

## **BAB V PENUTUP**

### **5.1. Kesimpulan**

Klasifikasi sentimen ulasan pengguna aplikasi PeduliLindungi dengan data sebanyak 1000 yang diberi dua label yaitu label positif dan negatif dengan jumlah label positif 250 dan label negatif 750. Data yang sudah dilabeli akan dilakukan praproses data dengan cara *data cleaning*, *case folding*, *spelling normalization*, *tokenizing*, *filtering - stopword removal*, dan *stemming*. Kemudian data tersebut akan dilakukan pembobotan dengan *Term Frequency – Inverse Document Frequency* (TF - IDF) supaya dapat diproses oleh model *machine learning*. Setelah itu akan dilakukan tiga klasifikasi dengan banyak data yang berbeda, namun sebelum dilakukan klasifikasi data akan di bagi menjadi data latih sebesar 80% dan data uji sebesar 20%. Ketiga jenis data tersebut antara lain yang pertama dengan menggunakan data yang tidak seimbang, lalu yang kedua menggunakan data yang seimbang dengan metode *undersampling* dan yang ketiga menggunakan data yang seimbang dengan metode *oversampling* SMOTE.

Dari ketiga hasil tersebut dapat disimpulkan

1. Sentimen ulasan pengguna aplikasi PeduliLindungi, lebih banyak sentimen negatif daripada sentimen positif yang mana terdapat 750 negatif dan 250 positif.
2. Performa klasifikasi sentimen ulasan pengguna aplikasi PeduliLindungi yang dilakukan menggunakan algoritma *K-Nearest Neighbor* memiliki performa yang paling baik ada pada penelitian yang menggunakan teknik SMOTE, pada  $K = 1$  dengan nilai akurasi sebesar 0.9766, nilai presisi sebesar 0.9691, nilai F1 score 0.9781, nilai spesifisitas sebesar 0.9645, dan nilai sensitivitas sebesar 0.9874.
3. Perbandingan performa dari ketiga data penelitian tersebut mendapatkan hasil rata-rata yang terus meningkat pada data yang seimbang dengan teknik *undersampling* dan hasil rata-rata terbaik pada teknik SMOTE. Ini menunjukkan bahwa performa klasifikasi sangat baik menggunakan data yang seimbang.

## 5.2. Saran

Saran untuk penelitian yang akan datang dengan metode atau teknik yang serupa sehingga dapat menghasilkan penelitian yang lebih baik :

1. Penambahan data yang lebih banyak lagi sesuai dengan prinsip model *machine learning*.
2. Penambahan label netral sehingga tidak hanya dua label yang diuji.
3. Bisa mencoba dengan algoritma klasifikasi lainnya, atau membandingkannya dengan algoritma *K-Nearest Neighbor* seperti yang dilakukan pada penelitian ini.
4. Bisa mencoba teknik lain dalam menyeimbangkan data seperti teknik *undersampling* lainnya atau teknik *oversampling* selain SMOTE.