

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan penelitian yang dilakukan untuk mengidentifikasi buzzer di platform *TikTok*. Dapat disimpulkan bahwa *Cosine Similarity* berhasil mengidentifikasi pola buzzer dengan mendeteksi kesamaan antar-komentar yang menunjukkan struktur dan pola komentar khas buzzer di *TikTok*. Komentar yang memiliki kemiripan tinggi kemudian dikelompokkan menggunakan algoritma K-Means, menghasilkan dua cluster, yaitu cluster 0 untuk komentar natural yang berjumlah 8655 dan cluster 1 untuk komentar buzzer berjumlah 441.

Selain itu, penerapan model *machine learning* seperti *Logistic Regression*, *Random Forest*, *Decision Tree*, dan *SVM* menunjukkan performa yang sangat baik, di mana rata-rata akurasi dari empat model yang digunakan berada di atas 90%. Model *Random Forest* menunjukkan performa terbaik dengan akurasi 99.96%, presisi 99.63%, *recall* 99.98%, dan *F1-score* 99.81%, yang menjadikannya model paling unggul dalam penelitian ini. Model *Decision Tree* juga menunjukkan performa yang sangat baik dengan akurasi 99.93%, diikuti oleh *SVM* dengan akurasi 99.85%. Sedangkan *Logistic Regression* sedikit lebih rendah namun tetap memiliki akurasi yang sangat baik yaitu 99.49%.

Pola data yang *linear* dan mudah dipahami oleh model menjadi salah satu faktor yang mendukung performa baik tersebut. Hasil validasi dengan *cross-validation* juga menunjukkan bahwa tidak ada indikasi *overfitting*, ditandai dengan standar deviasi yang rendah di setiap *fold*. Dengan hasil yang konsisten dan stabil, model ini dapat di gunakan untuk mendeteksi buzzer secara otomatis di platform *TikTok* dan siap diterapkan pada data baru.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat dipertimbangkan untuk pengembangan lebih lanjut. Pertama, mengkombinasikan metode lain seperti *sentiment analysis* atau *topic modeling*

dapat membantu memperkaya hasil identifikasi dengan melihat konteks emosional dan topik dari setiap komentar. Kedua, penambahan fitur interaksi pengguna, seperti jumlah *likes*, *replies*, dapat memberikan informasi tambahan tentang pola aktivitas akun. Ketiga, di masa mendatang, model *deep learning* seperti LSTM atau BERT bisa dicoba untuk melihat apakah model yang lebih kompleks mampu memberikan hasil yang lebih baik, terutama dalam menangani data teks yang lebih dinamis. Selain itu, penggunaan dataset yang lebih besar dan beragam dari platform lain seperti Twitter atau Instagram juga disarankan agar model yang dikembangkan lebih umum dan tidak hanya spesifik untuk *TikTok*. Terakhir, mengembangkan aplikasi yang lebih *user-friendly* dengan tampilan dan fitur visualisasi yang lebih interaktif dapat mempermudah pengguna dalam memahami hasil analisis dan mendeteksi *buzzer* dengan lebih efektif.

