

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat diambil beberapa kesimpulan, yaitu:

1. Klasifikasi ulasan pengguna aplikasi sambara dengan metode naïve bayes classifier dilakukan dengan mengumpulkan data ulasan aplikasi Sambara di *google play store* pada tanggal 20 april hingga 31 desember 2022 dengan versi aplikasi 4.0.7 sebanyak 1568 data ulasan. Kemudian dilakukan pelabelan oleh 3 *annotator* dan didapatkan hasil akhir pada masing-masing kelas yaitu positif 405 data, negatif 487 data, dan netral sebanyak 676 data. Setelah itu dilakukan penghapusan label netral, sehingga jumlah data yang diolah menjadi 892 data. Lalu dilakukan beberapa tahapan praproses data antara lain, *cleaning*, *case folding*, *spelling normalization*, *stopword removal*, *tokenizing*, dan *stemming*. Setelah itu dilakukan pemberian bobot pada tiap kata dengan metode TF-IDF, setelah itu hasil dari pembobotan TF-IDF digunakan untuk perhitungan seleksi fitur menggunakan metode *chi square*. Lalu dilakukan eksperimen menggunakan data yang sudah dibersihkan dan diseleksi fitur *chi square* berdasarkan alpha 0,01 hingga 0,99 dengan kenaikan alpha sebesar 0,01, hasil dari seleksi fitur tersebut selanjutnya dilakukan pembagian data kedalam data latih dan data uji dengan rasio pembagian data sebesar 80:20, dimana 80% data akan digunakan sebagai data latih, dan 20% data sebagai data uji berdasarkan fitur yang telah terseleksi oleh *chi square*, data latih dipakai untuk membuat model prediksi, dan data uji dipakai untuk menguji performa model *naïve bayes*.

2. Performa metode *naïve bayes* dan seleksi fitur *chi square* dalam klasifikasi ulasan aplikasi sambara dengan menggunakan data uji sebanyak 179 data ulasan, mendapatkan nilai akurasi terbesar dengan menggunakan seleksi fitur *chi square* pada saat nilai kritis 0,837 yaitu 94.4%, *specificity* 93.5%, dan nilai *recall* 95.3%, sedangkan klasifikasi dengan tanpa menggunakan *chi square* mendapatkan akurasi sebesar 92%, *specificity* 94.3% dan *recall* sebesar 94%.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, dapat diberikan saran, yaitu:

1. Penelitian berikutnya diharapkan menggunakan metode seleksi fitur lainnya seperti *Information gain*, *Odds Ratio*, *Bi-Normal Separation*, dan lain sebagai pembanding dengan metode *chi square*.
2. Penelitian berikutnya diharapkan menggunakan label netral pada proses klasifikasinya.
3. Penelitian berikutnya diharapkan menggunakan metode klasifikasi lainnya seperti *Linear Regression*, *K-Nearest Neighbor*, *Support Vector Machine*, *Long Short Term Memory*.