ANALISIS ENGAGEMENT RATE UNGGAHAN INSTAGRAM MENGGUNAKAN DATA ANALYTICS

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh Gelar Sarjana Komputer (S.Kom) pada Program Studi Teknologi Informasi Fakultas Teknik Universitas Muhammadiyah Palembang

Oleh :

Kristiawan

162022012

PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH PALEMBANG

2026

HALAMAN PENGESAHAN PEMBIMBING

ANALISIS ENGAGEMENT RATE UNGGAHAN INSTAGRAM MENGGUNAKAN DATA ANALYTICS

Oleh :
Kristiawan
162022012

Menyetujui,

Pembimbing Utama,	**Pembimbing Pendamping,**
Dedi Haryanto, S.Kom., M.Kom	**Kemas Muhammad Wahyu Hidayat., M.Kom**
NBM/NIDN: 1337459/0201089001	**NBM/NIDN: 1255881/0225068904**

Disetujui	**Program Studi Teknologi Informasi**
Dekan Fakultas Teknik,	**Ketua Program Studi,**
Ir. Ahmad Junaidi, M.T	**Dr. Karnadi, S.Kom., M.Kom**
NBM/NIDN: 763050/0202026502	**NBM/NIDN: 1088893/0210038202**

HALAMAN PERSETUJUAN

Judul Skripsi : Analisis Engagement Rate Unggahan Instagram Menggunakan Data Analytics

Oleh **Kristiawan** dengan NIM **162022012** Laporan Penelitian Skripsi ini telah disetujui dan disahkan oleh Tim Penguji Program Studi Teknologi Informasi, Konsentrasi Manajemen Tata Kelola Teknologi Informasi Strata 1 Universitas Muhammadiyah Palembang, 28 Agustus 2026 dan telah dinyatakan **LULUS**

Palembang, 30 Agustus 2026
Mengetahui,

Ketua Program Studi Teknologi Informasi
Universitas Muhammadiyah Palembang

Dr. Karnadi, S.Kom., M.Kom
NBM/NIDN: 1088893/0210038202

Tim Penguji
Ketua Penguji

Dedi Haryanto, S.Kom., M.Kom
NBM/NIDN: 1337459/0201089001

Penguji 1,

Dr. Karnadi, S.Kom., M.Kom
NBM/NIDN: 1088893/0210038202

Penguji 2,

Jimmie, S.Kom., M.Kom
NBM/NIDN: 1340253/0222047702

HALAMAN PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama : Kristiawan

NIM : 162022012

Dengan ini menyatakan bahwa:

1. Karya tulis saya (Skripsi) merupakan sebuah karya asli serta belum pernah diajukan untuk memperoleh Gelar Akademik dengan baik (Sarjana) di Program Studi Teknologi Informasi Fakultas Teknik Universitas Muhammadiyah Palembang atau Perguruan Tinggi lainnya.
2. Karya tulis saya (Skripsi) merupakan hasil murni memiliki gagasan, pokok permasalahan, serta hasil penelitian saya sendiri, tanpa kerja sama terhadap pihak lain melainkan dengan arahan dosen pembimbing.
3. Karya tulis saya (Skripsi) tidak terdapat karya serta pendapat yang telah ditulis atau dipublikasikan orang lain kecuali serta tertulis dengan jelas dicantumkan nama pengarang serta dimasukkan ke dalam daftar ivustaka.
4. Karya tulis saya (Skripsi) yang dihasilkan telah melakukan pengecekan keaslian menggunakan Turnitin yang dipublikasikan melalui internet sehingga dapat diakses secara daring.
5. Dengan ini surat pernyataan yang saya buat secara sungguh-sungguh serta apabila terbukti terdapat penyimpangan serta ketidakbenaran dari pernyataan ini, maka saya bersedia menerima sanksi sesuai dengan peraturan serta perundang-undangan akademik Program Studi Teknologi Informasi Fakultas Teknik Universitas Muhammadiyah Palembang.

Demikian surat pernyataan ini saya buat agar dapat dipergunakan sebagaimana mestinya.

Palembang, 30 Agustus 2026



Kristiawan

162022012

MOTTO DAN PERSEMBAHAN

MOTTO

"Maka sesungguhnya bersama kesulitan ada kemudahan."

(QS. Asy-Syarh: 5)

PERSEMBAHAN

Dengan penuh rasa syukur kepada Allah SWT, karya ini saya persembahkan kepada:

1. Kedua orang tua tercinta yang selalu memberikan doa, kasih sayang, dukungan moral maupun material yang tiada henti.
2. Keluarga besar yang selalu memberikan semangat dan motivasi.
3. Dosen pembimbing yang telah memberikan arahan dan bimbingan dalam penyusunan skripsi ini.
4. Sahabat dan teman-teman seperjuangan Program Studi Teknologi Informasi Universitas Muhammadiyah Palembang.
5. Almamater tercinta Universitas Muhammadiyah Palembang.

ABSTRAK

Media sosial Instagram menjadi salah satu platform digital yang digunakan untuk berbagi informasi, promosi, dan pemasaran. Engagement rate merupakan salah satu indikator untuk mengukur tingkat interaksi audiens terhadap unggahan. Penelitian ini bertujuan menganalisis hubungan karakteristik akun dan unggahan Instagram dengan engagement rate menggunakan metode Multiple Linear Regression. Dataset yang digunakan adalah Instagram Analytics Dataset yang terdiri atas 29.999 data unggahan. Tahapan penelitian meliputi pengumpulan data, pembersihan data, transformasi data menggunakan One-Hot Encoding dan Min-Max Scaling, pembentukan model Multiple Linear Regression, serta pengujian menggunakan Variance Inflation Factor (VIF), uji F, uji t, dan koefisien determinasi (R^2). Analisis dilakukan menggunakan dua ukuran engagement rate, yaitu Engagement Rate by Follower (ERF) dan Engagement Rate by Reach (ERR). Hasil penelitian menunjukkan bahwa model ERF memiliki nilai R-squared sebesar 0,1078 dan signifikan secara simultan dengan nilai Prob. F-statistic sebesar 0,0000. Pada uji parsial, follower_count menjadi satu-satunya variabel yang signifikan terhadap ERF dengan koefisien -9,6122 dan p-value 0,0000. Sementara itu, model ERR memiliki nilai R-squared sebesar 0,0008 dan tidak signifikan secara simultan dengan nilai Prob. F-statistic sebesar 0,3396. Tidak terdapat variabel independen yang signifikan terhadap ERR secara parsial. Hasil tersebut menunjukkan bahwa variabel yang digunakan lebih mampu menjelaskan variasi ERF dibandingkan ERR.

Kata Kunci: Instagram, Engagement Rate, Multiple Linear Regression, Data Analytics, Media Sosial.

ABSTRACT

Instagram has become one of the digital platforms used for information sharing, promotion, and marketing. Engagement rate is an indicator used to measure the level of audience interaction with posts. This study aims to analyze the relationship between Instagram account and post characteristics and engagement rate using Multiple Linear Regression. The dataset used is the Instagram Analytics Dataset, consisting of 29,999 posts. The research stages include data collection, data cleaning, data transformation using One-Hot Encoding and Min-Max Scaling, Multiple Linear Regression model development, and testing using Variance Inflation Factor (VIF), F-test, t-test, and coefficient of determination (R^2). The analysis uses two engagement rate measures, namely Engagement Rate by Follower (ERF) and Engagement Rate by Reach (ERR). The results show that the ERF model has an R-squared value of 0.1078 and is statistically significant simultaneously, with a Prob. F-statistic value of 0.0000. In the partial test, follower_count is the only variable that has a significant effect on ERF, with a coefficient of -9.6122 and a p-value of 0.0000. Meanwhile, the ERR model has an R-squared value of 0.0008 and is not statistically significant simultaneously, with a Prob. F-statistic value of 0.3396. None of the independent variables has a significant effect on ERR in the partial test. These results indicate that the variables used in this study are better able to explain the variation in ERF than ERR.

Keywords: *Instagram, Engagement Rate, Multiple Linear Regression, Data Analytics, Social Media.*

KATA PENGANTAR

Puji syukur kehadirat Allah SWT atas segala rahmat, karunia, dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "Analisis Engagement Rate Unggahan Instagram Menggunakan Data Analytics" dengan baik. Shalawat serta salam semoga senantiasa tercurah kepada Nabi Muhammad SAW beserta keluarga, sahabat, dan seluruh pengikut beliau hingga akhir zaman.

Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer (S.Kom.) pada Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Muhammadiyah Palembang.

Penulis menyadari bahwa penyusunan skripsi ini tidak terlepas dari bantuan, bimbingan, dukungan, dan doa dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. Abid Djazuli, S.E., M.M. Selaku Rektor Universitas Muhammdiyah Palembang
2. Bapak Ir. A. Junaidi, M.T selaku Dekan Universitas Muhammadiyah Palembang.
3. Bapak Dr. Karnadi, S.Kom., M.Kom. Selaku ketua Program Studi Teknologi Informasi.
4. Bapak Dedi Haryanto, S.Kom., M.Kom. Selaku Dosen Pembimbing Utama yang telah dengan sabar dan ikhlas untuk meluangkan waktunya dan membimbing saya.
5. Bapak Kemas Muhammad Wahyu Hidayat, S.Kom., M.Kom. Selaku Dosen Pembimbing Pendamping yang telah dengan sabar dan ikhlas untuk meluangkan waktunya dan membimbing saya.
6. Kedua orang tua tercinta, yang telah memberikan dukungan penuh berupa kasih sayang, doa, motivasi, nasihat berharga, bantuan material, serta kehadiran yang selalu menjadi sumber semangat dalam setiap langkah penulis.

Penulis menyadari bahwa skripsi ini masih memiliki keterbatasan dan jauh dari kata sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun demi perbaikan dan pengembangan penelitian di masa mendatang.

Akhir kata, penulis berharap semoga skripsi ini dapat memberikan manfaat bagi pembaca serta dapat menjadi referensi bagi penelitian selanjutnya, khususnya di bidang analisis data dan pemanfaatan metode Multiple Linear Regression dalam menganalisis data media sosial.

Palembang, 30 Agustus 2026



Kristiawan
162022012

DAFTAR ISI

HALAMAN JUDUL i
HALAMAN PENGESAHAN PEMBIMBING ii
HALAMAN PERSETUJUAN iii
HALAMAN PERNYATAAN iv
MOTTO DAN PERSEMBAHAN v
ABSTRAK vi
ABSTRACT vii
KATA PENGANTAR viii
DAFTAR ISI x
DAFTAR GAMBAR xiii
DAFTAR TABEL xiv
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Identifikasi Masalah 6
1.3 Rumusan Masalah 6
1.4 Batasan Masalah 7
1.5 Tujuan Penelitian 8
1.6 Manfaat Penelitian 8
1.7 Sistematika Penulisan 9
BAB II TINJAUAN PUSTAKA 10
2.1 Teknologi Informasi 10
2.2 Data Analytics 12
2.3 Media Sosial 13
2.4 Instagram 15
2.5 Engagement 17
2.6 Engagement Rate 19
2.7 Multiple Linear Regression 22
2.7.1 Koefisien Determinasi (R^2) 24
2.7.2 Uji Simultan (Uji F) 24

2.7.3 Uji Parsial (Uji t).....25
2.7.4 Variance Inflation Factor (VIF).....26
2.8 One-Hot Encoding.....27
2.9 Min-Max Scaling.....29
2.10 Penelitian Terdahulu.....31
2.10.1 Analisis Penelitian Terdahulu.....40
2.10.2 Kesenjangan Penelitian.....41
2.10.3 Keterbaruan Penelitian.....42
2.11 Model Analisis.....43
2.12 Hipotesis Penelitian.....44
BAB III METODE PENELITIAN.....46
3.1 Jenis Penelitian.....46
3.2 Jadwal Penelitian.....47
3.3 Alat dan Bahan Penelitian.....49
3.3.1 Perangkat Keras (Hardware).....49
3.3.2 Perangkat Lunak (Software).....49
3.4 Sumber Data.....49
3.5 Metode Pengumpulan Data.....50
3.6 Metode yang Digunakan.....50
3.7 Perancangan Model Analisis.....51
3.8 Metode Pengujian.....52
BAB IV HASIL DAN PEMBAHASAN.....54
4.1 Pengumpulan Data.....54
4.1.1 Sumber Data dan Proses Pengunduhan.....54
4.1.2 Gambaran Umum Dataset.....55
4.1.3 Deskripsi Variabel.....56
4.2 Pembersihan Data.....58
4.2.1 Pemeriksaan Missing Value.....58
4.2.2 Pemeriksaan Data Duplikat.....59
4.3 Transformasi Data.....61
4.3.1 Pemilihan Variabel Penelitian.....61

4.3.2 One-Hot Encoding 62
4.3.3 Min-Max Scaling 65
4.3.4 Perhitungan Engagement Rate (ERF dan ERR) 67
4.4 Analisis Data 69
4.4.1 Pembentukan Model Multiple Linear Regression 69
4.4.2 Uji Multikolinearitas (VIF) 72
4.4.3 Uji Simultan (Uji F) 74
4.4.4 Uji Parsial (Uji t) 75
4.4.5 Koefisien Determinasi (R^2) 77
4.5 Interpretasi Hasil 79
4.5.1 Pembahasan Hasil Model ERF 79
4.5.2 Pembahasan Hasil Model ERR 81
4.5.3 Perbandingan ERF dan ERR 81
BAB V KESIMPULAN DAN SARAN 84
5.1 Simpulan 84
5.2 Saran 85
DAFTAR PUSTAKA
LAMPIRAN

DAFTAR GAMBAR

Gambar 2.1 Diagram Model Analisis 44
Gambar 3.1 Flowchart Perancangan Model Analisis.....52
Gambar 4.1 Download Dataset.....54
Gambar 4.2 Struktur Awal Dataset 55
Gambar 4.3 Pemeriksaan Missing Value 59
Gambar 4.4 Pemeriksaan Data Duplikat.....60
Gambar 4.5 Pemilihan Variabel.....61
Gambar 4.6 Perubahan Setelah Proses One-Hot Encoding 63
Gambar 4.7 Perubahan kategorikal menjadi nilai dummy.....64
Gambar 4.8 Perubahan Min-Max Scaling.....66
Gambar 4.9 Engagement Rate by Follower 68
Gambar 4.10 Engagement Rate by Reach.....68
Gambar 4.11 Koefisien Regresi Multiple Linear Regression terhadap ERF 71
Gambar 4.12 Koefisien Regresi Multiple Linear Regression terhadap ERR.....72
Gambar 4.13 Hasil Uji VIF 73
Gambar 4.14 Hasil Uji t Model ERF dan ERR 76

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu.....31
Tabel 3.1 Jadwal Penelitian 48
Tabel 4.1 Variabel Dataset 57
Tabel 4.2 Hasil Uji F 74
Tabel 4.3 Hasil Uji R^2.....78

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kemajuan di bidang teknologi informasi selama beberapa dekade terakhir ini benar-benar mengubah banyak sisi kehidupan masyarakat. Salah satu dampak yang paling terasa adalah cara orang mendapatkan informasi. Selain itu, cara mereka berinteraksi satu sama lain juga ikut berubah. Berbagi data secara elektronik pun jadi jauh lebih mudah dan cepat dari sebelumnya [1]. Kemajuan teknologi internet yang ditunjang oleh perkembangan perangkat komputer dan telepon pintar sudah membuat masyarakat bisa saling terhubung tanpa batasan ruang dan waktu. Perkembangan ini juga ikut mendorong lahirnya banyak platform media sosial. Platform-platform ini menjadi salah satu alat komunikasi yang paling sering dipakai di seluruh dunia. Media sosial tidak hanya berfungsi sebagai tempat interaksi antar individu. Lebih dari itu, media sosial juga dipakai untuk menyebarkan informasi, menjalankan pemasaran digital, mempromosikan produk, memberikan edukasi, sampai membangun hubungan antara organisasi dengan masyarakat. Karena jumlah data yang dihasilkan dari aktivitas pengguna media sosial sangat besar, platform ini jadi salah satu objek yang sering diteliti di bidang teknologi informasi. Khususnya di bidang data analytics - bidang yang fokus pada penggalian informasi dan pola dari kumpulan data dalam jumlah besar [2].

Salah satu media sosial yang memiliki jumlah pengguna aktif terbesar di dunia adalah Instagram [3]. Platform ini memberi penggunanya ruang untuk

menyebarkan aneka konten digital - seperti foto, video, carousel, dan reel. Semua itu bisa dijangkau oleh orang lain lewat beranda, halaman eksplorasi, profil, maupun umpan reel. Seiring makin banyaknya fitur yang ditambahkan, Instagram tidak cuma dipakai buat berbagi cerita pribadi. Platform ini juga tumbuh jadi tempat untuk pemasaran digital, membangun merek pribadi, promosi bisnis, kampanye sosial, dan penyebaran informasi dari banyak organisasi atau institusi [4]. Oleh karena itu, keberhasilan sebuah unggahan di Instagram menjadi aspek yang penting untuk dianalisis, terutama bagi individu maupun organisasi yang memanfaatkan platform tersebut sebagai media komunikasi dan penyampaian informasi kepada audiens.

Salah satu cara terbaik untuk mengukur keberhasilan unggahan Instagram adalah dengan melihat seberapa banyak interaksi yang dilakukan pengguna dengan konten yang diposting [5]. Interaksi ini disebut sebagai keterlibatan, yang mencakup berbagai tindakan dari pengguna seperti menyukai (like), berkomentar (comment), membagikan konten (share), maupun menyimpan konten (save). Semakin banyak interaksi yang diterima oleh sebuah konten, maka semakin besar juga keterlibatan audiens terhadap materi tersebut. Untuk mengukur tingkat keterlibatan tersebut secara lebih objektif digunakan indikator *engagement rate*, yaitu ukuran yang menunjukkan tingkat interaksi pengguna terhadap suatu unggahan berdasarkan karakteristik tertentu yang dimiliki oleh unggahan tersebut [6]. Dibandingkan hanya menggunakan jumlah *likes* atau komentar, *engagement rate* dinilai mampu memberikan gambaran yang lebih representatif mengenai

kinerja suatu unggahan karena mempertimbangkan tingkat keterlibatan audiens secara keseluruhan.

Instagram menjadi lebih populer sebagai platform komunikasi dan pemasaran online, sehingga kualitas setiap unggahan harus menjadi perhatian utama bagi pemilik akun dan organisasi [7]. Jumlah interaksi yang diterima oleh unggahan bukan satu-satunya faktor yang mempengaruhi kinerjanya. Berbagai karakteristik unggahan juga mempengaruhi kinerjanya [8]. Sejumlah ciri yang biasanya dianalisis dalam studi media sosial mencakup jumlah pengikut (*follower count*), jenis konten yang dipublikasikan (*media type*), kategori konten (*content category*), sumber lalu lintas (*traffic source*), penggunaan ajakan bertindak (*call to action*), waktu publikasi (*post hour*), panjang keterangan (*caption length*), serta jumlah tagar (*hashtags count*) [9]. Karakteristik-karakteristik tersebut diduga memiliki hubungan dengan tingkat keterlibatan audiens sehingga sering dijadikan variabel dalam penelitian mengenai performa konten di media sosial.

Walaupun demikian, keterkaitan antara fitur unggahan dan tingkat keterlibatan tidak selalu menghasilkan temuan yang seragam di setiap studi maupun setiap set data [10]. Perbedaan perilaku pengguna media sosial, jenis akun, karakteristik audiens, serta algoritma distribusi konten yang terus berkembang memungkinkan suatu variabel memberikan pengaruh yang berbeda pada kondisi yang berbeda pula [11]. Sebagai contoh, penggunaan *hashtags* yang dalam beberapa penelitian dianggap mampu meningkatkan jangkauan unggahan belum tentu memberikan pengaruh yang sama pada seluruh jenis akun. Sama halnya dengan panjang caption dan waktu publikasi yang dalam situasi tertentu bisa

menunjukkan hubungan yang tidak terlalu kuat dengan tingkat engagement. Karena itu, analisis perlu dilakukan berdasarkan data yang dipakai supaya kesimpulan yang didapat benar-benar sesuai dengan ciri khas dataset yang sedang diteliti.

Dengan kemajuan teknologi informasi, pengolahan data menjadi lebih populer sebagai metode pengolahan dan analisis data yang sangat besar. Dengan demikian, proses pengambilan keputusan dapat dibantu dengan informasi yang dihasilkan dari pengolahan data ini [12]. Dalam konteks media sosial, *data analytics* memungkinkan peneliti untuk mengidentifikasi pola hubungan antarvariabel, mengevaluasi karakteristik unggahan, serta memperoleh gambaran mengenai faktor-faktor yang diduga memengaruhi tingkat *engagement [13]*. Metode ini memungkinkan analisis bergantung pada pengamatan deskriptif dan didukung oleh proses pengolahan data statistik. Dengan demikian, hasil penelitian menjadi lebih objektif dan dapat dipertanggungjawabkan secara ilmiah.

Penelitian ini memanfaatkan dataset Instagram Analytics Dataset yang tersedia pada platform Kaggle dan berisi 29.999 data unggahan Instagram. Dataset tersebut memuat berbagai informasi mengenai karakteristik unggahan, seperti jumlah pengikut akun, jenis media, kategori konten, sumber lalu lintas, waktu publikasi, panjang *caption*, jumlah *hashtags*, serta nilai *engagement rate*. Ketersediaan data dalam jumlah besar dengan variasi karakteristik yang cukup beragam memberikan peluang untuk melakukan analisis terhadap hubungan antara karakteristik unggahan dengan *engagement rate*. Selain itu, penggunaan dataset publik juga memungkinkan penelitian ini untuk direplikasi maupun dibandingkan dengan penelitian lain yang menggunakan sumber data serupa.

Dalam penelitian ini, analisis terhadap hubungan antara karakteristik unggahan Instagram dengan *engagement rate* dilakukan menggunakan pendekatan *data analytics* melalui proses pengolahan data secara sistematis. Analisis terdiri dari beberapa proses, seperti menyiapkan data awal, mengidentifikasi variabel yang akan dipelajari, mengubah data kategori dengan teknik Encoding Satu Panas, menyesuaikan data angka dengan pengukuran Min-Max, dan menerapkan analisis statistik menggunakan regresi linier ganda. Pilihan regresi linier berganda didasarkan pada kemampuan untuk menganalisis pengaruh beberapa variabel independen terhadap satu variabel dependen secara bersamaan, sehingga metode ini dinilai sesuai untuk mengevaluasi hubungan antara karakteristik unggahan dengan tingkat keterlibatan. Untuk mengevaluasi model regresi, pengujian multikolinearitas juga dilakukan [14].

Setiap variabel memiliki karakteristik dan skala data unik, menurut penelitian awal yang dilakukan terhadap kumpulan data. Oleh karena itu, tahapan preprocessing sangat penting sebelum melakukan analisis. Oleh karena itu, untuk memastikan bahwa semua variabel independen dalam penelitian berada dalam rentang nilai yang sama, normalisasi diterapkan pada variabel numerik yang independen dengan menggunakan metode Min-Max scaling [15]. Langkah-langkah tersebut dilakukan untuk meningkatkan kestabilan model analisis dan memastikan bahwa hasil yang diperoleh benar-benar mewakili karakteristik data yang digunakan. Akibatnya, hasil penelitian diharapkan dapat memberikan gambaran yang lebih objektif tentang hubungan antara fitur unggahan Instagram dan tingkat keterlibatan berdasarkan data yang dianalisis.

Menurut uraian tersebut, tujuan penelitian ini adalah untuk mengevaluasi hubungan antara fitur unggahan Instagram dan tingkat keterlibatan dengan menggunakan pendekatan data analytics yang digunakan dengan dataset publik Kaggle. Hasil penelitian ini diharapkan dapat membantu memahami lebih lanjut tentang hubungan antara fitur unggahan Instagram dan tingkat keterlibatan, dan juga dapat berfungsi sebagai referensi untuk penelitian selanjutnya yang mengkaji analisis media sosial menggunakan pendeka Penelitian ini diharapkan dapat membantu praktisi dan pengelola akun media sosial memahami karakteristik unggahan yang terkait dengan tingkat keterlibatan audiens, menurut hasil analisis data.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah diuraikan, masalah penelitian ini dapat diidentifikasi sebagai berikut :

1. Belum optimalnya pemanfaatan data Instagram untuk analisis.
2. Belum diketahui hubungan karakteristik unggahan dengan *Engagement Rate*.
3. Diperlukan pendekatan *Data Analytics* untuk menganalisis hubungan antarvariabel.
4. Belum diketahui pengaruh masing-masing variabel terhadap *Engagement Rate* menggunakan Multiple Linear Regression.

1.3 Rumusan Masalah

1. Bagaimana karakteristik data unggahan Instagram berdasarkan dataset yang digunakan dalam penelitian ?

2. Apakah variabel seperti jumlah follower, jenis media, kategori konten, sumber trafik, has call to action, jam posting, panjang caption, dan jumlah hashtags berdampak pada tingkat keterlibatan Instagram?
3. Bagaimana hasil analisis pengaruh masing-masing faktor terhadap engagement rate berdasarkan Multiple Linear Regression?

1.4 Batasan Masalah

Ruang lingkup penelitian dibatasi sebagai berikut untuk membuatnya lebih terfokus:

1. Dengan 29.999 data unggahan, dataset Instagram Analytics yang digunakan diambil dari platform Kaggle.
2. Jumlah pengikut, jenis media, kategori konten, sumber trafik, has call to action, jam posting, panjang caption, dan jumlah hashtag adalah variabel independen yang dianalisis.
3. Variabel dependen yang digunakan adalah engagement rate.
4. Untuk variabel kategorikal, one-hot encoding dan min-max scaling untuk variabel numerik adalah dua proses preprocessing data.
5. Proses analisis dilakukan menggunakan bahasa pemrograman Python di platform Google Colaboratory.
6. Regresi linier berganda atau juga dikenal sebagai Multiple Linear Regression digunakan sebagai metode analisis.

Penelitian ini tidak mengumpulkan data secara langsung dari platform Instagram, hanya memanfaatkan kumpulan data yang ada saat ini.

1.5 Tujuan Penelitian

Tujuan penelitian ini, berdasarkan rumusan masalah yang telah dibuat, adalah:

1. Menganalisis karakteristik data unggahan Instagram berdasarkan dataset yang digunakan.
2. Menganalisis dampak jumlah pengikut, jenis media, kategori konten, sumber lalu lintas, adanya panggilan untuk bertindak, waktu posting, panjang keterangan, dan jumlah hashtag terhadap tingkat keterlibatan dari unggahan Instagram dengan menggunakan metode Analisis Data.

1.6 Manfaat Penelitian

1. Manfaat Teoritis

Dengan menggunakan pendekatan data analytics, penelitian ini diharapkan dapat menambah pengetahuan ilmiah tentang bagaimana karakteristik unggahan berhubungan dengan tingkat keterlibatan.

2. Manfaat Praktis

1. Bagi akademisi, penelitian ini diharapkan berfungsi sebagai referensi untuk penelitian terkait analisis media sosial.
2. Penelitian ini diharapkan dapat membantu para profesional pemasaran digital memahami analisis data konten Instagram..
3. Bagi penulis, penelitian ini akan membantu mereka menerapkan pengetahuan yang mereka pelajari selama kuliah, terkhusus dalam bidang data analysis dan statistical analysis.

1.7 Sistematika Penulisan

Studi ini terbagi ke dalam lima bagian yang dapat dijelaskan sebagai berikut.

BAB I PENDAHULUAN

Bagian ini mencakup latar belakang, pengidentifikasian masalah, batasan analisis, tujuan penelitian, manfaat yang diperoleh dari studi ini, serta struktur penulisan.

BAB II TINJAUAN PUSTAKA

Seksi ini berisi teori dasar yang berkaitan dengan penelitian, studi yang telah dilakukan sebelumnya, kerangka pemikiran, dan hipotesis yang diusulkan.

BAB III METODE PENELITIAN

Bagian ini menggambarkan jenis penelitian, sumber data yang digunakan, variabel yang dianalisis, langkah-langkah dalam pengolahan data, teknik analisis yang diterapkan, serta perangkat lunak yang digunakan.

BAB IV HASIL DAN PEMBAHASAN

Bagian ini menyampaikan temuan dari analisis data, penjelasan terkait hasil regresi linier ganda, serta diskusi yang didasarkan pada teori dan penelitian yang telah dilakukan sebelumnya.

BAB V KESIMPULAN DAN SARAN

Bagian ini menyajikan ringkasan dan rekomendasi berdasarkan hasil dari penelitian yang telah dilaksanakan..

DAFTAR PUSTAKA

[1] H. Jiao, T. Wang, D. Libaers, J. Yang, and L. Hu, "The relationship between digital technologies and innovation: A review, critique, and research agenda," *Journal of Innovation and Knowledge*, vol. 10, no. 1, Jan. 2025, doi: 10.1016/j.jik.2024.100638.

[2] A. A. Vărzaru and C. G. Bocean, "Digital Transformation and Innovation: The Influence of Digital Technologies on Turnover from Innovation Activities and Types of Innovation," *Systems*, vol. 12, no. 9, Sep. 2024, doi: 10.3390/systems12090359.

[3] D. F. Cordeiro, M. Vállez, C. I. Font-Julian, and J. Guallar, "Instagram Engagement and Content Strategies of US and UK Legacy Media: A Quantitative Analysis of Five Leading News Outlets," *Journalism and Media*, vol. 6, no. 2, Jun. 2025, doi: 10.3390/journalmedia6020089.

[4] M. Sohaib, M. A. Ali, and M. Ahmad-ur-Rehman, "Social Media Marketing Drives Brand Love and Customer Behavioral Engagement: Gender as a Moderator," *Journal of Theoretical and Applied Electronic Commerce Research* , vol. 20, no. 4, Dec. 2025, doi: 10.3390/jtaer20040344.

[5] M. Trunfio and S. Rossi, "Conceptualising and measuring social media engagement: A systematic literature review," Sep. 01, 2021, *Springer International Publishing*. doi: 10.1007/s43039-021-00035-8.

[6] W. N. A. Rahman, D. S. Mutum, and E. M. Ghazali, "Consumer Engagement With Visual Content on Instagram: Impact of Different Features of Posts by Prominent Brands," *International Journal of E-Services and Mobile Applications*, vol. 14, no. 1, 2022, doi: 10.4018/IJESMA.295960.

[7] F. Herle, "The Effects Of Social Media Content Characteristics On Consumer Engagement-A Systematic Literature Review."

[8] S. P. Eslami, M. Ghasemaghaei, and K. Hassanein, "Understanding consumer engagement in social media: The role of product lifecycle," *Decis. Support Syst.*, vol. 162, Nov. 2022, doi: 10.1016/j.dss.2021.113707.

[9] M. Bonilla-Quijada, J. L. Del Olmo, D. Andreu, and J. Ripoll, "Customer engagement on Instagram for luxury fashion brands: An empirical comparative analysis," *Cogent Soc. Sci.*, vol. 9, no. 1, 2023, doi: 10.1080/23311886.2023.2235169.

[10] W. M. Lim and T. Rasul, "Customer engagement and social media: Revisiting the past to inform the future," *J. Bus. Res.*, vol. 148, pp. 325–342, Sep. 2022, doi: 10.1016/j.jbusres.2022.04.068.

[11] P. K. Theodoridis and D. C. Gkikas, “Maximizing Social Media User Engagement Through Predictive Analytics in Retail Tourism: Identifying Key Performance Indicators That Trigger User Interactions,” *Applied Sciences (Switzerland)*, vol. 15, no. 21, Nov. 2025, doi: 10.3390/app152111720.

[12] A. A. Kharlamov and M. Pilgun, “Data Analytics for Predicting Situational Developments in Smart Cities: Assessing User Perceptions,” *Sensors*, vol. 24, no. 15, Aug. 2024, doi: 10.3390/s24154810.

[13] S. Batistič and P. van der Laken, “History, Evolution and Future of Big Data and Analytics: A Bibliometric Analysis of Its Relationship to Performance in Organizations,” *British Journal of Management*, vol. 30, no. 2, pp. 229–251, Apr. 2019, doi: 10.1111/1467-8551.12340.

[14] M. Tranmer, J. Murphy, M. Elliot, and M. Pampaka, “Multiple Linear Regression (2 nd Edition),” 2020. [Online]. Available: https://hummedia.manchester.ac.uk/institutes/cmist/a

[15] S. Gopal, K. Patro, and K. Kumar Sahu, “Normalization: A Preprocessing Stage.” [Online]. Available: www.kiplinger.com,

[16] K. Ragazou, I. Passas, A. Garefalakis, E. Galariotis, and C. Zopounidis, “Big Data Analytics Applications in Information Management Driving Operational Efficiencies and Decision-Making: Mapping the Field of Knowledge with Bibliometric Analysis Using R,” *Big Data and Cognitive Computing*, vol. 7, no. 1, Mar. 2023, doi: 10.3390/bdcc7010013.

[17] Hamed Taherdoost, “The Role of Different Types of Management Information System Applications in Business Development: Concepts, and Limitations,” *Cloud Computing and Data Science*, pp. 31–48, Nov. 2022, doi: 10.37256/ccds.4120231959.

[18] R. Rossi and K. Hirama, “Characterizing Big Data Management,” 2015. [Online]. Available: http://iisit.org/Vol12/IISITv12p165-180Rossi1921.pdf

[19] I. H. Sarker, “Data Science and Analytics: An Overview from Data-Driven Smart Computing, Decision-Making and Applications Perspective,” Sep. 01, 2021, *Springer*. doi: 10.1007/s42979-021-00765-8.

[20] Hamed Taherdoost, “The Role of Different Types of Management Information System Applications in Business Development: Concepts, and Limitations,” *Cloud Computing and Data Science*, pp. 31–48, Nov. 2022, doi: 10.37256/ccds.4120231959.

[21] C. Li, “Preprocessing Methods and Pipelines of Data Mining: An Overview,” Jun. 2019, [Online]. Available: http://arxiv.org/abs/1906.08510

[22] I. C. Drivas, D. Kouis, D. Kyriaki-Manessi, and F. Giannakopoulou, “Social Media Analytics and Metrics for Improving Users Engagement,” *Knowledge*, vol. 2, no. 2, pp. 225–242, May 2022, doi: 10.3390/knowledge2020014.

[23] K. K. Kapoor, K. Tamilmani, N. P. Rana, P. Patil, Y. K. Dwivedi, and S. Nerur, “Advances in Social Media Research: Past, Present and Future,” *Information Systems Frontiers*, vol. 20, no. 3, pp. 531–558, Jun. 2018, doi: 10.1007/s10796-017-9810-y.

[24] O. Zimba and A. Y. Gasparyan, “Social media platforms: A primer for researchers,” *Reumatologia*, vol. 59, no. 2, pp. 68–72, 2021, doi: 10.5114/reum.2021.102707.

[25] G. Appel, L. Grewal, R. Hadi, and A. T. Stephen, “The future of social media in marketing,” *J. Acad. Mark. Sci.*, vol. 48, no. 1, pp. 79–95, Jan. 2020, doi: 10.1007/s11747-019-00695-1.

[26] Y. Hu, L. Manikonda, and S. Kambhampati, “What We Instagram: A First Analysis of Instagram Photo Content and User Types.” [Online]. Available: www.aaai.org

[27] L. D. Hollebeek, D. E. Sprott, and M. K. Brady, “Rise of the Machines? Customer Engagement in Automated Service Interactions,” Feb. 01, 2021, *SAGE Publications Inc.* doi: 10.1177/1094670520975110.

[28] V. A. Barger, J. W. Peltier, and D. E. Schultz, “Social media and consumer engagement: A review and research agenda †.”

[29] L. D. Hollebeek, M. S. Glynn, and R. J. Brodie, “Consumer brand engagement in social media: Conceptualization, scale development and validation,” *Journal of Interactive Marketing*, vol. 28, no. 2, pp. 149–165, 2014, doi: 10.1016/j.intmar.2013.12.002.

[30] M. Trunfio and S. Rossi, “Conceptualising and measuring social media engagement: A systematic literature review,” Sep. 01, 2021, *Springer International Publishing*. doi: 10.1007/s43039-021-00035-8.

[31] V. V. Pandora and Y. M. Djajalaksana, “Penerapan Digital Marketing Multichannel untuk Pemasaran Program Studi Sistem Informasi,” *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 8, no. 1, Apr. 2022, doi: 10.28932/jutisi.v8i1.4248.

[32] Yudhistira Akhmad Kurniawan and Widya Sastika, “Analisis Strategi Influencer Marketing Pada Instagram Studio Dapur Tahun 2024,” Jurnal Lentera Bisnis, vol. 13, no. 3, pp. 1538–1549, Sep. 2024, doi: 10.34127/jrlab.v13i3.1202.

[33] N. Roustaei, "Application and interpretation of linear-regression analysis," *Medical Hypothesis, Discovery, and Innovation in Ophthalmology*, vol. 13, no. 3, pp. 151–159, Sep. 2024, doi: 10.51329/mehdiophthal1506.

[34] O. Sureiman and C. M. Mangera, "F-Test of Overall Significance in Regression Analysis Simplified," *Journal of the Practice of Cardiovascular Sciences*, vol. 6, no. 2, pp. 116–122, May 2020, doi: 10.4103/jpcs.jpcs_18_20.

[35] Tae Kyun Kim, "T test as a parametric statistic," 2015. [Online]. Available: http://ekja.org

[36] J. H. Kim, "Multicollinearity and misleading statistical results," *Korean J. Anesthesiol.*, vol. 72, no. 6, pp. 558–569, Dec. 2019, doi: 10.4097/kja.19087.

[37] C. Herdian, A. Kamila, and I. G. Agung Musa Budidarma, "Studi Kasus Feature Engineering Untuk Data Teks: Perbandingan Label Encoding dan One-Hot Encoding Pada Metode Linear Regresi," *Technologia : Jurnal Ilmiah*, vol. 15, no. 1, p. 93, Jan. 2024, doi: 10.31602/tji.v15i1.13457.

[38] E. Poslavskaya and A. Korolev, "Encoding categorical data: Is there yet anything 'hotter' than one-hot encoding?," Dec. 2023, [Online]. Available: http://arxiv.org/abs/2312.16930

[39] R. Ramadhan Laska and A. Mudya Yolanda, "A Comparative Study of Z-Score and Min-Max Normalization for Rainfall Classification in Pekanbaru," 2024. [Online]. Available: http://ipublishing.intimal.edu.my/jods.html

[40] S. Sinsomboonthong, "Performance Comparison of New Adjusted Min-Max with Decimal Scaling and Statistical Column Normalization Methods for Artificial Neural Network Classification," *Int. J. Math. Math. Sci.*, vol. 2022, 2022, doi: 10.1155/2022/3584406.

[41] H. Arief, D. Putro Sarwo, E. Setiyawan Jullev, H. Yufit, and I. Husainy, "Measurement Of Engagement Rate On Instagram For Business Marketing (Case Study: MSME of Dowry In Jember)," 2020.

[42] D. A. R. K. M. Kristo Radion Purba, "Analysis and Prediction of Instagram Users Popularity using Regression Techniques based on Metadata, Media and Hashtags Analysis," 2020.

[43] S. Carta, A. S. Podda, D. R. Recupero, R. Saia, and G. Usai, "Popularity prediction of instagram posts," *Information (Switzerland)*, vol. 11, no. 9, Sep. 2020, doi: 10.3390/INFO11090453.

[44] P. P. Tricomi, M. Chilese, M. Conti, and A. R. Sadeghi, "Follow Us and Become Famous! Insights and Guidelines From Instagram Engagement

Mechanisms," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Apr. 2023, pp. 346–356. doi: 10.1145/3578503.3583623.

[45] E. Sanches and C. M. Q. Ramos, "Evaluating the Impact of Instagram Engagement Metrics on Corporate Revenue Growth: Introducing the Loyalty Rate," *Information (Switzerland)*, vol. 16, no. 4, Apr. 2025, doi: 10.3390/info16040287.

[46] C. Riis, D. K. Kowalczyk, and L. K. Hansen, "On the Limits to Multi-Modal Popularity Prediction on Instagram -- A New Robust, Efficient and Explainable Baseline," Feb. 2021, doi: 10.5220/0010377112001209.