondisi bersih dan konsisten.

Penerapan model KNN diterapkan pada data yang telah diolah dengan labeling lexicon menggunakan library *Vader Sentiment*, data diberi *compound scores* berdasarkan *polarity* nya, proses ini digunakan untuk menentukan sentimen masing-masing data. Sentimen tersebut diklasifikasikan ke dalam kategori positif atau negatif, yang dimana kategori ditetapkan berdasarkan ambang batas tertentu dari nilai *compound* dengan hasil sentimen terhadap film ini cenderung ke arah positif, dengan jumlah data 3484 data positif dan 753 data negatif, dilanjutkan dengan penerapan *cross-validation* untuk evaluasi performa model, hasil analisis menunjukkan bahwa nilai *k* terbaik adalah 1, dengan rata-rata akurasi yang mencapai 81,96%.

Hasil analisis keseluruhan yang menguji performa model KNN dengan berbagai rasio pembagian data 70:30, 80:20, dan 90:10, menunjukkan bahwa penggunaan ukuran data yang lebih kecil, yaitu 90:10 menghasilkan peningkatan performa model. Akurasi keseluruhan meningkat hingga 86,65%, *precision*, *recall*, dan *f1-score* juga mengalami peningkatan dibandingkan data uji lainnya, dimana ini menunjukkan bahwa model lebih optimal dalam mengidentifikasi komentar ketika proporsi data uji dikurangi.

## 5.2 Saran

Berdasarkan hasil dalam penelitian ini, beberapa rekomendasi untuk penelitian selanjutnya yang dapat meningkatkan kualitas dan cakupan analisis adalah sebagai berikut:

1. Mengoptimasi model menggunakan algoritma *Particle Swarm Optimization* (PSO) dalam model KNN agar meningkatkan akurasi dan efisiensi model dalam mengklasifikasikan data sentimen.
2. Menggunakan model *machine learning* lain selain KNN. Eksperimen dengan algoritma seperti *Support Vector Machines* (SVM), *Random Forest*, atau *Naïve Bayes*.
3. Menambahkan platform lain pada pengambilan sumber data, seperti *YouTube* atau *Instagram*. Hal ini dapat memberikan cakupan analisis yang lebih luas dan mendalam, mengingat karakteristik pengguna serta pola interaksi di masing-masing platform berbeda.