

BAB V

PENUTUP

5.1 Kesimpulan

Hasil penelitian ini adalah sistem analisis sentimen yang mampu membaca opini masyarakat tentang Universitas Dharma Persada di Instagram dan secara otomatis mengelompokkan komentar ke dalam tiga kategori: positif, netral, dan negatif. Sistem ini menggunakan model Transformer IndoBERT yang terbukti lebih akurat untuk bahasa Indonesia. Prosesnya meliputi pengumpulan komentar, pembersihan teks, pelatihan model, hingga penyajian hasil melalui tabel, diagram pie, dan word cloud. Dari total 1599 data, sistem berhasil mengklasifikasikan 1493 komentar dengan akurasi sebesar 93,37%, di mana sebagian besar komentar bersifat netral.

Penelitian ini memiliki kelebihan karena mampu menghasilkan sistem analisis sentimen dengan akurasi tinggi, klasifikasi yang jelas, dan visualisasi yang memudahkan pemahaman, sehingga bermanfaat bagi pihak kampus dalam mengevaluasi layanan secara real-time. Namun, penelitian ini juga memiliki keterbatasan, seperti banyaknya komentar yang menggunakan bahasa tidak baku atau emoji sehingga menyulitkan proses analisis, tahapan pembersihan teks yang cukup rumit, serta kendala teknis saat pelatihan model di Google Colab.

5.2 Saran

Selama proses penelitian ini, saya juga nemuin beberapa kendala. Salah satunya waktu proses scraping data komentar dari Instagram, ternyata nggak semua data bisa langsung diproses karena banyak komentar yang pakai bahasa gaul, singkatan, atau emoji yang bikin hasil analisis jadi kurang akurat. Selain itu, proses preprocessing juga agak ribet karena harus bener-bener bersihin data satu per satu supaya hasil klasifikasinya maksimal.

Kelemahan lain yang saya alami itu saat training model IndoBERT di Google Colab, terkadang koneksi atau runtime-nya suka putus di tengah jalan, jadi harus ulang proses dari awal. Selain itu, karena laptop yang saya pakai spesifikasinya standar, saya nggak bisa training langsung di lokal, harus pakai cloud service kayak Colab.

Solusi yang saya ambil, untuk data komentar yang terlalu singkat atau cuma emoji, saya buang dulu waktu proses cleaning. Terus, supaya proses training di Colab aman, saya bagi training jadi beberapa tahap, sambil sering save checkpoint-nya biar kalau runtime putus nggak mulai dari awal lagi. Terus buat kasus bahasa gaul, saya tambahkan daftar stopword sama kamus istilah gaul manual biar preprocessing-nya lebih rapi.

Intinya, meskipun ada beberapa hambatan, semua bisa diatasi pelan-pelan dengan ngulik solusi yang sesuai sama kondisi di lapangan. Hasilnya sistem bisa jalan lancar dan hasil akurasi model IndoBERT juga tetap bagus di atas 93%