

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1 Kesimpulan

Penelitian ini dilatarbelakangi oleh permasalahan bagaimana mengetahui dan menggambarkan persepsi serta tingkat kepuasan pengguna terhadap aplikasi Shopee di *Google Play Store* berdasarkan ulasan yang mereka berikan. Banyaknya ulasan yang tersedia membuat analisis secara manual menjadi kurang efisien, sehingga dibutuhkan metode otomatis untuk mengklasifikasikan sentimen pengguna. Untuk menjawab masalah tersebut, penelitian ini mengimplementasikan algoritma *Multinomial Naïve Bayes* yang dipadukan dengan pembobotan *Term Frequency – Inverse Document Frequency (TF-IDF)* guna mengklasifikasikan ulasan ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, diperoleh beberapa kesimpulan sebagai berikut :

1. Proses *pre-processing* data pada ulasan pengguna Shopee di *Google Play Store* telah dilakukan secara sistematis untuk memastikan data siap diolah menggunakan algoritma *Naïve Bayes*. Proses ini berhasil menghasilkan dataset yang bersih, konsisten, dan representatif, sehingga mendukung performa model secara optimal.
2. Pengujian model dengan pembagian data latih 70% dan data uji 30% menghasilkan akurasi sebesar 87,55%, yang menunjukkan mayoritas prediksi model sesuai dengan label sentimen yang sebenarnya.

3. Distribusi hasil klasifikasi menunjukkan bahwa sebagian besar ulasan pengguna termasuk kategori netral dan positif, yang mengindikasikan persepsi umum terhadap aplikasi Shopee cenderung baik hingga memuaskan.

## 5.2 Saran

Adapun saran dari peneliti untuk penelitian selanjutnya :

1. Pada penelitian selanjutnya dapat menambahkan metode seperti *Synthetic Minority Oversampling Technique* (SMOTE) atau *oversampling* untuk memperbaiki distribusi data antar kelas.
2. Pada penelitian selanjutnya dapat menggunakan dataset dari *platform e-commerce* yang berbeda untuk menguji kemampuan generalisasi model pada domain yang lebih luas.
3. Pada penelitian selanjutnya dapat menambahkan jumlah data pada setiap kategori sentimen agar hasil klasifikasi lebih akurat dan seimbang, khususnya untuk mengatasi permasalahan *class imbalance* pada kategori dengan jumlah data sedikit.