

TUGAS AKHIR
SKEMA MAGANG
ANALISIS DAN PREDIKSI *VIEWS* VIDEO PADA CHANNEL
YOUTUBE DHAKSINARGA TV MENGGUNAKAN
ALGORITMA *RANDOM FOREST REGRESSION* UNTUK
OPTIMALISASI STRATEGI KONTEN DENGAN
PENDEKATAN *CRISP-DM*



DWI INDAH ANDRIYANI

NIM : 215610052

PROGRAM STUDI SISTEM INFORMASI
PROGRAM SARJANA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA

2025

**TUGAS AKHIR
SKEMA MAGANG
ANALISIS DAN PREDIKSI *VIEWS* VIDEO PADA CHANNEL
YOUTUBE DHAKSINARGA TV MENGGUNAKAN
ALGORITMA *RANDOM FOREST REGRESSION* UNTUK
OPTIMALISASI STRATEGI KONTEN DENGAN
PENDEKATAN *CRISP-DM***

Diajukan sebagai salah satu syarat untuk menyelesaikan studi pada

**Program Sarjana
Program Studi Sistem Informasi
Fakultas Teknologi Informasi
Universitas Teknologi Digital Indonesia**

**Disusun Oleh
DWI INDAH ANDRIYANI
NIM : 215610052**

**PROGRAM STUDI SISTEM INFORMASI
PROGRAM SARJANA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN UJIAN TUGAS AKHIR

Judul : Analisis Dan Prediksi *Views* Video pada Channel Youtube Dhaksinarga TV Menggunakan Algoritma *Random Forest Regression* untuk Optimalisasi Strategi Konten dengan Pendekatan *CRISP-DM*

Nama : Dwi Indah Andriyani

NIM : 215610052

Program Studi : Sistem Informasi

Program : Sarjana


Semester : Ganjil

Tahun Akademik : 2024/2025

Telah diperiksa dan disetujui untuk diujikan di hadapan Dewan Penguji Tugas Akhir

Yogyakarta, 23 Januari 2025

Dosen Pembimbing,


Y. Yohakim Marwanta, S.Kom., M.Cs.
NIDN : 0026108101

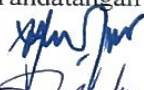


HALAMAN PENGESAHAN

ANALISIS DAN PREDIKSI *VIEW*S VIDEO PADA CHANNEL
YOUTUBE DHAKSINARGA TV MENGGUNAKAN
ALGORITMA *RANDOM FOREST REGRESSION* UNTUK
OPTIMALISASI STRATEGI KONTEN DENGAN
PENDEKATAN *CRISP-DM*

Telah dipertahankan di depan Dewan Penguji dan dinyatakan diterima untuk
memenuhi sebagian persyaratan guna memperoleh

Gelar Sarjana Komputer
Program Studi Sistem Informasi
Fakultas Teknologi Informasi
Universitas Teknologi Digital Indonesia

Yogyakarta, 10 Februari 2025

Dewan Penguji	NIDN	Tandatangan
1. Deborah Kurniawati, S.Kom.,M.Cs.	0511107301	
2. Y. Yohakim Marwanta, S.Kom., M.Cs.	0026108101	
3. Emy Susanti, S.Kom., M.Cs.	0003037901	

Mengetahui
Ketua Program Studi

Deborah Kurniawati, S.Kom.,M.Cs.
NIDN : 0511107301

PERNYATAAN KEASLIAN TUGAS AKHIR

Dengan ini saya menyatakan bahwa naskah Tugas Akhir ini belum pernah diajukan untuk memperoleh gelar Sarjana Komputer di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara sah diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 15 Januari 2025



Dwi Indah Andriyani
NIM: 215610052

HALAMAN PERSEMBAHAN

Pertama saya ucapkan puji syukur kehadirat Allah SWT atas segala nikmat berupa kesehatan, kekuatan, dan inspirasi yang sangat banyak dalam proses penyelesaian tugas akhir ini. Shalawat serta salam selalu terlimpahkan pada Nabi Muhammad SAW. Tugas akhir ini saya persembahkan sebagai bukti semangat, usaha, serta cinta dan kasih sayangku kepada orang-orang yang sangat berharga dalam hidupku.

Untuk karya yang sederhana ini, maka penulis persembahkan untuk :

1. Kedua orang tua tercinta yang selalu menjadi sumber semangat dan doa tak berujung, segala perjuangan ini tidak akan berarti tanpa kasih sayang dan pengorbanan kalian.
2. Kakak saya dan seluruh keluarga, yang selalu mendoakan dan memberikan dukungan selama ini.
3. Bapak Y. Yohakim Marwanta, S.Kom., M.Cs., selaku dosen pembimbing yang telah memberikan arahan dan bimbingan, dalam penyusunan tugas akhir.
4. Orang-orang baik yang tidak dapat disebutkan satu per satu, yang telah berjasa dan memberikan dukungan dengan cara yang tak terhitung, yang membantu sampai bisa berada di jenjang ini.
5. Terakhir, untuk diri saya sendiri yang sudah berjuang dan mampu bertahan hingga saat ini.

PRAKATA

Puji syukur atas kehadiran Allah SWT atas segala rahmat, hidayah, dan karunia-Nya, yang telah melimpahkan berkah-Nya sehingga dapat menyelesaikan penulisan tugas akhir. Tugas akhir ini disusun dalam rangka untuk memenuhi persyaratan untuk memperoleh gelar Sarjana Komputer Program Studi Sistem Informasi Universitas Teknologi Digital Indonesia.

Dalam momen bersejarah ini, izinkan penulis menyampaikan rasa terima kasih dan penghargaan kepada mereka yang telah memberikan dukungan, bimbingan, dan inspirasi dalam perjalanan panjang menuju pencapaian gelar ini. Dalam kesempatan ini penulis mengucapkan terima kasih kepada :

1. Ibu Sri Redjeki, S.Si., M.Kom., PhD., selaku Rektor Universitas Teknologi Digital Indonesia.
2. Ibu Deborah Kurniawati, S.Kom., M.Cs., selaku Kepala Prodi Sistem Informasi.
3. Bapak Y. Yohakim Marwanta, S.Kom., M.Cs., selaku pembimbing yang telah meluangkan waktunya untuk membimbing penulis.
4. Seluruh dosen dan staf karyawan Universitas Teknologi Digital Indonesia yang selama ini telah memberikan ilmunya kepada penulis.
5. Bapak Didik Handoko ST., dan Bapak Isnan Setiawan S.kom., selaku mentor yang telah memberikan ilmu pengetahuan, pengalaman baru, dan motivasi untuk selalu semangat dalam belajar hal baru selama proses magang.
6. Keluarga dan Orang Tua, terima kasih atas doa, dukungan, dan semangat yang telah diberikan selama perjalanan menyelesaikan pendidikan ini.
7. Orang-orang baik, khususnya Taree Anisah Herdanawati, Riska Novita Situmorang, Nailah Jeania Misharah, serta banyak lainnya yang tidak bisa disebutkan satu per satu. Terima kasih atas semangat, dukungan, serta bantuan yang telah diberikan selama proses penyusunan tugas akhir ini. Kehadiran dan kontribusi yang diberikan tidak hanya membantu

menyelesaikan tugas ini, tetapi juga menjadi sumber motivasi di setiap langkah perjalanan.

Penulis menyadari bahwa tugas akhir ini jauh dari kesempurnaan, oleh karena itu segala masukan baik saran ataupun kritik dari pembaca akan sangat berharga untuk perbaikan di masa mendatang. Semoga tugas akhir ini dapat memberikan manfaat bagi pembaca dan kontribusi nyata bagi perkembangan ilmu pengetahuan. Terima kasih atas perhatian dan kesempatan yang diberikan.

Yogyakarta, 15 Januari 2025



Dwi Indah Andriyani
NIM: 215610052

DAFTAR ISI

Hal

HALAMAN JUDUL.....	i
HALAMAN PERSETUJUAN UJIAN TUGAS AKHIR	iii
HALAMAN PENGESAHAN.....	iv
PERNYATAAN KEASLIAN TUGAS AKHIR	v
HALAMAN PERSEMBAHAN.....	vi
PRAKATA	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
INTISARI	xiii
ABSTRACT.....	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Deskripsi Pekerjaan	2
1.3 Tujuan	3
1.4 Manfaat	4
BAB II PROFIL INSTANSI TEMPAT MAGANG.....	5
2.1 Profil Perusahaan	5
2.2 Sejarah Perusahaan.....	5
2.3 Visi, Misi dan Tujuan Perusahaan	7
2.4 Struktur Organisasi.....	7
2.5 Deskripsi Tugas Struktur Organisasi	8
BAB III DESKRIPSI KEGIATAN.....	9
3.1 Persoalan	9
3.2 Implementasi Produk.....	9
3.2.1 Pengumpulan dan Pemrosesan Data	9
3.2.2 Analisis Faktor Engagement.....	10
3.2.3 Prediksi Jumlah Views	10
3.2.4 Rekomendasi Strategi Konten	11
3.2.5 Visualisasi dan Dokumentasi Hasil	11
3.3 Analisis dan Rancangan	11
3.3.1 Analisis Kebutuhan	11
3.3.2 Prosedur Kerja dan Pengumpulan Data	12

3.3.3 Analisis dan Rancangan Sistem.....	15
3.4 Jadwal Kerja.....	17
BAB IV PEMBAHASAN.....	18
4.1 Hasil	18
4.1.1 Pemahaman Bisnis	18
4.1.2 Pemahaman Data.....	18
4.1.3 Persiapan Data.....	24
4.1.4 Pemodelan.....	28
4.1.5 Evaluasi.....	30
4.1.6 Deployment.....	31
4.2 Uji coba.....	32
4.2.1 Uji Coba Menggunakan R^2 Score	32
4.2.2 Uji Coba Menggunakan <i>MAPE</i>	33
4.2.3 Visualisasi Hasil Uji Coba.....	34
4.3 Pembahasan.....	35
4.3.1 Analisis Korelasi	35
4.3.2 Rata-rata Prediksi <i>Views</i> Berdasarkan Kategori	36
4.3.3 <i>Feature Importance</i>	38
BAB V PENUTUP	39
5.1 Simpulan	39
5.2 Saran	40
DAFTAR PUSTAKA	41
LAMPIRAN.....	42

DAFTAR GAMBAR

Gambar 2.1 Logo Perusahaan.....	5
Gambar 2.2 Struktur Organisasi	7
Gambar 3.1 Diagram Alur Penelitian	13
Gambar 3.2 Alur Proses Analisis.....	15
Gambar 4.1 Kode <i>import Library Python</i>	22
Gambar 4.2 Kode Program fungsi <i>dead</i>	22
Gambar 4.3 Kode Program fungsi <i>info</i>	23
Gambar 4.4 Kode Program fungsi <i>describe</i>	24
Gambar 4.5 Kode penanganan <i>missing value</i>	25
Gambar 4.6 Kode Program fungsi <i>drop</i>	26
Gambar 4.7 Kode Program Transformasi Data	26
Gambar 4.8 Kode Program <i>Feature Engineering</i>	27
Gambar 4.9 Kode Program Perbandingan Jumlah & Rata-rata Prediksi.....	29
Gambar 4.10 Hasil Perbandingan Jumlah & Rata-rata Prediksi.....	30
Gambar 4.11 Kode Program <i>Scatter Plot</i>	31
Gambar 4.12 Kode Program <i>R² Score</i>	30
Gambar 4.13 Hasil Uji Coba <i>R² Score</i>	30
Gambar 4.14 Hasil Uji Coba <i>MAPE</i>	33
Gambar 4.15 Kode Program Visualisasi Uji Coba	33
Gambar 4.16 Hasil Visualisasi Uji Coba.....	34
Gambar 4.17 Visualisasi Heatmap Korelasi	35
Gambar 4.18 Visualisasi Rata-rata Prediksi <i>Views</i>	36
Gambar 4.19 Visualisasi <i>Feature Importance</i>	37

DAFTAR TABEL

Table 4.1 Atribut Variabel Dataset.....	21
Table 4.2 Rata-rata prediksi <i>views</i>	37

INTISARI

YouTube merupakan salah satu platform berbagi video yang memiliki pengaruh besar dalam menyampaikan informasi kepada masyarakat. Dalam proyek ini, dilakukan prediksi jumlah views video pada kanal YouTube Dhaksinarga TV berdasarkan kategori konten, dengan tujuan untuk memberikan rekomendasi strategi produksi konten yang lebih efektif.

Metode yang digunakan dalam proyek ini adalah *Random Forest Regression* yang diterapkan menggunakan Python. Dataset yang digunakan berisi data metrik performa video dengan jumlah total 998 data, diambil dari kanal Dhaksinarga TV. Data dibagi menjadi 80% untuk pelatihan dan 20% untuk pengujian. Selain itu, dilakukan analisis korelasi antar fitur untuk memahami hubungan antara variabel seperti *likes*, *shares*, *views*, *comments added*, *dislikes*, *subscribers lost*, *subscriber gained*, *shown in feed*, *average vies duration*, *watch time(hours)*, *subscribers*, *impressions*, *impressions click-through rate (%)*, dan *kategori*. Evaluasi menggunakan R^2 Score sebesar 92.58% menunjukkan bahwa model mampu menjelaskan 92.58% variabilitas data, sedangkan MAPE sebesar 11.58% menandakan tingkat kesalahan prediksi yang rendah.

Hasil prediksi menunjukkan bahwa kategori "Pendidikan dan Pelatihan" memiliki rata-rata views tertinggi (455.94 *views*), diikuti oleh "Kebudayaan" (266.13 *views*), sedangkan kategori "Kesehatan dan Lingkungan" memiliki rata-rata *views* terendah (139.71 *views*). Rekomendasi yang diberikan adalah memprioritaskan produksi konten dengan kategori berprediksi tinggi guna meningkatkan engagement dan jangkauan audiens.

Proyek ini berhasil menunjukkan bahwa pemanfaatan algoritma *Random Forest Regression* dapat menjadi alat yang efektif untuk memprediksi performa konten video berdasarkan data historis, serta memberikan wawasan untuk optimalisasi strategi konten di platform digital.

Kata kunci: *prediksi views, machine learning, random forest regression, YouTube.*

ABSTRACT

YouTube is one of the video sharing platforms that has a major influence in conveying information to the public. In this project, the number of video views on the Dhaksinarga TV YouTube channel is predicted based on content categories, with the aim of providing recommendations for more effective content production strategies.

The method used in this project is Random Forest Regression implemented using Python. The dataset used contains video performance metric data with a total of 998 data, taken from the Dhaksinarga TV channel. The data is divided into 80% for training and 20% for testing. In addition, a correlation analysis between features is carried out to understand the relationship between variables such as likes, shares, views, comments added, dislikes, subscribers lost, subscribers gained, shown in feed, average views duration, watch time (hours), subscribers, impressions, impressions click-through rate (%), and category. Evaluation using the R^2 Score of 92.58% indicates that the model is able to explain 92.58% of the data variability, while the MAPE of 11.58% indicates a low level of prediction error.

The prediction results show that the "Education and Training" category has the highest average views (455.94 views), followed by "Culture" (266.13 views), while the "Health and Environment" category has the lowest average views (139.71 views). The recommendation given is to prioritize content production with high-prediction categories to increase engagement and audience reach.

This project has successfully demonstrated that the use of the Random Forest Regression algorithm can be an effective tool for predicting video content performance based on historical data, as well as providing insights for optimizing content strategies on digital platforms.

Keywords: *machine learning, prediction of views, random forest regression, YouTube*