

## **BAB V**

### **PENUTUP**

#### **5.1. Kesimpulan**

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

1. Dari 36 skenario yang dilakukan untuk melatih tiga model dalam penelitian ini yaitu model aspek produk, layanan dan fungsional diambil 1 model terbaik dari masing-masing model aspek ulasan Midi Kriing. Model produk dengan nilai *F1 Score* 75% dan nilai *Accuracy* 86%. Model Layanan dengan nilai *F1 Score* 84% dan nilai *Accuracy* 86%. Model Fungsional dengan nilai *F1 Score* 84% dan nilai *Accuracy* 89%.
2. Hasil Uji *F1 Score* dibandingkan dengan data validasi memperoleh hasil sebagai berikut :
  - a. Model Produk : Nilai *F1 Score* pada model sebesar 0.75, sedangkan pada data validasi memperoleh nilai *F1 Score* sebesar 0.74.
  - b. Model Layanan : Nilai *F1 Score* pada model sebesar 0.84, sedangkan pada data validasi memperoleh nilai *F1 Score* sebesar 0.79.
  - c. Model Fungsional : Nilai *F1 Score* pada model sebesar 0.84, sedangkan pada data validasi memperoleh nilai *F1 Score* sebesar 0.85.
3. Website yang dikembangkan dalam penelitian ini dapat mengklasifikasikan ulasan masuk dalam aspek seperti produk, layanan dan fungsional dan mengetahui sentimen pada masing-masing aspek ulasan tersebut. Selain itu, website ini dapat memproses analisis data dalam jumlah yang banyak dalam file .csv dan mampu menampilkan prediksi sentimen pada aspek ulasan dan grafik distribusi data.

#### **5.2. Saran**

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa aspek yang dapat dikembangkan lebih lanjut untuk meningkatkan kinerja sistem klasifikasi ulasan Midi Kriing, di antaranya:

1. Jumlah dan variasi komentar dalam penelitian ini tergolong masih terbatas, sehingga terjadi ketidakseimbangan label sentimen pada ulasan. Penelitian kedepannya dapat menambah data ulasan dalam pelatihan model.
2. Dalam penelitian ini, model aspek produk memiliki nilai *F1 Score* yang rendah

dikarenakan ketidakseimbangan pada data label sentimen, meskipun sudah dilakukan *resampling* menggunakan teknik *SMOTE*. Diharapkan penelitian selanjutnya mencoba menggunakan teknik *resampling* lain seperti undersampling pada kelas mayoritas atau penerapan class weight adjustment dalam model, untuk mengatasi masalah ketidakseimbangan data.

3. Dalam proses data *preprocessing*, terdapat beberapa tahapan seperti normalisasi dan penghapusan *stopword*. Ke depannya, untuk proses normalisasi yang mengubah kata tidak baku menjadi baku, perlu digunakan kamus yang lebih lengkap, yang mencakup lebih banyak kata tidak baku dan bentuk bakunya dalam data pelatihan model. Untuk perbaikan *stopword* removal, disarankan untuk memperluas daftar *stopword* dengan kata-kata spesifik domain, menggunakan teknik berbasis frekuensi kata, serta menerapkan pendekatan kontekstual atau pembelajaran mesin untuk menyesuaikan penghapusan *stopword* sesuai dengan konteks analisis sentimen.