

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, diperoleh beberapa kesimpulan sebagai berikut:

1. Identifikasi aspek utama pengguna dalam aplikasi OVO dilakukan menggunakan metode *Latent Dirichlet Allocation*. Aspek-aspek yang berhasil diidentifikasi adalah *Transaction Efficiency*, *User Experience*, *Account Access and Registration*, serta *Balance and Charges*. Namun, pada masing-masing aspek yang ditemukan, sebagian besar ulasan dari pengguna lebih banyak berisi keluhan atau persepsi negatif dibandingkan persepsi positif terhadap aplikasi OVO.
2. Evaluasi terhadap hasil pelabelan aspek otomatis menggunakan LDA menunjukkan bahwa metode ini mempunyai keterbatasan dalam memahami konteks teks ulasan. Hasil perbandingan antara pelabelan otomatis berdasarkan *keyword* LDA dengan pelabelan manual anotator hanya menunjukkan kesesuaian sebesar 11.46%. Setelah dilakukan perbaikan *keyword* pada LDA, tingkat kesesuaian meningkat menjadi 40.60%. Meskipun terjadi peningkatan, nilai tersebut masih menunjukkan bahwa metode LDA kurang akurat untuk menangkap makna konteks dari kalimat pengguna jika dijadikan sebagai hasil label final aspek, sehingga pelabelan manual oleh anotator diperlukan dan digunakan sebagai label final aspek.
3. Pembangunan model klasifikasi menggunakan algoritma *Support Vector Machine* dengan kernel *Radial Basis Function* (RBF) menghasilkan model yang mampu melakukan klasifikasi dengan cukup baik. Model terbaik yang dihasilkan mencapai nilai *macro average F1-Score* sebesar 0.715 dan *hamming loss* 9.90% pada data pemodelan klasifikasi, serta *F1-Score* meningkat menjadi 0.794 dan *hamming loss* sebesar 0.102 pada data validasi sistem, menunjukkan bahwa model memiliki tingkat kesalahan prediksi yang rendah serta generalisasi yang kuat terhadap data baru. Selain itu, model terbaik diperoleh tanpa menggunakan metode *oversampling*.

Penerapan metode *oversampling* menggunakan ML-SMOTE pada berbagai skenario pengujian tidak menunjukkan peningkatan performa yang signifikan terhadap akurasi model.

4. Implementasi sistem klasifikasi berbasis web memungkinkan pengguna untuk memprediksi aspek dan sentimen ulasan secara langsung melalui input teks maupun *file CSV*. Sistem ini tidak hanya mampu menghasilkan prediksi aspek dan sentimen, tetapi juga menyajikan hasil analisis dalam bentuk tabel, *bar chart*, dan *pie chart*, sehingga pengguna dapat memahami distribusi aspek dan sentimen dengan lebih informatif dan interaktif.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat dipertimbangkan untuk pengembangan penelitian selanjutnya:

1. Penelitian dengan topik serupa dapat menggunakan data dengan jumlah yang lebih besar dan lebih beragam, misalnya dengan mengambil ulasan dari berbagai platform media sosial, forum diskusi, atau aplikasi dompet digital lain. Dengan variasi sumber data yang lebih luas, model klasifikasi yang dibangun diharapkan dapat memiliki cakupan konteks yang lebih baik serta meningkatkan kemampuan generalisasi terhadap berbagai jenis ekspresi pengguna.
2. Representasi teks pada penelitian ini hanya menggunakan metode TF-IDF. Penelitian selanjutnya dapat mempertimbangkan penggunaan metode *word embedding* yang lebih interpretatif, seperti Word2Vec, GloVe, atau BERT *embeddings*, untuk menangkap hubungan makna antar kata secara lebih mendalam dan memperkaya informasi fitur yang digunakan dalam klasifikasi.
3. Algoritma *Support Vector Machine* telah digunakan dalam penelitian ini untuk membangun model klasifikasi. Untuk penelitian berikutnya, disarankan mengeksplorasi algoritma lain yang mampu menangani klasifikasi multi-label dan ketidakseimbangan data secara lebih efektif, seperti *Random Forest*, *Logistic Regression*, *XGBoost*, atau bahkan

pendekatan berbasis *deep learning*, guna membandingkan dan meningkatkan performa model.

4. Penelitian ini telah menerapkan teknik *oversampling* menggunakan ML-SMOTE, namun tidak memberikan peningkatan performa yang signifikan. Untuk itu, penelitian berikutnya dapat mencoba teknik *resampling* lainnya, seperti ML-RUS (*multi-label random undersampling*), ML-ROS (*multi-label random oversampling*), serta mempertimbangkan pendekatan *cost-sensitive learning* agar model dapat lebih sensitif terhadap kelas minoritas.
5. Kesalahan prediksi yang disebabkan oleh *typo* atau penggunaan bahasa informal menunjukkan perlunya perbaikan pada tahapan *text pre-processing*. Penelitian selanjutnya disarankan untuk mengembangkan *pipeline pre-processing* yang lebih optimal, seperti penggunaan *spell checker* otomatis, normalisasi kosakata, atau teknik *text augmentation* untuk meningkatkan kemampuan model dalam memahami variasi bahasa pengguna.

Halaman ini sengaja dikosongkan