

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan tahapan penelitian, pengujian, dan pembahasan yang telah dilakukan, diperoleh kesimpulan sebagai berikut.

1. Model klasifikasi ulasan palsu pada website Female Daily berhasil dibangun menggunakan algoritma Random Forest dengan pendekatan semi-supervised pseudo-labeling. Penelitian menggunakan 5.724 ulasan produk Implora Day to Day Series yang terdiri atas 600 data berlabel dan 5.124 data tidak berlabel. Pelabelan terhadap 600 data dilakukan oleh tiga anotator dan menghasilkan 176 ulasan asli serta 424 ulasan palsu. Data teks melalui tahap *preprocessing*, kemudian dilakukan perbandingan representasi TF-IDF, FastText, dan Transformer/Sentence Embedding. FastText dipilih sebagai representasi teks karena memperoleh rata-rata *validation Macro-F1* tertinggi pada eksperimen ekstraksi fitur. Representasi FastText selanjutnya digabungkan dengan fitur nonteks berupa *Stars*, *Profile Age*, *Usage Period*, *Recommend*, *Nama Produk*, *Shade Produk*, dan *Purchase Point*. Pada tahap pembangunan model, Random Forest *supervised* dibandingkan dengan empat strategi *pseudo-labeling*, yaitu Standard, Confidence, Uncertainty, serta Confidence dan Uncertainty. Hasil perbandingan menunjukkan bahwa Uncertainty Pseudo-Label menjadi metode terpilih. Konfigurasi acuan menggunakan skenario 70:30 dengan *uncertainty quantile* Q50 dan *uncertainty threshold* 0,041456. Model akhir menggunakan 300 *estimator* dengan gabungan 300 fitur FastText dan 72 fitur nonteks sehingga menghasilkan 372 fitur. Model kemudian diimplementasikan pada aplikasi berbasis Flask yang dapat melakukan klasifikasi secara manual maupun melalui file CSV.
2. Performa model menunjukkan bahwa Uncertainty Pseudo-Label memberikan hasil terbaik berdasarkan evaluasi pada validation set, tetapi kemampuan generalisasi model menurun ketika digunakan pada data yang

memiliki karakteristik berbeda dari data pelatihan. Metode Uncertainty Pseudo-Label memperoleh rata-rata *validation Accuracy* sebesar 0,8464, *Balanced Accuracy* sebesar 0,8101, *Precision Macro* sebesar 0,8158, dan *Macro-F1* sebesar 0,8126. Nilai *Macro-F1* tersebut merupakan yang tertinggi dibandingkan Standard Pseudo-Label sebesar 0,8103, Supervised Random Forest sebesar 0,8041, serta Confidence dan Confidence + Uncertainty yang masing-masing memperoleh 0,7972. Pada pengujian eksternal terhadap 92 ulasan Implora, model memperoleh *Accuracy* 81,52%, *Balanced Accuracy* 72,07%, dan *Macro-F1* 65,21%. Pada 132 ulasan multiproduk, model memperoleh *Accuracy* 78,03%, *Balanced Accuracy* 60,90%, dan *Macro-F1* 59,72%. Pada subset 40 ulasan produk baru, performa turun menjadi *Accuracy* 70,00%, *Balanced Accuracy* 49,10%, dan *Macro-F1* 48,05%. Hasil tersebut menunjukkan bahwa model mampu melakukan klasifikasi dengan lebih baik pada data yang memiliki karakteristik serupa dengan domain pelatihan, sedangkan perbedaan produk, distribusi kelas, serta pola ulasan pada data baru memengaruhi kemampuan generalisasi model.

5.2 Saran

Berdasarkan hasil penelitian dan keterbatasan yang ditemukan selama proses pengembangan serta pengujian model, beberapa saran untuk penelitian selanjutnya adalah sebagai berikut.

1. Penelitian selanjutnya dapat memperluas data pelatihan dengan menggunakan ulasan dari lebih banyak merek, seri, dan kategori produk kecantikan. Penambahan variasi domain diperlukan karena hasil pengujian menunjukkan penurunan performa pada produk yang tidak terdapat dalam data pelatihan. Dengan distribusi produk yang lebih beragam, model diharapkan mampu mempelajari karakteristik ulasan yang lebih umum dan meningkatkan kemampuan generalisasi pada produk baru.
2. Distribusi kelas pada data berlabel dapat dibuat lebih representatif terhadap kondisi data aktual. Data berlabel pada penelitian ini terdiri atas 176 ulasan asli dan 424 ulasan palsu, sedangkan data eksternal memiliki proporsi kelas asli yang lebih besar. Perbedaan distribusi tersebut dapat

memengaruhi kecenderungan prediksi model. Penelitian berikutnya dapat menambah data berlabel dari kedua kelas dan melakukan eksperimen terhadap teknik penanganan ketidakseimbangan kelas.

Halaman ini sengaja dikosongkan