

Pengaruh Konten YouTube terhadap Penerimaan Mahasiswa Baru: Studi Analisis Sentimen BERT dan Korelasi Spearman

Edi Widodo^{1*}, Eka Putri Rachmawati², Pratama Angga Buana³

^{1,2,3} Program Studi Sistem Informasi, Fakultas Teknologi Informasi dan Komunikasi,
Universitas Semarang

*Email: ediwidodo@usm.ac.id

Abstrak

Penelitian ini bertujuan menganalisis hubungan antara jumlah video, views, komentar positif, negatif, dan netral terhadap jumlah mahasiswa baru di institusi pendidikan tinggi selama 2021–2024. Data diperoleh dari tiga sumber: (1) data kuantitatif jumlah mahasiswa baru dari LLDIKTI, (2) data jumlah video dan views dari YouTube, dan (3) data sentimen komentar yang diklasifikasikan menggunakan algoritma BERT menjadi positif, negatif, dan netral. Metode penelitian menggunakan pendekatan kuantitatif dengan model regresi Random Forest untuk mengevaluasi pengaruh variabel independen yaitu jumlah video, views, dan sentimen, terhadap variabel dependen yaitu jumlah mahasiswa baru. Hasil analisis menunjukkan korelasi positif signifikan antara jumlah views dan sentimen positif terhadap jumlah mahasiswa baru, sementara sentimen negatif berkorelasi negatif. Namun, hubungan ini tidak sepenuhnya linier dengan kemungkinan adanya faktor lain yang berpengaruh, hal ini ditunjukkan pada hasil nilai R^2 sebesar 32.1%. Analisis korelasi Spearman juga mengonfirmasi hubungan kuat antara jumlah video, jumlah views dan sentimen positif terhadap jumlah mahasiswa baru dengan nilai korelasi 0.7-0.8. Temuan ini menyoroti pentingnya strategi pemasaran digital, seperti meningkatkan jumlah publikasi, views dan menciptakan konten yang menghasilkan sentimen positif, untuk menarik lebih banyak mahasiswa baru.

Kata kunci: Analisis Sentimen, BERT, Random Forest, Spearman, Youtube.

Abstract

This study aims to analyze the relationship between the number of videos, views, positive comments, negative comments, and neutral comments with the number of new students at higher education institutions during 2021–2024. The data were obtained from three sources: (1) quantitative data on the number of new students from LLDIKTI, (2) data on the number of videos and views from YouTube, and (3) sentiment data from comments classified using the BERT algorithm into positive, negative, and neutral categories. The research employed a quantitative approach using the Random Forest regression model to evaluate the influence of independent variables, namely, the number of videos, views, and sentiment on the dependent variable, which is the number of new students. The analysis results showed a significant positive correlation between the number of views and positive sentiment with the number of new students, while negative sentiment had a negative correlation. However, this relationship is not entirely linear, as indicated by an R^2 value of 32.1%, suggesting the possibility of other influencing factors. Spearman correlation analysis also confirmed a strong relationship between the number of video, views and positive sentiment with the number of new students, with a correlation value of 0.7-0.8. These findings highlight the importance of digital marketing strategies, such as increasing publications, views and creating content that generates positive sentiment, to attract more new students.

Keywords: Sentiment Analysis, BERT, Random Forest, Spearman, Youtube

PENDAHULUAN

Perkembangan teknologi digital telah mengubah paradigma pemasaran pendidikan tinggi, universitas semakin bergantung pada platform media sosial seperti YouTube untuk membangun citra institusi dan menjangkau calon mahasiswa (Moenggah et al., 2022)(J. Xie et al., 2023). YouTube, sebagai salah satu platform terbesar untuk berbagi video, tidak hanya menjadi sarana promosi tetapi juga wadah interaksi antara universitas dengan calon mahasiswa dan alumni melalui fitur seperti komentar, likes, dan shares (Fritz & Smith, 2024). Interaksi ini memberikan peluang untuk menganalisis sentimen publik terhadap branding universitas, yang dapat menjadi indikator penting dalam memahami minat dan preferensi calon mahasiswa (Moenggah et al., 2022).

Namun, meskipun banyak universitas telah memanfaatkan YouTube sebagai alat pemasaran, tingkat keterlibatan dan efektivitas konten dalam menarik minat calon mahasiswa masih bervariasi (Fritz & Smith, 2024) (Alsufyan & Aloud, 2017). Studi sebelumnya menunjukkan bahwa konten video yang autentik dan interaktif cenderung lebih efektif dalam membangun citra positif universitas (Aazam et al., 2024; Ann Voss & Kumar, 2013). Di sisi lain, analisis sentimen terhadap komentar YouTube dapat memberikan wawasan mendalam tentang persepsi publik terhadap branding universitas, yang belum banyak dieksplorasi dalam konteks penerimaan mahasiswa baru (Ngoc et al., 2021; Nguyen et al., 2021).

Penerimaan mahasiswa baru merupakan indikator penting bagi keberlanjutan dan reputasi sebuah universitas. Dalam era digital, faktor-faktor seperti aktivitas media sosial dan sentimen publik terhadap branding universitas semakin berpengaruh terhadap keputusan calon mahasiswa (Nguyen et al., 2021). YouTube, sebagai platform yang memungkinkan interaksi langsung melalui komentar, menjadi sumber data yang kaya untuk menganalisis sentimen publik. Namun, penelitian sebelumnya cenderung terfokus pada analisis kuantitatif seperti jumlah views dan likes, sementara analisis kualitatif terhadap komentar masih kurang dieksplorasi (Moenggah et al., 2022) (Fritz & Smith, 2024).

Komentar yang ditulis dalam bentuk kalimat panjang, diperlukan metode khusus untuk menyimpulkan, apakah kalimat tersebut memiliki makna positif, negatif atau kalimat yang bermakna netral. Algoritma BERT telah terbukti efektif dalam menganalisis sentimen komentar panjang, terutama dalam konteks pendidikan (Ngoc et al., 2021). Secara komparatif, BERT unggul dibanding algoritma machine learning tradisional dalam menangkap nuansa bahasa media sosial dan umpan balik siswa (Bikku et al., 2023; Vega et al., 2024; P. Xie et al., 2024). Dalam konteks umpan balik dari pelayanan institusi pendidikan, BERT terbukti efektif dalam melakukan analisis sentimen.

Media sosial memiliki dampak signifikan terhadap tingkat pendaftaran mahasiswa di institusi pendidikan tinggi. Berbagai penelitian menunjukkan bahwa platform media sosial semakin banyak digunakan sebagai alat pemasaran dan perekrutan mahasiswa baru. Strategi yang efektif meliputi kampanye media sosial, *electronic word-of-mouth* (e-WOM), tur virtual kampus, serta sesi langsung di Facebook, yang terbukti berkontribusi positif terhadap peningkatan jumlah pendaftar (Constantinides & Zinck Stagno, 2011; Sandvig, 2016; Xiong et al., 2018). Validasi media sosial, seperti jumlah likes dan pengikut, juga memainkan peran penting dalam memperkuat efektivitas strategi pemasaran tersebut (Rutter et al., 2016). Selain itu, keterlibatan institusi dalam media sosial dan aktivitas e-WOM secara langsung membentuk persepsi calon mahasiswa terhadap citra merek perguruan tinggi, yang pada akhirnya memengaruhi keputusan mereka untuk mendaftar (Nguyen et al., 2021). Ulasan, peringkat, dan rekomendasi dari pengguna media sosial menjadi faktor penting dalam membentuk persepsi serta menentukan keputusan pendaftaran mahasiswa baru (Al-Dmour et al., 2024; Siji & Parsad, 2023).

TINJAUAN PUSTAKA

Beberapa studi telah menggunakan analisis sentimen untuk mengevaluasi branding universitas, seperti penelitian yang dilakukan oleh Moenggah et al. (2022) yang menganalisis sentimen terhadap slogan Telkom University di YouTube. Hasilnya menunjukkan bahwa analisis sentimen dapat memberikan akurasi tinggi

dalam mengklasifikasikan opini publik. Namun, penelitian tersebut belum mengintegrasikan hasil sentimen dengan data kuantitatif seperti jumlah mahasiswa baru, sehingga belum memberikan gambaran holistik tentang faktor-faktor yang memengaruhi penerimaan mahasiswa.

Algoritma machine learning BERT - *Bidirectional Encoder Representations from Transformers* telah terbukti efektif dalam analisis sentimen, terutama untuk teks berbahasa Indonesia (Ngoc et al., 2021). Algoritma BERT mampu menangkap konteks kata dalam kalimat, sehingga menghasilkan klasifikasi sentimen yang lebih akurat dibandingkan metode tradisional (Bikku et al., 2023).

Untuk memahami hubungan antara variabel-variabel digital, seperti jumlah video, views, likes, dan sentimen komentar, dengan jumlah mahasiswa baru penelitian ini menggunakan algoritma Random Forest, sebuah algoritma *ensemble learning* yang *robust* dalam menangani data kompleks dan variabel non-linear (Helaria Maria, 2024). Algoritma ini dipilih karena kemampuannya dalam menangani interaksi antar variabel terhadap variabel target dan memberikan hasil interpretatif melalui analisis feature importance (Miao & Xu, 2024).

Untuk memperkuat analisis, penelitian ini juga mengadopsi metode statistik non-parametrik korelasi Spearman. Korelasi Spearman dipilih karena mampu menangani data yang tidak terdistribusi normal dan mengukur kekuatan serta arah hubungan antara dua variabel ordinal atau kontinu (Thirumalai et al., 2017). Metode ini telah terbukti efektif dalam studi serupa, seperti analisis hubungan antara aktivitas media sosial dan keputusan konsumen (Al-Dmour et al., 2024). Dengan menggunakan korelasi Spearman, penelitian ini bertujuan untuk mengidentifikasi variabel-variabel kunci yang memiliki hubungan signifikan dengan fluktuasi jumlah mahasiswa baru ketika analisis random forest tidak menghasilkan akurasi model yang diharapkan.

Penelitian ini memiliki tiga tujuan utama. Pertama, menganalisis sentimen publik dengan menggunakan algoritma BERT untuk mengklasifikasikan sentimen yaitu, positif, negatif, netral, dari komentar YouTube di saluran resmi universitas. Kedua, mengidentifikasi hubungan antar variabel dengan

menganalisis korelasi antara aktivitas digital dalam variabel jumlah video, views, likes, dan sentimen komentar, dengan fluktuasi jumlah mahasiswa baru menggunakan pemodelan Random Forest dan korelasi Spearman. Ketiga, memberikan rekomendasi strategis bagi universitas dalam meningkatkan efektivitas konten YouTube berdasarkan hasil analisis sentimen dan pemodelan machine learning.

Penelitian ini berbeda dari studi-studi sebelumnya dalam beberapa aspek kunci. Penelitian seperti yang dilakukan oleh Moenggah et al. (2022) berfokus pada analisis sentimen terhadap slogan universitas menggunakan Naïve Bayes, penelitian ini mengimplementasikan algoritma BERT yang terbukti lebih unggul dalam menangkap nuansa bahasa alami, khususnya untuk teks dalam bahasa Indonesia (Ngoc et al., 2021; Bikku et al., 2023). Selain itu, berbeda dengan penelitian Fritz & Smith (2024) yang hanya menganalisis keterlibatan kuantitatif yaitu views, likes, penelitian ini mengintegrasikan data kualitatif berupa sentimen komentar dengan data kuantitatif eksternal yaitu data jumlah mahasiswa baru dari LLDIKTI. Pendekatan hybrid ini memungkinkan analisis yang lebih holistik. Lebih lanjut, penelitian ini tidak hanya menggunakan model prediktif Random Forest tetapi juga memperkuat temuan dengan analisis korelasi non-parametrik Spearman untuk memvalidasi kekuatan dan arah hubungan antar variabel, yang jarang dilakukan dalam penelitian sejenis. Dengan demikian, penelitian ini tidak hanya mengidentifikasi pola tetapi juga mengkonfirmasi hubungan sebab-akibat potensial antara aktivitas digital dan penerimaan mahasiswa baru.

METODE

Metode dalam penelitian ini dijelaskan pada Gambar 1, mencakup pengumpulan data kualitatif berupa komentar pada video dan short video YouTube, dan data kuantitatif jumlah mahasiswa baru tiap universitas. Data komentar dianalisis menggunakan algoritma BERT untuk menentukan sentimen positif, negatif, dan netral. Data kemudian dianalisis menggunakan algoritma random forest dan korelasi spearman untuk mendapatkan kesimpulan penelitian.

Berikut adalah penjelasan rinci setiap tahapannya.

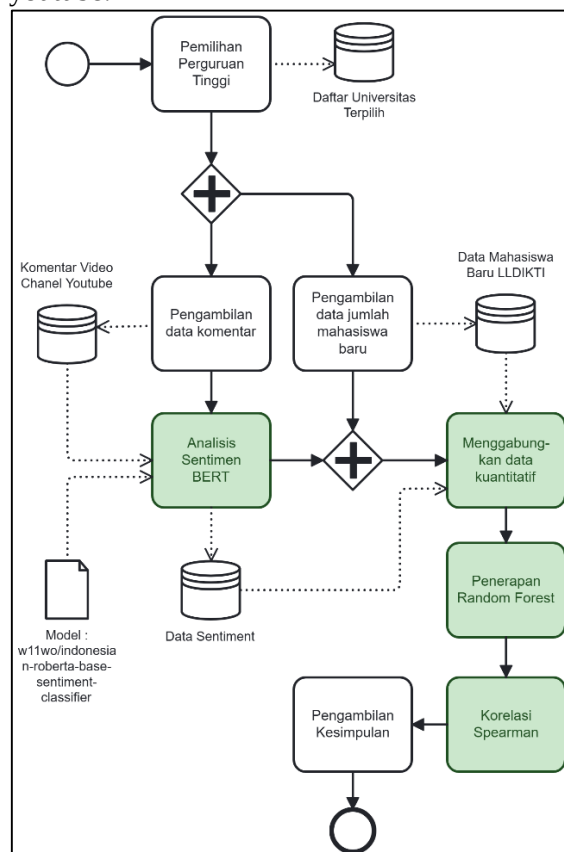
3.1. Pemilihan perguruan tinggi

Pemilihan 8 perguruan tinggi di wilayah LLDIKTI VI ini tidak menggunakan kriteria tertentu, hanya mempertimbangkan bahwa perguruan tinggi memiliki saluran youtube resmi. Daftar perguruan tinggi tersebut adalah:

1. Institut Teknologi dan Bisnis Semarang
2. STMIK Bina Patria Magelang
3. Universitas Dian Nuswantoro
4. Universitas Duta Bangsa Surakarta
5. Universitas Muria Kudus
6. Universitas Sahid Surakarta
7. Universitas Semarang
8. Universitas Wahid Hasyim Semarang

3.2. Pengambilan Data Komentar

Data komentar diambil dari setiap postingan video dan short video youtube pada saluran resmi perguruan tinggi. Proses ini dilakukan menggunakan pustaka request dari pemrograman python dengan menggunakan API (Application Programming Interface) dari youtube.



Gambar 1. Metode Penelitian

3.3. Pengambilan Data Mahasiswa Baru

Data jumlah mahasiswa baru dari rentang tahun 2021 s.d. tahun 2024, diambil dari laman

publikasi LLDIKTI VI untuk setiap perguruan tinggi terpilih.

3.4. Analisis Sentiment BERT

Sebelum proses analisis sentimen dilakukan, data hasil scrapping harus dilakukan pre-processing menggunakan pustaka sastrawi dan nltk dari python dengan tahapan:

- a. Menghapus komentar ganda.
- b. Pembersihan text, yang meliputi penghapusan mention, hashtag, retweet, link, angka, karakter spesial, spasi dan punctuation.
- c. Mengubah teks menjadi huruf kecil semua.
- d. Mengoreksi kata-kata slang dan menghapus stopwords.
- e. Melakukan tokenisasi kata.
- f. Menggabungkan kembali menjadi kalimat yang sudah bersih.

Setiap kalimat yang sudah bersih, siap dilakukan analisis sentiment menggunakan model pra-latih *w11wo/indonesian-roberta-base-sentiment-classifier* algoritma BERT yang sudah di *fine-tune* khusus untuk analisis sentimen dalam bahasa Indonesia. Hasil dari analisis sentimen ini adalah label untuk setiap komentar memiliki sentimen positif, negatif dan netral beserta tingkat kepercayaan hasil analisis.

3.5. Penggabungan Data Kuantitatif

Hasil analisis sentimen komentar pada data kualitatif, dihitung untuk diubah menjadi data kuantitatif dan digabungkan dengan data kuantitatif lainnya yang terdiri atas variabel *dependent* jumlah mahasiswa dan variabel *independent*, video, view, positif, negatif, netral, semua data dikelompokkan berdasarkan tahun 2021 s.d. 2024.

3.6. Analisis Prediksi Random Forest

Model *Random Forest Regressor* digunakan untuk memprediksi jumlah mahasiswa baru berdasarkan variabel input. Rumus matematis untuk prediksi menggunakan Random Forest Regressor adalah sebagai berikut:

$$\hat{y} = \frac{1}{N} \sum_{i=1}^N T_i(X) \quad (1)$$

$\hat{y}$ adalah nilai prediksi akhir (regresi). N adalah jumlah total pohon keputusan dalam random forest dan $T_i(X)$ adalah prediksi yang dihasilkan oleh pohon ke- i untuk input X .

Untuk menilai model yang dihasilkan dari analisis random forest, digunakan perhitungan MSE (*Mean Squared Error*). MSE

adalah metrik yang digunakan untuk mengevaluasi performa model regresi.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

Dalam konteks *Random Forest*, MSE menghitung rata-rata dari selisih kuadrat antara nilai prediksi $\hat{y}$ dan nilai sebenarnya y pada dataset uji. MSE memberikan gambaran tentang seberapa jauh prediksi model dari nilai sebenarnya.

3.7. Korelasi Spearman

Untuk meningkatkan kepercayaan pada hasil prediksi random forest, dilakukan analisis korelasi spearman. Tujuan dari analisis korelasi spearman adalah melihat secara visual hubungan antar variable independent dengan variable dependent.

3.8. Pengambilan Keputusan

Hasil analisis random forest dan korelasi spearman digunakan untuk memberikan rekomendasi kepada perguruan tinggi.

HASIL DAN PEMBAHASAN

4.1. Data Kuantitatif

Tabel 1. Memuat data jumlah mahasiswa baru pada 8 perguruan tinggi terpilih yang diambil dari laman LLDIKTI wilayah VI dalam rentang waktu tahun 2021 s.d. 2024

Tabel 1 Mahasiswa Baru 2021 s.d. 2024

Perguruan Tinggi	2021	2022	2023	2024
ITB Semarang	515	254	359	374
STMIK Bina Patria Magelang	108	149	270	92
Universitas Dian Nuswantoro	3,892	3,542	2,649	3,231
Universitas Duta Bangsa Surakarta	1,355	1,177	1,547	1,670
Universitas Muria Kudus	2,705	2,603	2,648	2,387
Universitas Sahid Surakarta	808	494	499	517
Universitas Semarang	4,301	2,474	4,515	4,387
Universitas Wahid Hasyim Semarang	2,076	2,167	2,207	2,405

Tabel 2 Jumlah video dan view pada saluran youtube perguruan tinggi

Id	Perguruan Tinggi	2021		2022		2023		2024	
		video	view	video	view	video	view	video	view
1	ITB Semarang	-	-	27	11,523	36	13,747	2	243
2	STMIK Bina Patria	20	5,002	31	3,655	16	3,861	17	1,250
3	Univ. Dian Nuswantoro	1,748	341,829	1,178	3,115,230	527	352,371	417	634,355
4	Universitas Duta Bangsa	84	98,604	39	49,851	34	18,649	48	16,483
5	Universitas Muria Kudus	63	51,724	47	46,004	67	79,600	67	90,018
6	Universitas Sahid Surakarta	43	13,991	76	32,087	47	15,172	70	29,485
7	Universitas Semarang	-	-	4	7,803	350	203,669	2,386	596,912
8	Universitas Wahid Hasyim	175	103,509	323	257,431	124	5,578,459	120	68,181

Tabel 3 Komentar pada saluran youtube perguruan tinggi

comment_id	comment
1	Selamat untuk wisuda ke 2....sukses selalu
...	...
33674	Salam sehat salam sejahtera buat semua trutama warga kab.kendal..terkait dengan jaminan kesehatan ..mohon izin saya sebagai warga Kendal mau tnya ketika warga Kendal sakit dan tidak punya BPJS baik BPJS PBI maupun non PBI itu bgmn solusinya sejak SKTM di hapus keluarga pasien kesulitan untuk mendapatkan jaminan kesehatan ketika kluarganya ada yg sakit dn tidak punya BPJS Baznas selama ini yg telah bntu mengaver biaya pasien tidak mampu di rumh sakit TPI hnya brapa persen dri total biaya keseluruhan... Harapan kami semoga siapapun yg terpilih nanti ..bisa menjalankan program programnya dengan NYATA dan bisa memberikan jaminan kesehatan 100% kepada warga kendal yg tidak mampu dan jangan. Lupa juga terkait jaminan pendidikan untuk warga Kendal yg kurang mampu jangan sampai masih ada ijazah yg tertahan di sekolah hanya KRNA blm melunasi kekurangan administrasi di sekolah ijasah itu hak siswa bos... SuwunðŸ™ • ðŸ • ¾

Tabel 2 memuat jumlah video dan jumlah view per tahun dari saluran youtube perguruan tinggi terpilih. Sedangkan Tabel 3 menampilkan beberapa komentar yang diambil

dari komentar pada video saluran youtube. Dari seluruh perguruan tinggi terpilih, didapatkan sejumlah 33.197 komentar, yang diambil dari 6,228 jumlah video dan *short video*, dengan total tayangan video keseluruhan mencapai 11,329,548 tayangan (*views*).

4.2. Analisis Sentiment BERT

Seluruh data komentar, ditandai dengan tanggal komentar dan *video_id*, untuk dianalisis menggunakan model pra-latih algoritma BERT *w11wo/indonesian-roberta-base-sentiment-classifier* yang sudah di *fine-tune* khusus untuk analisis sentimen dalam bahasa Indonesia. Setiap komentar akan dilakukan pre-processing menggunakan pustaka *sastrawi* dan *nlTK* dari *phyton*. Tabel 4 menampilkan hasil analisis sentimen untuk semua komentar. Hasil analisis dikategorikan menjadi 3, yaitu komentar positif, negatif dan netral.

Tabel 4. Jumlah komentar per tahun

Tahun	Jml Positif	Jml Negatif	Jml Netral
2021	1,316	311	1,149
2022	11,366	1,202	1,198
2023	7,582	900	1,094
2024	3,923	702	740
Total	24,187	3,115	4,181

4.3. Analisis Random Forest

Data komentar yang telah diubah menjadi data kuantitatif, digabungkan dengan data jumlah mahasiswa, untuk setiap perguruan tinggi, dalam rentang waktu pengamatan tahun 2021 s.d. 2024, ditampilkan dalam tabel 5.

Tabel 5. Data kuantitatif gabungan

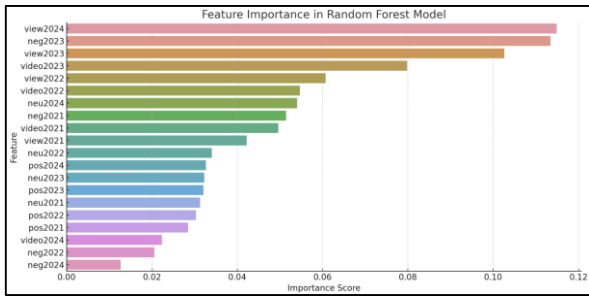
Id	Video	View	Pos	Neg	Net	Mhs
2021						
1	0	0	0	0	0	515
2	20	5,002	6	0	0	108
3	1,748	341,829	572	133	203	3,892
4	84	98,604	80	12	114	1,355
5	63	51,724	48	26	34	2,705
6	43	13,991	262	32	76	808
7	0	0	0	0	0	4,301
8	175	103,509	348	108	722	2,076
2022						
1	27	11,523	42	6	0	254
2	31	3,655	0	0	8	149
3	1,178	3,115,230	971	341	311	3,542
4	39	49,851	42	8	22	1,177
5	47	46,004	24	2	10	2,603
6	76	32,087	606	50	182	494
7	4	7,803	2	0	0	2,474
8	323	257,431	9,679	795	665	2,167

2023						
1	36	13,747	108	6	18	359
2	16	3,861	2	0	0	270
3	527	352,371	2,462	185	183	2,649
4	34	18,649	24	4	10	1,547
5	67	79,600	54	48	44	2,648
6	47	15,172	224	16	72	499
7	350	203,669	97	15	28	4,515
8	124	5,578,459	820	176	192	2,207
2024						
1	2	243	0	0	6	374
2	17	1,250	20	0	2	92
3	417	634,355	2,356	266	197	3,231
4	48	16,483	136	30	28	1,670
5	67	90,018	140	28	42	2,387
6	70	29,485	244	48	62	517
7	2,386	596,912	321	164	181	4,387
8	120	68,181	706	166	222	2,405

Hasil model Random Forest Regressor menunjukkan performa berikut:

- a. Mean Squared Error (MSE): 1,709,411.82
MSE mengukur rata-rata kesalahan prediksi. Semakin kecil nilainya, semakin baik modelnya.
- b. R-squared (R^2): 0.321
 R^2 menunjukkan seberapa besar variable target yaitu jumlah mahasiswa baru dapat dijelaskan oleh model. Nilai ini mengindikasikan bahwa model hanya mampu menjelaskan sekitar 32.1% variabilitas target.

Meskipun hasil dari model tidak menunjukkan angka yang tinggi, tetapi cukup baik jika mempertimbangkan aktifitas youtube hanya salah satu aspek yang mempengaruhi minat mahasiswa baru untuk mendaftar ke universitas. Gambar 2 menunjukkan hasil analisis feature importance dari setiap variable uji.



Gambar 2. Analisis Feature Importance

4.4. Analisis Korelasi SPEARMAN

Karena target penelitian adalah mencari hubungan antara publikasi sosial media melalui youtube, maka untuk membantu menganalisis model, digunakan tahapan lanjutan yaitu melakukan visualisasi pada gambar 3, untuk memahami hubungan antar variable.

Berikut beberapa korelasi antar variable yang teridentifikasi ($|\rho| \geq 0.7$):

1. Video vs Mahasiswa:

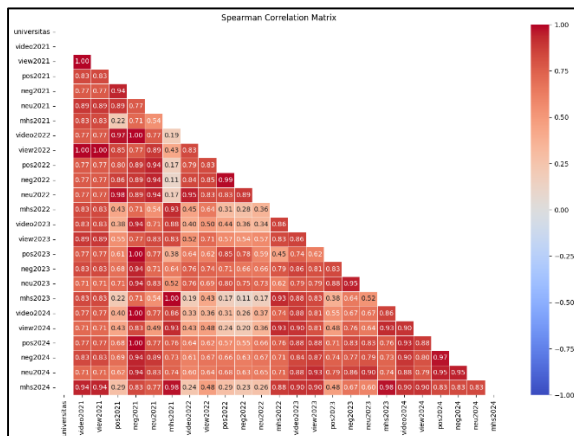
Korelasi kuat ($\rho = 0.77$) antara jumlah video yang diunggah dengan jumlah mahasiswa di tahun yang sama.

2. View vs Mahasiswa:

Korelasi kuat ($\rho = 0.8$) antara sentimen positif dan jumlah mahasiswa.

3. Positif vs Mahasiswa:

Korelasi positif sedang ($\rho = 0.65$) antara sentimen positif dan jumlah mahasiswa



Gambar 3. Matriks Korelasi Spearman

SIMPULAN

Hasil penelitian ini menunjukkan bahwa, hubungan antara aktivitas pemasaran digital di YouTube dengan jumlah mahasiswa baru yang mendaftar di perguruan tinggi, memiliki korelasi yang cukup baik untuk dijadikan landasan dalam strategi pemasaran dalam upaya meningkatkan jumlah mahasiswa baru pada perguruan tinggi.

Model Random Forest Regressor yang digunakan menunjukkan performa yang moderat, dengan nilai R-squared (R^2) sebesar 32.1%. Hal ini menunjukkan bahwa model hanya mampu menjelaskan sebagian kecil variabilitas jumlah mahasiswa baru. Salah satu faktor penyebab adalah tidak dimasukkannya variabel lain seperti biaya kuliah atau jarak kampus, karena fokus penelitian ini terbatas pada pengaruh postingan YouTube sebagai

media promosi. Meskipun begitu, penelitian tetap memberikan wawasan penting tentang pola hubungan antara variabel yang dianalisis.

Untuk meningkatkan kepercayaan terhadap model, dilakukan analisis visual pada feature importance untuk memahami kontribusi tiap variabel, serta analisis korelasi menggunakan korelasi Spearman. Hasilnya menunjukkan bahwa jumlah video, views dan komentar positif merupakan variabel yang paling berpengaruh terhadap jumlah mahasiswa baru. Korelasi yang signifikan dan konsisten antara kedua variabel tersebut menunjukkan bahwa sebaran tayangan tinggi terhadap video serta sentimen positif dari audiens memiliki dampak besar pada keputusan calon mahasiswa.

Hasil dari penelitian ini menegaskan pentingnya aktivitas media sosial, salah satunya platform YouTube, sebagai strategi pemasaran digital untuk meningkatkan jumlah mahasiswa baru.

Perguruan tinggi dapat mengambil langkah-langkah berikut:

1. Menjaga konsistensi dan jumlah konten, untuk memastikan unggahan video dilakukan secara berkala demi menjaga jangkauan dan membangun audiens.
2. Mengunggah video yang berkualitas, informatif, menarik, dan relevan dengan kebutuhan calon mahasiswa untuk meningkatkan jumlah views.
3. Mengelola komentar dan menciptakan interaksi positif di media sosial untuk memperkuat citra positif dari institusi.
4. Merespons komentar negatif secara profesional untuk memitigasi dampak buruk terhadap persepsi publik.

Strategi berbasis media sosial ini dapat menjadi bagian integral dalam upaya perguruan tinggi untuk menarik lebih banyak mahasiswa baru, terutama di era digital. Media online memainkan peran utama dalam membentuk keputusan calon mahasiswa.

Untuk penelitian selanjutnya, disarankan untuk memperluas cakupan variabel dengan memasukkan faktor-faktor lain yang mungkin memengaruhi keputusan calon mahasiswa, seperti biaya kuliah, akreditasi program studi, atau jarak geografis. Selain itu, penggunaan model machine learning yang lebih kompleks, seperti Gradient Boosting atau Neural Networks, dapat dieksplorasi untuk meningkatkan akurasi prediksi. Periode

penelitian yang lebih panjang juga diperlukan untuk mengamati tren yang lebih stabil dan dampak jangka panjang dari strategi digital. Terakhir, penelitian mendatang dapat memperluas analisis ke platform media sosial lainnya seperti Instagram dan TikTok, untuk mendapatkan pemahaman yang lebih komprehensif tentang strategi pemasaran digital di pendidikan tinggi.

DAFTAR PUSTAKA

- Aazam, F., Ang, P. S., & Isa, N. A. N. M. (2024). Exploring X usage and engagement strategies in higher education: A comparative study of Pakistani and Malaysian universities. *SEARCH Journal of Media and Communication Research*, 16(3), 31 – 46. <https://doi.org/10.58946/search-16.3.P3>
- Al-Dmour, R., Al-Dmour, H., & Al-Dmour, A. (2024). The role of marketing mix and social media strategies in influencing international students' university choices in Jordan. *Journal of International Students*, 14(4), 642–663. <https://doi.org/10.32674/jis.v14i4.6407>
- Alsufyan, N. K., & Aloud, M. (2017). The state of social media engagement in Saudi universities. *Journal of Applied Research in Higher Education*, 9(2), 267–303. <https://doi.org/10.1108/JARHE-01-2016-0001>
- Ann Voss, K., & Kumar, A. (2013). The value of social media: are universities successfully engaging their audience? *Journal of Applied Research in Higher Education*, 5(2), 156–172. <https://doi.org/10.1108/JARHE-11-2012-0060>
- Bikku, T., Jarugula, J., Kongala, L., Tummalala, N. D., & Vardhani Donthiboina, N. (2023). Exploring the Effectiveness of BERT for Sentiment Analysis on Large-Scale Social Media Data. *2023 3rd International Conference on Intelligent Technologies (CONIT)*, 1–4. <https://doi.org/10.1109/CONIT59222.2023.10205600>
- Constantinides, E., & Zinck Stagno, M. C. (2011). Potential of the social media as instruments of higher education marketing: a segmentation study. *Journal of Marketing for Higher Education*, 21(1), 7–24. <https://doi.org/10.1080/08841241.2011.573593>
- Fritz, A. M., & Smith, A. M. (2024). Marketing higher education on YouTube: a content analysis of college promotional videos. *Journal of Marketing for Higher Education*, 1–22. <https://doi.org/10.1080/08841241.2024.2400088>
- Helaria Maria. (2024). Hybrid Lexicon and Transformer-Based Sentiment Analysis of Student Feedback for Faculty Evaluation: A Speech-to-Text Approach. *Communications on Applied Nonlinear Analysis*, 32(2), 610–624. <https://doi.org/10.52783/cana.v32.1857>
- Miao, Y., & Xu, Y. (2024). Random Forest-Based Analysis of Variability in Feature Impacts. *2024 IEEE 2nd International Conference on Image Processing and Computer Applications (ICIPCA)*, 1130–1135. <https://doi.org/10.1109/ICIPCA61593.2024.10708791>
- Moenggah, R. F., Richasdy, D., & Purbolaksono, M. D. (2022). Telkom University Slogan Analysis on YouTube Using Naïve Bayes. *2022 International Conference on Data Science and Its Applications (ICoDSA)*, 283–288. <https://doi.org/10.1109/ICoDSA55874.2022.9862818>
- Ngoc, T. V., Thi, M. N., & Thi, H. N. (2021). Sentiment Analysis of Students' Reviews on Online Courses: A Transfer Learning Method. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, 306 – 314. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85126042003&partnerID=40&md5=e9808b6e91aaec399c81a9af86dd42ff>
- Nguyen, L., Lu, V. N., Conduit, J., Tran, T. T. N., & Scholz, B. (2021). Driving enrolment intention through social media engagement: a study of Vietnamese prospective students. *Higher Education Research & Development*, 40(5), 1040–1055. <https://doi.org/10.1080/07294360.2020.1798886>

- Rutter, R., Roper, S., & Lettice, F. (2016). Social media interaction, the university brand and recruitment performance. *Journal of Business Research*, 69(8), 3096–3104. <https://doi.org/10.1016/j.jbusres.2016.01.025>
- Sandvig, J. C. (2016). The role of social media in college recruiting. *International Journal of Web Based Communities*, 12(1), 23. <https://doi.org/10.1504/IJWBC.2016.074273>
- Siji, S., & Parsad, C. (2023). Navigating the Stars: The Moderating Effect of Social Media Usage on the Admission Intention of B-schools. *FIIB Business Review*. <https://doi.org/10.1177/23197145231185391>
- Thirumalai, C., Chandhini, S. A., & Vaishnavi, M. (2017). Analysing the concrete compressive strength using Pearson and Spearman. *2017 International Conference of Electronics, Communication and Aerospace Technology (ICECA)*, 215–218. <https://doi.org/10.1109/ICECA.2017.8212799>
- Vega, I., Valencia, J., Arcos, Á., Navarrete, D., & Baldeon-Calisto, M. (2024). A Comparison Between Transformers and Foundation Models in Sentiment Analysis of Student Evaluation of Teaching. *2024 12th International Symposium on Digital Forensics and Security (ISDFS)*, 1–7. <https://doi.org/10.1109/ISDFS60797.2024.10527264>
- Xie, J., Chai, Y., & Liu, X. (2023). Unbox the Black-Box: Predict and Interpret YouTube Viewership Using Deep Learning. *Journal of Management Information Systems*, 40(2), 541–579. <https://doi.org/10.1080/07421222.2023.2196780>
- Xie, P., Gu, H., & Zhou, D. (2024). Modeling Sentiment Analysis for Educational Texts by Combining BERT and FastText. *2024 6th International Conference on Computer Science and Technologies in Education (CSTE)*, 195–199. <https://doi.org/10.1109/CSTE62025.2024.00044>
- Xiong, L., Alsadoon, A., Maag, A., Prasad, P. W. C., Hoe, L. S., & Elchouemi, A. (2018). Rise of Social Media Marketing: A Perspective on Higher Education. *2018 13th International Conference on Computer Science & Education (ICCSE)*, 1–6. <https://doi.org/10.1109/ICCSE.2018.8468683>