

Penerapan Metode K-Means Clustering untuk Pengelompokan Musik Berdasarkan Karakteristik Audio

Shofi Shulhiyana¹, Apriade Voutama²

Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang

Penulis Korespondensi: shofishulhiyana6@gmail.com¹,

apriade.voutama@staff.unsika.ac.id²

Informasi Artikel:

Received: 13 Mei 2025, Accepted: 23 Mei 2025, Published: 28 Mei 2025

Abstract

Perkembangan teknologi digital telah menyebabkan lonjakan besar dalam jumlah data musik yang tersedia, khususnya pada platform streaming seperti Spotify. Penelitian ini menerapkan metode K-Means Clustering guna mengelompokkan lagu berdasarkan kemiripan fitur audio seperti danceability, energy, valence, acousticness, dan tempo. Proses penelitian menggunakan metodologi SEMMA (Sample, Explore, Modify, Model, Assess), dimulai dari pengumpulan data lagu terpopuler Spotify tahun 2024 yang berjumlah 6.513 lagu dengan 19 atribut. Setelah dilakukan eksplorasi, modifikasi, normalisasi, dan pemodelan, jumlah klaster optimal ditentukan menggunakan elbow method dan hasil evaluasi menggunakan silhouette score sebesar 0,435. Hasilnya, terbentuk tiga klaster musik yang merepresentasikan perbedaan karakteristik audio. Penelitian ini menunjukkan bahwa metode K-Means efektif dalam mengelompokkan musik, dan dapat dikembangkan lebih lanjut dengan algoritma alternatif untuk hasil yang lebih presisi.

Keywords: Data Mining, K-Means Clustering, Semma, Python, Fitur Audio

1. Pendahuluan

Musik merupakan elemen yang tak terpisahkan dari aktivitas sehari-hari banyak individu. Musik memiliki beragam karakteristik, genre, dan pola tersendiri. Peran musik tidak hanya menjadi sekadar hiburan, tetapi juga sebagai sarana ekspresi dan pelengkap aktivitas sehari-hari. Dahulu, musik dinikmati melalui media seperti CD, televisi, dan radio. Namun, dengan adanya perkembangan teknologi digital, berbagai platform streaming kini tersedia secara gratis dan dapat diakses oleh jutaan pengguna di seluruh dunia [1]. Hal ini memungkinkan siapa saja untuk mendengarkan musik kapan pun dan di mana pun hanya melalui satu platform. Meskipun begitu, dengan semakin pesatnya digitalisasi, jumlah data musik yang tersedia juga terus bertambah secara signifikan. Jika data ini tidak dimanfaatkan dengan baik, maka potensi besar yang terkandung di dalamnya akan terabaikan. Seiring dengan perkembangan zaman dan teknologi digital, jumlah data mengenai musik terus mengalami peningkatan yang signifikan [2].

Saat ini, banyak sekali platform yang menyediakan layanan streaming musik. Beberapa diantaranya seperti Spotify, Apple Musik, Joox, dan Resso telah menarik banyak sekali pendengar dan pertumbuhan musik yang tinggi. Dari banyaknya platform streaming musik

yang ada, spotify saat ini menduduki peringkat tertinggi dengan pengguna terbanyak di seluruh dunia [3], [4]. Spotify sebagai salah satu platform paling berpengaruh, tentu membutuhkan analisis lebih lanjut guna menjaga persaingan kualitas dan pengalaman pengguna [5].

Setiap tahunnya, ribuan hingga jutaan lagu baru diproduksi oleh berbagai musisi, penulis lagu, produser, dan ahli lainnya yang masing-masing memiliki karakteristik audio yang berbeda. Perbedaan ini terlihat dari berbagai aspek seperti tempo, ritme, intensitas beat, serta elemen musikal lainnya. Oleh karena itu, pengelompokan musik berdasarkan karakteristik audionya menjadi penting untuk membantu pendengar menemukan lagu yang sesuai dengan preferensi mereka. Dalam penelitian ini, metode K-Means Clustering digunakan untuk mengelompokkan musik berdasarkan kesamaan fitur audio tertentu.

Beberapa penelitian sebelumnya telah mengkaji penerapan algoritma K-Means dalam analisis pengelompokkan musik dengan pendekatan yang berbeda [6], [7], [8].

Salah satu penelitian yang mengimplementasikan metode ini adalah penelitian yang dilakukan oleh Lia Nurhalimah, Teguh Iman Hermanto, dan Ismi Kaniawulan dengan judul penelitian “Analisis Prediksi Mood Genre Musik Pop Menggunakan Algoritma K-Means dan C4.5”. mendapatkan hasil penelitian berupa penggolongan empat jenis suasana hati atau mood, yaitu marah, sedih, senang, dan bersemangat. Penelitian ini menghasilkan akurasi diatas 90% untuk seluruh class precisionnya [9].

Tinjauan penelitian berikutnya dilakukan oleh Noor Ravi Fitradhi, Moch.Firman Hidayat, Tegar Widi Saputro, dkk dengan judul penelitian “Laporan Final Project Machine Learning rekomendasi Musik Spotify Menggunakan Metode K-Means”. Hasil penelitian yang didapatkan ialah bahwa metode k-means clustering adalah metode yang efektif untuk digunakan dalam memberikan wawasan mengenai mengenai karakteristik audio [1].

Penelitian selanjutnya dilakukan oleh Muhammad Ikhsan Firmansyah, Ramdhan Saepul Rohman, dan Evi Masusanti dengan judul penelitian “Penerapan Algoritma Clustering K-Means Untuk Fitur Atribut Pada Layanan Streaming Musik Spotify”. Dari penelitian ini mendapatkan hasil bahwa algoritma K-Means berhasil digunakan dengan maksimal untuk pengelompokkan cluster data dengan menghasilkan sembilan cluster [10].

Hasil dari penelitian ini digunakan sebagai salah satu bahan pertimbangan untuk mengelompokkan musik berdasarkan karakteristik audio. Selain itu, hasil penelitian ini juga dapat dijadikan sebagai referensi, khususnya untuk pengklasifikasian atau pun sistem rekomendasi musik

2. Tinjauan Literatur

3.1 Spotify

Spotify adalah platform layanan untuk mendengarkan musik secara digital yang menyediakan akses ke jutaan lagu, podcast, dan konten audio lainnya dari berbagai artis dan seniman di seluruh dunia. Spotify memanfaatkan big data untuk menganalisis data preferensi pengguna [2].

3.2 Clustering

Clustering adalah salah satu teknik mengolah data dalam konsep unsupervised learning yang digunakan untuk mengelompokkan data berdasarkan kesamaan karakteristik tanpa adanya label [11].

3.3 Metode K-Means

Published by:

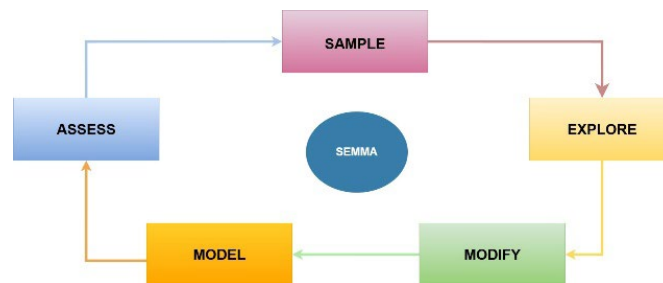
Metode K-Means adalah salah satu metode dalam konsep unsupervised learning dimana data dikelompokkan berdasarkan jarak terdekat ke titik pusat cluster atau centroid [12].

3.4 Python

Python adalah salah satu Bahasa pemrograman yang banyak digunakan dalam penerapan data mining karena Bahasa pemrograman python memiliki library dan framework yang mendukung, contohnya seperti pandas yang digunakan untuk manipulasi dan analisis data dan NumPy yang digunakan untuk komputasi numerik [13].

3. Metode Penelitian

Metode yang digunakan adalah SEMMA (Sample, Explore, Modify, Model, dan Assess). Sesuai dengan namanya, metode ini terdiri dari lima tahapan utama, yaitu *sample* (pengumpulan data), *explore* (eksplorasi data), *modify* (modifikasi data), *model* (pemodelan data), dan *assess* (evaluasi data). Berikut adalah penjelasan lebih rinci mengenai masing-masing tahapan:



Gambar 1. Tahapan Alur SEMMA

3.1 Sample

Dalam tahap pengumpulan data, dilakukan pengambilan dataset yang berisi fitur-fitur audio yang menjadi atribut utama dalam penelitian ini. Setelah dataset diperoleh, data tersebut kemudian dimuat dan diproses lebih lanjut melalui tahap preprocessing untuk memastikan kualitas dan konsistensi data sebelum dianalisis.

3.2 Explore

Pada tahap ini, dilakukan eksplorasi terhadap dataset yang telah diperoleh pada tahap sebelumnya. Dataset yang telah diperoleh, kemudian dianalisis lebih dalam untuk memahami struktur dan karakteristiknya. Proses ini mencakup deskripsi karakteristik data, serta visualisasi korelasi antar atribut dalam dataset. Visualisasi ini bertujuan untuk menyajikan hubungan antar data secara lebih jelas melalui gambar atau grafik.

3.3 Modify

Pada tahap ini, dilakukan modifikasi data untuk memastikan data sudah siap digunakan untuk tahap selanjutnya. Modifikasi yang dilakukan meliputi pemilihan, perubahan, penambahan, maupun penghapusan atribut atau variabel yang dianggap kurang relevan. Tahap ini memastikan bahwa data yang digunakan sesuai dengan kebutuhan analisis.

3.4 Model

Pada penelitian ini, pemodelan yang digunakan adalah K-Means Clustering. Data yang telah melalui tahap modifikasi, kemudian dikelompokkan berdasarkan karakteristik yang serupa untuk mengidentifikasi pola atau segmen tertentu. Proses pemodelan dilakukan dengan

menggunakan bahasa pemrograman Python. Sebelum melakukan pemodelan, digunakan Elbow Method untuk menentukan jumlah K yang optimal berdasarkan total inertia (*Sum of Squared Errors – SSE*).

3.5 Asses

Tahap ini merupakan langkah terakhir yang berfokus pada penilaian hasil pemodelan. Evaluasi dilakukan menggunakan Sillhouette Score untuk Mengukur seberapa baik data dikelompokkan dalam sebuah cluster.

4. Hasil dan Pembahasan

4.1 Sample

Proses pengumpulan data atau sampling dilakukan dengan menggunakan data yang bersumber dari platform Kaggle. Dataset yang diambil bernama “Spotify Top Songs Audio Features”. Dataset bersifat mentah dan berisikan keseluruhan atribut yang ada pada lagu di Spotify seperti `artist_name`, `artist_id`, `energy`, `track_name` hingga `key_mode`. Data lagu yang diambil berdasarkan lagu terpopuler pada platform Spotify pada tahun 2024 yang menghasilkan data sebanyak 6.513 data dan dilengkapi dengan 19 atribut.



	id	artist_names	track_name	source	key	mode	time_signature	danceability	energy	speechiness	acousticness
0	000xQL6tZNLJzirtlgxqSl	ZAYN, PARTYNEXTDOOR	Still Got Time (feat. PARTYNEXTDOOR)	RCA Records Label	G	Major	4 beats	430	408	0.0639	1249
1	003eolwxETJuiVWmNFMoZy	Alessia Cara	Growing Pains	Def Jam Recordings	C#/Db	Minor	4 beats	41	536	0.7330	1076
2	003vwx7Niy0yvvhH4a68B	The Killers	Mr. Brightside	Island Records	C#/Db	Major	4 beats	40	692	0.0747	93
3	00B7TZ0Xawar6NZ0UJFomN	Cardi B, Chance the Rapper	Best Life (feat. Chance The Rapper)	Atlantic/KSR	A	Major	4 beats	302	406	0.5530	1405
4	00Blm7zeNqgYLPtW6zg8cj	Post Malone, The Weeknd	One Right Now (with The Weeknd)	Republic Records	C#/Db	Major	4 beats	369	562	0.0530	683

Gambar 2. Dataset Spotify Top Songs Audio Feature

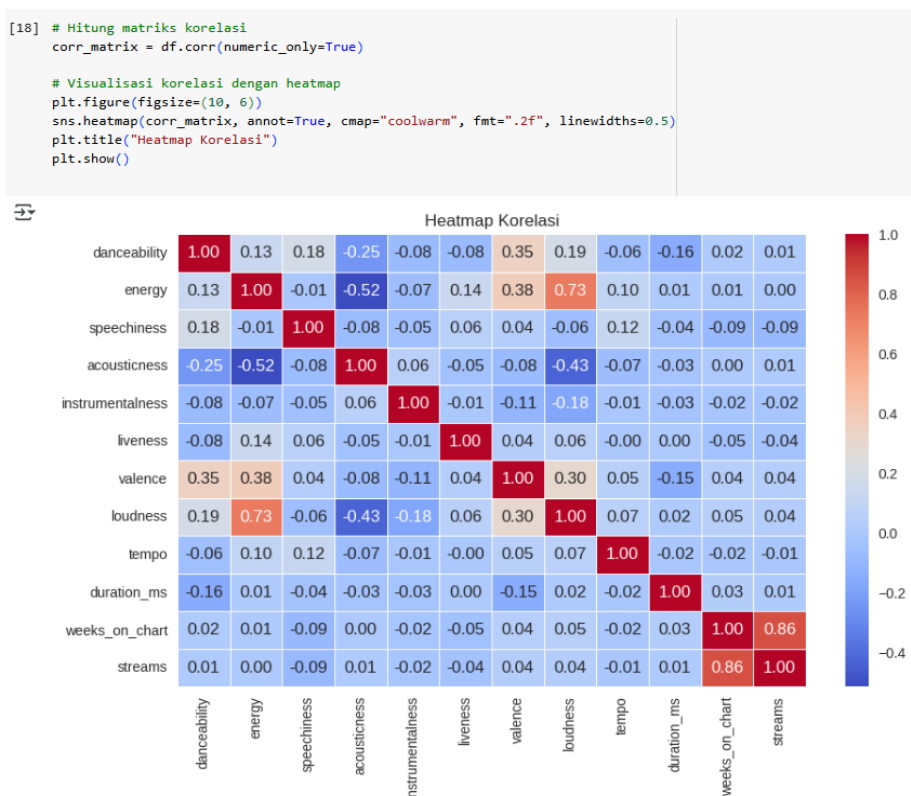
4.2 Explore

Dari dataset yang telah diperoleh, diketahui bahwa dataset terdiri dari 19 atribut yang saling berkaitan. Berikut adalah keterangan dan penjelasan lebih lanjut mengenai masing-masing atribut dalam dataset:

No	Nama Atribut	Keterangan
1.	Id	ID unik untuk setiap lagu.
2.	Artist_names	Nama artis atau penyanyi dari lagu tersebut.
3.	Track_name	Judul lagu.
4.	Source	Asal atau sumber data lagu.
5.	Key	Nada dasar lagu, direpresentasikan dengan angka
6.	Mode	Mode lagu
7.	Time_signature	Pola ketukan lagu
8.	Danceability	Tingkat kesesuaian lagu untuk menari.
9.	Energy	Tingkat intensitas atau semangat lagu.
10.	Speechiness	Seberapa banyak elemen bicara dalam lagu.
11.	Acousticness	Seberapa besar kemungkinan lagu bersifat akustik.
12.	Instrumentalness	Kemungkinan lagu tidak memiliki vokal.
13.	Liveness	Menunjukkan apakah lagu terdengar seperti rekaman live.
14.	Valence	Tingkat keceriaan/positivitas lagu.
15.	Loudness	Tingkat kekerasan suara rata-rata lagu dalam desibel.
16.	Tempo	Kecepatan lagu dalam beat per minute.
17.	Duration_ms	Durasi lagu dalam milidetik.
18.	Weeks_on_chart	Lama lagu bertahan di tangga lagu.
19.	Streams	Jumlah total streaming lagu.

Gambar 3. Penjelasan Atribut Pada Dataset

Setelah memahami maksud dan keterangan dari masing-masing atribut dalam dataset, berikutnya dilakukan analisis hubungan antar fitur yang relevan dengan menggunakan heatmap korelasi guna mengetahui sejauh mana keterkaitan antar atribut, sehingga dapat memberikan gambaran awal mengenai pola dalam data. Berikut adalah hasil visualisasinya:

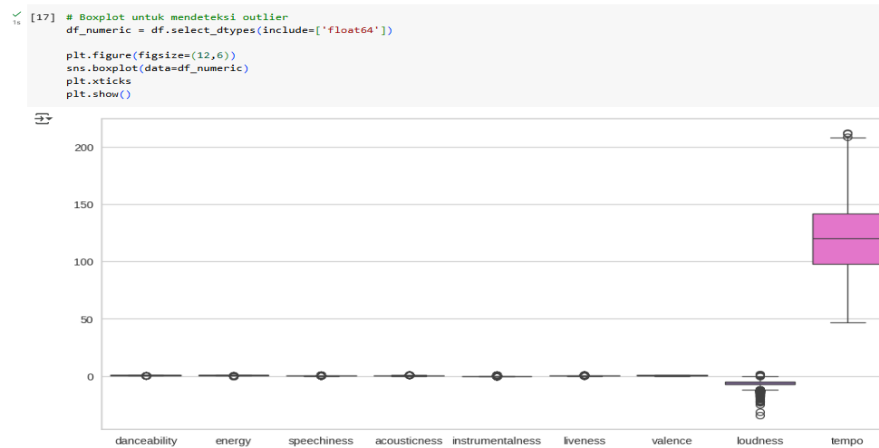


Gambar 4. Heatmap Korelasi Antar Variabel

Analisis korelasi hanya dilakukan pada fitur-fitur numerik yang dianggap relevan. Atribut yang digunakan dalam visualisasi diatas adalah fitur-fitur audio yang berpengaruh dalam

karaktersistik audio, diantaranya terdiri dari atribut danceability, valence, energy, acousticness, instrumentalness, liveness, speechiness, dan tempo. Heatmap menunjukkan hubungan antar fitur audio musik. Danceability cenderung meningkat seiring dengan valence, sedangkan energy berbanding terbalik dengan acousticness. Lalu, beberapa atribut seperti acousticness dan energy tidak memiliki korelasi kuat dengan fitur lain. Namun analisis korelasi awal ini akan dibuktikan lebih lanjut menggunakan model K-Means clustering.

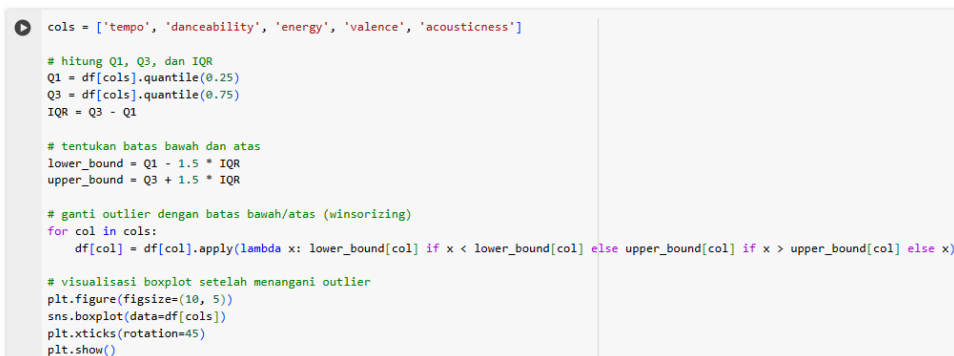
Pada tahap ini juga dilakukan pendeteksian outlier pada data. Dari hasil visualisasi, bahwa hampir semua fitur yang akan digunakan mengandung outlier. Keberadaan outlier ini perlu untuk diatasi karena dapat memengaruhi hasil clustering secara signifikan. Oleh karena itu, fitur-fitur numerik yang mengandung outlier akan dinormalisasi pada tahap berikutnya.



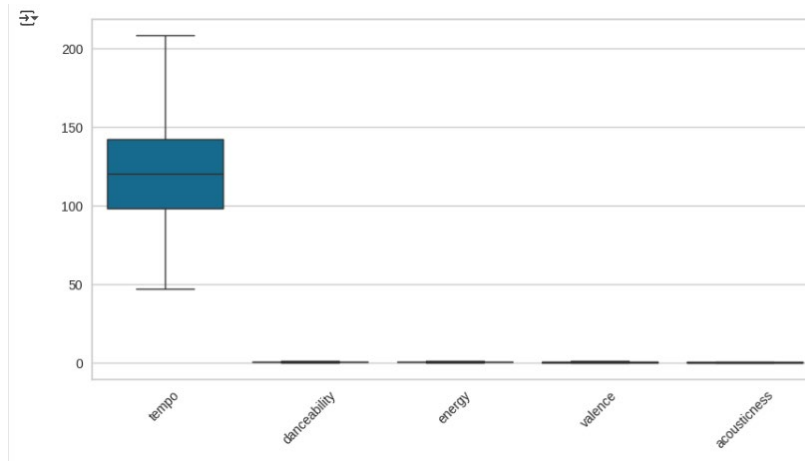
Gambar 5. Boxplot Untuk Mendeteksi Outlier

4.3 Modify

Pada tahap ini, dilakukan penanganan terhadap data outlier pada beberapa fitur menggunakan teknik *winsorizing*, yaitu mengganti nilai outlier dengan batas bawah atau atas tertentu. Proses ini diterapkan pada lima variabel utama, yaitu danceability, tempo, energy, valence, dan acousticness.



Gambar 6. Penerapan Winsorizing Untuk Penanganan Outlier



Gambar 7. Boxplot Setelah Dilakukan Penanganan Outlier

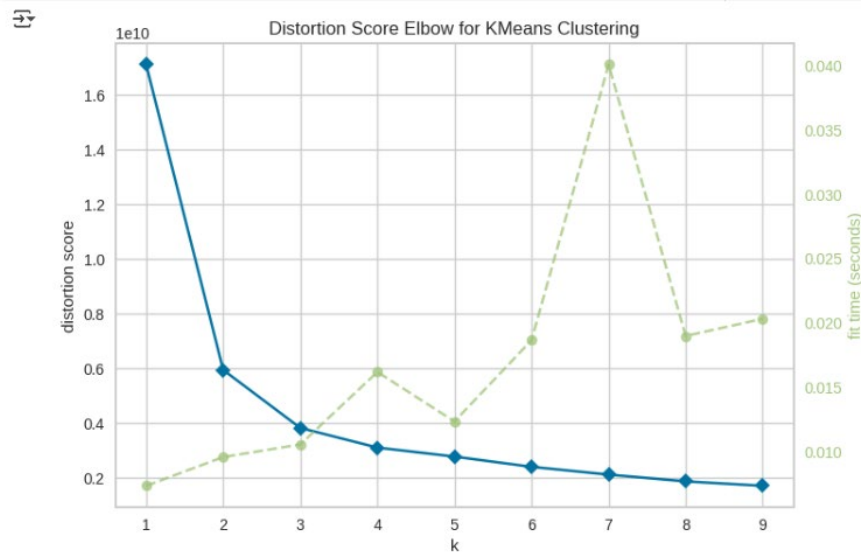
Setelah dilakukan proses *winsorizing*, hasil visualisasi menunjukkan bahwa kelima fitur tersebut sudah tidak lagi menunjukkan adanya outlier. Selanjutnya, dilakukan proses normalisasi menggunakan metode *MinMaxScaler* agar seluruh fitur numerik berada dalam rentang nilai antara 0 hingga 1. Normalisasi ini penting dilakukan untuk menyeragamkan skala data sehingga memudahkan dalam proses clustering menggunakan algoritma K-Means. Berdasarkan gambar yang dihasilkan, dapat terlihat bahwa normalisasi berhasil diterapkan pada kelima fitur tersebut.



Gambar 8. Normalisasi Dengan Menggunakan Minmaxscaler

4.4 Model

Pada tahap ini dilakukan proses penentuan jumlah kluster (K) yang paling sesuai menggunakan metode elbow. Dari grafik elbow yang dibuat, terlihat adanya penurunan tajam pada nilai inertia hingga K = 3, kemudian penurunan tersebut mulai melandai. Hal ini mengindikasikan bahwa tiga kluster merupakan jumlah yang optimal, sehingga dalam penelitian ini digunakan K = 3 sebagai jumlah kluster yang akan dibentuk.



Gambar 9. Penentuan Nilai K Optimal

mengetahui jumlah K yang optimal, model *clustering* kemudian dibangun dengan menggunakan nilai K tersebut, yaitu K = 3.

```

k = 3
model_kmeans = KMeans(n_clusters=k, random_state=42)
df['Cluster'] = model_kmeans.fit_predict(features) # Simpan hasil cluster ke df asli

# Analisis karakteristik masing-masing cluster
def analyze_clusters(df):
    print("Cluster Analysis:\n")
    for cluster_id in range(k):
        cluster_data = df[df['Cluster'] == cluster_id]

        # Hitung rata-rata untuk fitur numerik
        mean_danceability = cluster_data['danceability'].mean()
        mean_energy = cluster_data['energy'].mean()
        mean_valence = cluster_data['valence'].mean()
        mean_acousticness = cluster_data['acousticness'].mean()
        mean_tempo = cluster_data['tempo'].mean()

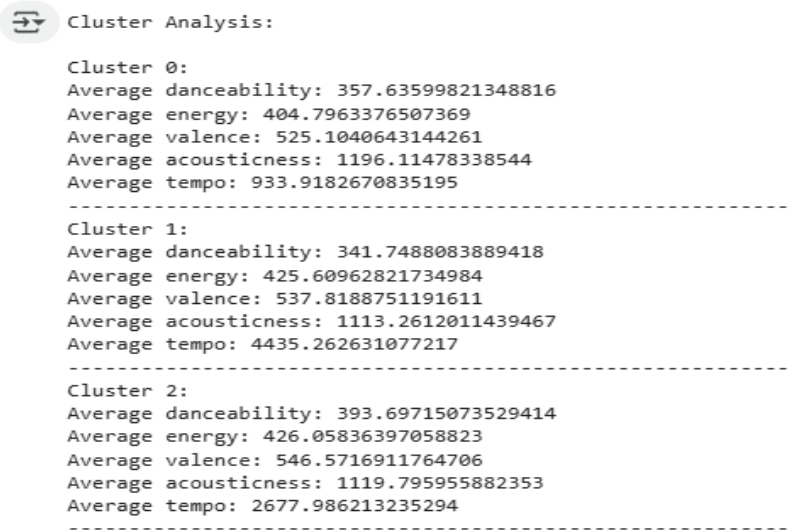
        print(f"Cluster {cluster_id}:")
        print(f"Average danceability: {mean_danceability}")
        print(f"Average energy: {mean_energy}")
        print(f"Average valence: {mean_valence}")
        print(f"Average acousticness: {mean_acousticness}")
        print(f"Average tempo: {mean_tempo}")
        print("-" * 60)

# Jalankan analisis cluster
analyze_clusters(df)

```

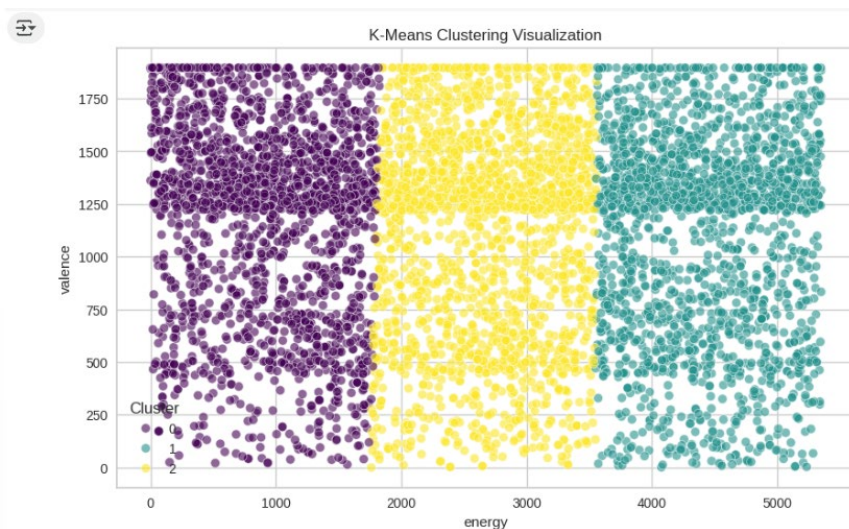
Gambar 10. Kode Python Untuk Analisis Hasil Cluster

Dari clustering yang telah dilakukan, didapatkanlah hasil clustering sebagai berikut



Gambar 11. Analisis Hasil Cluster

Hasil dari proses clustering ini menghasilkan tiga kluster yang berbeda, yang masing-masing memiliki karakteristik tertentu. Visualisasi juga dibuat untuk menunjukkan perbandingan antar fitur, khususnya antara valence dan energy, sehingga pola antar kluster dapat diamati dengan lebih jelas.



Gambar 12. Visualisasi Hasil Clustering Dengan K-Means

4.5 Assess

Evaluasi terhadap hasil clustering dilakukan menggunakan skor silhouette, yang berfungsi untuk mengukur sejauh mana pemisahan antar kluster terbentuk dengan baik. Berdasarkan perhitungan, diperoleh nilai silhouette score sebesar 0,416.

```
# Evaluasi dengan Silhouette Score
silhouette_avg = silhouette_score(features, df['Cluster'])
print(f'Silhouette Score: {silhouette_avg}')
```

➡ Silhouette Score: 0.416394114982518

Gambar 13. Evaluasi Dengan Silhouette Score

Setelah semua fitur dinormalisasi dengan MinMaxScaler, nilai-nilainya kemudian dikembalikan ke skala aslinya untuk keperluan analisis lanjutan. Dari proses ini, diperoleh hasil analisis sebagai berikut:

➡ Cluster Analysis (After Inverse Transform):

```
Cluster 0:
Average danceability: 1913352.5904421618
Average energy: 262712.82313532825
Average valence: 394353.152300134
Average acousticness: 1210468.1607860653
Average tempo: 1770709.034390353
-----
Cluster 1:
Average danceability: 1828356.124880839
Average energy: 276220.6487130601
Average valence: 403901.97521448997
Average acousticness: 1126620.335557674
Average tempo: 8409257.948522402
-----
Cluster 2:
Average danceability: 2106279.7564338236
Average energy: 276511.87821691175
Average valence: 410475.3400735294
Average acousticness: 1133233.5073529412
Average tempo: 5077461.860294118
-----
```

Gambar 14. Analisis Hasil Cluster Setelah Dilakukan Inverse-Transform

Dengan menggunakan jumlah cluster sebanyak 3, diperoleh hasil interpretasi sebagai berikut:

cluster	Avg. Danceability	Avg. Energy	Avg. Valence	Avg. Acousticness	Avg. Tempo	Analisis
0	1.913.352	262.712	394.353	1.210.468	1.770.709	Lagu dengan energi rendah dan acousticness sedang, cocok untuk suasana santai. Danceability cukup tinggi, kemungkinan besar lagu chill atau mellow.
1	1.828.356	276.220	403.901	1.126.620	8.409.257	Tempo sangat tinggi dan acousticness rendah, mengarah ke lagu elektronik atau dance cepat seperti EDM/techno.
2	2.106.279	276.511	410.475	1.133.233	5.077.461	Danceability paling tinggi, valence dan energy tinggi juga. Cocok untuk lagu-lagu pop upbeat yang ceria dan aktif.

Gambar 15. Tabel Interpretasi Hasil Cluster

5. Kesimpulan

Dari hasil penelitian yang dilakukan, metode K-Means berhasil mengelompokkan lagu-lagu berdasarkan kemiripan fitur audio. Proses clustering dilakukan dengan menggunakan jumlah cluster (k) sebanyak 3, kemudian menghasilkan silhouette score sebesar 0,435. Angka ini menunjukkan bahwa pemisahan antar kelompok cukup baik dan stabil. Setiap cluster yang terbentuk memiliki karakteristik yang berbeda, terutama dari segi energi, tempo, dan

danceability, sehingga dapat membantu pengguna dalam memilih musik sesuai suasana hati atau preferensi pribadi. Sebagai saran, penelitian ini dapat dikembangkan lebih lanjut untuk hasil clustering yang lebih akurat dan representatif, dengan menggunakan algoritma alternatif seperti DBSCAN atau Hierarchical Clustering, yang mungkin lebih baik dalam mengidentifikasi pola-pola tersembunyi dalam data musik.

Referensi

- [1] N. R. FITRADHI, M. Firman Hidayat, T. W. Saputro, M. G. Alifian, and A. P. Sari, "Laporan Final Project Machine Learning Rekomendasi Musik Spotify Menggunakan Metode K-Means," *Semin. Nas. Inform. Bela Negara*, vol. 3, pp. 2747–0563, 2023.
- [2] H. Tannady, J. Fernandes, A. Honni, and S. Lee, "Analisis Big Data Spotify dengan Metode Data Mining Analysis of Spotify Big Data with Data Mining Method," vol. 7, no. 2, pp. 52–59, 2024.
- [3] F. S. Farhani and A. Qoiriah, "Klasterisasi Musik Menggunakan K-Means," *JINACS ((Journal Informatics Comput. Sci.)*, vol. 06, pp. 566–572, 2024.
- [4] R. Maringka, A. Khoirunnita, R. Maringka, and E. Utami, "Analisa Perkembangan Musik Pada Spotify Menggunakan Structured Query Language (SQL) Analysis Of The Evolution Of Music On Spotify Using Structured Query Language (SQL)," *Cogito Smart J.*, vol. 7, no. 1, pp. 1–14, 2021.
- [5] P. Issn, A. R. Zaidah, C. I. Septiarani, S. Nisa, and N. Wahyudi, "Komparasi Algoritma K-Means , K-Medoid , Agglomeartive Clustering Terhadap Genre Spotify," *J. Ilm. Ilmu Komput. Fak. Ilmu Komput. Univ. Al Asyariah Mandar*, vol. 7, no. 1, pp. 49–54, 2021.
- [6] I. P. B. W. Brata and I. D. M. B. A. Darmawan, "Mood Classification of Balinese Songs with the K-Means Clustering Method Based on the Audio-Content Feature," *JELIKU (Jurnal Elektron. Ilmu Komput. Udayana)*, vol. 9, no. 3, p. 331, 2021, doi: 10.24843/jlk.2021.v09.i03.p03.
- [7] U. L. Musyarofah, S. N. Alima, and D. S. Y. Kartika, "Klasifikasi Top 50 Spotify Tahun 2010-2019 Menggunakan Metode K-Means Clustering," *Pros. Semin. Nas. Teknol. dan Sist. Inf.*, vol. 2, no. 1, pp. 215–220, 2022, doi: 10.33005/sitasi.v2i1.300.
- [8] A. V. Muhammda Daffa Rachman, "Implementasi Algoritma K-Means Dalam Sistem Rekomendasi Musik Menggunakan Python.pdf," *JATI (Jurnal Mhs. Tek. Inform.,* vol. 8, no. 3, pp. 3857–3862, 2024.
- [9] L. Nurhalimah, T. I. Hermanto, and I. Kaniawulan, "Analisis Prediksi Mood Genre Musik Pop Menggunakan Algoritma K-Means dan C4.5," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 4, p. 1006, 2022, doi: 10.30865/jurikom.v9i4.4597.

- [10] M. Ikhsan Firmansyah, R. Saepul Rohman, and E. Marsusanti, “Penerapan Algoritma Klastering K-Means Untuk Fitur Atribut Pada Layanan Streaming Musik Spotify,” *Indones. J. Comput. Sci.*, vol. 2, no. 2, pp. 58–66, 2023, doi: 10.31294/ijcs.v2i2.2465.
- [11] N. A. Hidayatullah and W. Prihartono, “Clustering Algoritma K-Means Untuk Meningkatkan Efektivitas Program Sosial Di Kota / Kabupaten Cirebon,” *JITET (Jurnal Inform. dan Tek. Elektro Ter.)*, vol. 13, no. 1, pp. 629–636, 2025.
- [12] I. Fitrianti, A. Voutama, and Y. Umaidah, “Clustering Film Populer Pada Aplikasi Netflix Dengan Menggunakan Algoritma K-Means Dan Metode CRISP-DM Clustering Popular Movies on Netflix App Using K-Means Algorithm and CRISP-DM Method,” *Jtsi*, vol. 4, no. 2, pp. 301–311, 2023.
- [13] C. Syaefulloh and D. G. Hutabarat, Rafida NL, “Optimalkan Peran Dinas Sosial Melalui Penggunaan Python Untuk Data Mining Pusdatin Kemensos Dalam Pelaksanaan Validasi Dan Verifikasi DtkS Di Provinsi Maluku Pendahuluan Kesejahteraan sosial adalah keadaan saat manusia merasa bahagia , aman serta terpenuhi,” *Explor. (Jurnal Inform. dan Komputer)*, vol. 12, no. 2, pp. 156–168, 2022.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>)

