

BAB V

KESIMPULAN

5.1. KESIMPULAN

Kesimpulan yang dapat diambil dari hasil penelitian yang telah dilakukan yaitu sebagai berikut:

1. Pada penelitian ini data yang di proses untuk *setiment analysis* menggunakan *K-Nearest Neighbor* dan *Naïve Bayes* sebanyak 4402 data dengan jumlah *sentiment positive* sebanyak 3513 data, *sentiment* netral sebanyak 169 data, serta *sentiment negative* sebanyak 720 data.
2. Pada penelitian ini, dilakukan penerapan klasifikasi menggunakan metode *K-Nearest Neighbor* dan *Naïve Bayes* dengan *train_test_split* yang dimulai dari 20% data uji hingga 90% data uji. Hasilnya menunjukkan:
 - a. Metode *K-Nearest Neighbor* memiliki akurasi rata-rata 78%, sedangkan metode *Naïve Bayes* mencapai 81.25%, menunjukkan bahwa metode *Naïve Bayes* memiliki performa lebih baik dalam akurasi.
 - b. Dalam penelitian ini, metode *K-Nearest Neighbor* memiliki *precision* lebih tinggi untuk *sentiment positive* (82%), dengan *recall* yang juga tinggi (90.25%).

- c. Metode *Naïve Bayes* lebih baik mengenali *sentiment negative* dengan *precision* 63.25%, meskipun *recall*-nya lebih rendah (17.13%).
 - d. Secara umum, *K-Nearest Neighbor* memiliki *f1-score* konsisten untuk semua sentimen, sedangkan *Naïve Bayes f1-score* lebih tinggi untuk *sentiment positive*.
3. Dari uji *confusion matrix K-Nearest Neighbor* dan *Naïve Bayes* (20-90% data uji), keduanya menunjukkan keunggulan dalam mengenali *sentiment positive*, namun kesulitan terjadi pada *sentiment negative* dan netral. Permasalahan ini disebabkan oleh ketidakseimbangan data *sentiment*, di mana jumlah data *sentiment positive* lebih banyak dibandingkan dengan *sentiment negative* dan netral. Kemudian nilai akurasi *Naïve Bayes* menunjukkan nilai akurasi yang lebih tinggi dibandingkan *K-Nearest Neighbor*, dengan rata-rata akurasi *K-Nearest Neighbor* sebesar 75.42% dan rata-rata akurasi *Naïve Bayes* sebesar 81.35%.
 4. Dalam uji *k-fold cross-validation* menggunakan metode *K-Nearest Neighbor* dan *Naïve Bayes* dengan pengujian sebanyak 5 kali pada rentang data uji dari 20% hingga 90%, diperoleh hasil akurasi sebesar 77% untuk metode *K-Nearest Neighbor* dan 81% untuk metode *Naïve Bayes*
 5. Dengan penelitian ini, dapat dilihat bahwa kedua metode menghasilkan akurasi yang baik, namun metode *Naïve Bayes*

memberikan akurasi yang lebih unggul. Selain itu, dapat diperhatikan bahwa aplikasi *TikTok* di *Google Play Store* mendapatkan lebih banyak *sentiment positive* dari pengguna, sebagaimana terlihat dari *word cloud positive* yang mencakup kata-kata seperti '*tiktok*,' '*good*,' dan '*love*'.

5.2 SARAN

Berdasarkan penelitian yang telah dilakukan maka dapat di kemukakan saran-saran sebagai berikut:

1. Sebaiknya pastikan data yang akan digunakan memiliki jumlah *sentiment* yang seimbang antara *sentiment positive*, *negative*, serta netral untuk meningkatkan kualitas *sentiment analysis*.
2. Selain metode *K-Nearest Neighbor* dan *Naïve Bayes*, pertimbangkan penggunaan metode klasifikasi lain yang mungkin lebih sesuai untuk penelitian *sentiment analysis* ini.
3. Tingkatkan proses *preprocessing data* untuk menghasilkan hasil *sentiment analysis* yang lebih baik.
4. Berdasarkan hasil penelitian, aplikasi *TikTok* mendapatkan lebih banyak *sentiment positive* dibandingkan dengan *sentiment negative* maupun netral. Hal ini menunjukkan bahwa aplikasi *TikTok* memberikan kepuasan kepada pengguna. Oleh karena itu, sebaiknya pengembang aplikasi *TikTok* tetap mempertahankan kualitasnya dan terus mengembangkan layanan agar pengguna tetap puas.

