

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai klasifikasi sentimen dan emosi pada komentar YouTube menggunakan model IndoBERT, maka diperoleh beberapa kesimpulan sebagai berikut:

1. Berdasarkan hasil penelitian, model IndoBERT mampu mengklasifikasikan sentimen dan emosi pada komentar YouTube yang membahas RUU Perampasan Aset. Pada tugas klasifikasi sentimen, performa terbaik diperoleh oleh varian *indoBenchmark/indoBERT-base-p2* dengan rasio pembagian data 80:20, yang mencapai nilai *accuracy* sebesar 0,9249 dan *weighted average F1-score* sebesar 0,9253. Varian yang sama juga memberikan hasil terbaik pada klasifikasi emosi dengan *accuracy* sebesar 0,8503 dan *weighted average F1-score* sebesar 0,8504. Selain itu, hasil analisis memperlihatkan bahwa sentimen negatif merupakan kategori yang paling banyak ditemukan pada komentar yang dianalisis, sedangkan pada klasifikasi emosi kategori marah menjadi emosi yang paling banyak ditemukan. Selain itu, rasio pembagian data 80:20 secara umum memberikan hasil yang lebih optimal dibandingkan rasio 70:30 karena jumlah data latih yang lebih besar membantu Model mampu mempelajari pola pada data dengan lebih baik.
2. Model terbaik hasil penelitian kemudian berhasil diimplementasikan ke dalam aplikasi berbasis web menggunakan *framework Flask*. Sistem yang dibangun mampu menerima input berupa teks secara langsung maupun file berformat CSV atau Excel untuk dianalisis secara otomatis. Hasil analisis yang ditampilkan meliputi label sentimen, label emosi. Selain itu, aplikasi juga dilengkapi dengan berbagai visualisasi data seperti *pie chart*, *bar chart*, tren analisis, dan *wordcloud* berdasarkan kategori sentimen maupun emosi. Implementasi tersebut membuktikan bahwa model yang telah dikembangkan dapat diintegrasikan ke dalam sebuah aplikasi untuk melakukan analisis sentimen dan emosi pada data teks secara otomatis. Selain itu, hasil analisis juga disajikan dalam bentuk visualisasi.

5.2 Saran

Berdasarkan hasil penelitian dan berbagai keterbatasan yang ditemukan selama proses penelitian, beberapa saran untuk penelitian selanjutnya dapat disampaikan sebagai berikut:

1. Penelitian selanjutnya diharapkan dapat memanfaatkan dataset dengan jumlah yang lebih besar serta mengombinasikan data dari berbagai platform media sosial, seperti X, Instagram, Facebook, dan TikTok. Dengan demikian, analisis yang dihasilkan diharapkan mampu memberikan gambaran opini publik yang lebih luas, beragam, dan representatif..
2. Penelitian ini masih menghadapi ketidakseimbangan distribusi data, terutama pada kategori emosi cinta dan sedih. Oleh sebab itu, penelitian berikutnya disarankan menerapkan teknik penyeimbangan data, seperti *Synthetic Minority Over-sampling Technique (SMOTE)*, *class weighting*, maupun augmentasi data. Penerapan metode tersebut diharapkan dapat meningkatkan kemampuan model dalam mengklasifikasikan kelas minoritas secara lebih optimal.
3. Penelitian selanjutnya disarankan untuk melakukan penyaringan atau deteksi terhadap komentar yang bersifat bot dan spam sebelum proses analisis. Hal ini bertujuan untuk meminimalkan keberadaan komentar yang tidak autentik atau mengandung upaya pembentukan opini tertentu, sehingga kualitas dataset dan hasil klasifikasi sentimen maupun emosi menjadi lebih representatif.
4. Penelitian selanjutnya dapat mempertimbangkan eksplorasi *hyperparameter tuning* yang lebih beragam, seperti penyesuaian *learning rate*, *batch size*, jumlah *epoch*, maupun penggunaan *optimizer* yang berbeda. Langkah tersebut diharapkan dapat membantu menemukan konfigurasi model yang lebih optimal sehingga performa klasifikasi sentimen dan emosi berpotensi menjadi lebih baik.