



**PREDIKSI VIEWERS CHANNEL YOUTUBE WARGANET LIFE  
OFFICIAL MENGGUNAKAN METODE RANDOM FOREST**

**SKRIPSI**

**MEIZA ALLIANSA**

**2110512115**

**UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA  
FAKULTAS ILMU KOMPUTER  
PROGRAM STUDI SISTEM INFORMASI**

**2025**



**PREDIKSI VIEWERS CHANNEL YOUTUBE WARGANET LIFE  
OFFICIAL MENGGUNAKAN METODE RANDOM FOREST**

**SKRIPSI**

**Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar  
Sarjana Komputer**

**MEIZA ALLIANSA**

**2110512115**

**UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA**

**FAKULTAS ILMU KOMPUTER**

**PROGRAM STUDI SISTEM INFORMASI**

**2025**

## PERNYATAAN ORISINALITAS

### PERNYATAAN ORISINALITAS

Skripsi ini merupakan hasil karya sendiri serta semua sumber referensi yang dikutip maupun yang dirujuk telah saya nyatakan benar.

Nama : Meiza Alliansa  
NIM : 2110512115  
Tanggal : 30 Januari 2025

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan berlaku.

Jakarta, 30 Januari 2025 Yang

Menyatakan,



Meiza Alliansa

## PERNYATAAN PERSTUJUAN PUBLIKASI SKRIPSI

### PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Meiza Alliansa  
NIM : 2110512115  
Fakultas : Ilmu Komputer  
Program Studi : S1 Sistem Informasi

Demi pembangunan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

#### **Prediksi Viewers Channel Youtube Warganet Life Official Menggunakan Metode Random Forest**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sesungguhnya.

Dibuat di : Jakarta

Pada Tanggal : 30 Januari 2025

Yang Menyatakan,



Meiza Alliansa

# LEMBAR PENGESAHAN

## LEMBAR PENGESAHAN

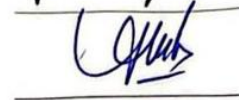
Judul : Prediksi Viewers Channel Youtube Warganet Life Official  
Menggunakan Metode Random Forest  
Nama : Meiza Alliansa  
NIM : 2110512115

Disetujui oleh :

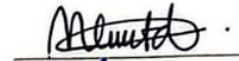
Penguji 1:  
Kraugusteeliana, S.Kom., M.Kom., M.M.



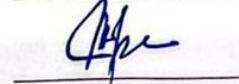
Penguji 2:  
Ati Zaidiah, S.Kom., M.Kom., MTI.



Pembimbing 1:  
Nur Hafifah Matondang, S.Kom., M.M., MTI.



Pembimbing 2:  
Rifka Dwi Amalia, S.Pd., M.Kom.



Diketahui oleh:

Koordinator Program Studi:  
Anita Muliawati, S.Kom., MTI.  
NIP. 197005212021212002



Dekan Fakultas Ilmu Komputer:  
Prof. Dr. Ir. Supriyanto, S.T., M.Sc., IPM  
NIP. 197605082003121002



Tanggal Ujian Tugas Akhir :  
13 Januari 2025

# **PREDIKSI VIEWERS CHANNEL YOUTUBE WARGANET LIFE OFFICIAL MENGGUNAKAN METODE RANDOM FOREST**

**Meiza Alliansa**

## **ABSTRAK**

Pesatnya pertumbuhan platform digital seperti *YouTube* menciptakan tantangan bagi pemilik *channel* dalam mengevaluasi kinerja konten mereka. *Channel YouTube "Warganet Life Official"* menghadapi kesulitan dalam menentukan apakah jumlah penayangan sebuah video memiliki performa yang baik atau tidak dari segi jumlah penonton. Untuk mengatasi masalah ini, penelitian ini menawarkan solusi berupa sistem prediksi jumlah penonton menggunakan algoritma *Random Forest*. Data yang digunakan berasal dari statistik *YouTube Studio channel* dengan atribut seperti jumlah suka, tidak suka, pembagian, waktu tonton, dan lainnya. Penelitian ini mengadopsi metode *Cross-Industry Standard Process for Data Mining (CRISP-DM)* yang melibatkan *preprocessing* data, pemodelan, evaluasi, dan implementasi sistem berbasis *website*. Sistem ini dirancang menggunakan *framework Streamlit* untuk mempermudah pengguna mengakses hasil prediksi. Hasil pengujian dengan data sebanyak 487 baris menunjukkan bahwa model prediksi dengan rasio pembagian data 70:30 menghasilkan akurasi terbaik, yaitu 87,22% dengan nilai MAPE sebesar 12,78% dan RMSE 196069.21. Model prediksi berbasis web ini mampu menampilkan prediksi jumlah penonton secara interaktif, baik melalui input manual maupun unggahan *file CSV*. Implementasi ini mempermudah pemilik *channel* dalam memberikan gambaran berapa banyak penayangan yang dihasilkan dalam waktu tertentu dan menentukan apakah konten memiliki performa baik atau tidak. Jika hasil performanya buruk, maka direkomendasikan untuk dievaluasi kontennya. Evaluasi maupun strategi konten yang akan dilakukan, akan diserahkan kepada pihak *channel*.

**Kata Kunci:** *YouTube, Random Forest, Sistem Prediksi, Prediksi Penayangan.*

**PREDICTING VIEWERS OF WARGANET LIFE OFFICIAL YOUTUBE  
CHANNEL USING THE RANDOM FOREST METHOD**

**Meiza Alliansa**

**ABSTRACT**

*The rapid growth of digital platforms like YouTube has created challenges for channel owners in evaluating the performance of their content. The YouTube channel "Warganet Life Official" faces difficulties in determining whether a video's view count performs well in terms of audience numbers. To address this issue, this research offers a solution in the form of a view count prediction system using the Random Forest algorithm. The data used comes from the YouTube Studio channel statistics, including attributes such as likes, dislikes, shares, watch time, and others. This study adopts the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology, which involves data preprocessing, modeling, evaluation, and the implementation of a web-based system. The system is designed using the Streamlit framework to facilitate users in accessing prediction results. Testing with 487 data entries showed that the prediction model with a 70:30 data split ratio yielded the best accuracy, achieving 87.22% with a MAPE of 12.78% and RMSE of 196069.21. The web-based prediction model allows for interactive view count predictions through both manual input and CSV file uploads. This implementation helps channel owners gain insights into how many views will be generated within a certain timeframe and determine whether the content is performing well or not. If the performance is poor, content evaluation is recommended. The evaluation or content strategy to be implemented will be handed over to the channel management team.*

**Keywords:** *YouTube, Random Forest, Prediction System, View Prediction.*

## KATA PENGANTAR

Puji syukur kehadiran Allah SWT, Tuhan Yang Maha Esa, atas limpahan rahmat, hidayah, serta karunia-Nya yang tak terhingga, sehingga penulis dapat menyelesaikan skripsi ini sebagai salah satu syarat untuk menyelesaikan studi di Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jakarta.

Dalam proses penyusunan skripsi ini, banyak pihak yang telah memberikan dukungan, bantuan, dan doa yang tulus. Oleh karena itu, dengan segala hormat, penulis mengucapkan terima kasih kepada:

1. Kedua orang tua saya yang senantiasa memberikan dukungan tanpa henti, baik secara moral maupun materil, serta menjadi sumber motivasi terbesar dalam perjalanan ini.
2. Bapak Prof. Dr. Ir. Supriyanto, ST., M.Sc., IPM, selaku Dekan Fakultas Ilmu Komputer.
3. Ibu Anita Muliawati, S.Kom., MTI., selaku Ketua Program Studi S1 Sistem Informasi.
4. Ibu Nur Hafifah Matondang, S.Kom., MM., MTI, selaku dosen pembimbing 1, yang telah memberikan arahan, masukan, dan koreksi dengan penuh kesabaran selama penulisan skripsi ini.
5. Ibu Rifka Dwi Amalia, S.Pd., M.Kom., selaku dosen pembimbing 2, yang juga memberikan dukungan dan panduan berharga dalam penyelesaian penelitian ini.
6. Teman-teman seperjuangan yang telah menjadi tempat berbagi semangat, ide, serta dukungan dalam menghadapi berbagai tantangan selama proses penulisan skripsi.
7. Inisial ANKZ yang selalu menemani dan memberikan dukungan dari awal sampai seterusnya.
8. Seluruh pihak Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jakarta yang tidak dapat saya sebutkan satu per satu tanpa mengurangi rasa hormat.



Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi perbaikan di masa mendatang. Semoga hasil dari penelitian ini dapat memberikan manfaat, baik bagi pembaca maupun dunia akademik.

Jakarta, 30 Januari 2025

Meiza Alliansa

## DAFTAR ISI

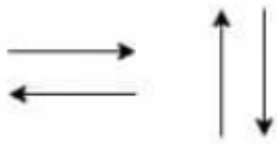
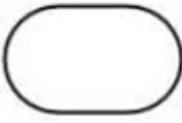

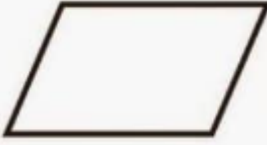
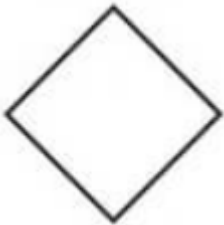
PERNYATAAN ORISINALITAS .....	iii
PERNYATAAN PERSTUJUAN PUBLIKASI SKRIPSI.....	iv
LEMBAR PENGESAHAN .....	v
ABSTRAK .....	vi
<i>ABSTRACT</i> .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI.....	x
DAFTAR SIMBOL.....	xiii
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR .....	xv
DAFTAR LAMPIRAN.....	xvi
DAFTAR RUMUS .....	xvii
BAB I.....	1
1.1. Latar Belakang .....	1
1.2. Rumusan Masalah .....	4
1.3. Batasan Penelitian .....	4
1.4. Tujuan Penelitian.....	4
1.5. Manfaat Penelitian.....	5
1.6. Luaran yang diharapkan .....	5
1.7. Sistematika Penulisan.....	5
BAB II.....	7
2.1. <i>YouTube</i> .....	7
2.2. <i>Channel YouTube</i> .....	7
2.3. Strategi Konten.....	8
2.4. Prediksi.....	9
2.5. <i>Data Mining</i> .....	9
2.6. Algoritma <i>Random Forest</i> .....	12
2.7. <i>Python</i> .....	16
2.8. Evaluasi Model.....	17
2.9. <i>Streamlit</i> .....	19
2.10. <i>Blackbox Testing</i> .....	20

2.1.1. Penelitian yang Relevan .....	20
BAB III .....	27
3.1. Alur Penelitian.....	27
3.1.1 Identifikasi Masalah .....	28
3.1.2 Persiapan Data .....	28
3.1.3 Pemahaman Data .....	28
3.1.4 <i>Peprocessing</i> Data .....	29
3.1.5 <i>Modeling</i> .....	30
3.1.6 Evaluasi Model.....	30
3.1.7 Implementasi .....	30
3.1.8 <i>Testing</i> .....	31
3.2 Alat Bantu Penelitian .....	31
3.3 Jadwal Rencana Penelitian .....	32
BAB IV .....	34
4.1. Persiapan Data.....	34
4.2. Pemahaman Data.....	35
4.3. <i>Preprocessing</i> Data .....	36
4.3.1 <i>Data Cleansing</i> .....	36
4.3.2 <i>Data Transformation</i> .....	38
4.3.3 <i>Feature Selection</i> .....	39
4.4. Modeling .....	41
4.4.1 Pembagian Data.....	41
4.4.2 Implementasi Model <i>Random Forest</i> .....	42
4.5 Evaluasi Model.....	50
4.5.1 Performa Model Rasio 70% - 30% .....	50
4.5.2 Performa Model Rasio 80% - 20% .....	51
4.5.3 Performa Model Rasio 90% - 10% .....	52
4.6 Implementasi .....	54
4.6.1 Menu Input Manual .....	54
4.6.2 Menu Input Unggah CSV .....	57
4.7 Testing.....	62
4.8 Perbandingan Dengan Penelitian Terdahulu .....	63

BAB V.....	66
5.1. Kesimpulan.....	66
5.2. Saran.....	67
DAFTAR PUSTAKA .....	68
LAMPIRAN.....	73

## DAFTAR SIMBOL

### 1. Notasi Diagram Alur Penelitian

Simbol	Nama	Keterangan
	<i>Flow</i>	Garis yang menggambarkan arah aktivitas atau proses.
	<i>Terminator</i>	Simbol untuk mengawali ( <i>start</i> ) dan mengakhiri ( <i>stop</i> ) sebuah <i>flowchart</i>
	<i>Process</i>	Simbol yang menjelaskan sebuah proses atau aktivitas.
	<i>Input/Output</i>	Simbol yang menggambarkan masukan ( <i>input</i> ) dan/atau keluaran ( <i>output</i> ) tanpa tergantung peralatan.
	<i>Decision</i>	Simbol yang menggambarkan percabangan pilihan dalam melakukan pengambilan keputusan antara <i>true</i> atau <i>false</i> .

## DAFTAR TABEL

Tabel 1 Penelitian yang Relevan.....	20
Tabel 2 Jadwal Rencana Penelitian.....	32
Tabel 3 Daftar Atribut pada Dataset .....	34
Tabel 4 Informasi Nilai dan Tipe Data Setiap Atribut.....	35
Tabel 5 Jumlah Baris Bernilai <i>Null</i> .....	36
Tabel 6 Perbandingan Rasio Pembagian Data <i>Training</i> dan Data <i>Testing</i> .....	41
Tabel 7 Parameter dan Nilai pada Algoritma <i>Random Forest</i> .....	42
Tabel 8 Asumsi 5 Data Teratas Sebagai Pengukuran Akurasi Prediksi.....	47
Tabel 9 Perhitungan Kuadrat dari Selisih .....	49
Tabel 10 Hasil Performa Model dengan Rasio 70% - 30% .....	50
Tabel 11 Hasil Performa Model dengan Rasio 80% - 20% .....	51
Tabel 12 Hasil Performa Model dengan Rasio 90% - 10% .....	52
Tabel 13 Evaluasi Model pada Masing-Masing Rasio .....	53
Tabel 14 <i>Blackbox Testing</i> Oleh Penguji 1 .....	104
Tabel 15 <i>Blackbox Testing</i> Oleh Penguji 2 .....	105
Tabel 16 <i>Blackbox Testing</i> Oleh Penguji 3 .....	106

## DAFTAR GAMBAR

Gambar 2.1 Tahapan <i>Knowledge Discovery in Databases</i> .....	11
Gambar 2.2 <i>Random Forest Tree</i> .....	14
Gambar 3.1 Alur Penelitian.....	27
Gambar 4.1 Contoh Data <i>Null</i> .....	37
Gambar 4.2 Preview Data <i>Clean</i> .....	38
Gambar 4.3 Dataset setelah transformasi kolom ‘Waktu publikasi video’ .....	39
Gambar 4.4 Heatmap korelasi antar atribut .....	40
Gambar 4.5 Contoh Pohon Keputusan pada Algoritma <i>Random Forest</i> .....	44
Gambar 4.6 Contoh Sampel Data Hasil <i>Bootstrapping</i> .....	45
Gambar 4.7 Feature Importance pada model <i>Random Forest</i> .....	46
Gambar 4.8 Perbandingan Data Aktual dengan Data Prediksi .....	51
Gambar 4.9 Perbandingan Data Aktual dengan Data Prediksi .....	52
Gambar 4.10 Perbandingan Data Aktual dengan Data Prediksi .....	53
Gambar 4.11 Tampilan Web Pada Menu <i>Input Manual</i> Bagian 1 .....	55
Gambar 4.12 Tampilan Web Pada Menu <i>Input Manual</i> Bagian 2.....	55
Gambar 4.13 Contoh Data Untuk Input Atribut Pada Menu <i>Input Manual</i> .....	56
Gambar 4.14 Tampilan <i>Web</i> Setelah Melakukan Prediksi Penayangan Pada Input Manual.....	56
Gambar 4.15 Tampilan <i>Web</i> Setelah Input Nilai Aktual Penayangan .....	57
Gambar 4.16 Tampilan <i>Web</i> Pada Menu Input Unggah CSV Bagian 1 .....	58
Gambar 4.17 Tampilan <i>Web</i> Pada Menu Input Unggah CSV Bagian 2 .....	58
Gambar 4.18 Tampilan <i>Web</i> Setelah Mengunggah <i>File</i> CSV .....	59
Gambar 4.19 5 Sampel Data Pada File CSV yang Diunggah .....	59
Gambar 4.20 Tabel Hasil Prediksi Pada 5 Sampel Data .....	60
Gambar 4.21 Visualisasi Data Hasil Prediksi Pada Menu Unggah CSV .....	61
Gambar 4.22 Tampilan Saat Mengunduh Hasil Prediksi Dalam Bentuk <i>Excel</i> .....	61
Gambar 4.23 Hasil Unduhan <i>File Excel</i> Bagian 1 .....	61
Gambar 4.24 Hasil Unduhan <i>File Excel</i> Bagian 2 .....	62

## DAFTAR LAMPIRAN

Lampiran 1 Surat izin riset .....	74
Lampiran 2 Bukti pengiriman data oleh pihak Warganet <i>Life Official</i> melalui <i>email</i> .....	75
Lampiran 3 Hasil wawancara .....	75
Lampiran 4 <i>Source code</i> prediksi <i>Random Forest</i> pada penayangan <i>channel</i> <i>YouTube</i> “Warganet <i>Life Official</i> ” .....	75
Lampiran 5 <i>Source code deploy</i> model pada <i>website</i> menggunakan <i>streamlit</i> .....	82
Lampiran 6 Dataset detail .....	85
Lampiran 7 Testing <i>website</i> prediksi oleh pihak <i>channel</i> .....	91
Lampiran 8 Hasil cek plagiasi Turnitin .....	92
Lampiran 9 Surat keterangan penelitian .....	103
Lampiran 10 Hasil lengkap pengujian <i>Black Box Testing</i> . .....	104



## DAFTAR RUMUS

Rumus 1 Rumus <i>Gini Impurity</i> .....	15
Rumus 2 Rumus Informasi <i>Gain</i> .....	15
Rumus 3 Rumus MAPE .....	17
Rumus 4 Rumus Akurasi .....	18
Rumus 5 Rumus RMSE .....	18