

BAB 5 PENUTUP

5.1. Kesimpulan

Berdasarkan hasil dan pembahasan yang diperoleh dari penelitian analisis sentimen aplikasi Tribunnews.com pada *Google Play* dapat diambil beberapa kesimpulan bahwa:

1. Tahapan pertama untuk pembuatan model klasifikasi yaitu dilakukan pengambilan data ulasan dengan teknik *scrapping*, data yang digunakan pada penelitian yaitu *review* pengguna aplikasi Tribunnews.com pada *Google Play* kemudian dilakukan pelabelan data secara manual oleh tiga anatator dan didapatkan kelas positif berjumlah 821 dan kelas negatif berjumlah 215. Setelah dilakukan pelabelan kemudian akan dilakukan praproses data yaitu *case folding*, data *cleaning*, normalisasi data, *stopword removal*, *stemming* dan *tokenizing* sebelum dilakukan pembobotan, data yang telah bersih kemudian dilakukan pembobotan terlebih dahulu menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF) untuk selanjutnya dilakukan seleksi fitur menggunakan metode *Information Gain*. Selanjutnya proses *split* data menggunakan *K-fold Cross Validation* dimana data latih dan data uji terbagi menjadi 80% banding 20%, selanjutnya fitur yang telah terseleksi digunakan untuk melakukan pelatihan data latih dalam pelatihan model klasifikasi, selanjutnya setelah dilakukan pembangunan model akan dilakukan pengujian model *Naïve Bayes Classifier* yang telah dibuat lalu akan dievaluasi menggunakan *confusion matrix*.
2. Performa pada model *Naïve Bayes Classifier* yang telah dilatih dibagi menjadi empat, model terbaik didapatkan menggunakan model *Naïve Bayes* menggunakan seleksi fitur *Information Gain* dan model di *oversampling* *undersampling* memperoleh *accuracy* sebesar 96%, *precision* sebesar 97%, *recall* sebesar 98%, *specificity* sebesar 88% dan AUC sebesar 93%.
3. Performa model *Naïve Bayes* menggunakan seleksi fitur *Information Gain* memperoleh nilai *specificity* sebesar 63% dan setelah di *oversampling*

menjadi 88% dimana model Naïve Bayes yang tidak menggunakan seleksi fitur *Information Gain* hanya memperoleh *specificity* sebesar 58%.

5.2. Saran

Berdasarkan hasil penelitian yang telah dilakukan didapatkan beberapa saran untuk perbaikan penelitian yang akan mendatang yakni:

1. Pada penelitian selanjutnya untuk membandingkan algoritma Naïve Bayes dengan algoritma lain seperti Random Forest, SVM atau KNN.
2. Pada penelitian selanjutnya untuk membandingkan hasil performa seleksi fitur dengan seleksi fitur lainnya seperti *Particle Swarm Optimization* (PSO) atau *Chi-Square*.
3. Meningkatkan hasil performa dengan melakukan normalisasi data dan seleksi data saat pengambilan agar tidak terjadi *imbalanced data*.