

# PEMODELAN PREDIKTIF TRAFIK WEBSITE BERDASARKAN VOLUME KONTEN: PENDEKATAN REGRESI LINEAR SEDERHANA

Hasanatul Iftitah<sup>1</sup>, Nindy Raisa Hanum<sup>2</sup>

<sup>12</sup>Fakultas Sains dan Teknologi, Universitas Jambi, Kota Jambi

E-mail: <sup>1</sup>hasanatul.iftitah@unja.ac.id, <sup>2</sup>nindyraisahanum01@unja.ac.id

**Abstract** In today's digital landscape, a website's performance serves as a key metric of an institution's online presence and communication strategy. This research focuses on forecasting website performance by analyzing the relationship between the number of published articles and the volume of page views using a simple linear regression approach. Monthly data was obtained from the official website of the Faculty of Science and Technology at Universitas Jambi, comprising content publication frequency and corresponding traffic. This data was collected manually and via the Google Analytics platform for the period from January 2024 to March 2025. The analysis reveals a strong positive correlation, where each additional published article contributes to a notable increase in page views. The regression model yields a coefficient of 103.75 with an  $R^2$  value of 0.7278, indicating that over 72% of traffic variation is attributable to content volume. These results emphasize the importance of consistent content production in enhancing web visibility and provide valuable insights for content strategy development.

**Keywords:** Web performance, content strategy, linear regression model, page view analysis, digital content optimization

## I. PENDAHULUAN

### 1.1. Latar Belakang

Di era digital saat ini, website menjadi salah satu media utama dalam penyebaran informasi dan representasi identitas suatu institusi. Website yang aktif, informatif, dan terstruktur dengan baik dapat meningkatkan kredibilitas dan daya tarik suatu organisasi, termasuk dalam konteks pendidikan, bisnis, maupun pemerintahan. Salah satu elemen penting dalam pengelolaan website adalah konsistensi dan kualitas konten yang diterbitkan. Konten yang relevan, *up to date*, dan mudah diakses akan menarik lebih banyak pengunjung, meningkatkan engagement, serta memperpanjang durasi kunjungan di halaman website (Järvinen & Karjaluoto, 2015).

Jumlah konten yang dipublikasikan dalam suatu periode diyakini memiliki keterkaitan dengan trafik website, khususnya dalam jumlah *views* atau kunjungan halaman. Namun, hubungan antara frekuensi penerbitan konten dengan jumlah *views* tidak selalu bersifat linier. Dalam beberapa studi ditemukan bahwa efektivitas konten sangat dipengaruhi oleh konteks, strategi SEO, dan keterlibatan audiens, bukan hanya dari kuantitas konten yang diunggah (Plaza, 2011). Oleh karena itu, diperlukan kajian berbasis data untuk melihat pola hubungan antara jumlah konten terbitan dan performa website, khususnya dari sisi jumlah *views* yang diperoleh.

### 1.2. Rumusan Masalah

Berdasarkan latar belakang tersebut, maka rumusan masalah dalam penelitian ini adalah:

1. Apakah terdapat hubungan antara jumlah konten yang diterbitkan dengan jumlah *views* pada website?
2. Sejauh mana jumlah konten dapat memengaruhi kinerja website dari sisi kunjungan halaman?
3. Bagaimana model regresi dapat digunakan untuk memprediksi trafik website berdasarkan jumlah konten yang diterbitkan?

### 1.3. Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Menganalisis hubungan antara jumlah konten yang diterbitkan dengan jumlah *views* pada website.
2. Membangun model prediktif berbasis regresi linier untuk memperkirakan kinerja website berdasarkan jumlah konten terbitan.
3. Memberikan insight strategis bagi pengelola website dalam mengoptimalkan frekuensi dan jumlah penerbitan konten guna meningkatkan performa website secara keseluruhan.

## II. TINJAUAN PUSTAKA

### 2.1 Teori Regresi Linier Sederhana

Regresi linier sederhana adalah metode statistik yang digunakan untuk memodelkan hubungan antara satu variabel independen (*predictor*) dengan satu variabel dependen (*response*) dalam bentuk persamaan linier. Regresi linier digunakan secara luas dalam bidang ilmu sosial, ekonomi, serta data digital karena kemudahannya dalam interpretasi dan implementasi.

Dalam konteks analisis performa website, regresi linier sederhana dapat digunakan untuk

mengevaluasi apakah terdapat hubungan antara jumlah konten yang dipublikasikan dan jumlah kunjungan yang diterima, sehingga dapat menjadi alat bantu dalam pengambilan keputusan berbasis data.

#### Persamaan Umum Regresi Linier Sederhana:

$$Y = a + bX + \varepsilon$$

Keterangan:

- $Y$  : variabel dependen (jumlah *views*)
- $X$  : variabel independen (jumlah konten)
- $a$  : *intercept* (nilai  $Y$  saat  $X = 0$ )
- $b$  : koefisien regresi (kemiringan garis regresi)
- $\varepsilon$  : error atau residual

#### Perhitungan Koefisien Regresi:

$$b = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2}$$

$$a = \bar{Y} - b\bar{X}$$

Keterangan:

- $\bar{X}$  : rata-rata dari variabel  $X$
- $\bar{Y}$  : rata-rata dari variabel  $Y$

#### Evaluasi Model: Koefisien Determinasi ( $R^2$ )

$$R^2 = 1 - \frac{\sum(Y_i - \check{Y}_i)^2}{\sum(Y_i - \bar{Y})^2}$$

$R^2$  menunjukkan proporsi variabilitas  $Y$  yang bisa dijelaskan oleh  $X$ . Nilai  $R^2$  mendekati 1 menandakan model prediksi yang baik. Dalam penelitian ini, pemilihan  $R^2$  (koefisien determinasi) sebagai metrik evaluasi model regresi linier sederhana dinilai paling sesuai karena sifat model yang hanya melibatkan satu variabel independen dan fokus analisis pada kekuatan hubungan linier antara jumlah konten dan jumlah *views*.  $R^2$  mampu menunjukkan proporsi variasi pada variabel dependen yang dapat dijelaskan oleh variabel independen, sehingga memberikan interpretasi yang lebih intuitif dan langsung terkait efektivitas model dalam menjelaskan fenomena yang diteliti. Dibandingkan dengan metrik lain seperti MAE, MSE, atau RMSE yang lebih menekankan pada kesalahan prediksi individual,  $R^2$  lebih tepat digunakan ketika tujuan utama penelitian adalah untuk memahami seberapa kuat pengaruh suatu variabel terhadap variabel lain dalam konteks hubungan linier, bukan untuk akurasi prediktif jangka panjang atau generalisasi model pada data baru. Oleh karena itu,  $R^2$  menjadi indikator evaluasi yang paling relevan dalam studi ini.

## 2.2 Studi Relevan tentang *Digital Content* dan *Traffic*

Hubungan antara jumlah konten digital dan trafik website telah menjadi fokus perhatian banyak penelitian. Semakin banyak konten yang dipublikasikan, terutama yang relevan dan berkualitas, umumnya akan meningkatkan visibilitas situs web dalam mesin pencari dan menarik lebih banyak pengunjung (Pulizzi, 2014). Strategi ini dikenal sebagai *content marketing*, yang menjadi fondasi utama dalam pertumbuhan trafik organik.

Penelitian oleh Ahmed et al. (2018) menemukan bahwa terdapat korelasi positif antara frekuensi publikasi konten dan peningkatan trafik website, terutama pada situs-situs edukatif dan berita. Dalam studi mereka, setiap peningkatan jumlah artikel per minggu berkontribusi pada kenaikan trafik hingga 15% dalam periode satu bulan. Temuan serupa diungkapkan oleh Kumar et al. (2020) yang menunjukkan bahwa halaman dengan update konten secara rutin lebih mungkin untuk mendapatkan pengindeksan ulang oleh mesin pencari, sehingga meningkatkan performa SEO dan trafik pengunjung.

Di sisi lain, Chaffey & Ellis-Chadwick (2019) menekankan pentingnya kualitas dan relevansi konten. Mereka berargumen bahwa kuantitas konten tidak selalu menjamin peningkatan trafik, terutama jika konten tidak sesuai dengan kebutuhan atau minat pengunjung. Oleh karena itu, mereka menyarankan pendekatan berbasis data (*data-driven content strategy*) dalam pengelolaan situs web.

Beberapa platform besar seperti HubSpot dan Moz juga menggarisbawahi bahwa konten yang disertai dengan visualisasi interaktif, judul yang menarik, dan optimasi SEO cenderung mendapatkan lebih banyak trafik dibandingkan konten statis biasa. Hal ini menunjukkan pentingnya tidak hanya jumlah, tetapi juga kualitas dan strategi publikasi.

## 2.3 Peran *Data Analytics* dalam Optimasi Website

Penggunaan *data analytics* dalam pengelolaan website telah menjadi praktik standar dalam transformasi digital organisasi. Melalui alat seperti *Google Analytics*, pengelola dapat mengakses berbagai metrik performa, seperti *page views*, *bounce rate*, *time on site*, hingga asal trafik. Informasi ini memungkinkan dilakukannya pengambilan keputusan berbasis data dalam menentukan strategi konten, desain halaman, serta efektivitas kampanye digital (Chaffey & Ellis-Chadwick, 2019).

Data analytics juga memungkinkan dilakukannya prediksi performa website menggunakan pendekatan *machine learning* dan statistik. Hal ini selaras dengan temuan Järvinen &

Karjaluoto (2015), yang menyatakan bahwa penggunaan data analitik terbukti meningkatkan efisiensi strategi pemasaran digital dan keterlibatan pengguna..

### III. HASIL DAN PEMBAHASAN

#### 3.1 Sumber Data

Penelitian ini menggunakan data internal dari situs resmi Fakultas Sains dan Teknologi Universitas Jambi (<https://fst.unja.ac.id>). Data diperoleh langsung oleh penulis selaku salah satu admin website, yang memiliki akses terhadap histori jumlah konten yang diterbitkan dan jumlah kunjungan (*views*) yang tercatat tiap bulan. Data ini dikumpulkan secara manual dan melalui *platform Google Analytics* untuk periode Januari 2024 hingga Maret 2025.

#### 3.2 Variabel Penelitian

- Variabel independen (X): Jumlah konten yang diterbitkan setiap bulan.
- Variabel dependen (Y): Jumlah *views* atau kunjungan halaman website per bulan.

#### 3.3 Alat dan Teknik Analisis

Analisis data dilakukan menggunakan *Python* dan *Microsoft Excel*. Teknik utama yang digunakan adalah regresi linier sederhana, yaitu metode untuk menemukan hubungan linier antara dua variabel kuantitatif.

#### 3.4 Langkah Kerja

Berikut ini adalah Langkah-langkah kerja penelitian:

1. **Pengumpulan Data:** Data jumlah konten dan jumlah *views* dikumpulkan per bulan.
2. **Visualisasi:** Data divisualisasikan dengan *scatter plot*.
3. **Pembangunan Model:** Model regresi linier sederhana dibangun menggunakan pustaka *scikit-learn*.
4. **Evaluasi:** Performa model dianalisis menggunakan nilai  $R^2$ .

#### 3.5 Analisis Pemodelan

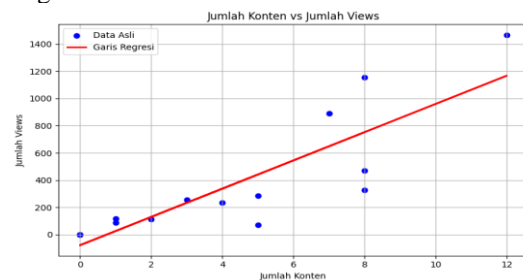
Pada Tabel 1 menunjukkan data histori jumlah konten dan jumlah *views* per bulan yang diambil secara manual.

Tabel 1. Data Jumlah Konten dan Jumlah *Views*

Bulan	Jumlah Konten	Jumlah Views
Januari 2024	3	256
Februari 2024	7	891
Maret 2024	8	326

April 2024	8	469
Mei 2024	1	87
Juni 2024	0	0
Juli 2024	0	0
Agustus 2024	1	117
September 2024	12	1467
Oktober 2024	8	1156
November 2024	2	112
Desember 2024	4	234
Januari 2025	5	284
Februari 2025	0	0
Maret 2025	5	71

Gambar di bawah menunjukkan *scatter plot* antara jumlah konten yang diterbitkan per bulan dengan jumlah *views* pada website FST UNJA. Titik-titik biru merepresentasikan data aktual, sementara garis merah merupakan garis regresi linier yang dibentuk dari model prediktif. Pola data menunjukkan tren kenaikan, di mana secara umum semakin banyak jumlah konten yang dipublikasikan, jumlah *views* juga cenderung meningkat, meskipun terdapat beberapa titik anomali yang menyimpang dari garis tren.



Gambar 1. *scatter plot* Data asli dan Garis Regresi

Model regresi linier sederhana yang dihasilkan dari data memiliki bentuk sebagai berikut:

$$\hat{Y} = 103.75X - 78.01$$

dengan:

- $\hat{Y}$  : Prediksi jumlah *views*
- $X$  : Jumlah konten yang diterbitkan dalam sebulan
- 103.75: Koefisien regresi
- -78.01: *Intercept* (nilai *views* saat jumlah konten = 0)

Interpretasi dari koefisien regresi ini adalah: setiap penambahan 1 konten diprediksi akan meningkatkan jumlah *views* sebanyak 103.75 kunjungan, dengan asumsi variabel lain tetap.

Analisis regresi linier sederhana yang dilakukan menunjukkan adanya hubungan positif yang signifikan antara jumlah konten yang diterbitkan per bulan dengan jumlah kunjungan (*views*) yang diterima oleh website FST UNJA. Model menghasilkan nilai koefisien regresi sebesar

103,75, yang mengindikasikan bahwa setiap penambahan satu konten berpotensi meningkatkan jumlah views sebanyak rata-rata 103,75 kunjungan, dengan asumsi faktor lain tetap.

Nilai koefisien determinasi ( $R^2 = 0,7278$ ) menunjukkan bahwa sekitar 72,78% variasi jumlah views dapat dijelaskan oleh jumlah konten yang diterbitkan. Ini merupakan indikator bahwa jumlah konten merupakan prediktor yang cukup kuat terhadap performa website dalam hal trafik. Nilai ini cukup tinggi untuk ukuran model sederhana, sehingga dapat dikatakan bahwa model cukup baik dalam merepresentasikan hubungan antara kedua variabel. Hasil ini juga mendukung pandangan dalam literatur bahwa peningkatan jumlah konten secara konsisten dapat meningkatkan visibilitas digital dan menjangkau audiens yang lebih luas.

Meski demikian, terdapat beberapa outlier dalam data yang mengindikasikan bahwa hubungan antara jumlah konten dan jumlah views tidak selalu bersifat linier sempurna. Misalnya, pada bulan-bulan tertentu terdapat konten dalam jumlah sedikit namun memperoleh views yang relatif tinggi, dan sebaliknya. Hal ini mencerminkan adanya variabel lain di luar jumlah konten yang turut memengaruhi kinerja website, seperti kualitas konten, faktor SEO, tren topik, waktu publikasi, serta keterlibatan pengguna.

Model yang dibangun juga memberikan dasar prediktif yang dapat digunakan untuk menyusun strategi publikasi konten. Sebagai contoh, jika target kunjungan dalam satu bulan adalah 1.000 views, maka berdasarkan model regresi, dibutuhkan penerbitan sekitar 10 konten untuk mendekati target tersebut. Dengan demikian, hasil analisis ini dapat berkontribusi langsung pada pengambilan keputusan strategis dalam pengelolaan konten digital secara lebih efisien dan berbasis data.

### 3.6 Implikasi Strategis

- Penerbitan konten yang konsisten dan cukup banyak dapat meningkatkan trafik.
- Namun, kuantitas perlu diiringi kualitas dan relevansi konten agar setiap konten memberikan dampak maksimal.
- Evaluasi rutin terhadap performa konten dan analisis data bulanan dapat membantu dalam optimalisasi pengelolaan website berbasis data.

## IV. PENUTUP

### 4.1. Kesimpulan

Penelitian ini menunjukkan adanya hubungan linier positif yang signifikan antara jumlah konten yang diterbitkan dengan jumlah kunjungan (views) pada website Fakultas Sains dan Teknologi Universitas Jambi. Melalui penerapan model regresi

linier sederhana, diperoleh persamaan:

$$\hat{Y} = 103.75X - 78.01$$

Yang mengindikasikan bahwa setiap penambahan satu unit konten secara umum berkontribusi terhadap peningkatan jumlah views website. Nilai koefisien determinasi ( $R^2$ ) sebesar 0,7278 menunjukkan bahwa sekitar 72,78% variasi dalam jumlah views dapat dijelaskan oleh jumlah konten yang diterbitkan. Temuan ini menggarisbawahi pentingnya strategi pengelolaan konten yang terencana dan berbasis data guna meningkatkan performa website secara keseluruhan. Dengan demikian, model ini dapat dijadikan dasar awal dalam merancang strategi pengelolaan konten berbasis data yang lebih sistematis dan terukur.

### 4.2. Saran

1. Optimalisasi konten berkualitas. Selain meningkatkan kuantitas, tim pengelola website perlu memperhatikan kualitas dan relevansi konten agar lebih berdampak terhadap peningkatan trafik.
2. Integrasi Analisis SEO. Penelitian lanjutan dapat memasukkan metrik SEO seperti keyword density, bounce rate, dan CTR sebagai variabel tambahan dalam model prediksi.
3. Perluasan Model Prediktif. Untuk meningkatkan akurasi, disarankan untuk mengeksplorasi model statistik atau machine learning lain seperti regresi berganda, decision tree, atau random forest dengan melibatkan lebih banyak variabel.

## DAFTAR REFERENSI

- Ahmed, S., Lodhi, M. S., & Ahmad, M. (2018). *Impact of Website Content on Traffic Generation*. International Journal of Computer Science and Network Security, 18(2), 17–21.
- Chaffey, D., & Ellis-Chadwick, F. (2019). *Digital Marketing* (7th ed.). Pearson Education.
- Chatterjee, S., & Hadi, A. S. (2015). *Regression Analysis by Example* (5th ed.). Wiley.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning with Applications in R*. Springer.
- Järvinen, J., & Karjaluoto, H. (2015). *The use of Web analytics for digital marketing performance measurement*. Industrial Marketing Management, 50, 117–127. <https://doi.org/10.1016/j.indmarman.2015.04.009>
- Kalhor, B., & Nikravanshalmani, A. (2020). Correlation between Content and Traffic of the Universities Website. *arXiv: Digital Libraries*. <https://arxiv.org/abs/2003.07097>

- Kumar, R., Singh, D., & Suri, P. K. (2020). *Predictive Analytics for Web Content Effectiveness: A Review*. *Procedia Computer Science*, 167, 2267–2276. <https://doi.org/10.1016/j.procs.2020.03.277>
- Mission, R. (2023). *Website Traffic Patterns and User Behavior: A Comprehensive Study of Visitor Interactions and Engagement Metrics*. <https://doi.org/10.69478/jitc2023v5n1a02>
- Molchanova, R. V. (2024). *Analysis of the influence of factors on traffic: from seo to content marketing*. *Ekonomika i Upravljenje: Problemy, Rešeniâ*, 10/8(151), 111–117. <https://doi.org/10.36871/ek.up.p.r.2024.10.08.015>
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis* (5th ed.). Wiley.
- Nindy, NRH., Ifitah, H., Waladi, A., Perdana, Y. (2024). *Implementation of Machine Learning for Stock Price Prediction Using the LSTM Algorithm*. *Media Journal of Information System and Informatic* 1 (1), 31-37. <https://doi.org/10.62205/hfvfmj88>
- Plaza, B. (2011). *Google Analytics for measuring website performance*. *Tourism Management*, 32(3), 477–481. <https://doi.org/10.1016/j.tourman.2010.03.015>
- Pulizzi, J. (2014). *Epic Content Marketing: How to Tell a Different Story, Breakthrough the Clutter, and Win More Customers by Marketing Less*. McGraw-Hill Education.
- Waladi, A., Perdana, Y., Ifitah, H., & Hanum, NR. (2024) *Stock Price Prediction Using Machine Learning-Based on RNN Algorithms*. *Media Journal of Information System and Informatic* 1 (1), 1-8. <https://doi.org/10.62205/dg2h0j98>

#### IDENTITAS PENULIS

Nama : Hasanatul Ifitah  
NIDN/NIP : 199604022024062002  
TTL : Jambi, 2 April 1996  
Golongan / Pangkat : IIIb/ Penata Muda Tingkat I  
Alamat Rumah : Perumahan Permata Kenali Asri

Nama : Nindy Raisa Hanum  
NIDN/NIP : 199607052024062001  
TTL : 5 Juli 1996  
Golongan / Pangkat : IIIb/ Penata Muda Tingkat I  
Alamat Rumah : Mayang, Kota Jambi