

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, didapatkan beberapa kesimpulan sebagai berikut:

1. Model klasifikasi emosi berhasil dibangun secara optimal menggunakan arsitektur mBERT, dimana performa terbaik diperoleh pada model dengan skenario pra-pemrosesan normalisasi tanpa fitur *treatment* setelah melalui proses *hyperparameter tuning*. Model terbaik ini mencatatkan nilai *Accuracy* sebesar 0.72 dan *Macro F1-Score* sebesar 0.59. Berdasarkan evaluasi menggunakan *confusion matrix*, konfigurasi model hasil *tuning* mengindikasikan model lebih mampu mengenali kelas minoritas dibandingkan konfigurasi *baseline*. Selanjutnya, model BERTopic berhasil melakukan pemodelan topik permasalahan dengan parameter paling ideal pada *minimum cluster size* sebesar 17, menghasilkan nilai *Coherence* sebesar 0.6175 dan *Stability* sebesar 0.6253.
2. Analisis distribusi teks ulasan menunjukkan adanya ketimpangan kelas (*class imbalance*) yang signifikan, dimana persepsi pengguna sangat didominasi oleh ekspresi emosi negatif, khususnya *sadness* dan *anger*. Melalui perhitungan *cosine similarity*, sebaran topik tersebut berhasil diklasifikasikan dan digabungkan secara manual ke dalam aspek-aspek keluhan dominan seperti *App Performance*, *Pricing*, *Billing Deduction*, dan *Loyalty Program*. Hal ini membuktikan bahwa variasi bahasa pengguna yang penuh dengan *code-mixing* umumnya difokuskan untuk menyampaikan keluhan teknis layanan dibandingkan kepuasan layanan.
3. Hasil klasifikasi emosi dan pemodelan topik telah berhasil disajikan secara informatif melalui pengembangan aplikasi *web* menggunakan *framework* Flask. Aplikasi tersebut mampu menyajikan visualisasi data, seperti grafik distribusi emosi dan pemetaan topik, sekaligus menyediakan fitur prediksi langsung berbasis teks maupun *file* yang

dapat digunakan oleh pihak *stakeholder* untuk memantau performa layanan dengan lebih mudah.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa aspek yang dapat dikembangkan lebih lanjut untuk meningkatkan performa sistem klasifikasi emosi dan pemodelan topik, diantaranya:

1. Menambahkan lebih banyak variasi data ulasan pada kelas emosi minoritas (seperti *fear*, *surprise*, dan *neutral*). Dengan begitu, model bisa lebih baik dan seimbang dalam memahami konteks kalimat untuk berbagai macam emosi, sehingga dapat mengurangi bias prediksi yang saat ini cenderung mengarah pada kelas mayoritas.
2. Metode penanganan *class imbalance* dalam penelitian ini menggunakan teknik *resampling* konvensional seperti *Random Oversampling* dan SMOTE, namun metode tersebut justru terbukti menurunkan akurasi dan memicu tumpang tindih batas semantik antarkelas. Sebagai alternatif, untuk memperbanyak sampel emosi minoritas secara natural, teknik augmentasi seperti *back-translation* atau augmentasi *Large Language Model* dapat diterapkan agar model terhindar dari tingginya FP.
3. Teknik pengelompokan topik dalam penelitian ini masih memerlukan tahapan penggabungan (*merging*) kluster secara manual berdasarkan metrik *cosine similarity*. Untuk meningkatkan akurasi model terhadap aspek layanan secara otomatis, metode *Guided BERTopic* atau *Zero-Shot Topic Modelling* dapat diterapkan agar topik permasalahan lebih presisi sejak awal klusterisasi.
4. Penelitian ini telah menerapkan model mBERT dan teknik *Hyperparameter Tuning*, namun eksplorasi metode lain masih terbuka untuk pengembangan lebih lanjut. Metode *deep learning* yang menggabungkan *embeddings* dari model *transformer* dengan arsitektur sekuensial seperti Bi-LSTM atau CNN dapat dipertimbangkan untuk meningkatkan kemampuan model dalam menangkap aturan konteks dan fitur lokal pada kalimat ulasan yang kompleks.