

Penerapan K-Means Clustering dan Heatmap untuk Menganalisis Peminat Animasi pada Forum MyAnimeList

Naufal Septianto¹, Muhammad Ikhsan²

^{1,2}Ilmu Komputer, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia, 20353
e-mail: ¹naufalseptianto@uinsu.ac.id ²mhd.ikhsan@uinsu.ac.id

Submitted Date: September 19th, 2023
Revised Date: October 15th, 2023

Reviewed Date: September 22th, 2023
Accepted Date: October 18th, 2023

Abstract

MyAnimeList, as one of the primary forums for anime enthusiasts, faces challenges in managing and understanding the complex preferences of its growing user base. This research focuses on this issue and applies the K-Means Clustering and Heatmap methods to analyze anime enthusiasts on the MyAnimeList forum. The identified problem involves the complexity and diversity of user preferences, making it challenging for MyAnimeList to present content that aligns with each user's taste. K-Means Clustering is employed to group users based on their preferences, while Heatmap is used to visualize viewing patterns. This approach is applied to the MyAnimeList user dataset, encompassing 20,535,533 votes, with a focus on clustering based on similarity in anime ratings and preferences. The analysis results reveal four distinct user clusters, each marked by characteristic preferences. This clustering provides a better understanding of viewer preferences and tendencies, aiding the platform in crafting more accurate content recommendations. By implementing this approach, the research offers a deeper understanding of the diversity of MyAnimeList user preferences and addresses the complexity issue in delivering tailored content. The analysis results can serve as a strategic foundation to enhance user experience and improve content recommendation accuracy on similar platforms

Keywords: Data Visualization; Anime Preferences; MyAnimeList; Heatmap; K-Means Clustering

Abstrak

Platform komunitas online telah menjadi pusat interaksi dan pertukaran informasi. MyAnimeList, sebagai salah satu forum utama bagi penggemar anime, menghadapi tantangan dalam mengelola dan memahami preferensi kompleks pengguna yang terus bertambah. Penelitian ini fokus pada masalah tersebut dan menerapkan metode K-Means Clustering dan Heatmap untuk menganalisis peminat animasi di forum MyAnimeList. Masalah yang diidentifikasi melibatkan kompleksitas dan keragaman preferensi pengguna, menyulitkan MyAnimeList dalam menyajikan konten yang sesuai dengan selera masing-masing. K-Means Clustering digunakan untuk mengelompokkan pengguna berdasarkan preferensi mereka, sementara Heatmap digunakan untuk visualisasi pola tontonan. Pendekatan ini diterapkan pada dataset pengguna MyAnimeList yang mencakup 20.535.533 suara, dengan fokus pada pengelompokan berdasarkan kesamaan dalam rating dan preferensi anime. Hasil analisis menunjukkan empat kluster pengguna yang berbeda, masing-masing ditandai dengan preferensi yang khas. Klusterisasi ini memberikan gambaran yang lebih baik tentang preferensi dan kecenderungan penonton, yang dapat membantu platform dalam menyusun rekomendasi konten yang lebih akurat. Dengan menerapkan pendekatan ini, penelitian ini memberikan pemahaman yang lebih mendalam tentang diversitas preferensi pengguna MyAnimeList dan mengatasi masalah kompleksitas dalam menyajikan konten yang disesuaikan. Hasil analisis dapat digunakan sebagai landasan strategis untuk meningkatkan pengalaman pengguna dan meningkatkan akurasi rekomendasi konten di platform sejenis.

Kata kunci: Visualisasi Data; Preferensi Anime; MyAnimeList; Heatmap; K-Means Clustering

1. Pendahuluan

Dalam era digital ini, komunitas online memainkan peran vital sebagai wadah bagi interaksi dan pertukaran informasi di berbagai bidang. Salah satu aspek menarik adalah peran platform komunitas dalam menggali preferensi penggunanya (Kurniawan et al., 2022). Dalam konteks ini, penelitian ini menyoroti keberlanjutan pertumbuhan platform komunitas, khususnya pada MyAnimeList, sebuah forum populer yang mendukung pertukaran preferensi dan ulasan mengenai anime. Studi menunjukkan bahwa penggunaan metode analisis data untuk mengidentifikasi pola preferensi pengguna dapat memberikan wawasan yang berharga dalam mengelola dan menyajikan konten yang lebih relevan (Nasution et al., 2020).

Namun, sistem saat ini di MyAnimeList masih menghadapi tantangan dalam menyediakan rekomendasi yang akurat dan relevan bagi penggunanya. Meskipun terdapat berbagai metode rekomendasi yang telah diterapkan, kompleksitas preferensi pengguna, seringkali heterogen dan berubah dinamis, membuatnya sulit untuk memberikan rekomendasi yang sesuai (Platers, 2022). Penelitian terkini menunjukkan bahwa penerapan metode K-Means Clustering dan Heatmap dalam menganalisis pola preferensi pengguna dapat meningkatkan akurasi rekomendasi (Rosadi et al., 2023).

Masalah pada sistem saat ini juga melibatkan kebutuhan untuk memahami lebih baik karakteristik pengguna dan preferensi mereka di MyAnimeList (Platers, 2022). Hal ini sesuai dengan temuan yang menegaskan bahwa pengelompokan pengguna berdasarkan preferensi dapat meningkatkan pemahaman atas preferensi individu dan memungkinkan penyajian konten yang lebih personal (Rusydiyah et al., 2021).

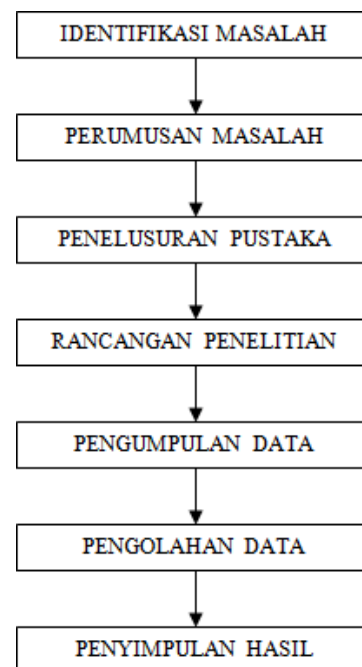
Untuk mengatasi masalah tersebut, penelitian ini mengusulkan penerapan metode K-Means Clustering dan Heatmap untuk menganalisis peminat animasi pada MyAnimeList. Kombinasi kedua metode ini dapat memberikan pemahaman yang lebih mendalam tentang preferensi pengguna dan membantu meningkatkan akurasi rekomendasi konten. Penerapan metode ini akan dilakukan melalui pendekatan yang sistematis dan terukur (Rahmalinda et al., 2022).

Rencana penerapan melibatkan pengumpulan data dari MyAnimeList, implementasi algoritma K-Means Clustering untuk mengelompokkan pengguna, dan visualisasi hasil analisis menggunakan Heatmap (Platers, 2022). Langkah-langkah ini sejalan dengan penelitian yang menunjukkan efektivitas K-Means Clustering dalam analisis preferensi pengguna (Simanjuntak et al., 2023).

Dengan melakukan penelitian ini, kontribusi signifikan dapat diberikan pada pengembangan sistem rekomendasi di MyAnimeList dan platform sejenis. Pendekatan ini dapat meningkatkan pengalaman pengguna dengan memberikan rekomendasi konten yang lebih sesuai dengan preferensi individu mereka.

2. Metodologi Penelitian

Dalam kerangka penelitian ini, melibatkan beberapa tahapan yang saling berhubungan antara satu tahap dengan tahap lainnya. Peneliti akan melalui beberapa tahapan dalam melakukan perancangan model analisis peminat animasi pada forum MyAnimeList. Tahapan ini dilakukan untuk memastikan bahwa penelitian dapat berlangsung secara efisien dan memberikan hasil yang sesuai dengan tujuan sebenarnya (Mulyani, 2023).



Gambar 1. Tahapan Penelitian (Muningsih, 2021)

Menyusun gambaran dari Ilustrasi pada gambar 1, penelitian diawali dengan tahap mengidentifikasi masalah dan diakhiri dengan tahap menyimpulkan hasil. Tahap penyimpulan ini dilakukan pada akhir proses penelitian, bertujuan untuk menghasilkan temuan yang dapat memberikan wawasan dan pemahaman terkait dengan permasalahan yang diselidiki (Mulyani et al., 2019).

2.1 Pengumpulan Data

Dalam kerangka penelitian ini, data yang diperlukan mencakup dataset pengguna MyAnimeList. Dataset ini diperoleh secara online melalui akses langsung ke forum MyAnimeList.

Berdasarkan tabel 1, variabel di atas merupakan data genre, jumlah anime, jumlah pengguna dan judul anime yang paling banyak ditonton per Summer 2023. Variabel-variabel tersebut akan dipakai sebagai faktor-faktor yang memengaruhi variabel target atau hasil.

2.2 Pengolahan dan Analisis Data

Sebelum mengolah dan menganalisis data pada tabel 1, peneliti mengambil data variabel pada forum MyAnimeList berdasarkan judul anime yang paling banyak ditonton bersama-sama dengan genre yang dibawa judul tersebut.

Variabel untuk data forum MyAnimeList ini meliputi jumlah pengguna, judul dan genre. Yang

di mana genre dan judul diambil berdasarkan pengguna ulasan yang paling banyak.

Tabel 1. Data Variabel pada Forum MyAnimeList berdasarkan Genre dengan Judul yang Paling Banyak Ditonton per Summer 2023

Genre	User	Title (Most)
Action	3.8M	Shingeki no Kyojin
Adventure	3.2M	Fullmetal Alchemist: Brotherhood
Avant Garde	1.7M	Neon Genesis Evangelion
Award Winning	3.8M	Shingeki no Kyojin
Boys Love	546K	Given
Comedy	3.1M	One Punch Man
Drama	3.8M	Shingeki no Kyojin
Fantasy	3.2M	Fullmetal Alchemist: Brotherhood
Girls Love	503K	Citrus
Gourmet	1.6M	Shokugeki no Souma
Horror	2.7M	Tokyo Ghoul
Mystery	2.0M	Boku dake ga Inai Machi
Romance	3.0M	Sword Art Online
Sci-Fi	2.5M	Steins;Gate
Slice of Life	1.3M	Hyouka
Sports	1.9M	Haikyuu!!
Supernatural	3.8M	Death Note
Suspense	3.8M	Shingeki no Kyojin

Tabel 2. Data Variabel pada Forum MyAnimeList berdasarkan Judul Anime yang Paling Banyak Ditonton per Summer 2023

User	Title	Genre
3.8M	Shingeki no Kyojin	Action, Award Winning, Drama, Suspense
3.8M	Death Note	Supernatural, Suspense
3.2M	Fullmetal Alchemist: Brotherhood	Action, Adventure, Drama, Fantasy
3.1M	One Punch Man	Action, Comedy
3.0M	Sword Art Online	Action, Adventure, Fantasy, Romance
2.7M	Tokyo Ghoul	Action, Fantasy, Horror
2.5M	Steins;Gate	Drama, Sci-Fi, Suspense
2.0M	Boku dake ga Inai Machi	Mystery, Supernatural, Suspense
1.9M	Haikyuu!!	Sports
1.7M	Neon Genesis Evangelion	Action, Avant Garde, Award Winning, Drama, Sci-Fi, Suspense
1.6M	Shokugeki no Souma	Gourmet
1.3M	Hyouka	Mystery, Slice of Life
546K	Given	Boys Love, Drama
503K	Citrus	Drama, Girls Love

Sesuai dengan tabel 1, tabel 2 menunjukkan seberapa antusias peminat animasi terhadap judul anime pada setiap genre yang mereka pilih. Tabel 2

menjelaskan bahwa urutan data variabel tersebut diambil berdasarkan sortir genre dengan judul

anime yang paling banyak ditonton pada musim panas 2023.

Tabel 3. Data Variabel pada Forum MyAnimeList berdasarkan Judul Anime yang Paling Disukai Penonton per Summer 2023

Score	Title	Genre
9.10	Fullmetal Alchemist: Brotherhood	Action, Adventure, Drama, Fantasy
9.07	Steins;Gate	Drama, Sci-Fi, Suspense
8.62	Death Note	Supernatural, Suspense
8.54	Shingeki no Kyojin	Action, Award Winning, Drama, Suspense
8.50	One Punch Man	Action, Comedy
8.44	Haikyuu!!	Sports
8.35	Neon Genesis Evangelion	Action, Avant Garde, Award Winning, Drama, Sci-Fi, Suspense
8.32	Given	Boys Love, Drama
8.30	Boku dake ga Inai Machi	Mystery, Supernatural, Suspense
8.15	Shokugeki no Souma	Gourmet
8.07	Hyouka	Mystery, Slice of Life
7.79	Tokyo Ghoul	Action, Fantasy, Horror
7.20	Sword Art Online	Action, Adventure, Fantasy, Romance
6.45	Citrus	Drama, Girls Love

Sesuai tabel 1, tabel 3 menunjukkan seberapa puas peminat animasi terhadap judul anime pada setiap genre yang mereka pilih. Tabel 3 menjelaskan bahwa urutan data variabel tersebut diambil berdasarkan sortir genre dengan judul anime yang paling disukai penonton di musim panas tahun 2023.

Setelah melakukan klasifikasi seperti tabel 2 dan 3 di atas, maka peneliti meringkas data yang ada secara deskriptif. Adapun tabelnya ialah sebagai berikut:

Tabel 4. Analisis Deskriptif berdasarkan Top 14 Judul MyAnimeList

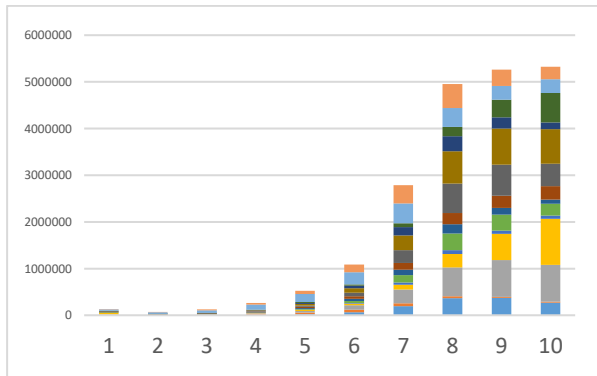
Var	Score	Min	Max	\bar{x}	σ
X1	10	2149	373253	131994.9	155083.6
X2	10	6625	63929	26738.9	19559.33
X3	10	3141	790259	265582.1	337052.1
X4	10	2617	987197	204659.5	328206.8
X5	10	519	80152	29861	34020.46
X6	10	1370	359486	118241.8	147262
X7	10	1307	195889	62363	70740.18
X8	10	4754	291366	104383	117337.6
X9	10	1961	663735	216074.6	274644.2
X10	10	3690	775974	269795.3	336216.3
X11	10	1740	322620	97757	117081.8
X12	10	2683	628046	135481.3	210895.6
X13	10	30130	424757	209378.1	150510.8
X14	10	6830	509908	181242.8	185063.2

Tabel 4 menunjukkan 14 judul teratas di forum MyAnimeList yang berdasarkan dari penonton terbanyak di setiap genrenya. Tabel 4 menjelaskan bahwa analisis deskriptif tersebut diambil berdasarkan 14 judul teratas di MyAnimeList.

Tabel 5. Analisis Deskriptif berdasarkan Score MyAnimeList

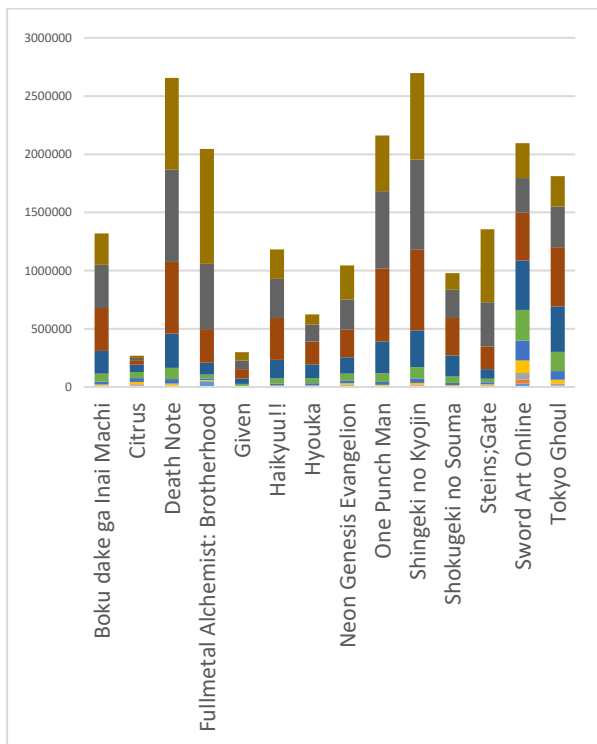
Var	Judul	Min	Max	\bar{x}	σ
X1	14	1399	48269	9747.714	13232.83
X2	14	519	32384	5202.071	8065.346
X3	14	918	57348	8814	14421.49
X4	14	2163	108298	18815.43	27021.21
X5	14	6361	170312	37629.64	41489.12
X6	14	15860	263262	77770.86	64984.15
X7	14	47025	424757	199068.8	121430.6
X8	14	38010	694081	353753.4	202111.9
X9	14	17917	790259	375882.3	242214
X10	14	19141	987197	380139.6	299121.4

Tabel 5 di atas menunjukkan bahwa Score MyAnimeList sebagai pijakan pada genre yang sering dipilih peminat animasi. Tabel 5 menjelaskan bahwa analisis deskriptif tersebut dinilai berdasarkan Score yang ada pada forum MyAnimeList.



Gambar 2. Grafik Peminat Animasi dan Scorenya

Terlihat pada gambar 2, bahwa Score tertinggi pada grafik tersebut adalah 10. Grafik peminat animasi dan scorenya merupakan gambaran dari besaran jumlah semua variabel yang digunakan di setiap score pada 14 judul teratas yang ada di forum MyAnimeList.



Gambar 3. Grafik Peminat Animasi dan Judulnya

Selanjutnya, peneliti melakukan penukaran pada variabel score dan judul menjadi grafik peminat animasi beserta judulnya, yang di mana merupakan gambaran dari besaran jumlah semua variabel yang digunakan di setiap judul pada 20.535.533 data score yang ada di forum MyAnimeList. Terlihat di gambar 3, Death Note

dan Shingeki no Kyojin menjadi judul yang paling populer per rilis Summer 2023.

Score dan judul dengan ulasan yang paling sedikit beserta yang paling banyak akan disortir sesuai data yang didapatkan dari gambar 2 dan 3. Adapun tabel nilai tertinggi dan terendah untuk di masing-masing pada setiap variabel adalah sebagai berikut:

Tabel 6. Nilai variabel Terendah dan Tertinggi berdasarkan Judul

Var	Terendah	Tertinggi
X1	Score 2	Score 9
X2	Score 2	Score 7
X3	Score 2	Score 9
X4	Score 3	Score 10
X5	Score 2	Score 8
X6	Score 2	Score 8
X7	Score 2	Score 8
X8	Score 2	Score 10
X9	Score 2	Score 9
X10	Score 2	Score 9
X11	Score 2	Score 8
X12	Score 2	Score 10
X13	Score 1	Score 7
X14	Score 1	Score 8

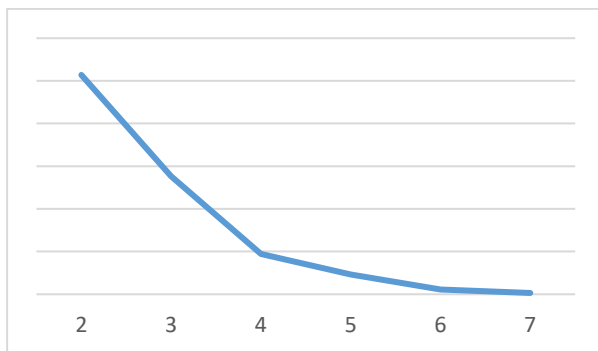
Tabel 7. Nilai Variabel Terendah dan Tertinggi berdasarkan Score

Var	Terendah	Tertinggi
X1	Hyouka	Fullmetal Alchemist: Brotherhood
X2	Given	Sword Art Online
X3	Given	Sword Art Online
X4	Given	Sword Art Online
X5	Given	Sword Art Online
X6	Given	Sword Art Online
X7	Given	Sword Art Online
X8	Citrus	Shingeki no Kyojin
X9	Citrus	Death Note
X10	Citrus	Fullmetal Alchemist: Brotherhood

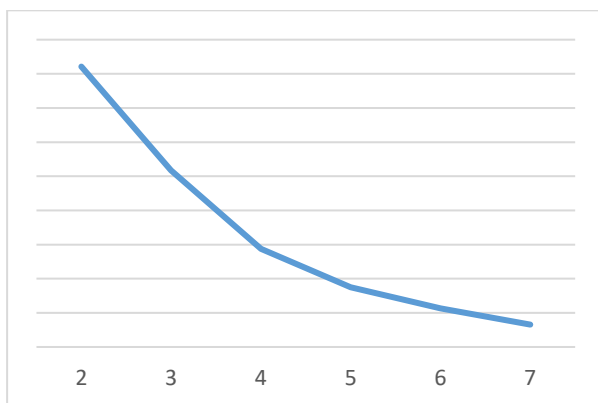
Berdasarkan tabel tersebut dapat diinterpretasikan bahwa Score 2 menjadi vote terendah pada judul Boku dake ga Inai Machi, Citrus, Death Note, Given, Haikyuu!!, Hyouka, Neon Genesis Evangelion, One Punch Man, Shingeki no Kyojin, Shokugeki no Souma dan Steins;Gate. Score 9 sebagai vote tertinggi pada judul Boku dake ga Inai Machi, Death Note, One Punch Man, dan Shingeki no Kyojin. Score 7 sebagai vote tertinggi pada judul Citrus dan Sword

Art Online. Score 3 sebagai vote terendah pada judul Fullmetal Alchemist: Brotherhood. Score 10 sebagai vote tertinggi pada judul Fullmetal Alchemist, Neon Genesis Evangelion dan Steins;Gate. Score 8 sebagai vote tertinggi pada judul Given, Haikyuu!!, Hyouka, Shokugeki no Souma dan Tokyo Ghoul. Score 1 sebagai vote terendah pada judul Sword Art Online dan Tokyo Ghoul.

Jadi didapatkan insight bahwa Hyouka merupakan judul yang memiliki vote Score 1 terendah. Kemudian, Fullmetal Alchemist: Brotherhood merupakan judul yang memiliki vote Score 1 tertinggi sekaligus vote Score 10 tertinggi di MyAnimeList. Given dan Sword Art Online merupakan judul yang memiliki vote Score 2 s/d 7 terendah dan tertinggi di MyAnimeList. Citrus merupakan judul yang memiliki Score 8 s/d 10 terendah di MyAnimeList. Shingeki no Kyojin menjadi judul dengan vote Score 8 tertinggi dan terakhir, Death Note menjadi judul yang memiliki vote Score 9 tertinggi di forum MyAnimeList.



Gambar 4. Elbow Method untuk dataset Judul



Gambar 5. Elbow Method untuk dataset Score

Untuk menentukan banyaknya kluster yang optimal maka perlu dilakukan Elbow Method, tahapan yang dilakukan adalah menghitung jarak

data dengan titik pusat klasternya, tahapan ini dilakukan terhadap jumlah kluster lainnya, setelah mendapat nilai jarak data dari masing-masing jumlah kluster, nilai tersebut disusun ke dalam grafik garis kemudian dilihat mana kurva yang bentuknya paling siku. Berdasarkan kedua gambar tersebut diketahui banyaknya kluster yang optimal dari Elbow Method terjadi patahan siku ketika keduanya sama-sama $K=4$. Berikut adalah rumus yang digunakan:

$$SSE = \sum_{k=1}^k \sum_{x_i \in S_k} (X_1 - C_k)^2$$

Pada tahapan ini analisis K-Means Clustering dilakukan dengan 2 cara, pertama perhitungan manual analisis K-Means Clustering dan yang kedua perhitungan otomatis K-Means Clustering menggunakan aplikasi RapidMiner Studio, sehingga hasil analisis dapat dinilai akurat. Tahapan yang dilakukan dalam perhitungan manual analisis K-Means Clustering yaitu, menentukan nilai K sebanyak jumlah cluster yang diinginkan. Pada tahap ini nilai K yang kita gunakan adalah $K=4$, nilai ini sudah dianalisis sebagai jumlah kluster yang optimal. Kemudian menentukan pusat cluster (centroid) secara acak dari dataset judul sebanyak jumlah K, adapun centroid yang digunakan, yaitu, Buku dake ga Inai Machi sebagai C1, Given sebagai C2, Shingeki no Kyojin sebagai C3 dan Tokyo Ghoul sebagai C4.

Selanjutnya, peneliti menentukan pusat centroid secara acak untuk dataset score sebanyak jumlah K, yaitu, Score 1 sebagai C1, Score 4 sebagai C2, Score 6 sebagai C3 dan Score 10 sebagai C4.

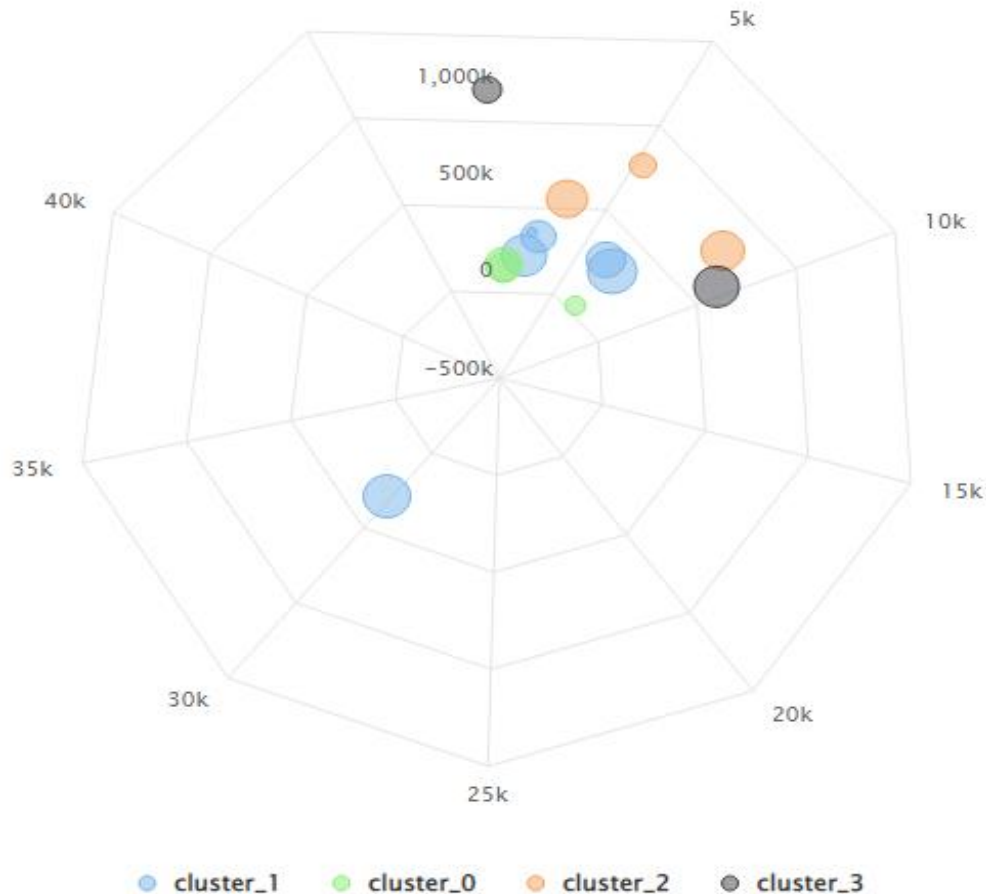
Setelah menentukan pusat centroid seperti di atas. Kemudian, tahap berikutnya adalah mengukur jarak antara objek dan setiap centroid dengan menggunakan rumus Euclidean Distance seperti berikut:

$$d_{(x_i, x_j)} = \sqrt{((x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{ip} - x_{jp})^2)}$$

Setelah mendapatkan hasilnya, kemudian objek dikelompokkan berdasarkan jarak terdekat dengan centroid. Selanjutnya, pada iterasi berikutnya, penentuan centroid baru harus dilakukan menggunakan rumus sebagai berikut:

$$C_{m(q)} = \frac{1}{n_m} \sum_{i=1}^{n_m} X_{i(q)}$$

Selanjutnya, lakukan perhitungan ulang terhadap jarak antara objek dengan masing-masing centroid baru tersebut dan klasifikasikan objek berdasarkan jarak terdekat pada iterasi berikutnya hingga tidak ada lagi objek yang berpindah klaster.



Gambar 6. Hasil K-Means Clustering berbentuk Radar Spider Web

Setelah melakukan analisis K-Means Clustering dengan menghitung secara manual, kemudian dilakukan kembali analisis K-Means Clustering dengan perhitungan otomatis. Dalam mengelompokkan data judul dan skor peminat animasi ke dalam klaster yang dibutuhkan aplikasi RapidMiner Studio. Tahapan yang dilakukan adalah menyiapkan data yang sudah ditentukan nilai K yang paling optimal, kemudian rancang tahapan-tahapan yang digunakan dalam mengklaster pada RapidMiner Studio, lalu jalankan. Maka didapatlah hasil klaster dalam bentuk jarring laba-laba seperti gambar 5 di atas.

Tabel 8. Anggota Klaster untuk dataset Judul

Klaster	Anggota Klaster	Jumlah Klaster
1	Citrus, Given, Hyouka	3
2	Boku dake ga Inai Machi, Haikyuu!!, Neon Genesis Evangelion, Shokugeki no Souma, Sword Art Online, Tokyo Ghoul	6
3	Death Note, One Punch Man, Shingeki no Kyojin	3
4	Fullmetal Alchemist: Brotherhood, Steins;Gate	2

Pada tabel berikut diketahui jumlah anggota kluster 1 yaitu 3 judul (Citrus, Given, Hyouka), kluster 2 terdapat 6 judul (Boku dake ga Inai Machi, Haikyuu!!, Neon Genesis Evangelion, Shokugeki no Souma, Sword Art Online, Tokyo Ghoul), kluster 3 terdapat 3 judul (Death Note, One Punch Man, Shingeki no Kyojin), dan kluster 4 terdapat 2 judul (Fullmetal Alchemist: Brotherhood, Steins;Gate). Disimpulkan, maka kluster 2 menjadi kluster yang mempunyai anggota kluster paling banyak untuk dataset Judul.

Tabel 9. Anggota Kluster untuk dataset Score

Kluster	Anggota Kluster	Jumlah Kluster
1	Score 1, Score 2, Score 3, Score 4, Score 5, Score 6	6
2	Score 10	1
3	Score 8, Score 9	2
4	Score 7	1

Pada tabel 9 diketahui jumlah anggota kluster 1 yaitu 6 skor (Score 1, Score 2, Score 3, Score 4, Score 5, Score 6), kluster 2 terdapat 1 skor (Score 10) saja, kluster 3 terdapat 2 judul (Score 8, Score 9), dan kluster 4 terdapat 1 skor (Score 7). Dapat disimpulkan, kluster 1 menjadi kluster yang memiliki anggota kluster paling banyak pada dataset Score.

Tabel 10. Mean Variabel Kluster untuk dataset Judul

Variabel	Kluster			
	1	2	3	4
X1	9651	51271	17603	57943
X2	8451	49660	8792	5926
X3	13953	88549	14731	6163
X4	29490	182447	38410	13069
X5	57059	332495	106360	30901
X6	112602	649476	265399	61315
X7	228436	1492260	880940	185327
X8	314051	2211723	1941080	485693
X9	237611	1854945	2229968	939828
X10	178325	1517150	2011237	1615243
Mean	118962.9	842997.6	751452	340140.8

Untuk merancang skema dari data heatmap, maka dilakukan penghitungan rata-rata variabel kluster, sehingga dari hasil hitung tersebut diketahui urutan dari tiap-tiap kluster. Pada tabel 10 dapat diinterpretasikan bahwa kluster 1 memiliki mean terendah. Pada kluster 2 memiliki mean tertinggi. Pada kluster 3 memiliki mean sedang untuk judul di genre Action dan Suspense. Terakhir, pada kluster 4 memiliki rata-rata sedang untuk judul di genre-bender.

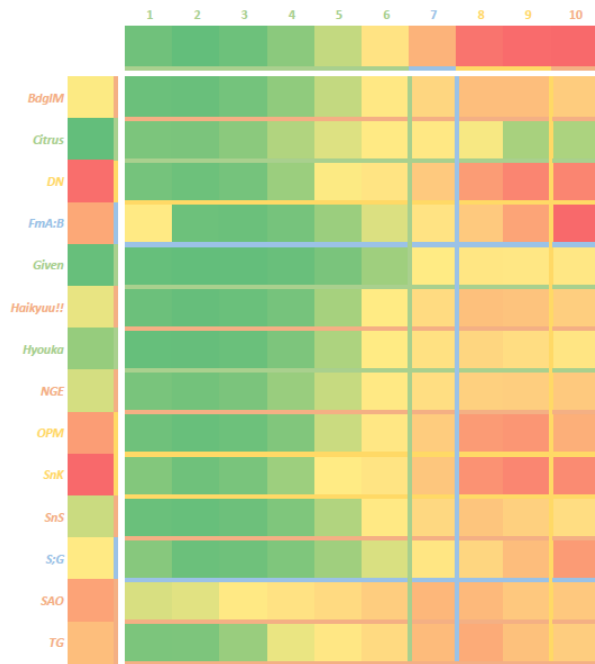
Tabel 11. Mean Variabel Kluster untuk dataset Score

Variabel	Kluster			
	1	2	3	4
X1	114387	269039	741353	195170
X2	128392	19141	55927	63929
X3	164731	789890	1408814	292386
X4	105207	987197	850715	103476
X5	27226	72227	152132	47025
X6	72312	252365	697408	160333
X7	75588	86957	343603	117482
X8	114577	291366	498310	139577
X9	116614	479502	1292179	272451
X10	169950	741845	1470055	316103
X11	90017	142011	563752	181790
X12	70110	628046	574806	81851
X13	661734	299288	708002	424757
X14	300871	263081	857843	390633
Mean	157979.7	380139.6	729635.6	199068.8

Jadi didapatkan insight bahwa kluster 1 menjadi urutan terendah. Kemudian, disusul kluster 4 sebagai urutan kedua, lalu kluster 2 sebagai urutan ketiga, dan terakhir kluster 3 sebagai urutan tertinggi pada ulasan peminat animasi di forum MyAnimeList.

3. Hasil dan Pembahasan

Dari data yang sudah dikumpulkan sebelumnya, telah diperoleh 14 judul dengan 10 skor berbeda. Kemudian, peneliti mengklasifikasikan 14 judul dan 10 skor tersebut menjadi 4 kluster yang berbeda juga seperti pada gambar di bawah.



Gambar 7. Heatmap Peminat Animasi di forum MAL

Di gambar 7 menjelaskan bahwa pada urutan vertikal adalah judul-judul yang mewakili setiap genre peminat animasi, sedangkan pada urutan horizontal adalah skor-skor yang diperoleh dari peminat animasi terhadap 14 judul di atas. Heatmap pada gambar 7 di atas sangat mudah untuk dibaca, hanya dengan melihat gradasi warna hijau terang (paling sedikit) ke merah terang (paling banyak). Lihat contoh di bawah:

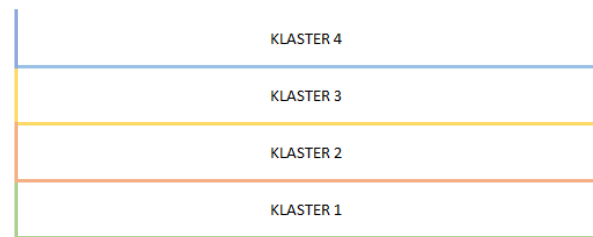


Gambar 8. Gradasi warna heatmap (Google, 2023)

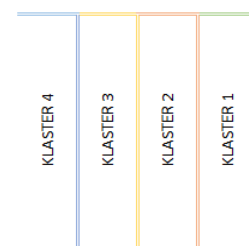
Dapat dilihat, Death Note dan Shingeki no Kyojin menjadi judul yang paling banyak disukai oleh peminat animasi, ditandai dengan kotak berwarna merah terang. Kemudian, Citrus dan Given menjadi judul yang paling sedikit disukai oleh peminat animasi, ditandai dengan kotak berwarna hijau terang. Selanjutnya, Score 1 sampai dengan 6 ditandai dengan klaster berwarna hijau (Klaster 1), Score 7 ditandai dengan klaster berwarna biru (Klaster 4), Score 8 dan 9 ditandai dengan klaster berwarna kuning (Klaster 3), lalu Score 10 ditandai dengan warna merah (Klaster 2).

Pada gambar 9 terdapat 4 warna klaster berbeda dengan karakteristik klaster yang berbeda pula. Hijau menunjukkan klaster 1, merah

menunjukkan klaster 2, kuning menunjukkan klaster 3, dan biru menunjukkan klaster 4.



Gambar 9. Klaster dengan variabel Judul



Gambar 10. Klaster dengan variabel Score

Terakhir, klaster pada judul-judul di atas juga ditandai dengan 4 warna klaster sesuai warna klaster yang ada di urutan Score. Klaster 1 yang ditandai dengan warna hijau (Citrus, Given, Hyouka) = (Score 1, 2, 3, 4, 5, 6) menunjukkan bahwa judul tersebut adalah judul yang paling sedikit disukai oleh kalangan peminat animasi di forum MyAnimeList. Begitu pula dengan klaster 3 (Death Note, One Punch Man, Shingeki no Kyojin) = (Score 8, 9) menunjukkan bahwa judul-judul itu adalah judul yang paling banyak disukai oleh peminat animasi di MyAnimeList.

4. Kesimpulan

Dari hasil analisis menggunakan K-Means disimpulkan terdapat 4 klaster yang terbentuk berdasarkan variabel yang digunakan dengan pembagian klaster yaitu klaster 1: Citrus, Given, Hyouka dengan Score 1 s/d 6. Klaster 2: Boku dake ga Inai Machi, Haikyuu!!, Neon Genesis Evangelion, Shokugeki no Souma, Sword Art Online, Tokyo Ghoul dengan Score 10. Klaster 3: Death Note, One Punch Man, Shingeki no Kyojin dengan Score 8-9. Klaster 4: Fullmetal Alchemist: Brotherhood, Steins;Gate dengan Score 7.

Dari hasil pemodelan data heatmap disimpulkan bahwa 4 klaster tersebut diklasifikasi berdasarkan data terendah hingga tertinggi, kemudian urutan klaster tersebut diberi warna yang berbeda. Dari keempat klaster tersebut memiliki

karakteristik masing-masing yang dapat dipahami oleh penonton animasi dengan mudah dan jelas sehingga membantu penonton baru dalam memilih dan memilah tontonan mana yang sesuai dengan minat mereka.

Daftar Pustaka

- Arfiani, I., Yuliansyah, H., & Suratin, M. D. (2022). *Implementasi Bee Colony Optimization Pada Pemilihan Centroid (Klaster Pusat) Dalam Algoritma K-Means*. Jurnal BITS, 3(4), 756-763. <https://doi.org/10.47065/bits.v3i4.1446>
- Aulia, L., Ikhsan, M., & Armansyah, A. (2023). Rancang Bangun Alat Bantu Pengenalan Warna Untuk Penyandang Buta Warna Menggunakan Metode Coloring Filters (Cf) Dan K-Means Clustering Berbasis Mikrokontroler. *Innovative: Journal Of Social Science Research* 3(4), 7079-7087. <https://doi.org/10.31004/innovative.v3i4.4348>
- Billah, M., Zartesyah, M. A., & Komalasari, D. (2021). *Penerapan Collaborative Filtering, PCA dan K-Means dalam Pembangunan Sistem Rekomendasi Ongoing dan Upcoming Film Animasi Jepang*. Senamika, 2(1), 606-615.
- Billah, M., Zartesyah, M. A., & Prasvita, D. S. (2021). *Penerapan Collaborative Filtering, PCA dan K-Means dalam Pembangunan Sistem Rekomendasi Film*. Senamika, 2(1), 579-587.
- Faujia, R. A., & Subarkah, M. Z. (2022). *Analisis Klaster K-Means Dan Visualisasi Data Spasial Berdasarkan Karakteristik Persebaran Covid-19 Dan Pelanggaran Protokol Kesehatan Di Jawa Tengah*. Seminar Nasional Official Statistics, 2022(1), 813-822. <https://doi.org/10.34123/semnasoffstat.v2022i1.1222>
- Google (2023). *Google Maps Android Heatmap Utility*, <https://developers.google.com/maps/documentation/android/utility/heatmap#introduction>. Diakses tanggal 15 Juni 2023.
- Kurniawan, R. A., Hasibuan, M. S., Piramida, P., & Ramadhan, R. S. (2022). *Penerapan Algoritma K-Means Untuk Clustering Tempat Makan Di Batubara*. Journal of Computer Science and Informatics Engineering (CoSIE), 10-18. <https://doi.org/10.55537/cosie.v1i1.27>
- Mulyani, A. (2023). *Visualisasi Data Ticketing Servicedesk Dengan Dashboard Pada PT Brantas Abipraya (PERSERO)*. 7(2), 289-300. <https://doi.org/10.52362/jisamar.v7i2.1074>
- Mulyani, E. D. S., Yusup, A. M., Tisna, A. K., Fauzi, F. A., Seta, I. B., Khairunas, R., & Ardiansyah, W. (2019). *Clustering Wilayah Dan Pelanggaran Berkendaraan Menggunakan Algoritma K-Means Pada Data Satlantas Polres Tasikmalaya Kota*. Jurnal Sistem Informasi dan Teknologi Informasi, 6(1), 1-7. <https://doi.org/10.36774/jusiti.v8i1.595>
- Muningsih, E., Maryani, I., & Handayani, V. R. (2021). *Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa*. Evolusi: Jurnal Sains dan Manajemen, 9(1), 95-100. <https://doi.org/10.31294/evolusi.v9i1.10428>
- Rahmalinda, N. A., & Jananto, A. (2022). *Penerapan Metode K-Means Clustering dalam Menentukan Strategi Promosi Berdasarkan Data Penerimaan Mahasiswa Baru*. Jurnal Teknokompak, 16(2), 163-175. <https://doi.org/10.33365/jtk.v16i2.1971>
- Nasution, M. Z., & Hasibuan, M. S. (2020). *Pendekatan Initial Centroid Search Untuk Meningkatkan Efisiensi Iterasi Klustering K-Means*. Techno. Com, 19(4), 341-352. <https://doi.org/10.33633/tc.v19i4.3875>
- Platers (2022). *MAL Map: Clustering and visualizing recommendations*, <https://myanimelist.net/forum/?topicid=1984781>. Diakses tanggal 12 Juni 2023.
- Rosadi, M., Nurhasanah, D. A., & Hasibuan, M. S. (2023). *Clustering Panjang Ruas Jalan di BBPJN Sumut Menggunakan Algoritma K-Means*. Journal of Computer Science and Informatics Engineering (CoSIE), 29-38. <https://doi.org/10.55537/cosie.v2i1.567>
- Rusydijah, E. F., Virgiannada, H. R., Ridwan, M., Nugroho, B. A., & Rahman, M. R. (2021). *Clustering of Learning Media User Data During Covid-19 Pandemic Using K-Means Method Based on Multicultural Culture in Indonesia*. Journal of Innovation in Educational and Cultural Research, 2(2), 67-76. <https://doi.org/10.46843/jiecr.v2i2.39>
- Simanjuntak, A., & Hasibuan, M. S. (2023). *Application of PCA and K-Means Clustering Methods to Identify Diabetes Mellitus Patient Groups Based on Risk Factors*. Prisma Sains: Jurnal Pengkajian Ilmu dan Pembelajaran Matematika dan IPA 11(4), 1002-1017. <https://doi.org/10.33394/jps.v11i4.9263>
- Septianingsih, A. (2022). *Analisis K-Means Clustering Pada Pemetaan Provinsi Indonesia Berdasarkan Indikator Rumah Layak Huni*. Jurnal Lebesgue, 1(1) 1-8. <https://doi.org/10.46306/lb.v3i1.116>
- Setyaningtyas, S., Nugroho, B. I., & Arif, Z. (2022). *Tinjauan Pustaka Sistematis Pada Data Mining: Studi Kasus Algoritma K-Means Clustering*. Jurnal Teknoif, 10(2), 52-61. <https://doi.org/10.21063/jtif.2022.V10.2.52-61>

