

Analisis Data *Viewer* Youtube Untuk Menentukan Waktu Terbaik *Live Streaming Game* Menggunakan Decision Tree

Akhmad Rizky Fahrozy¹, Heny Pratiwi², Hanifah Ekawati³

¹ Program Studi Teknik Informatika, STMIK Widya Cipta Dharma

² Program Studi Sistem Informasi, STMIK Widya Cipta Dharma

³ Program Studi Bisnis Digital, STMIK Widya Cipta Dharma

Jl. M. Yamin No.25 - Gn. Kelua - Kec. Samarinda Ulu - Kota Samarinda - Kalimantan Timur - Indonesia

2243094@wicida.ac.id, henypratiwi@wicida.ac.id, hanifah@wicida.ac.id

Abstrak— Permasalahan dalam penelitian ini adalah belum adanya acuan yang jelas dalam menentukan waktu optimal untuk melakukan live streaming yang dapat menghasilkan tingkat interaksi pengguna yang tinggi. Oleh karena itu, penelitian ini bertujuan untuk menentukan waktu terbaik dalam melakukan live streaming berdasarkan data aktivitas viewer pada platform YouTube Studio. Metode yang digunakan adalah Decision Tree dengan pendekatan CRISP-DM yang meliputi enam tahapan, yaitu business understanding, data understanding, data preparation, modeling, evaluation, dan deployment. Dataset yang digunakan berjumlah 50 data, kemudian setelah dilakukan proses pembersihan menjadi 46 data yang valid untuk dianalisis. Hasil penelitian ini menunjukkan bahwa model Decision Tree mampu mengklasifikasikan performa live streaming ke dalam kategori rendah, sedang, dan tinggi dengan tingkat akurasi sebesar 0.60. Variabel yang paling berpengaruh adalah interaksi pengguna seperti chat, reaksi, dan durasi tonton. Berdasarkan hasil analisis, diperoleh waktu optimal untuk melakukan live streaming yaitu pada hari Selasa pukul 00:00 dengan kategori performa tinggi. Dengan demikian, dapat disimpulkan bahwa waktu memiliki pengaruh terhadap tingkat engagement pengguna, dan penelitian ini mampu memberikan rekomendasi waktu live streaming yang optimal berbasis data untuk membantu content creator dalam meningkatkan performa siaran.

Kata Kunci— Pohon Keputusan, CRISP-DM, Siaran Langsung, YouTube, Waktu Optimal, Keterlibatan

Abstract— The problem addressed in this study is the absence of clear guidelines in determining the optimal time for live streaming that can generate high user engagement. Therefore, this study aims to determine the best time to conduct live streaming based on viewer activity data from YouTube Studio. The method used is the Decision Tree algorithm with the CRISP-DM approach, which consists of six stages: business understanding, data understanding, data preparation, modeling, evaluation, and deployment. The dataset used consists of 50 records, which were reduced to 46 valid data points after the data cleaning process. The results show that the Decision Tree model is capable of classifying live streaming performance into low, medium,

and high categories with an accuracy of 0.60. The most influential variables are user interactions such as chat, reactions, and watch duration. Based on the analysis, the optimal time for live streaming is identified as Tuesday at 00:00, which falls into the high-performance category. In conclusion, time has a significant influence on user engagement, and this study provides data-driven recommendations for optimal live streaming schedules to help content creators improve their streaming performance.

Keywords— Decision Tree, CRISP-DM, Live Streaming, YouTube, Optimal Time, Engagement

I. PENDAHULUAN

Perkembangan teknologi informasi telah mendorong perubahan signifikan dalam pola konsumsi media digital, khususnya melalui platform berbagi video seperti YouTube. Peningkatan jumlah pengguna internet di Indonesia yang mencapai 215,63 juta pengguna pada tahun 2023 [1] menunjukkan bahwa media digital telah menjadi bagian penting dalam kehidupan masyarakat. Salah satu fitur yang mengalami pertumbuhan pesat adalah *live streaming*, yaitu proses penyiaran konten secara langsung melalui internet yang memungkinkan interaksi *real-time* antara kreator dan audiens. *Live streaming* telah berkembang menjadi bagian dari strategi pemasaran digital di berbagai platform media sosial dan e-commerce di Indonesia [2]. Fenomena ini menjadi tren dalam media digital karena mampu menciptakan komunikasi dua arah yang lebih interaktif dibandingkan media konvensional [3].

Dalam beberapa tahun terakhir, *live streaming game* menjadi salah satu jenis konten yang paling diminati oleh pengguna. Hal ini disebabkan oleh karakteristiknya yang interaktif, di mana penonton dapat berpartisipasi secara langsung melalui fitur komentar, *live chat*, serta reaksi selama siaran berlangsung. Interaksi ini membentuk komunitas digital antara *streamer* dan audiens yang berperan penting dalam meningkatkan keterlibatan pengguna [4].

Tingkat keterlibatan atau *engagement* merupakan indikator utama dalam mengukur keberhasilan suatu konten *live streaming*. *Engagement* dapat diukur melalui berbagai metrik seperti jumlah *like*, komentar, penambahan *subscriber*, serta aktivitas interaksi lainnya. Penelitian menunjukkan bahwa *engagement* memiliki pengaruh signifikan terhadap minat menonton secara berkelanjutan (*continuous watching intention*) [5], serta berkontribusi terhadap loyalitas pengguna terhadap suatu kanal digital [6]. Selain itu, elemen interaktif seperti *live comment*, reaksi virtual, *likes*, dan sesi tanya jawab secara langsung terbukti memengaruhi respons kognitif, afektif, dan perilaku audiens selama siaran berlangsung [7].

Selain faktor interaksi, waktu pelaksanaan *live streaming* juga memiliki pengaruh terhadap jumlah penonton dan tingkat *engagement*. Pola aktivitas pengguna pada platform digital cenderung mengikuti waktu tertentu, seperti pada malam hari atau akhir pekan. Oleh karena itu, pemilihan waktu yang tepat menjadi strategi penting dalam meningkatkan performa *live streaming*.

Untuk mengolah data dalam jumlah besar dan mengidentifikasi pola yang tersembunyi, diperlukan pendekatan *data mining*. *Data mining* merupakan proses ekstraksi informasi yang bernilai dari data untuk mendukung pengambilan keputusan. Dalam konteks ini, data *viewer* YouTube yang terdiri dari berbagai indikator *engagement* dan waktu siaran dapat dianalisis untuk menemukan pola hubungan antara waktu dan performa *streaming*. Penelitian menunjukkan bahwa *data mining* dengan algoritma Decision Tree efektif dalam menganalisis pola perilaku pengguna pada platform digital serta mampu mengklasifikasikan data berdasarkan atribut numerik maupun kategorikal [8].

Beberapa penelitian terdahulu menunjukkan bahwa analisis data pada platform YouTube dapat digunakan untuk mengevaluasi performa konten secara kuantitatif. Penelitian oleh Rinaldo dan Irwansyah (2022) menyatakan bahwa *live streaming* merupakan fenomena yang berkembang pesat dalam media digital dan berpengaruh terhadap interaksi pengguna [3]. Selanjutnya, Rabbani dan Wati (2024) menunjukkan bahwa interaksi dalam *live streaming* mampu membangun komunitas digital yang aktif antara *streamer* dan audiens [4]. Penelitian lain juga membuktikan bahwa *engagement* berperan dalam membentuk loyalitas pengguna terhadap konten digital [6], serta memiliki pengaruh terhadap minat menonton secara berkelanjutan [5].

Berdasarkan permasalahan tersebut, diperlukan suatu metode yang mampu mengklasifikasikan dan menganalisis data *viewer* untuk menentukan waktu optimal dalam melakukan *live streaming*. Salah satu metode yang dapat digunakan adalah algoritma Decision Tree, yang merupakan teknik klasifikasi dalam *data mining* yang mampu menghasilkan model keputusan dalam bentuk struktur pohon yang mudah dipahami dan diinterpretasikan. Algoritma Decision Tree masih relevan dan efektif digunakan dalam klasifikasi data karena menghasilkan model keputusan yang mudah dipahami serta dapat diterapkan pada berbagai kasus analisis data [9].

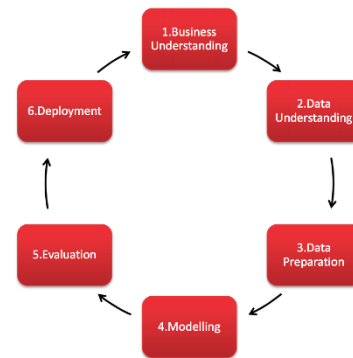
Selain itu, pendekatan CRISP-DM digunakan sebagai kerangka kerja dalam proses analisis data karena bersifat sistematis dan dapat diterapkan pada berbagai bidang. Pendekatan ini dapat dikombinasikan dengan algoritma Decision Tree untuk membangun model klasifikasi berbasis data [10][11], serta terbukti mampu menghasilkan model prediksi dan klasifikasi yang valid berdasarkan pola data yang tersedia.

Penelitian ini bertujuan untuk menganalisis data *viewer* YouTube guna menentukan waktu terbaik dalam melakukan *live streaming game* menggunakan algoritma Decision Tree. Data yang digunakan meliputi indikator *engagement* seperti *like*, komentar, *subscriber*, serta aktivitas interaksi lainnya, yang dikombinasikan dengan variabel waktu untuk menghasilkan model klasifikasi. Hasil penelitian ini diharapkan dapat memberikan rekomendasi waktu optimal yang dapat meningkatkan jumlah penonton dan interaksi selama *live streaming*, serta menjadi referensi dalam penerapan teknik *data mining* pada analisis media digital.

II. METODE PENELITIAN

2.1 Cross Industry Standard Process for Data Mining (CRISP-DM)

Penelitian ini menggunakan metode *Cross Industry Standard Process for Data Mining* (CRISP-DM) sebagai kerangka kerja dalam proses analisis data. Metodologi ini dikembangkan oleh CRISP-DM Consortium dan hingga saat ini masih digunakan secara luas karena mampu memberikan pendekatan sistematis dalam pengolahan data, mulai dari pemahaman masalah hingga evaluasi model [12].



Gbr. 1 Metode CRISP-DM (*Cross Industry Standard Process for Data Mining*)

Sumber : sv-europe.com

CRISP-DM terdiri dari enam tahapan utama, yaitu *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment* [12]. Dalam penelitian ini, tahapan tersebut diadaptasi sesuai kebutuhan analisis data *viewer* YouTube.

Tahap *business understanding* dilakukan dengan mengidentifikasi permasalahan utama, yaitu menentukan waktu terbaik dalam melakukan *live streaming* berdasarkan

data interaksi pengguna. Tahap *data understanding* untuk memahami karakteristik dataset yang digunakan, termasuk distribusi dan kualitas data. Dataset dalam penelitian ini diperoleh dari hasil ekspor data viewer pada platform YouTube Studio, dengan jumlah awal sebanyak 50 data dan setelah proses pembersihan menjadi 46 data yang valid untuk dianalisis [7].

Tahap *data preparation* dilakukan melalui proses pembersihan data (*data cleaning*), transformasi variabel waktu menjadi atribut hari dan jam, serta pengolahan variabel interaksi menjadi bentuk numerik. Proses ini penting karena kualitas data sangat mempengaruhi hasil model yang dibangun [7].

Tahap *modeling* dilakukan dengan menerapkan algoritma Decision Tree untuk membangun model klasifikasi. Tahap *evaluation* digunakan untuk mengukur performa model, umumnya menggunakan metrik akurasi untuk mengetahui tingkat ketepatan klasifikasi [10].

2.2 Decision Tree

Metode klasifikasi yang digunakan dalam penelitian ini adalah Decision Tree, yaitu algoritma yang digunakan untuk memodelkan proses pengambilan keputusan dalam bentuk struktur pohon. Menurut Quinlan, Decision Tree bekerja dengan membagi dataset ke dalam beberapa subset berdasarkan atribut tertentu hingga menghasilkan keputusan akhir dalam bentuk kelas [13].

Struktur Decision Tree terdiri dari root node, internal node, dan leaf node, yang masing-masing merepresentasikan proses pemilihan atribut dan hasil klasifikasi. Proses pembentukan pohon keputusan didasarkan pada konsep entropy dan information gain untuk menentukan atribut terbaik dalam pemisahan data [13].

Entropy digunakan untuk mengukur Tingkat ketidakpastian dalam dataset. Semakin kecil nilai entropy, maka data semakin homogen [13].

$$Entropy(S) = - \sum_{i=1}^n p_i \log_2(p_i)$$

Keterangan:

- p_i : proporsi data pada kelas ke-i
- S : himpunan data

Sedangkan information gain digunakan untuk menentukan atribut yang paling optimal dalam membagi data [13].

$$Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v)$$

Keterangan:

- A : atribut
- S_v : subset data berdasarkan nilai atribut

Atribut dengan nilai *information gain* tertinggi akan dipilih sebagai node dalam pohon keputusan. Penelitian terbaru juga menunjukkan bahwa Decision Tree masih efektif digunakan dalam klasifikasi data karena memiliki interpretasi yang jelas dan performa yang baik dalam berbagai kasus [9]. Dataset yang digunakan dalam penelitian ini berasal dari data *viewer* pada platform YouTube yang berkaitan dengan aktivitas *live streaming*. Data ini mencerminkan tingkat performa siaran berdasarkan interaksi pengguna serta waktu pelaksanaan *streaming*.

Variabel yang digunakan meliputi variabel waktu dan variabel interaksi pengguna. Variabel waktu terdiri dari hari dan jam pelaksanaan *streaming*, sedangkan variabel interaksi meliputi jumlah *like*, komentar, penambahan *subscriber*, aktivitas *chat*, dan reaksi pengguna. Pemilihan variabel ini didasarkan pada penelitian yang menyatakan bahwa indikator interaksi pengguna dapat digunakan untuk mengukur performa konten digital secara kuantitatif [7].

TABLE I
VARIABEL DATASET

Variabel	Tipe Data	Deskripsi
Hari	Kategorikal	Hari Streaming
Jam	Numerik	Waktu Streaming
Like	Numerik	Jumlah Suka
Komentar	Numerik	Jumlah Komentar
Subscriber	Numerik	Penambahan Subscriber
Chat	Numerik	Interaksi Chat
Reaksi	Numerik	Respons Pengguna

Sebelum dilakukan pemodelan, dataset melalui tahap *data preparation* yang meliputi proses pembersihan data, transformasi variabel, serta encoding data kategorikal. Proses ini bertujuan untuk memastikan bahwa data yang digunakan memiliki kualitas yang baik dan sesuai untuk proses analisis [2].

Selanjutnya, dilakukan proses klasifikasi dengan membagi data ke dalam tiga kategori, yaitu rendah, sedang, dan tinggi. Pembagian ini dilakukan berdasarkan distribusi nilai menggunakan pendekatan quantile, sehingga menghasilkan klasifikasi yang objektif dan sesuai dengan karakteristik data.

III. HASIL DAN PEMBAHASAN

3.1 Business Understanding

Pada tahap ini ditentukan tujuan penelitian, yaitu untuk mengetahui waktu terbaik dalam melakukan live streaming berdasarkan data aktivitas viewer pada platform YouTube Studio. Permasalahan utama dalam penelitian ini adalah belum diketahui waktu optimal yang dapat menghasilkan interaksi pengguna yang tinggi. Oleh karena itu, diperlukan analisis berbasis data untuk menentukan waktu dengan performa terbaik.

3.2 Data Understanding

Dataset yang digunakan merupakan data hasil ekspor dari YouTube Studio dengan jumlah awal 50 data dan 15 variabel.

Konten	Judul video	Waktu publikasi video	Durasi	Jam streaming	Suka	Reaksi	jumlah komentarnya	Subscriber yang diberikan	Pesan chat	Reaksi	Rate-rata durasi tonton	Persentase pemangkas rate-rata (%)	Pemangkas tak akuisisi	Rasio klik-tayang dari tayangan (%)	Waktu nilai streaming
0	intal	nan	nan	nan	1300.000	130.0	0.0	83.0	637	3856	0.00443	7.41	2654.0	7.20	2100
1	TROVCTPwDy	● LIVE NAWANG 1081148401 NEMH LINE AKANSI1518	Feb 13, 2026	25584.0	71507	4.0	1.0	2.0	6	0	00.0450	2.58	43.0	1.80	100
2	h1wpyP1P1P1g	● LINC AWANG11 JAWANTH11 AMBHE1P1P1P1P1P1P1	Feb 14, 2026	21603.0	60167	4.0	0.0	7.0	36	1	00.0450	1.73	98.0	3.41	300
3	whU1TR1g	● LINC KADU DAPAT DUSPEM-PANORAMA ROK'S1518	Dec 5, 2025	20280.0	96368	1.0	0.0	2.0	7	3	00.0856	2.05	193.0	5.00	100
4	Yuh1T51G1W1C	● LINC DOKSI1 SODUK1 HENK1D1R1 KAPR1S1 P1P1P1P1	Feb 6, 2026	19094.0	65269	3.0	0.0	0.0	23	0	00.2047	4.27	37.0	0.62	2000

Gbr. 2 Dataset Awal

Dataset awal terdiri dari 50 data yang merepresentasikan aktivitas live streaming. Variabel yang tersedia mencakup informasi waktu dan interaksi pengguna seperti suka, komentar, subscriber, chat, dan reaksi. Data ini cukup representatif untuk dilakukan analisis.

3.3 Data Prepration

Pada tahap ini dilakukan proses pembersihan dan transformasi data.

Jumlah Data Sebelum: (50, 18)

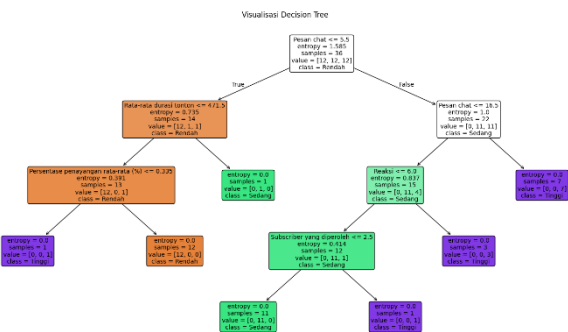
Jumlah Data Setelah: (46, 18)

Gbr. 3 Hasil Cleaning Data

Setelah dilakukan pembersihan data, jumlah data berkurang dari 50 menjadi 46 data. Hal ini disebabkan oleh adanya data yang memiliki nilai kosong sehingga harus dihapus agar tidak mempengaruhi hasil analisis. Selain itu, jumlah variabel meningkat dari 15 menjadi 18 variabel akibat proses *feature engineering*, seperti penambahan variabel waktu (hari dan jam) serta transformasi data.

3.4 Modeling

Pada tahap ini dilakukan pembangunan model menggunakan algoritma Decision Tree.



Gbr. 4 Hasil Model Decision Tree

Model Decision Tree yang dihasilkan menunjukkan bahwa variabel pesan chat menjadi faktor utama dalam proses klasifikasi. Selain itu, variabel seperti durasi tonton, reaksi, dan subscriber juga berpengaruh terhadap penentuan kategori performa. Hal ini menunjukkan bahwa tingkat interaksi pengguna menjadi indikator utama dalam menentukan keberhasilan live streaming.

3.5 Evaluation

Pada tahap ini dilakukan evaluasi model untuk mengetahui tingkat akurasi.

Accuracy: 0.6	precision	recall	f1-score	support
0	0.25	0.50	0.33	2
1	0.67	0.67	0.67	3
2	1.00	0.60	0.75	5
accuracy			0.60	10
macro avg	0.64	0.59	0.58	10
weighted avg	0.75	0.60	0.64	10

Gbr. 5 Hasil Akurasi Model

Model Decision Tree menghasilkan tingkat akurasi sebesar 0.60 (60%). Nilai ini menunjukkan bahwa model memiliki kemampuan yang cukup dalam mengklasifikasikan data, meskipun masih terdapat kesalahan prediksi.

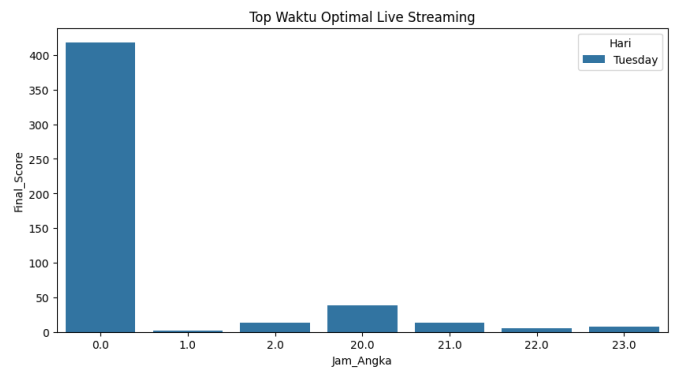
3.6 Deployment

Pada tahap ini model digunakan untuk menentukan waktu optimal live streaming.

Waktu optimal live streaming
 Hari : Tuesday
 Jam : 00:00
 Kategori : Tinggi

Gbr. 6 Hasil Waktu Optimal

Hasil analisis menunjukkan bahwa waktu terbaik untuk melakukan live streaming adalah pada hari Selasa pukul 00:00 dengan kategori performa tinggi. Hal ini menunjukkan bahwa pada waktu tersebut tingkat interaksi pengguna berada pada kondisi optimal.



Gbr. 7 Grafik Waktu Optimal

Grafik menunjukkan adanya perbedaan performa pada setiap waktu. Waktu tertentu memiliki nilai *final score* yang lebih tinggi, yang menunjukkan bahwa waktu tersebut lebih efektif untuk live streaming.

IV. KESIMPULAN DAN SARAN

4.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan menggunakan metode Decision Tree dengan pendekatan CRISP-DM, dapat disimpulkan sebagai berikut:

1. Dataset yang digunakan dalam penelitian ini berasal dari YouTube Studio dengan jumlah awal sebanyak 50 data dan 15 variabel. Setelah dilakukan proses pembersihan data, jumlah data yang valid menjadi 46 data, sehingga data yang digunakan dalam pemodelan telah memenuhi kualitas yang baik.

2. Proses *data preparation* yang meliputi pembersihan data, transformasi waktu, serta *feature engineering* berhasil menghasilkan variabel baru yang relevan, seperti hari dan jam streaming serta skor performa (*engagement*), yang digunakan sebagai dasar klasifikasi.

3. Model Decision Tree yang dibangun mampu mengklasifikasikan performa live streaming ke dalam tiga kategori, yaitu rendah, sedang, dan tinggi, dengan tingkat akurasi sebesar 0.60 (60%), yang menunjukkan bahwa model memiliki kemampuan yang cukup dalam mengenali pola data.

4. Berdasarkan hasil analisis, variabel interaksi pengguna seperti pesan chat, reaksi, dan durasi tonton merupakan faktor yang paling berpengaruh dalam menentukan performa live streaming.

5. Hasil akhir penelitian menunjukkan bahwa waktu optimal untuk melakukan live streaming adalah pada hari Selasa pukul 00:00, yang termasuk dalam kategori performa tinggi. Hal ini menunjukkan bahwa waktu memiliki pengaruh terhadap tingkat interaksi pengguna.

4.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat diberikan untuk pengembangan penelitian selanjutnya:

1. Penelitian selanjutnya disarankan menggunakan jumlah dataset yang lebih besar agar model dapat menghasilkan akurasi yang lebih tinggi dan representatif.

2. Dapat dilakukan perbandingan dengan algoritma lain seperti Random Forest atau Naive Bayes untuk mengetahui metode yang paling optimal dalam menentukan waktu live streaming.

3. Penambahan variabel lain seperti jenis konten, durasi streaming, dan target audiens dapat meningkatkan kualitas analisis.

4. Hasil penelitian ini dapat dikembangkan menjadi sistem rekomendasi berbasis web atau aplikasi yang dapat digunakan secara langsung oleh *content creator*.

streaming pada media sosial dalam perspektif social construction of technology," *Jurnal Komunikasi dan Desain*, vol. 2, no. 5, 2022.

- [4] F. Z. Rabbani and R. Wati, "Komunitas dan interaksi penonton pada siaran langsung di kanal YouTube Naplive," *Jurnal Sosiohumaniora Nusantara*, vol. 1, no. 3, pp. 406–415, 2023, doi: 10.62180/7x0vdr35.
- [5] A. Leiwakabessy, "Pengaruh live streaming engagement terhadap continuous watching intention," *Jurnal Ekonomi Kreatif dan Manajemen Bisnis Digital*, vol. 2, no. 1, pp. 64–77, 2022. [Online]. Available: <https://transpublika.co.id/ojs/index.php/JEKOMBITAL>
- [6] S. N. Hidayah, D. T. Alamanda, and D. Nurhayati, "Pengaruh customer engagement dan customer intimacy terhadap loyalitas pelanggan membership channel YouTube 'Mengkudu Official'," *Jurnal Edueco*, vol. 8, no. 2, pp. 365–376, 2025.
- [7] N. Fitriyani et al., "Pengaruh interaktivitas konten terhadap respon audiens (studi live streaming marathon by YB)," *Jurnal Sosial Humaniora dan Pendidikan*, vol. 5, no. 1, pp. 1–12, 2026.
- [8] F. Satria, "Analisis data mining strategi digital marketing terhadap keputusan pembelian mahasiswa," *JIKO (Jurnal Informatika dan Komputer)*, vol. 9, no. 2, p. 427, 2025, doi: 10.26798/jiko.v9i2.1910.
- [9] R. N. Ramadhon, A. Ogi, A. P. Agung, R. Putra, S. S. Febrihartina, and U. Firdaus, "Implementasi algoritma decision tree untuk klasifikasi pelanggan aktif atau tidak aktif pada data bank," *Karimah Tauhid*, vol. 3, no. 2, pp. 1860–1874, 2024, doi: 10.30997/karimahtauhid.v3i2.11952.
- [10] M. I. Nasution, "Application of data mining for diabetes mellitus risk prediction using the C4.5 method based on medical data," *Jurnal Multidisipliner Bharasumba*, vol. 4, no. 3, pp. 366–380, 2025. [Online]. Available: <https://azramedia-indonesia.azramediaindonesia.com/index.php/bharasumba/article/view/1651>
- [11] T. Q. H. Zatadini, A. F. Wajdi, I. M. Wiryana, and C. M. B. Zataamani, "YouTube analytics channel visualization results using Google Data Studio and Klipfolio," *Jurnal Educ*, vol. 6, no. 2, pp. 10909–10917, 2024, doi: 10.31004/joe.v6i2.4687.
- [12] P. Chapman et al., "CRISP-DM 1.0: Step-by-step data mining guide," SPSS Inc., 2000. [Online]. Available: <http://www.crisp-dm.org/CRISPWP-0800.pdf>
- [13] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, no. 1, pp. 81–106, 1986, doi: 10.1007/BF00116251.

DAFTAR PUSTAKA

- [1] Kominfo and Katadata, "Status Literasi Digital 2023," 2023.
- [2] A. Wiranda and R. S. Hayu, "Fitur Live Streaming sebagai Strategi Digital Marketing," *Oikos: Jurnal Kajian Pendidikan Ekonomi dan Ilmu Ekonomi*, vol. 4, no. 1, pp. 1–23, 2023.
- [3] E. Rinaldo and Irwansyah, "Fenomena tren live