

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, maka dapat ditarik beberapa kesimpulan sebagai berikut:

1. Identifikasi aspek utama dalam ulasan pengguna dilakukan menggunakan metode Latent Dirichlet Allocation (LDA) dan BERTopic. Berdasarkan evaluasi menggunakan *coherence score* (c_v), BERTopic menunjukkan kinerja yang lebih unggul dibandingkan LDA pada sebagian besar variasi jumlah topik. Pada jumlah topik 2 dan 3, LDA masih memperoleh nilai *coherence* yang lebih tinggi, masing-masing sebesar 0,5543 dan 0,5190, dibandingkan BERTopic sebesar 0,3754 dan 0,4900. Namun, mulai dari jumlah topik 4 hingga 10, BERTopic secara konsisten menghasilkan nilai *coherence* yang lebih tinggi dibandingkan LDA dengan selisih yang signifikan. Nilai *coherence* tertinggi diperoleh BERTopic pada jumlah 4 topik sebesar 0,6707, jauh lebih tinggi dibandingkan LDA pada jumlah topik yang sama yaitu 0,4495 (selisih 0,2211).
2. Algoritma Support Vector Machine (SVM) dan Random Forest mampu melakukan klasifikasi sentimen multilabel berbasis aspek dengan baik. Random Forest menunjukkan peningkatan performa yang signifikan setelah penerapan teknik SMOTE-Tomek Links, terutama dalam meningkatkan *recall* pada kelas minoritas *Service Quality Negative*. Sebaliknya, SVM cenderung lebih stabil pada data tanpa *balancing*, tetapi tidak menunjukkan peningkatan signifikan setelah penyeimbangan data. Hal ini menunjukkan bahwa Random Forest lebih adaptif terhadap distribusi data sintetis hasil *oversampling* dibandingkan SVM.
3. Kombinasi metode terbaik dalam penelitian ini adalah LDA + Random Forest dengan penerapan SMOTE-Tomek Links pada rasio pembagian data 80:20. Konfigurasi tersebut menghasilkan nilai *Macro F1-Score* sebesar 0.769 dan AUC sebesar 0.852, serta menunjukkan keseimbangan yang baik antara *precision* dan *recall*. Selain itu, model ini mampu meningkatkan kemampuan deteksi pada kelas minoritas tanpa menurunkan *precision* secara drastis,

sehingga memiliki performa yang stabil dan seimbang secara keseluruhan. Meskipun nilai *coherence score* tertinggi diperoleh oleh *topic modeling* BERTopic, tingkat koherensi topik yang tinggi tidak secara langsung menjamin performa klasifikasi yang lebih baik. Hal ini terlihat dari hasil pengujian model yang menunjukkan bahwa kombinasi berbasis LDA justru menghasilkan performa klasifikasi yang lebih optimal.

4. Sistem analisis sentimen berbasis website yang dibangun telah berhasil mengimplementasikan model terbaik secara fungsional dan mampu menghasilkan prediksi sentimen berbasis aspek secara otomatis. Sistem ini memungkinkan pengguna untuk melakukan prediksi baik melalui input tunggal berupa teks maupun input massal (*batch*) dalam bentuk *file* CSV. Selain menghasilkan prediksi aspek dan sentimen, sistem juga menyajikan hasil analisis dalam bentuk visualisasi interaktif, seperti diagram batang (*bar chart*) untuk distribusi sentimen pada setiap aspek, diagram lingkaran (*pie chart*) untuk proporsi sentimen secara keseluruhan, serta diagram garis (*line chart*) untuk menampilkan tren sentimen tiap aspek berdasarkan tanggal. Penyajian visual tersebut memberikan *insight* yang lebih detail dan informatif dalam memahami pola dan dinamika opini pelanggan.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, beberapa saran yang dapat diberikan untuk pengembangan penelitian selanjutnya maupun peningkatan sistem adalah sebagai berikut:

1. Algoritma LDA dan BERTopic telah digunakan dalam penelitian ini untuk pemodelan topik. Penelitian selanjutnya disarankan untuk mengeksplorasi metode *topic modeling* lainnya atau melakukan optimasi parameter yang lebih mendalam pada kedua algoritma tersebut guna memperoleh representasi topik yang lebih optimal. Selain itu, pendekatan berbasis transformer yang lebih mutakhir dapat dipertimbangkan untuk dianalisis pengaruhnya terhadap kualitas topik maupun performa klasifikasi.
2. Meskipun SMOTE-Tomek Links terbukti mampu meningkatkan performa model, khususnya pada kelas minoritas, penelitian selanjutnya dapat membandingkannya dengan teknik *balancing* lain seperti ADASYN,

Borderline-SMOTE, atau pendekatan *cost-sensitive learning* untuk memperoleh strategi yang lebih adaptif terhadap karakteristik data multilabel.

3. Pembangunan model klasifikasi dalam penelitian ini menggunakan algoritma SVM dan Random Forest. Penelitian selanjutnya dapat mempertimbangkan penerapan algoritma berbasis *deep learning* atau *pre-trained language models* seperti BERT untuk meningkatkan kemampuan pemahaman konteks, terutama dalam menangani negasi, ambiguitas kalimat, serta opini yang bersifat implisit.
4. Dataset penelitian dapat diperluas dengan periode waktu yang lebih panjang atau sumber data tambahan seperti media sosial lainnya untuk meningkatkan kemampuan generalisasi model. Selain itu, jumlah aspek analisis dapat diperluas agar sistem mampu menangkap dimensi sentimen yang lebih beragam.
5. Sistem berbasis *website* dapat dikembangkan lebih lanjut dengan penambahan fitur seperti dashboard interaktif *real-time*, fitur *filtering* berdasarkan periode waktu atau kategori aspek tertentu, serta integrasi API untuk otomatisasi pengambilan data ulasan. Hal ini akan meningkatkan kegunaan sistem sebagai alat monitoring sentimen yang lebih komprehensif dan aplikatif.