

ANALISIS POLA GENRE ANIME BERDASARKAN VIEW DAN LIKE MENGGUNAKAN ALGORITMA PARTITIONING AROUND MEDOIDS PADA CHANNEL YOUTUBE MUSE INDONESIA

Muhammad Daffa Rachman¹, Nina Sulistiyowati², Azhari Ali Ridha³

^{1,2}Prodi Sistem informasi, Universitas Singaperbangsa Karawang, Jl. H.S. Ronggowaluyo Kel. Puseurjaya Kec. Telukjambe Timur Kab. Karawang, Jawa Barat

Keywords:

YouTube, Anime, Data Mining, Partitioning Around Medoids, CRISP-DM

Correspondent Email:

2110631250052@student.unsika.ac.id

Abstrak: Penelitian ini dilatarbelakangi oleh tingginya jumlah konten anime pada Channel YouTube Muse Indonesia serta belum optimalnya pemanfaatan data interaksi penonton dalam menentukan strategi pemilihan konten. Ketergantungan pada survei dinilai kurang mampu merepresentasikan perilaku aktual penonton, sehingga berpotensi menyebabkan pemilihan anime yang tidak sesuai dan berdampak pada rendahnya performa video serta inefisiensi biaya lisensi. Oleh karena itu, penelitian ini bertujuan untuk menganalisis pola genre anime berdasarkan jumlah view dan like menggunakan algoritma Partitioning Around Medoids (PAM). Data diperoleh melalui proses crawling YouTube API pada periode 2020–2025 dan menghasilkan 220 judul anime setelah melalui tahap preprocessing. Fitur yang digunakan meliputi genre, view_count, dan like_count, dengan proses ekstraksi genre menggunakan regex, encoding menggunakan MultiLabelBinarizer, serta normalisasi menggunakan StandardScaler. Hasil clustering menunjukkan jumlah cluster optimal adalah $k = 2$ dengan nilai Silhouette Score sebesar 0,2169. Cluster pertama didominasi genre Action, Fantasy, dan Adventure dengan tingkat interaksi lebih tinggi, sedangkan cluster kedua didominasi genre Comedy, Drama, dan Romance dengan interaksi lebih rendah. Hasil ini menunjukkan adanya perbedaan pola preferensi penonton yang dapat dimanfaatkan sebagai dasar pengambilan keputusan strategis dalam pemilihan dan pengembangan konten anime.



Copyright © 2026 The Author(s), Published by Jurnal Informatika dan Teknik Elektro Terapan. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Abstract: This research was motivated by the large volume of anime content on the Muse Indonesia YouTube channel and the limited use of audience interaction data in content selection. Survey-based decisions may lead to inappropriate anime selection and inefficient licensing costs. This study analyzes anime genre patterns based on view and like counts using the Partitioning Around Medoids (PAM) algorithm. Data were collected through the YouTube API using web crawling from 2020 to 2025, resulting in 220 anime titles after preprocessing. Genre, view count, and like count were used as analysis features. Genre data were extracted using regular expressions (regex), encoded with MultiLabelBinarizer, and normalized using StandardScaler. The optimal clustering result was $k = 2$ with a Silhouette Score of 0.2169. The first cluster, dominated by Action, Fantasy, and Adventure, showed higher audience interaction than the second cluster, dominated by Comedy, Drama, and Romance. These findings support data-driven anime content selection and development.

1. PENDAHULUAN

Di Indonesia, YouTube telah menjadi salah satu platform media sosial yang sangat populer dengan jumlah pengguna yang terus berkembang pesat. Berdasarkan data terbaru yang dilansir dari evoca.tv, pada tahun 2024, jumlah pengguna YouTube di Indonesia telah mencapai 139 juta orang, menjadikannya salah satu negara dengan jumlah pengguna terbesar di dunia. Keberadaan YouTube di Indonesia tidak hanya sekadar sebagai sarana hiburan, tetapi juga telah menjadi platform utama bagi banyak orang untuk mencari informasi, berbagi pengetahuan, serta berinteraksi dengan berbagai komunitas online. Dengan pertumbuhan yang pesat, YouTube semakin menjadi bagian tak terpisahkan dari kehidupan digital masyarakat Indonesia, baik untuk keperluan pribadi maupun profesional.

YouTube menyediakan banyak sekali konten hiburan, salah satunya adalah anime. Di Indonesia, ada sebuah Channel YouTube terbesar di Indonesia yang menampilkan berbagai konten Anime, Channel tersebut adalah Muse Indonesia. Muse Indonesia merupakan bagian dari Muse Communication Co., Ltd. yang didirikan pada tahun 1992 sebagai perusahaan yang bergerak dalam distribusi animasi Jepang. Berbasis di Distrik Xinzhuang, Kota New Taipei, perusahaan ini juga mendistribusikan film Asia dan Eropa di Taiwan. Channel YouTube Muse Indonesia juga telah mencapai lebih dari 10 juta subscriber dan telah mengupload lebih dari 7 ribu video. Muse Indonesia pun menempati posisi pertama dalam pencarian anime dengan filter channel.

Channel YouTube Muse Indonesia merupakan salah satu channel YouTube distribusi resmi anime ber-subtitle Indonesia yang memiliki jumlah video dan penonton yang sangat besar. Setiap video anime yang diunggah memiliki kombinasi genre yang beragam serta tingkat interaksi pengguna yang berbeda-beda yang dapat dilihat dari jumlah view dan like. Beberapa video anime yang diunggah bahkan telah mencapai jutaan penayangan dan ribuan likes. Keberagaman genre seperti action, drama, school, dan genre lainnya menunjukkan bahwa setiap konten anime memiliki karakteristik yang berbeda serta berpotensi menarik minat penonton yang beragam. Namun, informasi genre tersebut hanya tercantum pada bagian deskripsi video sehingga belum dimanfaatkan

secara optimal untuk menghasilkan pemahaman mengenai pola preferensi genre yang diminati oleh penonton channel YouTube Muse Indonesia. Interaksi dari penonton seperti jumlah view dan like pada setiap video anime dapat menjadi indikator untuk mengetahui kecenderungan minat penonton terhadap genre anime tertentu.

Anime, singkatan dari "animation" dalam bahasa Inggris, adalah jenis animasi khas yang berasal dari Jepang. Lebih dari sekadar tontonan biasa, anime telah menjadi bagian esensial dari budaya populer Jepang dan secara bertahap meraih penggemar di seluruh dunia. Dengan visual yang khas, alur cerita yang mendalam, serta tema yang beragam mulai dari fantasi hingga fiksi ilmiah, anime menawarkan pengalaman yang berbeda dibandingkan animasi barat. Seperti halnya tren konten di YouTube, yang sering memengaruhi preferensi penonton, daya tarik anime juga menciptakan pola unik dalam konsumsi hiburan global [1]. Sebuah penelitian menemukan setidaknya 153 istilah genre pada anime, dengan 9 aspek yang setiap aspek mewakili dimensi yang berbeda dari setiap anime seperti aspek, target audiens, latar cerita, dll [2]. Keberagaman genre pada anime, membuat peminat anime terus meningkat [3]. Dengan berbagai genre yang ada, industri anime mampu menawarkan lebih banyak pilihan yang dapat mencapai berbagai minat dan preferensi penonton. Di Indonesia, saat ini telah menjadi negara dengan jumlah penggemar anime terbanyak di dunia. Berdasarkan data yang diperoleh dari Google Trends, tercatat bahwa Indonesia menduduki peringkat ketiga dengan jumlah pencarian terbanyak untuk kata "anime" di Google sejak tahun 2004 hingga saat ini [4]. Indonesia memiliki dua kota besar yang termasuk sebagai penggemar anime versi Google, yaitu Surabaya dan Jakarta. Kedua kota ini menempati posisi 13, setelah Riyadh, Arab Saudi di peringkat 12, dan di atas Santiago, Chili yang berada di peringkat 14. Saat ini ada banyak sekali genre anime, dibandingkan pada periode 1980-an.

Untuk menemukan pola genre anime berdasarkan jumlah view dan like pada platform channel YouTube Muse Indonesia, analisis pola data dengan memanfaatkan teknik data mining. Salah satunya adalah menggunakan metode algoritma unsupervised learning clustering. Pengelompokan (clustering) merupakan

metode untuk mengidentifikasi sekelompok objek yang saling mirip dan dapat mengungkapkan pola distribusi serta hubungan dalam data yang besar [5]. Algoritma Partitioning Around Medoids (PAM) merupakan salah satu metode algoritma clustering, yang dapat mencari hubungan pada data. Algoritma clustering Partitioning Around Medoid (PAM) digunakan untuk menganalisis kluster yang mempunyai kelebihan menghasilkan klustering yang tidak terpengaruh oleh urutan masuknya dataset. Selain itu, algoritma Partitioning Around Medoid (PAM) juga efektif untuk menangani sensitivitas terhadap noise dan outlier [6].

Penelitian yang dilakukan [7] yang menganalisa pola Youtube Trending Menggunakan Algoritma Partitioning Around Medoids dan Frequent Pattern Growth. Dari hasil penelitian tersebut, ditemukan adanya 12 pola dari sampel dataset YouTube Trending. Hasil tersebut dievaluasi menggunakan Davies Bouldin Index, dari hasil pengujian tersebut terhadap hasil dataset dilakukan 10 kluster, dengan total 3 kluster yang mampu memberikan hasil evaluasi yang mendekati 0 dengan hasil 0,1488 yang dianggap lebih baik dibanding kluster lainnya.

Oleh karena itu, diperlukan analisis untuk mengidentifikasi pola hubungan antara genre anime dengan tingkat interaksi pengguna. Salah satu pendekatan yang dapat digunakan untuk menemukan pola tersebut adalah teknik data mining dengan metode clustering. Dalam penelitian ini digunakan algoritma Partitioning Around Medoids (PAM) untuk mengelompokkan konten anime berdasarkan kesamaan karakteristik genre serta tingkat interaksi pengguna yang diukur melalui jumlah view dan like. Berdasarkan permasalahan di atas, maka akan dilakukan penelitian dengan judul: “Analisis Pola Genre Anime Berdasarkan Genre, View Dan Like Menggunakan Algoritma Partitioning Around Medoids Pada Channel Youtube Muse Indonesia”.

2. TINJAUAN PUSTAKA

2.1 YouTube

Youtube merupakan salah satu aplikasi media sosial yang populer karena dapat memenuhi kebutuhan seperti informasi, pengetahuan, konten hiburan, dan

hal-hal lainnya yang dibutuhkan oleh masyarakat dunia maya [8].

2.2 Views

Views pada YouTube merupakan jumlah berapa kali video pada YouTube telah ditonton [9]. Jumlah views menjadi salah satu indikator penting dalam mengukur tingkat popularitas, jangkauan, serta daya tarik sebuah video terhadap audiens. Semakin tinggi jumlah views yang diperoleh, maka semakin besar pula tingkat eksposur dan perhatian yang diterima oleh video tersebut dari pengguna YouTube. Selain itu, views juga sering digunakan sebagai parameter untuk menilai keberhasilan konten, efektivitas strategi promosi digital, serta tingkat keterlibatan pengguna terhadap informasi atau hiburan yang disajikan dalam video.

2.3 Likes

Menurut Shen, Like pada YouTube merupakan jumlah likes atau suka yang diterima pada video YouTube [9]. Fitur like digunakan sebagai bentuk respons atau apresiasi audiens terhadap konten yang dianggap menarik, bermanfaat, maupun menghibur. Jumlah like pada sebuah video dapat menjadi indikator tingkat penerimaan dan ketertarikan audiens terhadap konten yang disajikan. Semakin banyak like yang diperoleh, maka semakin tinggi pula tingkat interaksi pengguna terhadap video tersebut. Dalam konteks media digital, jumlah like juga sering digunakan untuk mengukur engagement atau keterlibatan audiens terhadap suatu konten di platform YouTube.

2.4 Anime

Di Jepang, istilah anime digunakan untuk merujuk pada semua jenis animasi, tanpa memandang asal negaranya. Namun, di luar Jepang, anime lebih sering diidentifikasi sebagai animasi yang secara khusus diproduksi di Jepang. Anime memiliki gaya visual yang khas, yang membedakannya dari animasi barat. Selain itu, anime mencakup berbagai genre, mulai dari aksi dan petualangan hingga drama, komedi, fantasi, horor, dan fiksi ilmiah [10].

2.5 Data Mining

Data mining adalah proses menggali informasi berharga dari suatu kumpulan data. Proses ini melibatkan metode yang kompleks

seperti penggunaan kecerdasan buatan, teknik statistik, ilmu matematika, dan pembelajaran mesin. Data mining merupakan bidang ilmu yang mengintegrasikan berbagai disiplin, seperti machine learning, pengenalan pola (pattern recognition), statistika, basis data, dan visualisasi data untuk mendukung proses ekstraksi informasi dari kumpulan data berukuran besar [11]. Dengan memanfaatkan teknik-teknik tersebut, data mining bertujuan untuk mengidentifikasi dan mengekstraksi informasi yang relevan serta berguna dari database berukuran besar [12]. Proses ini juga memungkinkan untuk mengungkap tren yang mungkin tidak terlihat sebelumnya, membantu untuk lebih memahami perilaku pengguna, memprediksi hasil yang mungkin terjadi, dan akhirnya meningkatkan efisiensi operasional serta daya saing

2.6 Clustering

Menurut Iddrus & Dewi Wulan Sari, teknik data mining clustering digunakan untuk mengelompokkan data berdasarkan kesamaan atribut untuk menemukan kelompok yang memiliki pola serupa. Algoritma yang populer termasuk K-Means dan Density-Based Clustering [13].

2.6 Partitioning Around Medoids

Algoritma PAM, yang dikembangkan oleh Leonard Kaufman dan J. Rousseeuw, memiliki kesamaan dengan algoritma K-Means. Keduanya termasuk algoritma partisional, yang berarti keduanya membagi dataset menjadi beberapa kelompok. Selain itu, kedua algoritma ini bekerja secara acak, diikuti dengan perhitungan ulang cost. Jika total cost yang dihasilkan lebih rendah dibandingkan dengan cost objek yang memiliki medoid lama, maka objek tersebut dapat dijadikan medoid baru. Proses iterasi akan berlanjut hingga tidak ada perubahan lebih lanjut pada cost yang dihasilkan oleh medoid yang baru [5]. Berikut adalah tahapan algoritma PAM menurut Abdurrahman [5]:

1. Secara acak, pilih k objek dari kumpulan objek untuk dijadikan medoid.
2. Lakukan langkah berikut secara berulang:
3. Tempatkan objek non-medoid ke dalam cluster yang paling dekat dengan medoid,

menggunakan perhitungan jarak Euclidean Distance.

Formula:

$$D(x,y)=\sum x_i y_i \quad (2.1)$$

Keterangan :

- $D(x,y)$: Jarak Euclidean antara dua objek x dan y.
 - x_i dan y_i : Nilai koordinat objek xxx dan yyy pada dimensi ke-i.
 - $\sum$: Merupakan penjumlahan atas seluruh dimensi (untuk ruang n-dimensi, jumlahkan perbedaan antara koordinat objek pada tiap dimensi).
4. Secara acak, pilih sebuah objek non-medoid, yang disebut Orandom.
 5. Hitung total biaya, S, dari pertukaran medoid Oj dengan Orandom.
 6. Jika $S < 0$, tukar Oj dengan Orandom untuk membentuk k objek baru sebagai medoid.
 7. Ulangi langkah 3 hingga 6 sampai tidak ada perubahan lebih lanjut.

2.7 Silhouette Coefficient

Silhouette Coefficient (SC) adalah metode evaluasi untuk mengukur kualitas sebuah cluster, yaitu seberapa baik objek dikelompokkan dalam cluster yang tepat [14]. Berikut adalah langkah-langkah dalam menghitung Silhouette Coefficient [14]:

1. Hitung rata-rata jarak suatu objek dengan semua objek lain dalam satu cluster.

$$a(i) = \frac{1}{|A|-1} \sum_{j \in A, j \neq i} d(i,j) \quad (2.2)$$

Keterangan :

- $a(i)$: rata-rata jarak objek iii ke semua objek lain dalam cluster yang sama.
 - $d(i,j)$: jarak antara objek iii dan objek jjj, dengan iii dan jjj berada dalam cluster yang sama.
 - $|A|$: total objek dalam cluster A
2. Hitung rata-rata jarak objek dengan semua objek dalam cluster lain, kemudian pilih nilai terkecil.

$$b(i) = \min_{C \neq A} \frac{1}{|C|} \sum_{j \in C} d(i,j) \quad (2.3)$$

Kemudian, ambil nilai terkecilnya dengan persamaan berikut :

$$b(i) = \min_C \left(\frac{1}{|C|} \sum_{j \in C} d(i,j) \right) \quad (2.4)$$

Keterangan :

- $b(i)$: rata-rata jarak objek i ke objek-objek dalam cluster C, yang bukan merupakan cluster A.

- $d(i,j)$: jarak objek i dengan objek j , di mana j adalah anggota cluster C .

3. Hitung nilai Silhouette Coefficient (SC) dengan rumus berikut:

$$SC(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (2.5)$$

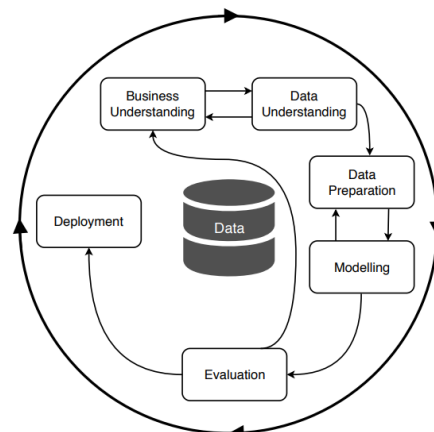
Nilai Silhouette Coefficient (SC) berada dalam rentang -1 hingga 1. Jika nilai SC mendekati 1, maka pengelompokan data dalam cluster sangat baik. Sebaliknya, jika SC mendekati -1, maka pengelompokan data dalam cluster dianggap buruk.

2.8 Google Colaboratory

Google Colaboratory adalah platform berbasis cloud yang memungkinkan pengguna untuk menulis dan menjalankan kode Python secara online melalui browser tanpa memerlukan instalasi perangkat lunak tambahan. Platform ini menawarkan keunggulan dalam hal kemudahan akses, kolaborasi secara langsung, dan integrasi dengan layanan Google lainnya seperti Google Drive dan Google Docs. Oleh karena itu, Google Colaboratory menjadi pilihan utama bagi para pendidik dan pelajar dalam menyampaikan materi serta melaksanakan praktikum di bidang ilmu computer [15].

3. METODE PENELITIAN

Pada penelitian ini, metode yang digunakan adalah metode *Cross-Industry Standard Process for Data Mining*. CRISP-DM merupakan sebuah model proses yang tidak terikat pada sektor industri tertentu, yang digunakan dalam data mining. Model ini mencakup enam tahap yang saling berulang, mulai dari business understanding hingga tahap deployment [16]. Menurut Afika Rianti berikut adalah tahapan dari metode Cross-Industry Standard Process for Data Mining [17] :



1 Business Understanding

Pada tahap ini, penelitian dimulai dengan memahami tujuan bisnis yang ingin dicapai. Peneliti mengidentifikasi kondisi saat ini untuk mengetahui permasalahan yang ada, kemudian merumuskan tujuan data mining yang sesuai. Setelah itu, disusun rencana proyek sebagai panduan pelaksanaan penelitian secara keseluruhan.

2. Data Understanding

Tahap ini berfokus pada pengenalan data yang akan digunakan. Dimulai dengan mengumpulkan data, kemudian mendeskripsikan karakteristik dari data yang didapat. Selanjutnya dilakukan eksplorasi data untuk menemukan pola awal, serta memverifikasi kualitas data guna memastikan data layak digunakan.

3. Data Preparation

Pada tahap ini, data dipersiapkan agar siap digunakan dalam proses modelling. Prosesnya meliputi pemilihan data yang relevan, pembersihan data dari noise atau missing value, pembentukan dan integrasi data dari berbagai sumber, serta pemformatan data agar sesuai dengan kebutuhan algoritma yang akan digunakan.

4. Modelling

Tahap modelling bertujuan untuk membangun model data mining. Peneliti memilih teknik yang sesuai, kemudian merancang desain pengujian. Setelah itu model dibangun dan diakses untuk melihat performa awal sebelum evaluasi lebih lanjut.

5. Evaluation

Pada tahap ini, model yang telah dibuat dievaluasi untuk memastikan apakah sudah memenuhi tujuan penelitian. Selain itu, seluruh proses ditinjau kembali untuk mengidentifikasi

kekurangan dan menentukan langkah berikutnya.

6. Deployment

Tahap terakhir adalah deployment, di mana hasil penelitian diimplementasikan atau disajikan. Peneliti merencanakan proses deployment dan monitoring, menyusun laporan akhir, serta melakukan peninjauan terhadap keseluruhan hasil proyek.

4. HASIL DAN PEMBAHASAN

4.1 Hasil

Penelitian ini bertujuan untuk menganalisis pola anime berdasarkan genre dan jumlah view dan like pada channel YouTube Muse Indonesia menggunakan teknik data mining dengan algoritma Partitioning Around Medoids (PAM). Data yang digunakan diperoleh melalui proses crawling menggunakan YouTube API dari channel Muse Indonesia dalam rentang waktu 22 September 2020 hingga 18 Juli 2025. Setelah melalui tahapan pembersihan data, penyaringan konten yang bukan merupakan episode anime, serta proses agregasi episode berdasarkan judul anime, diperoleh sebanyak 220 judul anime yang digunakan sebagai dataset dalam penelitian ini. Setiap data anime direpresentasikan menggunakan beberapa fitur utama, yaitu genre anime, jumlah view (view_count), dan jumlah like (like_count).

Informasi genre yang terdapat pada kolom deskripsi video terlebih dahulu diekstraksi menggunakan metode regular expression (regex). Karena satu anime dapat memiliki lebih dari satu genre, maka data genre bersifat multi-label. Oleh karena itu, dilakukan proses encoding menggunakan MultiLabelBinarizer untuk mengubah genre ke dalam bentuk numerik sehingga dapat digunakan dalam proses pemodelan clustering. Selain itu, fitur numerik view_count dan like_count dinormalisasi menggunakan StandardScaler untuk menyamakan skala data.

Proses clustering dilakukan menggunakan algoritma K-Medoids dengan metode Partitioning Around Medoids (PAM). Untuk menentukan jumlah cluster yang optimal, dilakukan pengujian nilai k pada rentang $k = 2$ hingga $k = 10$. Evaluasi dilakukan menggunakan Elbow Method dan Silhouette Score.

Berdasarkan hasil evaluasi tersebut, jumlah cluster optimal yang dipilih adalah $k = 2$, karena menghasilkan nilai Silhouette Score tertinggi dibandingkan jumlah cluster lainnya. Nilai Silhouette Score yang diperoleh sebesar 0,2169, yang menunjukkan bahwa struktur cluster yang terbentuk memiliki pemisahan yang masih relatif rendah namun tetap menunjukkan adanya pola pengelompokan pada data.

Hasil clustering menunjukkan bahwa data anime pada channel YouTube Muse Indonesia terbagi menjadi dua kelompok utama dengan karakteristik genre dan tingkat interaksi yang berbeda. Cluster pertama didominasi oleh anime dengan genre Action, Fantasy, dan Adventure, dengan rata-rata jumlah view sebesar 9.433.067 dan rata-rata like sebesar 203.745. Hal ini menunjukkan Cluster dengan engagement tinggi didominasi oleh anime yang memiliki kombinasi genre Action, Fantasy, dan Adventure. Sementara itu, cluster kedua didominasi oleh anime dengan genre Comedy, Drama, dan Romance, dengan rata-rata jumlah view sebesar 2.624.148 dan rata-rata like sebesar 65.003. Cluster ini menunjukkan tingkat interaksi pengguna yang relatif lebih rendah dibandingkan cluster pertama.

Secara umum, hasil penelitian ini menunjukkan bahwa terdapat perbedaan pola interaksi pengguna terhadap genre anime pada channel YouTube Muse Indonesia. Anime dengan genre aksi dan fantasi cenderung memperoleh jumlah view dan like yang lebih tinggi dibandingkan anime dengan genre komedi dan romansa.

4.2 Pembahasan

Berdasarkan data yang telah dikumpulkan dari channel YouTube Muse Indonesia, diperoleh dataset yang terdiri dari 7433 video dalam rentang waktu 22 September 2020 hingga 18 Juli 2025. Data yang digunakan dalam penelitian ini meliputi variabel genre, view_count, dan like_count yang merepresentasikan karakteristik konten serta tingkat interaksi penonton. Penelitian ini bertujuan untuk mengidentifikasi pola preferensi penonton terhadap genre anime berdasarkan performa video melalui pendekatan data mining menggunakan metode clustering.

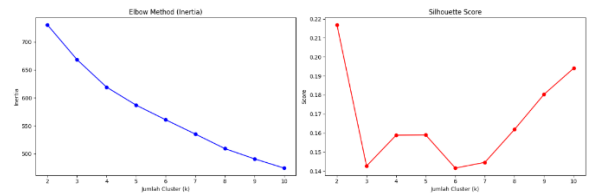
Penelitian ini menggunakan metodologi Cross-Industry Standard Process for Data Mining

(CRISP-DM), yang merupakan kerangka kerja standar dalam proses data mining yang bersifat independen terhadap domain serta banyak digunakan dalam industri dan penelitian [16]. Metodologi ini memberikan alur sistematis mulai dari pemahaman masalah hingga implementasi hasil, sehingga analisis dapat dilakukan secara terstruktur dan menghasilkan insight yang relevan. Hasil penerapan tahapan CRISP-DM dalam penelitian ini ditunjukkan pada Tabel 4.1 berikut:

Tahapan CRIPS-DM	Hasil
Business Understanding	Permasalahan pada Channel Muse Indonesia yang masih belum berbasis data dari interaksi penonton seperti jumlah views dan likes
Data Understanding	Diperoleh 7433 video dengan variabel genre, views, dan likes
Data Preparation	Data diproses menjadi 220 judul anime dengan 45 genre setelah cleaning dan transformasi
Modelling	Diperoleh jumlah cluster optimal $k = 2$ menggunakan algoritma K-Medoids (PAM)
Evaluation	Nilai Silhouette Score sebesar 0,2169 menunjukkan struktur cluster cukup terbentuk
Deployment	Hasil segmentasi digunakan sebagai dasar strategi pemilihan konten

Berdasarkan tahapan tersebut, data yang digunakan telah melalui proses pembersihan dan transformasi sehingga berada dalam kondisi terstruktur dan siap dianalisis. Proses transformasi seperti encoding genre dan normalisasi fitur numerik memastikan bahwa setiap variabel memiliki kontribusi yang seimbang dalam proses clustering.

Hasil clustering menggunakan algoritma K-Medoids dengan jumlah cluster optimal $k = 2$ menghasilkan dua kelompok anime dengan karakteristik yang berbeda. Hasil optimal dari nilai K didapat dari membandingkan hasil Silhouette Score dengan elbow-method.



Gambar 1. Nilai K Optimal

Perbandingan karakteristik kedua cluster ditunjukkan pada Tabel 4.2 berikut:

Aspek	Cluster 0	Cluster 1
Jumlah Data Anime	92	128
Rata-rata Views	9.433.067	2.624.148
Rata-Rata Likes	203.745	65.003
Genre Dominan	Action, Fantasy, Adventure	Comedy, Drama, Romance
Contoh Anime	Tensura, Shangri-La	Mieruko, Pacar Sewaan
Karakteristik	Interaksi tinggi, popular	Interaksi rendah, niche audiens

Berdasarkan hasil clustering yang diperoleh, Cluster 0 didominasi oleh kombinasi genre Action, Fantasy, dan Adventure, dengan proporsi genre Action mencapai 71,7%. Cluster ini memiliki rata-rata jumlah views dan likes yang lebih tinggi dibandingkan Cluster 1. Temuan ini menunjukkan bahwa anime yang memiliki dominasi kombinasi genre tersebut lebih sering ditemukan pada kelompok dengan tingkat engagement yang tinggi. Karakteristik ini mengindikasikan bahwa konten dalam Cluster 0 cenderung memiliki jangkauan audiens yang lebih luas dan mampu menarik perhatian penonton dalam jumlah yang lebih besar.

Sebaliknya, Cluster 1 didominasi oleh genre Comedy, Drama, dan Romance dengan rata-rata jumlah views dan likes yang lebih rendah. Meskipun jumlah anime dalam cluster ini lebih banyak dibandingkan Cluster 0, tingkat engagement yang diperoleh relatif lebih rendah. Hal ini menunjukkan bahwa anime dalam Cluster 1 cenderung menjangkau kelompok audiens yang lebih spesifik dengan preferensi terhadap unsur komedi, drama, atau hubungan

antar karakter. Oleh karena itu, karakteristik konten pada cluster ini dapat dikategorikan sebagai konten yang memiliki segmentasi audiens yang lebih khusus dibandingkan Cluster 0.

Perbedaan karakteristik antara kedua cluster menunjukkan adanya pola distribusi engagement yang berbeda pada data anime yang dianalisis. Anime dengan dominasi kombinasi genre Action, Fantasy, dan Adventure lebih sering muncul pada cluster dengan engagement tinggi, sedangkan anime yang didominasi genre Comedy, Drama, dan Romance lebih banyak ditemukan pada cluster dengan engagement yang relatif lebih rendah.

Evaluasi hasil clustering dilakukan menggunakan metode Silhouette Score, dengan nilai yang diperoleh sebesar 0,2169. Nilai ini menunjukkan bahwa struktur cluster yang terbentuk, bahwa pemisahan antar cluster belum terlalu kuat namun tetap memiliki struktur yang dapat diinterpretasikan. Nilai evaluasi yang tidak terlalu tinggi ini dapat disebabkan oleh beberapa faktor, di antaranya adalah karakteristik data yang bersifat multi-label pada genre, serta distribusi popularitas yang tidak merata. Selain itu, adanya variasi preferensi penonton terhadap genre tertentu juga menyebabkan batas antar cluster tidak terbentuk secara tegas. Meskipun demikian, nilai positif menunjukkan bahwa proses clustering masih mampu mengelompokkan data dengan cukup baik.

5. KESIMPULAN

5.1 Kesimpulan

1. Penerapan Algoritma Partitioning Around Medoids (PAM) dalam mengelompokkan data dilakukan melalui pengelompokan data anime berdasarkan variabel genre, jumlah views, dan jumlah likes. Proses clustering dilakukan terhadap 220 judul anime hasil agregasi dari 7433 video yang diperoleh melalui proses crawling menggunakan YouTube API. Hasil clustering menunjukkan bahwa pengelompokan data mampu membedakan anime dengan tingkat interaksi tinggi dan rendah.

2. Identifikasi karakteristik genre yang dominan dan pola genre pada tiap-tiap cluster menghasilkan dua cluster genre anime berdasarkan tingkat engagement (view dan like). Cluster pertama didominasi oleh genre

Action, Fantasy, dan Adventure dengan rata-rata views sebesar 9.433.067 dan likes sebesar 203.745. Sementara itu, cluster kedua didominasi oleh genre Comedy, Drama, dan Romance dengan rata-rata views sebesar 2.624.148 dan likes sebesar 65.003. Hasil tersebut menunjukkan bahwa anime dengan dominasi kombinasi genre Action, Fantasy, dan Adventure lebih sering ditemukan pada cluster dengan engagement yang lebih tinggi dibandingkan cluster yang didominasi genre Comedy, Drama, dan Romance. Temuan ini menunjukkan adanya pola karakteristik anime yang berbeda pada masing-masing cluster.

3. Evaluasi hasil performa klasterisasi Algoritma Partitioning Around Medoids (PAM) menggunakan metrik evaluasi Silhouette Coefficient menghasilkan nilai sebesar 0,2169 dengan jumlah cluster optimal sebanyak dua cluster ($k = 2$). Nilai tersebut menunjukkan bahwa struktur cluster yang terbentuk masih memiliki pemisahan yang relatif rendah, namun tetap mampu menunjukkan pola pengelompokan data yang dapat diinterpretasikan. Dengan demikian, algoritma PAM masih dapat digunakan untuk melakukan segmentasi data anime berdasarkan karakteristik genre dan tingkat interaksi penonton sehingga dapat memastikan pengelompokan penonton Channel YouTube Muse Indonesia yang terbentuk sesuai dengan tujuan penelitian.

5.2 Saran

1. Penelitian selanjutnya dapat mempertimbangkan penerapan teknik penyeimbangan data (data balancing) atau teknik resampling seperti Synthetic Minority Oversampling Technique (SMOTE) atau metode resampling data lainnya untuk mengatasi ketimpangan distribusi data popularitas antar anime. Dengan distribusi data yang lebih seimbang, diharapkan proses pembentukan cluster dapat menghasilkan pemisahan cluster yang lebih jelas dan meningkatkan kualitas evaluasi clustering.

2. Metode clustering lain seperti K-Means, DBSCAN, atau Hierarchical Clustering dapat digunakan sebagai pembanding untuk melihat perbedaan hasil segmentasi data dan memperoleh model clustering yang lebih optimal.

3. Penggunaan dataset yang lebih luas lagi yang ada di YouTube yang tidak terbatas pada Channel YouTube Muse Indonesia bisa menjadi insight tersendiri untuk memberikan Gambaran pola preferensi penonton anime pada YouTube.

4. Penelitian selanjutnya dapat mempertimbangkan penggunaan metadata tambahan yang hanya dapat diakses oleh pemilik channel YouTube atau melalui YouTube Studio. Metadata tersebut seperti jumlah komentar, durasi waktu tonton (watch time), rasio klik tayang (click-through rate/CTR), serta data retensi penonton. Informasi tersebut dapat memberikan gambaran yang lebih komprehensif mengenai tingkat keterlibatan penonton terhadap konten anime. Dengan menambahkan metadata tersebut sebagai fitur dalam proses analisis, diharapkan hasil clustering dapat menggambarkan perilaku penonton secara lebih akurat dibandingkan hanya menggunakan jumlah view dan like.

UCAPAN TERIMA KASIH

Terimakasih kepada para dosen pembimbing saya yang telah membantu saya menyelesaikan penelitian ini. Terimakasih juga kepada pihak Jurnal Informatika dan Teknik Elektro Terapan (JITET) yang telah menjadi wadah untuk penerbitan penelitian ini.

DAFTAR PUSTAKA

- [1] T. Ahmad, A. Jamil, A. Zidan, and Y. Huda, "Komunitas Wibu Situbondo Sekai: Sebuah Identitas Kebudayaan Yang Terbentuk Dalam Masyarakat Modern," *AKADEMIK: Jurnal Mahasiswa Humanis*, vol. 4, no. 2, pp. 227–238, 2024, doi: 10.37481/jmh.v4i2.741.
- [2] H. Cho, T. Disher, W.-C. Lee, S. A. Keating, and J. H. Lee., "Facet Analysis of Anime Genres: The Challenges of Defining Genre Information for Popular Cultural Objects," *KNOWLEDGE ORGANIZATION*, vol. 47, no. 1, pp. 13–30, 2020, doi: 10.5771/0943-7444-2020-1-13.
- [3] Y. Toi, "Kepopuleran dan Penerimaan Anime Jepang Di Indonesia," *Ayumi : Jurnal Budaya, Bahasa dan Sastra*, vol. 7, no. 1, Jul. 2020, doi: 10.25139/ayumi.v7i1.2808.
- [4] N. Andari and A. G. M. Litaay, "Anime Trend as An Advert Genre in Indonesian Youtube," *Journal of Development Research*, vol. 6, no. 2, pp. 164–170, Nov. 2022, doi: 10.28926/jdr.v6i2.245.
- [5] D. D. Abdurrahman, F. Agus, and G. M. Putra, "Implementasi Algoritma Partitioning Around Medoids (PAM) untuk Mengelompokkan Hasil Produksi Komoditi Perkebunan (Studi Kasus: Dinas Perkebunan Provinsi Kalimantan Timur)," *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, vol. 16, no. 2, p. 130, Oct. 2021, doi: 10.30872/jim.v16i2.6520.
- [6] S. Wulandari and N. Dwitianti, "IMPLEMENTASI ALGORITMA CLUSTERING PARTITIONING AROUND MEDOID (PAM) DALAM CLUSTERING VIRUS MERS-CoV," Aug. 2020. doi: <https://www.journal.lppmunindra.ac.id/index.php/STRING/article/view/6469>.
- [7] S. Seimahuira, "Analisa Pola Pengguna Youtube Trading Menggunakan Algoritma Partitioning Around Medoids (PAM) dan FP-Growth," *Jurnal Penelitian Inovatif*, vol. 2, no. 1, pp. 59–66, Jan. 2022, doi: 10.54082/jupin.38.
- [8] Wardah and M. J. Reza, "Persepsi Mahasiswa terhadap Penggunaan Youtube sebagai Media Konten Video Kreatif," *Jurnal Komunikasi dan Organisasi (J-KO)*, vol. 3, no. 1, pp. 39–46, 2021.
- [9] J. Shen, "The research of the factors that influence the popularity of YouTube videos," *Theoretical and Natural Science*, vol. 51, no. 1, pp. 187–193, Nov. 2024, doi: 10.54254/2753-8818/51/2024ch0200.
- [10] K. Pristianti, "Pengaruh Menonton Animasi Jepang Di Internet Terhadap Gaya Hidup Wibu Di Yogyakarta," *Commsphere: Jurnal Mahasiswa Ilmu Komunikasi*, vol. 2, no. 1, pp. 93–121, 2024, doi: 10.37631/commsphere.v2i1.1353.
- [11] Meri Fitriani, Gigih Forda Nama, and Mardiana, "Implementasi Association Rule Dengan Algoritma Apriori Pada Data Peminjaman Buku Upt Perpustakaan Universitas Lampung Menggunakan Metodologi CRISP-DM," *Jurnal Informatika dan Teknik Elektro Terapan (JITET)*, vol. 10, Jan. 2022.
- [12] B. G. Sudarsono, M. I. Leo, A. Santoso, and F. Hendrawan, "Analisis Data Mining Data Netflix Menggunakan Aplikasi Rapid Miner," *JBASE - Journal of Business and Audit Information Systems*, vol. 4, no. 1, 2021, doi: 10.30813/jbase.v4i1.2729.
- [13] Iddrus and Dewi Wulan Sari, "Penerapan Data Mining Menggunakan Algoritma Decision Tree C4.5 Untuk Memprediksi Mahasiswa Drop Out Di Universitas Wiraraja," *Jurnal Advance Research Informatika*, Jun. 2023.
- [14] A. P. Rengganis, D. Arifianto, and H. Oktavianto, "Algoritma Partitioning Around

Medoids (PAM) dalam Mengelompokkan Kecamatan di Kabupaten Jember Berdasarkan Produktivitas Tanaman Padi Partitioning Around Medoids (PAM) Algorithm in Classifying Districts in Jember Regency Based on Rice Productivity,” Mar. 2021. [Online]. Available:
<http://jurnal.unmuhjember.ac.id/index.php/JS>
T

- [15] Haana Udtari Anjani, Vitriani Vitriani, and Mulya Hastuti, “Pemanfaatan Media Google Colaboratory Pada Mata Pelajaran Informatika di SMA Negeri 5 Pekanbaru,” *SOKO GURU: Jurnal Ilmu Pendidikan*, vol. 4, no. 1, pp. 101–108, 2024, doi: 10.55606/sokoguru.v4i1.3613.
- [16] V. Plotnikova, M. Dumas, and F. Milani, “Adaptations of data mining methodologies: A systematic literature review,” *PeerJ Comput. Sci.*, vol. 6, pp. 1–43, 2020, doi: 10.7717/PEERJ-CS.267.
- [17] Afika Rianti, Nuur Wachid Abdul Majid, and Ahmad Fauzi, “CRISP-DM: Metodologi Proyek Data Science,” Jul. 2023.