

## BAB V

### PENUTUP

#### 5.1 Simpulan

Berdasarkan penelitian yang sudah dilaksanakan, dapat diambil simpulan sebagai berikut:

1. Implementasi kombinasi *Naïve Bayes Classifier* dengan *Information Gain* dalam analisis sentimen komentar terhadap aplikasi *dating online*, dapat ditarik beberapa kesimpulan bahwa metode ini membantu dalam mengambil fitur-fitur yang paling informatif dan relevan untuk melakukan klasifikasi sentimen. Selain itu, penelitian ini juga menggunakan teknik *10-fold cross-validation*. Metode ini membagi dataset menjadi 10 subset atau *fold*. Dengan demikian, model dievaluasi pada setiap *fold*, sehingga memberikan estimasi yang lebih akurat tentang kinerja model pada data yang tidak terlihat sebelumnya. Kombinasi penggunaan *Naïve Bayes Classifier* dan *Information Gain* berhasil menghasilkan model yang mampu melakukan uji sentiment analysis pada dataset topik yang berkaitan dengan penggunaan aplikasi *dating online* melalui *Twitter*.
2. Tingkat akurasi yang didapatkan pada kombinasi *Naïve Bayes Classifier* dan *Information Gain* dalam analisis sentimen terhadap penggunaan aplikasi *dating online* melalui *twitter* didapatkan sebesar 65%, *precision* 65% dan *recall* 64%. Sedangkan tingkat akurasi pada *Naïve Bayes Classifier* saja hanya sebesar 52%, *precision* 51% dan *recall* 51% hal ini tentunya terdapat perbandingan yang cukup signifikan dengan selisih 13%. Dimana pada proses ini, data yang sudah melalui tahap *preprocessing* langsung diklasifikasi menggunakan *Naïve Bayes Classifier*.

## 5.2 Saran

Saran, yang diberikan untuk penelitian selanjutnya adalah sebagai berikut:

1. Mencoba untuk menggunakan metode untuk mengatasi masalah *imbalanced data*, karena pada penelitian ini tidak memperhitungkan *imbalanced data* yang kemungkinan menjadi penyebab kurang optimalnya akurasi klasifikasi.
2. Peneliti berikutnya dapat menggunakan algoritma *feature extraction* TF-IDF dalam mengidentifikasi kata-kata yang paling signifikan karena ketika hanya menggunakan *information gain*, akurasi yang didapatkan masih rendah karena pada tahap ini hanya memperhitungkan seberapa informatif suatu kata atau fitur. Oleh karena itu salah satu alternatif yang bisa dilakukan adalah menggunakan algoritma *feature extraction* TF-IDF. Hal ini juga diperkuat oleh hasil penelitian (Isnanda dkk., 2021) yang melakukan klasifikasi menggunakan *naïve bayes* dan *information gain* serta menerapkan TF-IDF, tingkat akurasi yang didapatkan cukup tinggi.

