

BAB V PENUTUP

5.1 Kesimpulan

Penelitian ini telah mengevaluasi beberapa pendekatan analisis sentimen terhadap komentar video YouTube, yaitu menggunakan metode supervised learning dan pendekatan leksikon. Dari hasil pengujian, model *Linear Support Vector Classification* (*Linear SVC*) memberikan performa terbaik dalam hal akurasi dengan *F1-Score* sebesar 0.7063, sementara *Multinomial Naïve Bayes* (MNB) mencatat *F1-Score* sebesar 0.6555 dengan keunggulan utama pada efisiensi dan kecepatan prediksi. Di sisi lain, metode leksikon seperti VADER dan TextBlob menghasilkan *F1-Score* yang lebih rendah, masing-masing 0.4945 dan 0.4445 dan relatif lambat, namun memiliki kelebihan dalam hal kemampuan menganalisis berbagai bahasa tanpa perlu proses pelatihan model terlebih dahulu. Hal ini menjadikan metode leksikon relevan dalam konteks tertentu, terutama jika diperlukan fleksibilitas bahasa dan instalasi yang ringan.

Pengembangan sistem dilakukan melalui pendekatan prototyping yang terdiri dari dua iterasi besar. Sistem *backend* dirancang menggunakan *framework Flask* yang menangani proses *crawling* komentar, *preprocessing*, klasifikasi sentimen, serta pengolahan hasil. Pada sisi *frontend*, antarmuka dirancang agar sederhana dan informatif, menampilkan visualisasi persentase sentimen dan daftar komentar terklasifikasi. Hasil dari prototipe pertama kemudian disempurnakan berdasarkan umpan balik pengguna, yang berujung pada peningkatan tampilan, penambahan fitur, serta perbaikan kecepatan respons sistem pada prototipe kedua. Sistem juga telah *di-deploy* dan diuji secara lokal maupun melalui server online.

Secara keseluruhan, penelitian ini berhasil menjawab seluruh tujuan yang dirumuskan pada Bab I. Sistem berhasil dikembangkan dan diimplementasikan secara fungsional dalam bentuk aplikasi web. Sistem ini mampu menyajikan ringkasan sentimen dari komentar YouTube secara otomatis, memudahkan pengguna dalam mengevaluasi kualitas video tanpa perlu membaca komentar satu per satu. Selain itu, penelitian ini juga berhasil membandingkan performa berbagai metode sentimen dan menentukan pendekatan yang paling tepat berdasarkan metrik *F1-Score* dan efisiensi penggunaan. Hasil dan sistem yang dikembangkan diharapkan dapat menjadi kontribusi nyata dalam pengembangan teknologi analisis sentimen berbasis web yang praktis dan responsif.

5.2 Saran

Penelitian ini masih memiliki ruang pengembangan yang cukup luas untuk ditindaklanjuti pada penelitian selanjutnya. Pertama, disarankan untuk membandingkan lebih banyak model klasifikasi, seperti model *deep learning* seperti LSTM atau BERT, guna melihat sejauh mana peningkatan akurasi dan pemahaman konteks dapat dicapai dalam klasifikasi komentar yang lebih kompleks. Kedua, iterasi pengembangan sistem dapat terus dilanjutkan dengan melibatkan lebih banyak responden dari latar belakang pengguna yang beragam, agar desain sistem dan fitur yang dikembangkan semakin relevan dan sesuai kebutuhan nyata di lapangan. Selain itu, penting pula untuk menambahkan fitur deteksi komentar spam atau bot yang sering muncul di YouTube, karena komentar semacam ini dapat memengaruhi hasil analisis sentimen dan mengurangi keakuratan sistem dalam menilai opini asli pengguna. Pengembangan modul ini bisa menjadi fokus utama dalam penelitian lanjutan, baik melalui pendekatan berbasis pola teks maupun teknik deteksi akun otomatis. Terakhir, perlu dipertimbangkan untuk memperluas dukungan bahasa dalam sistem agar tidak hanya terbatas pada komentar berbahasa Inggris, melainkan juga mencakup bahasa Indonesia dan lainnya dengan memanfaatkan model multibahasa.