

# BAB V

## PENUTUP

### 5.1. Kesimpulan

Penelitian ini menggunakan 1000 data ulasan dari aplikasi Maxim, yang dilabeli oleh tiga anotator dan menghasilkan 842 data berlabel positif dan 158 data berlabel negatif. Data tersebut kemudian melewati tahap praproses dan pembobotan dengan TF-IDF. Pada data positif, seringnya kata yang ada adalah 'Bagus', sedangkan pada data negatif, yang paling sering ada adalah 'Maxim'. Penelitian ini melakukan klasifikasi menggunakan dua jenis data, yaitu data yang tidak seimbang dan data yang seimbang dengan metode SMOTE. Berikut adalah hasil penelitiannya:

1. Proses pembangunan model klasifikasi sentimen pada data ulasan aplikasi Maxim memakai algoritma *K-Nearest Neighbor*, Hasil penelitian menggunakan data tidak seimbang dengan data positif sebesar 842 dan data negatif 158, percobaan yang dilakukan menggunakan nilai  $K = 1, 2, 3, 4,$  dan  $5$ . Didapatkan nilai tertinggi pada nilai  $K = 2$ , yaitu akurasi sebesar  $0.875$ , presisi  $0.877$ , *f1-score*  $0.932$ , spesifisitas  $0.111$ , dan sensitivitas  $0.994$ . Walau mendapatkan hasil akurasi yang cukup tinggi namun pada nilai spesifisitas dan sensitivitas memiliki selisih yang cukup tinggi sehingga dapat dinyatakan bahwa penelitian kurang baik karena perbedaan jumlah data negatif dan positif yang sangat signifikan.
2. Menangani ketidak seimbangan data klasifikasi sentimen pada ulasan aplikasi MAXIM dengan menerapkan teknik oversampling menggunakan SMOTE agar data dapat seimbang yang kemudian data positif menjadi 842 dan data negatif juga 842, percobaan ini juga dilakukan dengan menggunakan nilai  $K = 1, 2, 3, 4,$  dan  $5$ . Hasilnya tertinggi didapatkan pada nilai  $K = 1$  dan  $2$  dengan nilai yang sama yaitu akurasi  $0.985$ , presisi  $1$ , *f1-score*  $0.986$ , spesifisitas  $1$  dan sensitivitas  $0.973$ . Penelitian dengan data yang sudah melalui proses SMOTE ini mendapatkan hasil yang sangat baik dapat dilihat dari nilai spesifisitas dan

sensitivitas yang memiliki selisih tipis, menunjukkan bahwa data penelitian yang digunakan seimbang. Sehingga dapat disimpulkan penelitian yang menggunakan data yang seimbang ini lebih baik daripada yang tidak seimbang.

## 5.2. Saran

Tindakan yang disarankan untuk penelitian serupa atau untuk mengembangkan penelitian ini lebih lanjut adalah :

1. Menggunakan teknik *oversampling* lainnya agar lebih mengetahui mana teknik yang lebih baik
2. Data yang digunakan bisa lebih banyak agar menguji seberapa efektif algoritma ini bekerja
3. Dapat menggunakan beberapa metode klasifikasi lainnya agar mendapatkan perbandingan metode mana yang lebih baik.