



Department of Digital Business

Journal of Artificial Intelligence and Digital Business (RIGGS)

Homepage: <https://journal.ilmudata.co.id/index.php/RIGGS>

Vol. 5 No. 1 (2026) pp: 1497-1504

P-ISSN: 2963-9298, e-ISSN: 2963-914X

Analisis Waktu Optimal Unggah Instagram Menggunakan *K-Means Clustering* pada Data Sintetis

Fadhlan Ihsan Lubis, Muhammad Iqbal

Magister Teknologi Informasi, Universitas Pembangunan Panca Budi

fihsanlubis@gmail.com, muhammadiqbal@dosen.pancabudi.ac.id

Abstrak

Instagram merupakan salah satu platform media sosial yang banyak dimanfaatkan dalam strategi pemasaran digital karena kemampuannya dalam menjangkau audiens secara luas dan cepat. Efektivitas suatu unggahan pada Instagram dipengaruhi oleh berbagai faktor, salah satunya adalah waktu publikasi, yang berperan penting dalam menentukan tingkat jangkauan (*reach*) dan interaksi (*engagement*) pengguna. Penelitian ini bertujuan untuk menganalisis waktu unggah optimal Instagram berdasarkan performa *reach* dan *engagement* menggunakan metode *K-Means Clustering* pada data sintetis. Dataset yang digunakan terdiri dari 29.999 data posting sintetis yang merepresentasikan aktivitas unggahan media sosial, dengan variabel waktu unggah, jenis konten, *reach*, dan *engagement*. Tahapan penelitian mengacu pada kerangka Knowledge Discovery in Databases (KDD) yang meliputi proses seleksi data, pembersihan data, transformasi data, serta proses clustering untuk menemukan pola tersembunyi dalam data. Penentuan jumlah cluster dilakukan menggunakan Elbow Method dan menghasilkan tiga cluster yang merepresentasikan tingkat performa rendah, sedang, dan tinggi. Hasil analisis menunjukkan bahwa metode *K-Means* mampu mengelompokkan data posting Instagram secara efektif berdasarkan kesamaan karakteristik performanya. Berdasarkan analisis centroid, diperoleh satu cluster dengan nilai *reach* dan *engagement* tertinggi yang merepresentasikan waktu unggah paling optimal, yaitu hari Sabtu pukul 09.00 WIB. Temuan ini menunjukkan bahwa faktor hari memiliki pengaruh signifikan terhadap performa unggahan. Hasil penelitian ini diharapkan dapat menjadi dasar rekomendasi strategis bagi pengelola media sosial dalam menentukan waktu unggah yang optimal berbasis analisis data serta mendukung pengembangan strategi pemasaran digital yang lebih efektif.

Kata kunci: Instagram, Waktu Unggah Optimal, Data Sintetis, *K-Means Clustering*, *Reach*, *Engagement*.

1. Latar Belakang

Perkembangan teknologi informasi dan komunikasi yang pesat telah mendorong peningkatan penggunaan media sosial sebagai sarana utama dalam penyebaran informasi, komunikasi, serta pemasaran digital. Media sosial tidak hanya digunakan oleh individu, tetapi juga oleh pelaku usaha, instansi, dan organisasi sebagai media promosi dan interaksi dengan audiens. Salah satu faktor penting yang menentukan keberhasilan suatu unggahan di media sosial adalah waktu posting, karena waktu yang tepat dapat memengaruhi tingkat jangkauan (*reach*) dan interaksi (*engagement*) pengguna secara signifikan [1].

Seiring dengan meningkatnya volume data aktivitas media sosial, pendekatan *data mining* menjadi salah satu solusi yang tepat untuk menggali informasi tersembunyi dari kumpulan data yang besar. Salah satu teknik dalam *data mining* yang banyak digunakan untuk analisis pola adalah *clustering*, yaitu proses pengelompokan data berdasarkan tingkat kemiripan karakteristiknya [2]. Metode *K-Means Clustering* dipilih dalam penelitian ini karena memiliki keunggulan dalam hal kesederhanaan, kecepatan komputasi, serta efektivitas dalam menangani data berukuran besar [3].

Penelitian ini menggunakan data sintetis media sosial yang merepresentasikan aktivitas posting dengan variabel waktu posting, jenis konten, *reach*, dan *engagement*. Penggunaan data sintetis dilakukan untuk mensimulasikan kondisi nyata aktivitas media sosial tanpa melibatkan data pribadi pengguna. Melalui proses *clustering* menggunakan metode *K-Means*, data akan dikelompokkan ke dalam beberapa *cluster* berdasarkan kesamaan performanya, sehingga dapat diketahui kelompok waktu dengan tingkat *reach* dan *engagement* tertinggi [4].

Diharapkan hasil penelitian ini dapat memberikan rekomendasi waktu unggah yang optimal bagi pengelola media sosial serta menjadi referensi ilmiah dalam penerapan metode *K-Means* pada analisis performa media sosial. Selain itu, penelitian ini juga diharapkan dapat menjadi dasar pengembangan sistem pendukung keputusan dalam strategi pemasaran digital berbasis data.

2. Metode Penelitian

Penelitian ini dilaksanakan melalui serangkaian tahapan yang sistematis untuk mengidentifikasi pola interaksi pengguna secara komprehensif. Proses penelitian diawali dengan pengumpulan data aktivitas pengguna pada media sosial, kemudian dilanjutkan dengan penentuan jumlah sampel yang representatif guna menjamin keandalan hasil analisis. Selanjutnya, dilakukan pemodelan data dengan menerapkan metode *K-Means Clustering* untuk mengelompokkan waktu unggah yang optimal berdasarkan pola *engagement* yang terbentuk. Hasil pemodelan tersebut diharapkan mampu mengungkap keterkaitan antara waktu unggahan dan tingkat interaksi pengguna, sehingga dapat dijadikan dasar dalam perumusan strategi pengelolaan konten yang lebih efektif dan berbasis data.

2.1. Metode Pengumpulan Data

Data dalam penelitian ini diperoleh melalui metode pengumpulan data sekunder yang bersumber dari platform Kaggle. Pengumpulan data sekunder merupakan proses pemanfaatan data yang telah dihimpun dan diolah oleh pihak lain, serta memiliki relevansi dengan permasalahan yang diteliti, tanpa melibatkan pengambilan data secara langsung dari sumber utama [5]. Dataset yang digunakan memuat informasi terkait waktu unggah dan tingkat *engagement* pada media sosial. Data tersebut mencakup berbagai atribut, antara lain platform media sosial, jenis konten, waktu posting, karakteristik audiens (usia dan gender), sentimen audiens, minat audiens, lokasi audiens, serta jumlah interaksi yang meliputi like, share, dan komentar. Pemanfaatan dataset sekunder yang telah terstruktur dan siap dianalisis memungkinkan proses penelitian berlangsung lebih efisien serta mendukung keberagaman sampel yang merepresentasikan pola *engagement* pengguna secara nyata.

2.2. Sumber dan Jenis Data

Data yang digunakan dalam penelitian ini berupa data sintesis media sosial yang merepresentasikan aktivitas unggah pada platform media sosial. Data disimpan dalam format CSV dengan jumlah sebanyak 29.999 data unggah. Setiap data memuat beberapa atribut utama yang digunakan dalam proses analisis, yaitu: Waktu unggah (tanggal dan jam), Jenis konten, *Reach*, *Engagement* (jumlah *like*, komentar, *share*, dan *save*).

Penggunaan data sintesis bertujuan untuk mensimulasikan kondisi nyata aktivitas media sosial tanpa melibatkan data pribadi pengguna.

2.3. Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri dari variabel independen: waktu posting (jam dan hari) dan jenis konten, variabel dependen: *reach* dan *engagement*.

Nilai *engagement* dihitung sebagai akumulasi dari jumlah *like*, *comment*, *share*, dan *save* pada setiap unggahan.

2.4. Tahapan Penelitian

Tahapan penelitian mengacu pada konsep *Knowledge Discovery in Databases* (KDD) yang terdiri dari beberapa langkah sebagai berikut:

2.4.1 Data Selection

Pada tahap ini dilakukan pemilihan atribut yang relevan dengan tujuan penelitian, yaitu waktu posting, jenis konten, *reach*, dan *engagement* [6].

2.4.2 Data Cleaning

Proses pembersihan data dilakukan untuk menghilangkan data yang tidak lengkap (*missing value*), data duplikat, serta nilai-nilai ekstrem (*outlier*) yang berpotensi mengganggu hasil *clustering* [7].

2.4.3 Data Transformation

Data yang telah dibersihkan kemudian ditransformasi agar dapat diproses oleh algoritma *K-Means*. Proses transformasi meliputi, ekstraksi jam dan hari dari atribut waktu unggah, proses *encoding* pada data jenis konten agar berbentuk numerik, normalisasi data pada variabel reach dan engagement menggunakan metode *Min-Max Normalization* agar seluruh atribut berada pada skala yang sama [8].

2.4.4 Data Mining (Clustering dengan K-Means)

Clustering merupakan salah satu teknik dalam analisis data yang digunakan untuk mengidentifikasi kesamaan dan membentuk kelompok-kelompok data berdasarkan karakteristik tertentu [9]. Tujuan utama dari teknik *clustering* adalah mengungkap pola atau struktur tersembunyi di dalam kumpulan data sehingga dapat memberikan pemahaman yang lebih baik mengenai pembagian kelompok atau kategori yang ada. Salah satu metode *clustering* yang paling banyak diterapkan dalam analisis data adalah *K-Means Clustering*. Algoritma *K-Means* bekerja dengan cara mengelompokkan data ke dalam sejumlah K klaster yang telah ditentukan sebelumnya, di mana nilai K merepresentasikan jumlah klaster yang diinginkan [10]. Proses *clustering* dilakukan menggunakan algoritma *K-Means* dengan langkah-langkah seperti, Menentukan jumlah cluster (K) menggunakan pendekatan *Elbow Method*, Menentukan *centroid* awal secara acak, Menghitung jarak setiap data ke pusat *cluster* menggunakan *Euclidean Distance*, Mengelompokkan data ke dalam *cluster* terdekat, Memperbarui *centroid* hingga diperoleh kondisi konvergen [11].

2.4.5 Evaluation

Evaluasi dilakukan dengan menganalisis nilai *centroid* dari masing-masing cluster untuk mengetahui karakteristik setiap kelompok waktu unggah. *Cluster* dengan nilai rata-rata reach dan engagement tertinggi ditetapkan sebagai kelompok waktu posting paling efektif [12,13].

2.4.6 Knowledge Discovery

Pada tahap akhir, dilakukan interpretasi hasil *clustering* untuk menghasilkan pengetahuan baru berupa pola waktu unggah yang optimal berdasarkan data yang telah dianalisis [14,15].

2.5. Metode Perhitungan

Algoritma *K-Means* menggunakan perhitungan jarak *Euclidean* untuk menentukan kedekatan data dengan pusat *cluster*, yang dirumuskan seperti pada rumus 1.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Sedangkan perhitungan pusat *cluster* (*centroid*) dilakukan dengan menggunakan rata-rata nilai seluruh data dalam satu *cluster*, yang dirumuskan seperti pada rumus 2.

$$C_k = \frac{1}{n} \sum_{i=1}^n x_i \quad (2)$$

2.6. Perangkat Lunak yang Digunakan

Proses pengolahan dan analisis data dalam penelitian ini dilakukan dengan bantuan perangkat lunak berbasis Python menggunakan pustaka pendukung seperti *pandas*, *scikit-learn*, dan *numpy* untuk proses normalisasi dan *clustering* menggunakan algoritma *K-Means*.

3. Hasil dan Diskusi

3.1. Deskripsi Data

Dataset yang digunakan dalam penelitian ini berjumlah 29.999 data posting Instagram yang terdiri dari atribut waktu unggah, jenis konten, jumlah *like*, *comment*, *share*, *save*, serta *reach*. Seluruh data telah melalui tahap *data cleaning* untuk menghilangkan duplikasi dan nilai yang tidak valid. Selanjutnya dilakukan ekstraksi fitur waktu berupa jam unggah (0–23) dan hari unggah (0–6), serta perhitungan *engagement* total yang merupakan akumulasi dari *like*, komentar, *share*, dan *save*.

Sebelum proses *clustering* dilakukan, seluruh atribut numerik dinormalisasi menggunakan metode *Min-Max Normalization* agar berada dalam rentang nilai yang sama, sehingga tidak terjadi dominasi antar variabel dalam proses perhitungan jarak.

3.2. Penentuan Jumlah Cluster

Penentuan jumlah *cluster* (K) pada penelitian ini dilakukan dengan menggunakan pendekatan *Elbow Method*, yaitu suatu teknik evaluasi dalam proses *clustering* yang bertujuan untuk memperoleh jumlah *cluster* yang paling optimal berdasarkan karakteristik distribusi data. Pendekatan ini dilakukan dengan cara menganalisis perubahan nilai *Sum of Squared Error* (SSE) terhadap variasi jumlah *cluster* yang diuji secara bertahap. Nilai SSE merupakan ukuran total jarak kuadrat antara setiap data terhadap *centroid cluster* tempat data tersebut berada. Semakin kecil nilai SSE yang dihasilkan, maka semakin baik tingkat homogenitas data dalam suatu *cluster*, karena jarak antar data dengan pusat *cluster* menjadi semakin dekat.

Dalam implementasinya, pengujian dilakukan dengan mengiterasikan nilai K dalam beberapa rentang tertentu, kemudian menghitung nilai SSE untuk masing-masing jumlah *cluster* yang diuji. Selanjutnya, hasil perhitungan divisualisasikan dalam bentuk grafik hubungan antara jumlah *cluster* (K) dengan nilai SSE. Secara umum, grafik SSE akan menunjukkan pola penurunan yang cukup signifikan pada awal penambahan jumlah *cluster*, kemudian mengalami perlambatan penurunan pada titik tertentu. Titik perlambatan penurunan tersebut dikenal sebagai titik siku (*elbow point*), yang mengindikasikan bahwa penambahan jumlah *cluster* setelah titik tersebut tidak lagi memberikan peningkatan kualitas *clustering* yang signifikan.

Berdasarkan hasil pengujian yang telah dilakukan pada dataset penelitian, diperoleh bahwa titik siku paling optimal berada pada nilai K = 3. Pada titik ini, terjadi keseimbangan antara kompleksitas model *clustering* dengan kualitas pengelompokan data yang dihasilkan. Jika jumlah *cluster* ditambahkan lebih dari tiga, maka penurunan nilai SSE yang terjadi relatif kecil dan tidak memberikan perbedaan signifikan terhadap kualitas pengelompokan data. Sebaliknya, jika jumlah *cluster* kurang dari tiga, maka tingkat variasi data dalam setiap *cluster* masih tergolong tinggi, sehingga *cluster* yang terbentuk belum mampu merepresentasikan karakteristik data secara optimal.

Dengan demikian, pada penelitian ini digunakan tiga *cluster* utama yang masing-masing merepresentasikan kelompok waktu unggah berdasarkan tingkat performa yang berbeda, yaitu *cluster* dengan performa rendah, *cluster* dengan performa sedang, serta *cluster* dengan performa tinggi. Pembagian ini memungkinkan proses analisis menjadi lebih terstruktur, karena setiap *cluster* dapat dianalisis berdasarkan karakteristik *centroid* yang dimilikinya, khususnya pada variabel performa *reach* dan *engagement*.

Pemilihan jumlah *cluster* yang optimal melalui pendekatan *Elbow Method* juga berperan penting dalam meningkatkan validitas hasil *clustering*. Dengan jumlah *cluster* yang tepat, proses segmentasi data menjadi lebih representatif terhadap pola distribusi data yang sebenarnya. Hal ini mendukung proses interpretasi hasil *clustering* dalam kerangka *Knowledge Discovery in Databases* (KDD), sehingga pengetahuan yang dihasilkan dari proses data mining dapat memiliki tingkat akurasi dan reliabilitas yang lebih baik.

Secara metodologis, penerapan *Elbow Method* dalam penelitian ini menunjukkan bahwa proses penentuan parameter dalam algoritma *K-Means* tidak dilakukan secara subjektif, melainkan berdasarkan pendekatan kuantitatif yang terukur. Hal ini memperkuat bahwa model *clustering* yang dihasilkan memiliki dasar analisis yang kuat dan dapat dipertanggungjawabkan secara ilmiah, khususnya dalam konteks analisis pola waktu unggah optimal pada media sosial berbasis data.

3.3. Hasil *Clustering* Menggunakan *K-Means*

Proses *clustering* dengan algoritma *K-Means* menghasilkan tiga kelompok waktu unggah, dengan karakteristik *centroid* seperti pada Tabel 1.

Tabel 1. Hasil pengelompokan waktu unggah

Cluster	Jam Posting	Hari	Rata-rata Reach	Rata-rata Engagement
C0	09.00	Rabu	992.897	115.002
C1	09.00	Sabtu	1.001.917	115.222
C2	09.00	Selasa	996.819	114.516

Berdasarkan tabel tersebut, dapat dilihat bahwa *Cluster 1* (C1) memiliki nilai rata-rata *reach* dan *engagement* tertinggi, sehingga ditetapkan sebagai *cluster* terbaik.

3.4. Identifikasi Waktu Unggah Paling Optimal

Cluster dengan performa tertinggi (C1) didominasi oleh posting yang dilakukan pada hari Sabtu pukul 09.00 WIB.

Nilai rata-rata *reach* pada *cluster* ini mencapai $\pm 1.001.917$ akun, sedangkan nilai *engagement* total mencapai ± 115.222 interaksi. Nilai ini lebih tinggi dibandingkan dua *cluster* lainnya, yaitu *cluster 0* (Rabu) dan *cluster 2* (Selasa), meskipun ketiganya berada pada jam posting yang relatif sama, yaitu pukul 09.00 WIB.

Temuan ini menunjukkan bahwa hari unggah memiliki pengaruh yang lebih signifikan dibandingkan jam unggah, karena meskipun jam relatif sama, perbedaan hari menghasilkan performa *reach* dan *engagement* yang berbeda.

3.5. Pembahasan

Hasil penelitian menunjukkan bahwa metode *K-Means Clustering* mampu mengelompokkan data unggah Instagram berdasarkan tingkat performa *reach* dan *engagement* secara efektif. Kemampuan algoritma dalam mengelompokkan data berdasarkan kesamaan karakteristik menunjukkan bahwa pendekatan *clustering* dapat digunakan untuk mengidentifikasi pola performa unggahan secara objektif. Diperolehnya *cluster* terbaik pada hari Sabtu pukul 09.00 WIB mengindikasikan bahwa aktivitas pengguna media sosial pada waktu tersebut berada pada tingkat yang relatif tinggi, sehingga peluang konten untuk memperoleh jangkauan dan interaksi yang lebih besar juga meningkat. Hal ini menunjukkan bahwa kombinasi faktor waktu dan perilaku pengguna memiliki pengaruh langsung terhadap performa distribusi konten pada media sosial.

Secara perilaku pengguna, hari Sabtu cenderung menjadi periode dengan tingkat aktivitas digital yang lebih tinggi dibandingkan hari kerja. Hal ini disebabkan oleh berkurangnya aktivitas formal seperti pekerjaan atau kegiatan akademik, sehingga pengguna memiliki lebih banyak waktu luang untuk mengakses media sosial. Selain itu, waktu pagi sekitar pukul 09.00 WIB dapat dikategorikan sebagai waktu transisi setelah aktivitas pagi dimulai, seperti setelah sarapan atau persiapan kegiatan harian. Pada periode ini, pengguna cenderung membuka media sosial untuk mencari informasi, hiburan, maupun berinteraksi dengan konten digital. Kondisi tersebut berpotensi meningkatkan peluang konten untuk muncul pada *timeline* pengguna dan meningkatkan probabilitas interaksi.

Dari sisi analisis data, hasil penelitian ini menunjukkan bahwa variabel hari memiliki pengaruh yang lebih dominan dibandingkan jam posting, meskipun pada hasil *clustering* terlihat bahwa jam optimal relatif sama. Hal ini menunjukkan bahwa perilaku pengguna media sosial lebih dipengaruhi oleh pola aktivitas mingguan dibandingkan pola aktivitas harian. Dengan demikian, strategi waktu unggah sebaiknya tidak hanya mempertimbangkan jam posting, tetapi juga mempertimbangkan karakteristik hari tertentu.

Hasil penelitian ini juga sejalan dengan beberapa penelitian sebelumnya yang menyatakan bahwa periode akhir pekan dan waktu pagi memiliki kecenderungan menghasilkan tingkat interaksi yang lebih tinggi dibandingkan waktu lainnya. Keselarasan hasil ini memperkuat bahwa pendekatan data mining, khususnya metode clustering, dapat digunakan sebagai dasar dalam menentukan strategi waktu unggah yang optimal secara sistematis dan berbasis data.

Selain itu, penggunaan data sintetis dalam penelitian ini menunjukkan bahwa simulasi data dapat digunakan untuk merepresentasikan pola perilaku pengguna media sosial secara umum tanpa melibatkan data pribadi pengguna. Hal ini penting dalam konteks etika penelitian dan perlindungan data pribadi. Data sintetis memungkinkan peneliti untuk melakukan eksplorasi pola data dalam skala besar tanpa melanggar privasi pengguna, sekaligus memberikan fleksibilitas dalam pengembangan model analisis.

Dari sisi metodologi, keberhasilan penerapan K-Means pada data sintetis menunjukkan bahwa algoritma ini memiliki tingkat generalisasi yang baik, sehingga dapat digunakan tidak hanya pada data nyata, tetapi juga pada data simulasi untuk keperluan pengujian model dan perancangan sistem rekomendasi. Hal ini membuka peluang pengembangan sistem pendukung keputusan berbasis kecerdasan buatan yang mampu memberikan rekomendasi waktu unggah secara otomatis berdasarkan pola data historis.

Secara praktis, hasil penelitian ini dapat dimanfaatkan oleh pengelola media sosial, pelaku bisnis digital, maupun praktisi pemasaran digital dalam menyusun strategi distribusi konten yang lebih efektif. Dengan memanfaatkan pendekatan analisis data, keputusan penentuan waktu unggah tidak lagi bersifat subjektif, tetapi dapat didasarkan pada pola perilaku pengguna yang terukur secara statistik.

3.6. Implikasi Hasil Penelitian

Berdasarkan hasil proses clustering yang telah dilakukan, penelitian ini memberikan beberapa implikasi penting baik dari sisi praktis, metodologis, maupun akademis. Secara praktis, hasil penelitian menunjukkan bahwa waktu unggah konten memiliki pengaruh signifikan terhadap performa jangkauan (*reach*) dan interaksi (*engagement*) pengguna. Temuan bahwa waktu unggah paling optimal berada pada hari Sabtu pukul 09.00 WIB dapat dijadikan sebagai acuan strategis bagi pengelola media sosial dalam merancang jadwal publikasi konten, khususnya untuk konten yang bersifat promosi, kampanye pemasaran, maupun informasi penting yang membutuhkan tingkat visibilitas tinggi. Dengan memanfaatkan waktu unggah yang optimal, peluang konten untuk menjangkau audiens yang lebih luas serta memperoleh interaksi yang lebih tinggi dapat meningkat secara signifikan.

Dari sisi pengembangan sistem, penelitian ini menunjukkan bahwa metode K-Means Clustering dapat digunakan sebagai fondasi dalam pengembangan sistem rekomendasi waktu unggah berbasis data. Implementasi metode ini memungkinkan pengelola media sosial atau pelaku bisnis digital untuk melakukan analisis performa konten secara otomatis dan berkelanjutan. Dengan integrasi ke dalam sistem pendukung keputusan, hasil clustering dapat dimanfaatkan untuk memberikan rekomendasi waktu unggah secara dinamis berdasarkan pola data terbaru, sehingga strategi pemasaran digital dapat menjadi lebih adaptif terhadap perubahan perilaku pengguna.

Selain itu, secara metodologis penelitian ini memperkuat bahwa pendekatan data mining, khususnya teknik clustering, dapat digunakan untuk mengidentifikasi pola perilaku pengguna media sosial secara efektif. Hal ini menunjukkan bahwa analisis berbasis data memiliki potensi besar dalam mendukung pengambilan keputusan strategis pada bidang pemasaran digital, manajemen konten, serta analisis perilaku pengguna.

Dari sisi akademis, penelitian ini dapat menjadi referensi ilmiah dalam pengembangan penelitian selanjutnya yang berkaitan dengan analisis performa media sosial menggunakan metode data mining. Penggunaan data sintetis dalam penelitian ini juga memberikan kontribusi dalam menunjukkan bahwa simulasi data dapat digunakan untuk merepresentasikan pola perilaku pengguna secara umum tanpa melibatkan data pribadi, sehingga dapat mendukung penelitian yang berfokus pada keamanan dan privasi data.

Lebih luas lagi, hasil penelitian ini menunjukkan bahwa pemanfaatan teknologi analitik data dalam pengelolaan media sosial dapat menjadi bagian dari transformasi digital, khususnya dalam meningkatkan efektivitas strategi komunikasi digital dan pemasaran berbasis data. Dengan demikian, penelitian ini tidak hanya memberikan kontribusi pada aspek teknis analisis data, tetapi juga pada penerapan strategi digital yang lebih terukur, sistematis, dan berbasis bukti empiris.

4. Kesimpulan

Berdasarkan keseluruhan rangkaian penelitian yang telah dilakukan melalui penerapan metode K-Means Clustering terhadap data sintesis media sosial yang berjumlah 29.999 data posting, dapat disimpulkan bahwa pendekatan clustering berbasis data mining mampu memberikan gambaran yang sistematis dan terukur dalam mengidentifikasi pola performa unggahan media sosial. Proses pengolahan data yang mengacu pada tahapan Knowledge Discovery in Databases (KDD), yang meliputi seleksi data, pembersihan data, transformasi data, proses data mining, serta interpretasi hasil, telah menghasilkan model analisis yang mampu mengekstraksi pengetahuan baru dari kumpulan data yang dianalisis. Hasil pengolahan data menunjukkan bahwa metode K-Means Clustering berhasil mengelompokkan data unggah media sosial ke dalam tiga cluster utama berdasarkan kesamaan karakteristik variabel waktu posting, jenis konten, serta performa *reach* dan *engagement*. Setiap cluster yang terbentuk merepresentasikan tingkat performa yang berbeda, yaitu cluster dengan performa rendah, cluster dengan performa sedang, serta cluster dengan performa tinggi. Pembentukan cluster ini menunjukkan bahwa pola performa unggahan media sosial tidak bersifat acak, melainkan memiliki kecenderungan tertentu yang dapat diidentifikasi melalui analisis berbasis algoritma clustering. Berdasarkan analisis nilai centroid dari masing-masing cluster, diperoleh satu cluster yang memiliki nilai rata-rata *reach* dan *engagement* paling tinggi dibandingkan cluster lainnya. Cluster tersebut merepresentasikan kelompok waktu unggah dengan tingkat efektivitas paling optimal dalam menghasilkan jangkauan audiens yang luas serta tingkat interaksi pengguna yang tinggi. Hasil analisis lebih lanjut menunjukkan bahwa waktu unggah paling optimal berdasarkan data penelitian adalah hari Sabtu pukul 09.00 WIB. Temuan ini mengindikasikan bahwa pada waktu tersebut aktivitas pengguna media sosial berada pada kondisi yang relatif tinggi, sehingga memungkinkan konten yang dipublikasikan memiliki peluang lebih besar untuk memperoleh eksposur serta keterlibatan pengguna secara maksimal. Dari sisi perilaku pengguna media sosial, temuan ini menunjukkan adanya kecenderungan peningkatan aktivitas digital pada akhir pekan, khususnya pada hari Sabtu, yang secara umum dikaitkan dengan meningkatnya waktu luang pengguna dibandingkan hari kerja. Selain itu, periode pagi hari sekitar pukul 09.00 WIB dapat dikategorikan sebagai waktu transisi setelah aktivitas pagi dimulai, di mana pengguna cenderung mengakses media sosial untuk memperoleh hiburan, informasi, maupun interaksi sosial secara daring. Kondisi ini secara tidak langsung mempengaruhi tingkat visibilitas konten yang dipublikasikan pada waktu tersebut. Secara metodologis, penelitian ini memperkuat bahwa algoritma K-Means Clustering merupakan salah satu metode clustering yang efektif dalam melakukan segmentasi data berdasarkan tingkat kemiripan karakteristik menggunakan pendekatan perhitungan jarak Euclidean. Kemampuan algoritma ini dalam mengelompokkan data secara terstruktur memungkinkan proses interpretasi hasil menjadi lebih mudah dilakukan, sehingga dapat mendukung proses pengambilan keputusan berbasis data. Selain itu, pemanfaatan metode Elbow Method dalam menentukan jumlah cluster optimal juga terbukti mampu menghasilkan pembagian cluster yang representatif terhadap distribusi data. Penggunaan data sintesis dalam penelitian ini juga menunjukkan bahwa simulasi data mampu merepresentasikan pola umum perilaku pengguna media sosial tanpa melibatkan data riil yang bersifat sensitif atau mengandung informasi pribadi. Hal ini menunjukkan bahwa pendekatan berbasis data sintesis dapat menjadi alternatif yang efektif dalam penelitian data mining, khususnya dalam konteks pengujian model, simulasi sistem rekomendasi, serta pengembangan metode analisis tanpa melanggar aspek etika dan privasi data pengguna. Secara keseluruhan, penelitian ini menunjukkan bahwa pendekatan data mining dengan metode K-Means Clustering mampu menghasilkan informasi yang bernilai dalam mengidentifikasi pola waktu unggah optimal pada media sosial. Hasil penelitian ini juga memperlihatkan bahwa analisis berbasis clustering memiliki potensi yang luas untuk diterapkan dalam analisis performa konten digital, pengembangan strategi pemasaran berbasis data, serta pengembangan sistem analisis perilaku pengguna media sosial secara lebih komprehensif dan terstruktur.

Referensi

1. R. Wahid and M. Wadud, "Social media marketing on Instagram: When is the most effective posting timing?," *EPRA International Journal of Multidisciplinary Research (IJMR)*, pp. 312–321, 2020, doi: 10.36713/epra4834.
2. W. A. Kusuma and A. Ramadhani, "Analisis waktu posting optimal media sosial menggunakan metode clustering," *Jurnal Informatika*, vol. 7, no. 2, pp. 145–153, 2020.
3. F. Ramadhan and R. Hidayat, "Analisis engagement media sosial Instagram menggunakan data mining," *Jurnal Teknologi dan Sistem Informasi*, vol. 2, no. 1, pp. 66–74, 2021.
4. N. A. Basulina, S. Gustiani, S. Amalia, A. Sabila, and A. Ibrahim, "Identifikasi waktu optimal posting terhadap pola engagement sosial media menggunakan metode K-Means clustering," *Jurnal Mahasiswa Teknik Informatika*, vol. 9, no. 4, 2025.
5. F. A. Tanjung, A. P. Windarto, and M. Fauzan, "Penerapan metode K-Means pada pengelompokan pengangguran di Indonesia," *Jurnal Riset Sistem Informasi dan Teknik Informatika (JURASIK)*, vol. 6, no. 1, pp. 61–74, 2021. [Online]. Available: <https://tunasbangsa.ac.id/ejournal/index.php/jurasik>
6. F. Khalish, N. M. Piranti, and O. Martadireja, "Implementasi data mining menggunakan teknik clustering dengan metode K-Means," *Jurnal Ilmiah Ilmu Pendidikan (JIIP)*, vol. 8, no. 3, pp. 1760–1767, 2025.

7. M. A. Aldi and Z. Fatah, "Implementasi K-Means clustering dalam pengelompokan data menggunakan metodologi Knowledge Discovery in Databases (KDD)," *Jurnal Manajemen Informatika dan Sistem Informasi*, vol. 4, no. 2, pp. 85–93, 2025.
8. M. A. Irwansyah, R. Rinaldi, and S. Putra, "Analisis K-Means clustering menggunakan RapidMiner berbasis Knowledge Discovery in Databases," *ZONASI: Jurnal Sistem Informasi*, vol. 6, no. 1, pp. 45–53, 2024.
9. T. M. Dista and F. F. Abdulloh, "Clustering pengunjung mall menggunakan metode K-Means dan Particle Swarm Optimization," *Jurnal Media Informatika Budidarma*, vol. 6, no. 3, p. 1339, Jul. 2022, doi: 10.30865/mib.v6i3.4172.
10. N. Hendrastuty, "Penerapan data mining menggunakan algoritma K-Means clustering dalam evaluasi hasil pembelajaran siswa," *Jurnal Ilmiah Informatika dan Ilmu Komputer (JIMA-ILKOM)*, vol. 3, no. 1, pp. 46–56, Mar. 2024, doi: 10.58602/jima-ilkom.v3i1.26.
11. M. Fatkuroji, Fajrizal, Taslim, E. Sabna, and K. W. Ningsih, "Optimasi nilai K pada algoritma K-Means untuk klasterisasi data pasien COVID-19," *Indonesian Journal of Computer Science*, vol. 11, no. 2, Jul. 2022.
12. L. Barokah, "K-Means clustering dengan pendekatan Knowledge Discovery in Databases (KDD)," *Jurnal Processor*, 2025.
13. M. Mahda, R. Kurniawan, and T. Suprpti, "Optimalisasi algoritma K-Means dengan evaluasi Davies–Bouldin Index dalam proses Knowledge Discovery in Databases (KDD)," *Jurnal Teknologi Sistem Informasi*, 2025.
14. M. Noval, "Implementasi algoritma K-Means clustering dengan Knowledge Discovery in Databases (KDD)," *Jurnal Sistem Informasi*, 2025.
15. M. Rochmawati, G. W. C. Bagaskara, *et al.*, "Implementasi algoritma K-Means dalam klasterisasi penjualan menggunakan metodologi Knowledge Discovery in Databases," *Jurnal SISTEMASI*, 2023.
16. M. R. Syahkur, D. Hartama, and S. Solikhun, "Evaluasi jumlah cluster pada K-Means++ menggunakan metode Silhouette dan Elbow dengan validasi Davies–Bouldin Index," *JST (Jurnal Sains dan Teknologi)*, 2024.