

BAB V

PENUTUP

5.1 Simpulan

Berdasarkan implementasi yang telah dilakuakn mengenai prediksi jumlah views video pada kanal Youtube Dhaksinarga TV menggunakan algoritma *Random Forest Regression*, kesimpulan yang diperoleh dari seluruh proses pada implementasi yang dilakukan adalah sebagai berikut :

1. Dari 998 dataset yang telah dilakukan proses data splitting dengan menggunakan perbandingan sebesar 80:20 menghasilkan 800 data latih dan 200 data uji.
2. Berdasarkan hasil uji coba menggunakan metrik evaluasi R^2 Score dan Mean Absolute Percentage Error (MAPE), model *Random Forest Regression* yang diterapkan menunjukkan kinerja yang sangat baik. R^2 Score yang diperoleh sebesar 0.9258 atau 92.58%, yang mengindikasikan bahwa model mampu menjelaskan 92.58% variansi data secara akurat. Selain itu, nilai MAPE sebesar 11.58% menunjukkan bahwa kesalahan rata-rata prediksi model terhadap nilai aktual hanya sebesar 11.58%, yang termasuk dalam kategori akurasi tinggi.
3. Analisis korelasi antar fitur mengungkapkan bahwa likes dan share memiliki hubungan yang signifikan dengan jumlah views. Fitur *likes* menjadi faktor yang paling mempengaruhi jumlah *views* secara keseluruhan.
4. Hasil prediksi menunjukkan bahwa kategori "Pendidikan dan Pelatihan" memiliki rata-rata views tertinggi (455.94), sementara "Kesehatan dan Lingkungan" terendah (139.71). Optimalisasi konten, distribusi lebih luas, dan format interaktif dapat meningkatkan performa kategori tertinggi.

5.2 Saran

Hasil dari proyek ini masih memiliki kekurangan serta kelemahan, sehingga perlu adanya perbaikan untuk penelitian berikutnya agar memperoleh hasil yang lebih baik. Beberapa saran yang dapat dijadikan acuan untuk meningkatkan penelitian ini antara lain :

1. Disarankan untuk mempertahankan fokus pada fitur yang relevan dengan tujuan prediksi, namun dapat mempertimbangkan untuk meninjau kembali fitur-fitur yang dihapus atau mengembangkan cara untuk membuat fitur baru yang relevan. Hal ini bertujuan untuk memastikan bahwa model dapat menangkap pola yang lebih signifikan dan meningkatkan kualitas prediksi jumlah views video.
2. Lakukan tuning parameter pada algoritma *Random Forest Regression*, seperti jumlah pohon (*n_estimators*), kedalaman pohon (*max_depth*), dan *min_samples_split*. Teknik seperti *Grid Search* atau *Random Search* dapat digunakan untuk menemukan kombinasi parameter terbaik sehingga meningkatkan performa model.
3. Gunakan teknik validasi silang (*k-fold cross-validation*) untuk memastikan hasil evaluasi model lebih akurat dan tidak bergantung pada satu set data tertentu. Hal ini dapat meningkatkan kepercayaan terhadap performa model pada data yang belum pernah dilihat sebelumnya.
4. Kategori dengan prediksi views rendah, seperti "Kesehatan dan Lingkungan," dapat ditingkatkan dengan strategi seperti konten lebih interaktif, pemanfaatan tren, optimasi SEO YouTube, dan kolaborasi dengan pihak terkait.