

## ABSTRAK

Virtual YouTuber (VTuber) merupakan salah satu bentuk konten digital yang berkembang pesat di Indonesia. Tingkat *engagement* video menjadi indikator penting dalam menilai keberhasilan sebuah konten VTuber karena mencerminkan keterlibatan penonton melalui interaksi seperti *views*, *likes*, dan komentar. Pemahaman terhadap tingkat *engagement* penting bagi VTuber dan agensi untuk mengevaluasi performa konten serta merancang strategi pengembangan channel. Namun, penelitian yang secara khusus membangun model prediksi untuk mengukur tingkat *engagement* video VTuber Indonesia masih terbatas. Penelitian ini bertujuan untuk menerapkan algoritma *Random Forest* dalam memprediksi tingkat *engagement* video VTuber Indonesia. *Dataset* terdiri dari 37.406 data video VTuber Indonesia pada periode 2020-2024 yang dikumpulkan dari 50 *channel* yang diurutkan berdasarkan total *views channel*. Pengumpulan data dilakukan dengan metode *web scraping* menggunakan YouTube API, dengan fitur yang mencakup metrik interaksi video, durasi, usia video, usia *channel*, jumlah *subscribers*, waktu unggah, status *livestream*, kategori konten, panjang judul, jumlah tag, serta afiliasi agensi. Tahapan penelitian meliputi *data cleaning*, *data transformation*, *one-hot encoding* untuk fitur kategorikal, klasifikasi nilai *engagement* menggunakan batas kuartil dari metode IQR, pemilihan fitur, serta pembagian data latih dan data uji. Model *Random Forest* kemudian dilatih dengan pengaturan *hyperparameter*  $n\_estimators = 300$ ,  $max\_depth = 20$ ,  $min\_sample\_split = 5$ ,  $min\_sample\_leaf = 1$ , dan  $class\_weight = 'balanced'$ . Model menghasilkan akurasi sebesar 80,23%, dengan *macro average F1-score* 0,80 dan *weighted average F1-score* 0,81. Model mampu mengklasifikasikan tiga kelas *engagement* dengan baik, di mana kelas *Low* memiliki *F1-score* tertinggi (0,84), sedangkan kelas *High* menunjukkan performa terendah (0,76). Analisis *feature importance* mengidentifikasi *video\_age*, *duration*, *subscribers*, dan *channel\_age\_months* sebagai fitur yang paling berpengaruh dalam prediksi *engagement*.

**Kata kunci** : Virtual YouTuber, YouTube, *engagement*, *Random Forest*, klasifikasi