

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan terkait klasifikasi *multi-label* faktor pembentuk *IS Continuance Intention* pada ulasan aplikasi Blu by BCA Digital menggunakan pendekatan *IndoBERT* dan NER, diperoleh beberapa kesimpulan sebagai berikut:

1. Model klasifikasi *intent multi-label* berbasis *IndoBERT* dengan pendekatan *Single Task* menghasilkan performa terbaik pada skenario *Focal Loss* dan *Oversampling* dengan rasio data 80:20, dengan *F1-Score* sebesar 0,797, *Precision* 0,794, *Recall* 0,801, dan *Hamming Loss* sebesar 0,131. Model ini terbukti mampu mengklasifikasikan keempat faktor pembentuk *IS Continuance Intention* berdasarkan kerangka ECM, yaitu *Confirmation*, *Perceived Usefulness*, *E-satisfaction*, dan *Perceived Security*, di mana label *E-Satisfaction* mencapai performa tertinggi dengan *F1-Score* 0,83 dan *Perceived Security* menjadi label dengan tantangan terbesar akibat data yang lebih terbatas. Sementara itu, model NER *Single Task* dengan konfigurasi *Token Augmentation* dan CRF pada rasio 80:20 mencapai *F1-Score* sebesar 0,812, *Precision* 0,802, dan *Recall* 0,823, dengan entitas ORG dan PERF sebagai yang paling akurat, sementara COST dan SERV masih memerlukan peningkatan. Secara komparatif, pendekatan *Single Task* mengungguli pendekatan *Multi-Task*, mengindikasikan bahwa pelatihan model secara independen lebih efektif untuk data ulasan aplikasi perbankan digital berbahasa Indonesia dengan karakteristik *multi-label*.
2. Dashboard berbasis web yang dikembangkan berhasil mengimplementasikan hasil prediksi model *IndoBERT* untuk klasifikasi *intent multi-label* dan NER secara visual. Website mampu menerima *input* berupa teks tunggal maupun file *CSV*, menampilkan prediksi empat faktor ECM pembentuk *IS Continuance Intention*, yaitu *Confirmation*, *Perceived Usefulness*, *E-Satisfaction*, dan *Perceived Security*, menampilkan entitas hasil ekstraksi NER, serta menyajikan visualisasi distribusi hasil prediksi.

5.2 Saran

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

1. Penelitian selanjutnya disarankan untuk menggunakan dataset dengan jumlah yang lebih besar dan distribusi label yang lebih seimbang, khususnya untuk entitas *NER* yang jarang muncul. Hal ini penting untuk meningkatkan kemampuan generalisasi model serta mengurangi kesalahan prediksi pada
2. Perlu dilakukan perbaikan pada tahap *pre-processing*, khususnya dalam menangani *typo*, bahasa informal, serta variasi penulisan pada ulasan pengguna. Kesalahan prediksi yang ditemukan menunjukkan bahwa model masih sensitif terhadap bentuk kata yang tidak baku.
3. Penelitian berikutnya dapat mengeksplorasi arsitektur seperti *IndoBERT-large* atau model berbasis LLM (*Large Language Model*) untuk bahasa Indonesia. Selain itu, penelitian selanjutnya juga dapat mengembangkan strategi *multi-task learning* yang lebih adaptif, seperti penggunaan mekanisme *task-specific attention* atau *adversarial training*, untuk mengevaluasi efektivitas representasi bersama antara tugas *intent classification* dan *Named Entity Recognition (NER)*.
4. Website yang telah dikembangkan dapat diperluas dengan penambahan fitur evaluasi otomatis, yaitu membandingkan hasil prediksi dengan label asli secara langsung melalui input file, sehingga proses validasi model dapat dilakukan dengan lebih cepat.