

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Berdasarkan hasil penelitian yang telah dilakukan terhadap analisis sentimen ulasan pengguna game Hay day di google playstore menggunakan algoritma Naïve Bayes, Support Vector Machine (SVM), dan K-Nearest Neighbor (KNN) dapat disimpulkan bahwa:

1. Penelitian dilakukan menggunakan dua skenario ukuran data yaitu 1000 data dan seluruh dataset sebanyak 3493 data. Hasil menunjukkan bahwa penggunaan seluruh data tidak selalu menghasilkan akurasi yang lebih tinggi. Model Naive Bayes justru memperoleh akurasi tertinggi 86% saat menggunakan 1000 data dengan rasio data latih 90%. Hal ini menunjukkan bahwa kualitas distribusi data latih lebih berpengaruh daripada sekadar kuantitas data.
2. Penggunaan data latih sebesar 90% memberikan performa yang lebih baik dibanding 80% pada semua model. Ini disebabkan model memiliki informasi pelatihan yang lebih lengkap, sehingga mampu mengenali pola lebih akurat saat mengklasifikasi data uji.
3. Setiap model memberikan hasil yang berbeda tergantung parameter yang digunakan. Model Naive Bayes menghasilkan performa paling tinggi dan konsisten, dengan akurasi terbaik mencapai 86% saat menggunakan 1000 data dan data latih 90%. Sementara itu, model Support Vector Machine dengan kernel RBF menunjukkan performa yang stabil di kisaran 79–81%, meskipun tidak mampu melampaui model Naive Bayes. Terakhir model K-Nearest Neighbor menghasilkan akurasi terendah dari keseluruhan eksperimen, menandakan bahwa algoritma ini masih memiliki keterbatasan dalam menangani pola yang kompleks pada dataset ini.

#### **5.2 Saran**

Berdasarkan hasil analisis sentimen ulasan aplikasi game Hay Day yang telah dilakukan, penelitian ini memiliki kekurangan dan masih banyak yang dapat

dikembangkan dan di eksplorasi. Oleh karena itu berikut saran yang dapat diterapkan pada penelitian selanjutnya:

1. Penelitian selanjutnya dapat menggunakan dataset dari sumber lain, seperti instagram, X, Facebook, atau platform ulasan lainnya.
2. penelitian selanjutnya dapat meningkatkan optimasi model melalui tuning hyperparameter, misalnya dengan GridSearchCV atau RandomizedSearchCV, khususnya untuk KNN yang sangat sensitif terhadap jumlah tetangga ( $n\_neighbors$ )
3. penelitian selanjutnya dapat mengeksplorasi model *deep learning* seperti LSTM atau BERT untuk analisis sentimen guna menangkap konteks yang lebih baik dalam teks.

