

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, penelitian ini berhasil membangun model *forecasting* frekuensi kemunculan topik, *hashtag*, dan kata di platform X serta membangun model klasifikasi tingkat interaksi *tweet* sebagai analisis pendukung dalam optimalisasi strategi *social media marketing campaign*. Secara keseluruhan, dapat disimpulkan bahwa:

1. Proses pengumpulan dan pengolahan data telah berhasil dirancang secara sistematis melalui tahapan *scraping*, *preprocessing*, dan ekstraksi fitur. Proses pengumpulan data dilakukan dengan *scraping* platform X menggunakan *tools Tweet Harvest* dalam rentang waktu 18 bulan, yaitu dari 27 Mei 2024 hingga 27 November 2025, menghasilkan total 3.549 *tweet* yang setelah melalui seluruh tahapan *preprocessing* menjadi 3.110 data bersih. Data teks kemudian dibersihkan dan dinormalisasi untuk menghilangkan *noise*. Data dilengkapi dengan metadata pengguna seperti jumlah *followers*, *following*, dan status verifikasi yang dikumpulkan secara manual.
2. Model *forecasting* frekuensi topik, *hashtag*, dan kata berhasil dikembangkan sebagai analisis utama dengan memanfaatkan BERTopic berbasis IndoBERTweet untuk ekstraksi topik dan LightGBM untuk peramalan deret waktu. Pendekatan ini mampu mengidentifikasi tren dan memprediksi kemunculan topik, *hashtag*, dan kata di masa depan secara lebih terstruktur dan berbasis data. Namun, model ini masih memiliki keterbatasan, seperti hanya berfokus pada frekuensi kemunculan tanpa mempertimbangkan konteks semantik yang lebih dalam serta sensitivitas terhadap fluktuasi data yang tinggi pada media sosial yang bersifat dinamis.
3. Model klasifikasi tingkat interaksi *tweet* berhasil dibangun sebagai analisis pendukung dengan pendekatan yang menggabungkan representasi teks menggunakan IndoBERTweet, analisis sentimen menggunakan RoBERTa, serta metadata *tweet* sebagai fitur numerik. Proses pelabelan dilakukan

menggunakan *Gaussian Mixture Model* (GMM) untuk mengelompokkan *tweet* berdasarkan tingkat interaksinya. Seluruh fitur tersebut kemudian diproses menggunakan *Multi-Layer Perceptron* (MLP) yang mampu menangkap pola kompleks yang memengaruhi tingkat interaksi. Model ini mendukung optimalisasi strategi *social media marketing campaign* dengan memberikan estimasi potensi interaksi *tweet* sebelum dipublikasikan, termasuk prediksi jumlah *like*, *retweet*, *quote*, dan *reply*.

4. Integrasi hasil *forecasting* frekuensi topik, *hashtag*, dan kata dengan model klasifikasi tingkat interaksi *tweet* menghasilkan *insight* yang komprehensif dalam mendukung pengambilan keputusan untuk optimalisasi *strategi social media marketing campaign* di platform X. Hasil *forecasting* berperan sebagai strategi penentu isi konten melalui identifikasi tren topik, *hashtag*, dan kata yang diprediksi akan populer, sedangkan model klasifikasi tingkat interaksi berfungsi sebagai validasi tambahan untuk mengevaluasi apakah konten yang telah dirancang berpotensi memperoleh interaksi tinggi beserta estimasi jumlah *engagement*-nya. Integrasi ini memungkinkan perumusan strategi konten yang lebih optimal, mulai dari tahap perencanaan (pemilihan topik dan *keyword*), pembuatan konten (penyesuaian gaya bahasa dan sentimen), hingga penentuan waktu publikasi (*timing tweeting*), sehingga strategi *social media campaign* menjadi lebih efektif, adaptif, dan tepat sasaran.
5. Kinerja model yang diusulkan menunjukkan hasil yang cukup bervariasi pada model *forecasting* dan baik pada model klasifikasi. Pada model *forecasting*, performa terbaik ditunjukkan pada data *hashtag* dengan RMSE 0,5470, MAE 0,3683, dan RMSSE 0,7554, diikuti data topik dengan RMSE 1,5999, MAE 0,5783, dan RMSSE 0,7754, yang keduanya menunjukkan tingkat error rendah dan prediksi yang cukup akurat karena mampu mengungguli *baseline naive forecast*. Namun, pada data kata, performa model masih kurang optimal dengan RMSE 3,2833, MAE 1,3928, dan RMSSE 1,4273, yang mengindikasikan bahwa model belum mampu menangkap pola data yang sangat fluktuatif akibat pengaruh faktor eksternal seperti kampanye, *event*, dan tren viral sesaat. Pada model klasifikasi tingkat interaksi *tweet* mencapai akurasi 0,8358, *precision*

0,8351, *recall* 0,8358, dan *F1-score* 0,8355, yang menunjukkan performa klasifikasi yang baik, seimbang, dan mampu melakukan generalisasi dengan baik. Secara keseluruhan, hasil model dapat dikatakan cukup efektif untuk *forecasting* frekuensi topik dan *hashtag*, meskipun masih memiliki keterbatasan pada *forecasting* kata, serta cukup andal untuk klasifikasi tingkat interaksi *tweet*.

5.2. Saran Pengembangan

Berdasarkan hasil penelitian dan kesimpulan yang telah dipaparkan sebelumnya, terdapat beberapa saran pengembangan yang dapat dilakukan untuk meningkatkan kinerja serta memperluas pemanfaatan model yang telah dibangun. Saran-saran ini diharapkan dapat menjadi acuan bagi penelitian selanjutnya dalam mengembangkan model yang lebih akurat, adaptif, dan aplikatif dalam mendukung optimalisasi strategi *social media marketing* berbasis data. Berikut saran pengembangan yang dapat diajukan.

1. Pengembangan selanjutnya dapat difokuskan pada peningkatan kualitas dan keberagaman data, seperti memperluas periode pengambilan data, menambahkan data lintas konteks, serta mempertimbangkan data multimodal (teks, gambar, dan video) agar model mampu menangkap faktor-faktor yang memengaruhi viralitas secara lebih komprehensif.
2. Pada aspek pengembangan model, dapat dilakukan eksplorasi metode yang lebih adaptif dan kompleks, seperti penggunaan model *time series*, contohnya LSTM atau *Transformer-based forecasting* untuk meningkatkan performa, terutama pada prediksi *forecasting* kata yang masih memiliki tingkat error tinggi, serta arsitektur berbasis transformer untuk klasifikasi.
3. *Framework* VIRAL-Net dapat dikembangkan menjadi aplikasi yang lebih interaktif dan *real-time*, dengan integrasi API platform X serta implementasi berbasis *website*, sehingga dapat digunakan secara langsung oleh praktisi dalam merancang dan mengoptimalkan strategi *social media marketing campaign* berbasis data.

Halaman ini sengaja dikosongkan