

Perbandingan Metode K-Means dan Hierarchical Clustering pada Rekomendasi Musik Berbasis Audio Spotify Features Sistem

Farrel Reyhan Putra¹, Samuel Mahesa Sinulingga², Andhika Dwi Rachmawanto³, Michael Jeconiah Yonathan⁴, Valentino Wijaya⁵, Vitri Tundjungsari⁶

^{1,2,3,4,5,6}Universitas Esa Unggul

Jl Arjuna Utara No.9, Kebon Jeruk, Jakarta, Indonesia

e-mail: ¹sunmgaming@student.esaunggul.ac.id, ²sinulingga848@student.esaunggul.ac.id,
³dwirahmawanto39@student.esaunggul.ac.id, ⁴michaeljeconiah09@student.esaunggul.ac.id,
⁵valentinowijaya30@student.esaunggul.ac.id, ⁶vitri.tundjungsari@esaunggul.ac.id

Artikel Info : Diterima : 21-01-2026 | Direvisi : 00-00-0000 | Disetujui : 23-01-2026

Abstrak - Penelitian ini membandingkan metode K-Means dan Hierarchical Clustering dalam sistem rekomendasi musik berbasis audio features Spotify. Dataset yang digunakan berasal dari Spotify Tracks Dataset yang terdiri dari sekitar 114.000 lagu, kemudian melalui tahap pra-pemrosesan diperoleh sekitar 81.000 lagu valid. Untuk efisiensi komputasi, digunakan 5.000 lagu sebagai data eksperimen. Clustering dilakukan menggunakan 6 cluster dengan sembilan atribut audio. Evaluasi menggunakan Silhouette Score dan Davies–Bouldin Index menunjukkan bahwa K-Means memperoleh nilai Silhouette Score 0,1900 dan Davies–Bouldin Index 1,4445, sedangkan Hierarchical Clustering memperoleh nilai Silhouette Score 0,1782 dan Davies–Bouldin Index 1,4522. Hasil ini menunjukkan bahwa K-Means menghasilkan cluster yang lebih kompak. Sistem rekomendasi yang dibangun mampu memberikan rekomendasi lagu yang relevan berdasarkan kemiripan karakteristik audio.

Kata Kunci : Sistem Rekomendasi Musik, K-Means Clustering, Hierarkis Clustering

Abstracts - This study compares the K-Means and Hierarchical Clustering methods in a music recommendation system based on Spotify audio features. The dataset used comes from the Spotify Tracks Dataset, which consists of approximately 114,000 songs, and after the preprocessing stage, about 81,000 valid songs were obtained. For computational efficiency, 5,000 songs were used as experimental data. Clustering was performed using 6 clusters with nine audio attributes. Evaluation using the Silhouette Score and Davies–Bouldin Index shows that K-Means achieved a Silhouette Score of 0.1900 and a Davies–Bouldin Index of 1.4445, while Hierarchical Clustering achieved a Silhouette Score of 0.1782 and a Davies–Bouldin Index of 1.4522. These results indicate that K-Means produces more compact clusters. The developed recommendation system is able to provide relevant song recommendations based on the similarity of audio characteristics.

Keywords : Music Recommendation System, K-Means Clustering, Hierarchical Clustering

PENDAHULUAN

Perkembangan layanan musik digital seperti Spotify menghasilkan banyak data musik dengan variasi audio yang berbeda (Marlia, Setiawan, & Juliane, 2024; Damayanti, Fajriana, Meilani, & Fatimah, 2024). Kondisi ini mendorong kebutuhan akan sistem rekomendasi musik. Sistem ini dapat membantu pengguna menemukan lagu yang sesuai dengan selera mereka dengan cepat dan efisien (Fitradhi, Hidayat, Saputro, Alifiant, & Sari, 2023; Farhani & Qoiriah, 2024).. Salah satu pendekatan yang umum digunakan dalam sistem rekomendasi adalah pendekatan berbasis konten. Pendekatan ini menggunakan karakteristik bawaan lagu dalam bentuk fitur audio (Munajad, Ridwan, & Pratama, 2025; Mukhopadhyay, Kumar, Parashar, & Singh, 2024). Spotify menawarkan berbagai fitur audio seperti *danceability*, *energy*, *tempo*, *valence*, dan *acousticness*. Fitur-fitur ini dapat menggambarkan karakteristik musik secara kuantitatif (Safitri & Ningsih, 2025; Branchris, Yech, & Wijaya, 2026). Data ini memungkinkan penerapan metode pembelajaran tanpa pengawasan, khususnya pengelompokan, untuk mengelompokkan lagu berdasarkan kesamaan karakteristik audio tanpa harus memakai label genre yang jelas (Tundjungsari, 2024; Firmansyah, Rohman, & Marsusanti, 2023). Pendekatan ini tepat untuk membangun sistem rekomendasi musik yang fleksibel dan tidak bergantung pada preferensi pengguna sebelumnya. (Hamid, Abdullah, Pradana, Purnama, 2025).. Metode *K-Means* dan *Hierarchical Clustering* adalah dua teknik



pengelompokan yang memiliki pendekatan berbeda (Sudirwo et al., 2024). K-Means mengelompokkan data berdasarkan titik pusat atau centroid. Metode ini efisien untuk dataset besar (Tundjungsari, 2024; Rachman & Voutama, 2024). Sementara itu, *Hierarchical Clustering* membangun struktur klaster berlapis yang menunjukkan hubungan antar data (Zaidah, Septiarani, Nisa, Yusuf, & Wahyudi, 2021). Perbedaan ini membuatnya menarik untuk dibandingkan dalam sistem rekomendasi musik yang menggunakan fitur audio (Zaidah et al., 2021; Mukhopadhyay et al., 2024). Kualitas pengelompokan sangat penting, namun pemahaman tentang sistem rekomendasi juga tidak kalah penting. Penelitian ini mengevaluasi pengelompokan secara kuantitatif dengan menggunakan *Silhouette Score* dan *Davies-Bouldin Index*. (Rohman & Wibowo, 2024; Nuriska, Irawan, Bahtiar, & Dikananda, 2023). Evaluasi ini dilengkapi dengan visualisasi data. Hal ini membantu kita mengenali karakteristik unik dari setiap klaster dengan lebih mudah. (Dharmasena & Pramatha, 2024). Dengan pendekatan ini, sistem rekomendasi diharapkan dapat memberikan pilihan lagu yang tidak hanya relevan, tetapi juga sesuai dan mudah dipahami oleh pengguna.

METODE PENELITIAN

1. K-Means

K-Means adalah metode clustering non-hierarkis yang mengelompokkan data ke dalam sejumlah klaster berdasarkan jarak antara fitur (Tundjungsari, 2024), (Sudirwo et al., 2024). Setiap data hanya bisa menjadi anggota satu klaster. Ini disebut hard clustering (Firmansyah et al., 2023). Tujuan utama K-Means adalah meminimalkan jarak dalam klaster. Data dikelompokkan ke dalam klaster yang memiliki pusat terdekat, yang disebut centroid (Tundjungsari, 2024).

Proses *K-Means* dimulai dengan menginisialisasi centroid secara acak. Selanjutnya, langkah-langkah iteratif dimulai. Ini mencakup penugasan data ke klaster terdekat dan pembaruan posisi centroid sampai kondisi konvergen tercapai (Branchris et al., 2026; Rachman & Voutama, 2024). Dalam penelitian ini, metode *K-Means* diterapkan pada data yang sudah dinormalisasi untuk memastikan setiap fitur memberi kontribusi yang seimbang dalam perhitungan jarak (Al Aini, Endynda, & Al Rosyid, 2026).

Metode ini unggul karena kesederhanaan dan kecepatan komputasinya, sehingga cocok untuk dataset besar (Rachman & Voutama, 2024). Namun, *K-Means* memiliki batasan saat memodelkan data dengan distribusi kompleks. Metode ini juga sangat terpengaruh oleh pemilihan centroid awal, terutama pada data audio yang memiliki karakteristik non-linear (Sudirwo et al., 2024).

2. Hierarchical Clustering

Hierarchical Clustering adalah metode pembelajaran tanpa pengawasan yang mengelompokkan data dengan membangun struktur hierarkis berdasarkan tingkat kemiripan antar data (Sudirwo et al., 2024). Metode ini tidak melakukan pengelompokan langsung. Sebaliknya, ia membentuk klaster secara bertahap, sehingga hubungan antar data dapat dianalisis lebih mendalam (Zaidah et al., 2021).

Penelitian ini menggunakan pendekatan *agglomerative hierarchical clustering*. Di sini, setiap data pada awalnya dianggap sebagai satu klaster, yang kemudian secara bertahap digabungkan sampai mencapai jumlah klaster yang diinginkan (Sudirwo et al., 2024). Proses penggabungan klaster ditentukan dengan metode Ward linkage. Tujuannya adalah untuk meminimalkan peningkatan varians dalam klaster agar klaster yang dihasilkan menjadi lebih kompak (Zaidah et al., 2021).

Hierarchical Clustering memiliki keunggulan dalam menggambarkan struktur hubungan antar data tanpa memerlukan inisialisasi centroid (Sudirwo et al., 2024). Namun, metode ini juga memiliki keterbatasan dalam efisiensi komputasi ketika diterapkan pada dataset besar (Tundjungsari, 2024). Oleh karena itu, metode ini digunakan sebagai pembandingan terhadap *K-Means* dalam sistem rekomendasi musik berbasis fitur audio *Spotify* (Mukhopadhyay et al., 2024).

3. Sumber Data

Data yang digunakan dalam penelitian ini berasal dari *Spotify Tracks* Dataset yang diperoleh melalui platform Kaggle. Dataset ini berisi sekitar 114.000 lagu Spotify, dan setiap lagu dilengkapi dengan berbagai fitur audio yang diekstraksi dari *Spotify Web API*. Fitur audio yang digunakan dalam penelitian ini antara lain *danceability*, *energy*, *loudness*, *speechiness*, *acousticness*, *instrumentalness*, *liveness*, *valence*, dan *tempo*. Fitur-fitur tersebut merepresentasikan karakteristik audio setiap lagu secara kuantitatif (Safitri & Ningsih, 2025; Farhani & Qoiriah, 2024). Definisi dari masing-masing atribut audio disajikan pada Tabel 1.

Sumber: <https://www.kaggle.com/datasets/maharshipandya/-spotify-tracks-dataset>

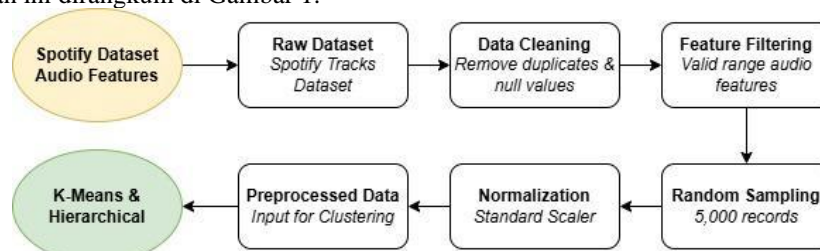
Tabel 1. Definisi Atribut Audio Spotify

No	Atribut	Deskripsi
1	danceability	Menggambarkan tingkat kecocokan sebuah lagu untuk menari berdasarkan kombinasi tempo, kestabilan ritme, kekuatan beat, dan regularitas keseluruhan. Nilai berkisar antara 0,0 hingga 1,0, di mana nilai yang lebih tinggi menunjukkan lagu lebih mudah untuk menari.
2	energy	Merepresentasikan intensitas dan tingkat aktivitas musik secara perseptual. Lagu dengan nilai energi tinggi umumnya terdengar cepat, keras, dan dinamis, sedangkan nilai rendah menunjukkan lagu yang lebih tenang dan lembut.
3	loudness	Menunjukkan tingkat kekerasan suara keseluruhan dari sebuah lagu yang diukur dalam satuan desibel (dB). Nilai ini merepresentasikan persepsi volume rata-rata sepanjang durasi lagu.
4	speechiness	Mengindikasikan keberadaan unsur ucapan dalam sebuah lagu. Nilai tinggi menunjukkan dominasi elemen berbicara seperti pada rap atau podcast, sedangkan nilai rendah menunjukkan dominasi unsur musik.
5	acousticness	Mengukur tingkat kepercayaan bahwa sebuah lagu bersifat akustik. Nilai mendekati 1,0 menunjukkan kemungkinan tinggi bahwa lagu menggunakan instrumen akustik tanpa banyak pemrosesan elektronik.
6	instrumentalness	Menunjukkan probabilitas bahwa sebuah lagu tidak mengandung vokal. Nilai yang tinggi mengindikasikan lagu bersifat instrumental, sedangkan nilai rendah menunjukkan keberadaan vokal yang dominan.
7	liveness	Mengindikasikan kemungkinan keberadaan audiens atau suasana pertunjukan langsung dalam rekaman lagu. Nilai tinggi menunjukkan kemungkinan besar lagu direkam secara live.
8	valence	Merepresentasikan tingkat emosi positif yang disampaikan oleh sebuah lagu. Lagu dengan nilai valence tinggi cenderung terdengar ceria dan positif, sedangkan nilai rendah mencerminkan nuansa sedih atau emosional.
9	tempo	Menunjukkan kecepatan dasar lagu yang diukur dalam satuan beats per minute (BPM), yang merepresentasikan seberapa cepat atau lambat sebuah komposisi musik.

Setelah melalui tahap pra-pemrosesan mulai dari menghapus data duplikat, menangani nilai yang kosong, hingga memvalidasi rentang fitur audio terkumpul sekitar 81.000 lagu yang layak diolah. Untuk mengoptimalkan kinerja komputasi sekaligus menjaga objektivitas dalam membandingkan metode, saya melakukan pengambilan sampel secara strategis. Hasilnya, 5.000 lagu terpilih dan digunakan sebagai dasar untuk tahap pengelompokan, evaluasi, dan pengembangan sistem rekomendasi musik.

4. Preprocessing Data

Tahapan ini bertujuan untuk mengubah data mentah menjadi format yang siap digunakan dalam pemodelan. Proses dimulai dengan membersihkan data duplikat dan data yang memiliki nilai hilang, kemudian dilanjutkan dengan menyaring nilai fitur agar berada dalam rentang yang wajar. Mengingat kompleksitas algoritma *Hierarchical Clustering*, dilakukan pengambilan sampel acak sebanyak 5.000 baris data untuk meningkatkan efisiensi komputasi tanpa mengurangi signifikansi statistik. Selanjutnya, seluruh fitur numerik dinormalisasi menggunakan Standard Scaler untuk memastikan kesetaraan skala antar fitur (Farhani & Qoiriah, 2024). Alur kerja pada tahapan ini dirangkum di Gambar 1.



Gambar 1. Alur Preprocessing Data

5. Evaluasi Clustering

Evaluasi model dilakukan dengan pendekatan validasi internal karena dataset yang digunakan tidak memiliki label atau bersifat tidak terawasi (Shulhiyana & Voutama, 2025). Untuk mengukur kualitas kluster secara kuantitatif, penelitian ini menggunakan dua metrik statistik utama. Metrik tersebut adalah *Silhouette Score* dan *Davies Bouldin Index* (DBI) (Firmansyah et al., 2023). *Silhouette Score* mengukur tingkat kepadatan dan keterpisahan antar kluster, dengan nilai ideal mendekati +1 (Rohman & Wibowo, 2024). Sementara itu, DBI mengevaluasi rasio penyebaran dalam kluster terhadap jarak antar kluster; nilai yang lebih rendah menunjukkan struktur kluster yang lebih baik (Nuriska et al., 2023).

Untuk melengkapi evaluasi numerik, analisis validitas visual dilakukan menggunakan *Principal Component Analysis* (PCA) (Dharmasena & Pramarta, 2024). Karena dataset terdiri dari fitur audio dengan dimensi tinggi, PCA digunakan untuk mengurangi dimensi data menjadi dua komponen utama yang merepresentasikan variasi terbesar (Rachman & Voutama, 2024). Transformasi ini memungkinkan pemetaan distribusi seluruh lagu ke dalam ruang dua dimensi. Ini memudahkan analisis pemisahan antara kluster dan posisi relatif lagu masukan secara spasial.

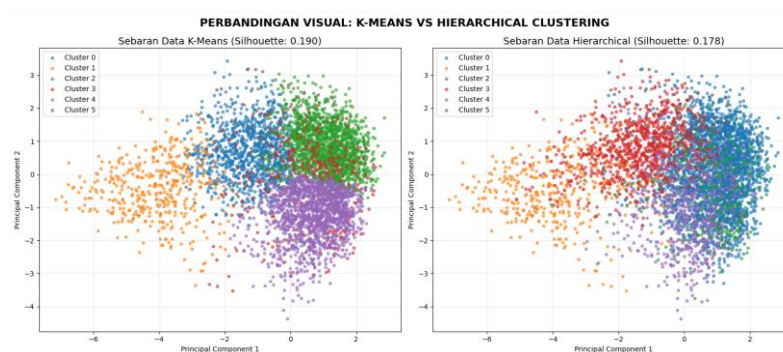
6. Tools dan Bahasa Pemrograman

Penelitian ini dikembangkan menggunakan bahasa pemrograman Python sebagai fondasi utamanya. Untuk manipulasi data matriks yang efisien, saya mengandalkan pustaka Pandas dan NumPy. Seluruh proses pemodelan algoritma clustering, pra-pemrosesan, hingga evaluasi kinerja dijalankan melalui ekosistem Scikit-Learn. Sementara itu, untuk mempermudah analisis data, hasil pengelompokan divisualisasikan menggunakan Matplotlib dalam bentuk Radar Chart dan pemetaan PCA guna memberikan gambaran struktur data yang lebih jelas.

HASIL DAN PEMBAHASAN

1. Hasil Clustering

Dalam penelitian ini, kami menggunakan metode clustering hierarkis dan K-Means. Kedua metode ini memiliki jumlah cluster yang sama, yaitu enam cluster. Visualisasi dilakukan dengan menggunakan analisis komponen utama (PCA) dua dimensi untuk mempermudah interpretasi hasil.



Gambar 2. Perbandingan visual hasil clustering menggunakan K-Means dan Hierarchical Clustering berdasarkan proyeksi PCA

Gambar 2 menunjukkan perbandingan visual hasil clustering hierarkis dan K-Means dengan proyeksi dua dimensi menggunakan Principal Component Analysis (PCA). Pola sebaran data dan pemisahan antara cluster yang dihasilkan oleh masing-masing metode terlihat dalam visualisasi ini.

Dari visualisasi di sisi kiri gambar, hasil clustering dengan metode K-Means menunjukkan bahwa sebagian besar cluster terbentuk lebih kompak. Batas antar cluster terlihat jelas. Ini menunjukkan bahwa K-Means dapat mengelompokkan lagu berdasarkan kesamaan fitur audio secara konsisten.

Di sisi lain, visualisasi di kanan gambar menunjukkan hasil *clustering* menggunakan *Hierarchical Clustering*. Di sini, sebaran cluster terlihat saling beririsan. Beberapa cluster memiliki area tumpang tindih yang besar, sehingga pemisahan antar cluster tidak sejelas pada metode *K-Means*.

Secara keseluruhan, perbandingan visual ini menunjukkan bahwa metode *K-Means* menghasilkan struktur cluster yang lebih kompak dan terpisah dibandingkan dengan *Hierarchical Clustering* pada dataset yang digunakan dalam penelitian ini.

2. Evaluasi Kualitas Clustering

Untuk melihat seberapa baik metode clustering hierarchial dan K-Means bekerja, kita mengevaluasi kualitas clustering. *Silhouette Score* dan *Davies Bouldin Index* adalah metrik evaluasi internal yang digunakan untuk menilai kualitas clustering. *Silhouette Score* mengukur tingkat kemiripan data dalam suatu kluster dibandingkan dengan kluster lain. Nilai yang lebih rendah menunjukkan kualitas clustering yang lebih baik.

Tabel 2. Hasil Evaluasi Kualitas Clustering

Metode	Silhouette Score	Davies-Bouldin Index
K-Means	0.1900	1.4445
Hierarchical	0.1782	1.4522

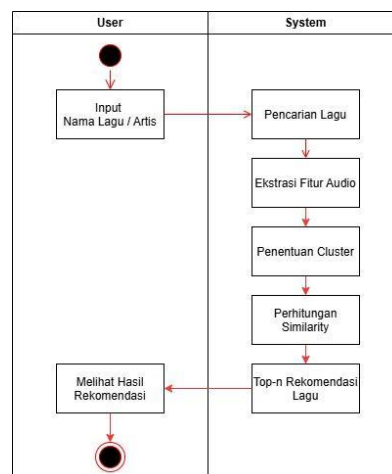
Untuk menguji efektivitas antara metode *Hierarchical Clustering* dan *K-Means*, evaluasi dilakukan menggunakan dua metrik internal utama: *Silhouette Score* dan *Davies-Bouldin Index*. Hasil evaluasi pada Tabel 2 menunjukkan bahwa *K-Means* sedikit lebih unggul dibandingkan metode lainnya. Capaian *Silhouette Score* yang lebih tinggi ini menunjukkan bahwa data dalam tiap kluster memiliki tingkat kepadatan atau kohesi yang lebih solid, mengindikasikan bahwa anggota kluster tersebut saling berkaitan erat satu sama lain (Rachman & Voutama, 2024; Al Aini et al., 2026).

Kinerja unggul *K-Means* juga diperkuat oleh nilai *Davies-Bouldin Index* yang lebih rendah, yang mencerminkan pemisahan antar kluster yang lebih tegas dan optimal (Firmansyah et al., 2023). Secara garis besar, meski selisih angkanya tidak terpaut jauh, *K-Means* terbukti memberikan hasil sedikit lebih baik dibandingkan *Hierarchical Clustering* dalam memproses *Spotify Tracks Dataset* pada penelitian ini.

3. Sistem Rekomendasi Musik Berbasis Clustering

Sistem rekomendasi musik dalam penelitian ini dibangun dengan pendekatan content-based filtering. Sistem ini memanfaatkan hasil clustering berdasarkan fitur audio. Setiap lagu dalam dataset telah dikelompokkan ke dalam kluster tertentu menggunakan metode K-Means dan Hierarchical Clustering. Proses rekomendasi difokuskan pada lagu-lagu yang memiliki karakteristik audio yang serupa.

Alur kerja sistem rekomendasi dimulai dari masukan pengguna yang berupa judul lagu atau nama artis. Sistem kemudian mencari lagu yang sesuai dalam dataset dan menentukan kluster tempat lagu tersebut berada. Selanjutnya, sistem menghitung tingkat kemiripan antara lagu masukan dan lagu-lagu lain dalam kluster yang sama. Ini menggunakan metode cosine similarity berdasarkan fitur audio yang telah dinormalisasi. Lagu-lagu dengan nilai kemiripan tertinggi kemudian dipilih sebagai rekomendasi dan disajikan kepada pengguna (Hamid et al., 2025).



Gambar 3. Alur sistem rekomendasi musik berbasis clustering

Alur lengkap sistem rekomendasi musik berbasis clustering yang digunakan dalam penelitian ini ditunjukkan pada Gambar 3. Dengan menggunakan hasil clustering, ruang pencarian rekomendasi menjadi lebih terbatas dan relevan. Sistem ini dapat menghasilkan rekomendasi lagu yang lebih terarah dan efisien tanpa bergantung pada data preferensi pengguna lain..

4. Studi Kasus Rekomendasi Lagu

Untuk menguji kinerja sistem rekomendasi musik yang dikembangkan, kami melakukan studi kasus dengan lagu dari artis Skillet, yaitu *Whispers in the Dark*. Lagu ini dipilih sebagai contoh dari lagu bergenre rock yang

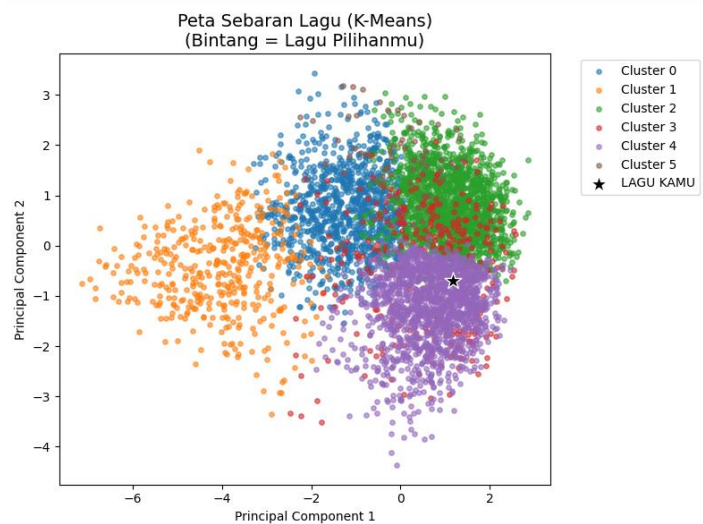
memiliki energi tinggi dan nuansa emosional yang kuat.

Tabel 3. Hasil Rekomendasi Lagu untuk “Whispers in the Dark”

Metode Clustering	Cluster	Lagu Rekomendasi	Artis	Similarity (%)
K-Means	4	Black Supremacy	Samael	94.5
K-Means	4	No Hemos Aprendido Nada	Viva Suecia	93.9
K-Means	4	COLLIDE	Emma Hewitt; Ben Nicky	93.8
K-Means	4	Dawn Of The Dead (Radio Edit)	Soltan; Shaquille O’Neal	93.5
K-Means	4	Feeling This Way	Ray Volpe	93.2
Hierarchical	0	Mann Lagyo Mero Yaar Fakiri Mein	Kabir Cafe	95.2
Hierarchical	0	Bodies	Drowning Pool	94.2
Hierarchical	0	Duş	Teoman	93.3
Hierarchical	0	Ruhum Bela	Can Bonomo	93.2
Hierarchical	0	Earth Is A Black Hole	Teenage Wrist	91.5

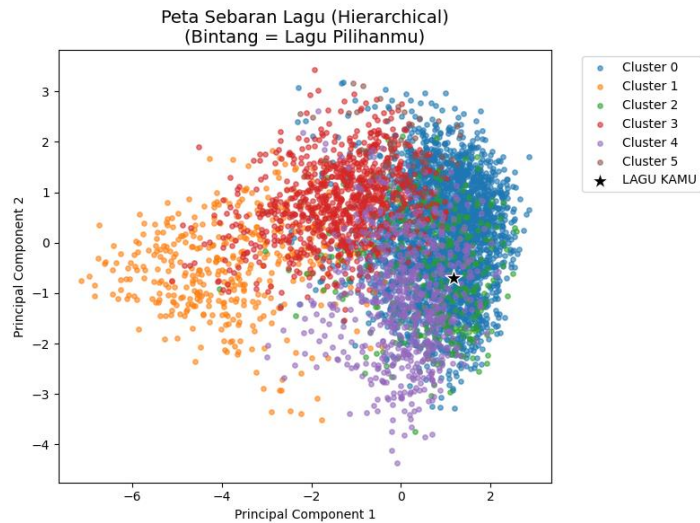
Berdasarkan hasil yang terangkum pada Tabel 3, sistem memberikan rekomendasi lagu yang cukup menarik dari kedua metode pengelompokan yang diuji. Pada model *K-Means*, lagu *Whispers in the Dark* masuk ke dalam *Cluster 4*, di mana lagu-lagu rekomendasinya didominasi oleh karakteristik energi tinggi dan intensitas musik yang sangat kuat.

Sementara itu, pendekatan *Hierarchical Clustering* menempatkan lagu tersebut di *Cluster 0*. Menariknya, meskipun strukturnya berbeda, rekomendasi yang dihasilkan tetap memiliki nuansa agresif dan dinamika yang sangat mirip. Hal ini membuktikan bahwa baik *K-Means* maupun *Hierarchical Clustering* sama-sama efektif dalam menyajikan rekomendasi yang relevan secara audio, terlepas dari perbedaan cara mereka membentuk kluster di balik layar.



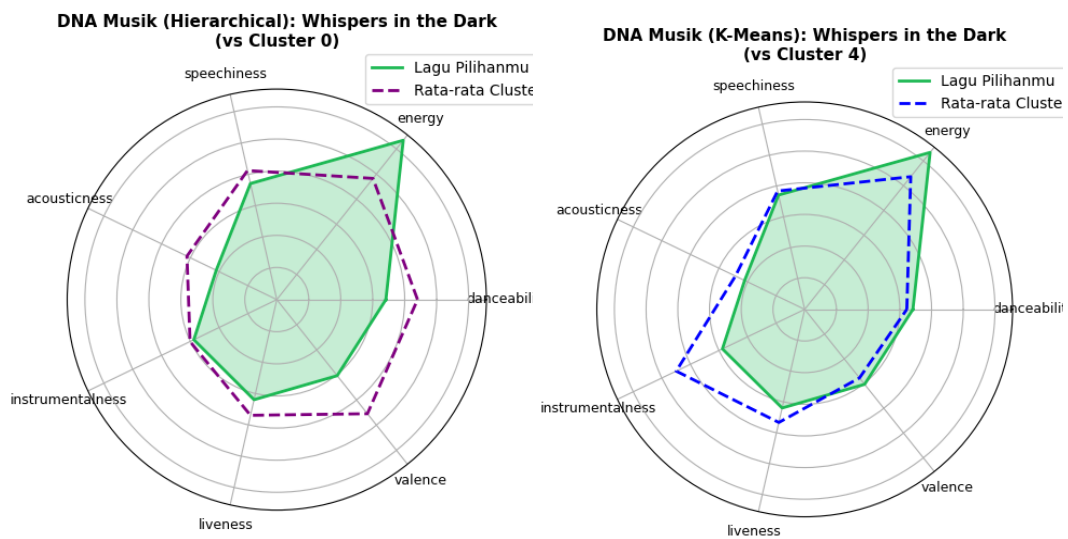
Gambar 4. PCA K-Means (Skillet)

Melalui metode *K-Means*, lagu *Whispers in the Dark* diklasifikasikan ke dalam *Cluster 4*. Menariknya, jika kita melihat visualisasi PCA pada Gambar 4, posisi lagu ini terletak pada koordinat kluster yang padat, yang menunjukkan adanya kemiripan karakteristik yang kuat dengan lagu-lagu lain di dalam kelompok tersebut. Ini menandakan adanya kemiripan karakteristik audio dengan lagu-lagu lain dalam cluster tersebut. Hasil rekomendasi *K-Means* didominasi oleh lagu-lagu dengan intensitas tinggi dan karakter musik agresif, seperti *Black Supremacy* dan *COLLIDE*, yang memiliki tingkat kemiripan audio di atas 93%.



Gambar 5. PCA Hierarchical (Skillet)

Sebagai perbandingan, hasil pengelompokan dengan metode Hierarchical Clustering menempatkan lagu yang sama ke dalam Cluster 0, seperti yang ditunjukkan pada Gambar 5. Dengan metode ini, sebaran cluster terlihat lebih tumpang tindih. Namun, sistem masih dapat memberikan rekomendasi lagu yang relevan, seperti Bodies dan Earth Is A Black Hole, yang memiliki karakteristik energi dan dinamika musik yang mirip.



Gambar 6. Radar Chart K-Means & Hierarchical

Guna memperkuat interpretasi terhadap hasil rekomendasi, kami menyajikan visualisasi grafik radar yang membandingkan profil fitur audio lagu masukan dengan rata-rata fitur dalam klaster. Jika mengamati Gambar 7, terlihat bahwa lagu *Whispers in the Dark* memiliki pola fitur yang sangat selaras dengan rata-rata klaster pada kedua metode, terutama untuk atribut *energy*, *speechiness*, dan *liveness*. Hal ini menegaskan bahwa sistem tidak bekerja secara acak, melainkan memberikan rekomendasi berdasarkan kemiripan karakteristik audio yang terukur secara kuantitatif (Chinta & Fibriani, 2025).

Secara keseluruhan, studi kasus ini membuktikan bahwa sistem rekomendasi musik berbasis *clustering* mampu menyaring lagu-lagu yang relevan sesuai karakteristik audionya. Dalam hal ini, metode K-Means cenderung menghasilkan klaster yang lebih kompak dan konsisten. Di sisi lain, metode *Hierarchical Clustering* menawarkan fleksibilitas yang berbeda dalam memetakan struktur data, namun keduanya tetap berhasil menangkap esensi musik yang serupa.

KESIMPULAN

Penelitian ini membandingkan metode K-Means dan Hierarchical Clustering dalam sistem rekomendasi musik berbasis fitur audio Spotify. Hasil evaluasi dengan menggunakan Silhouette Score dan Davies-Bouldin Index menunjukkan bahwa metode K-Means menghasilkan kluster yang lebih kompak dan terpisah dengan baik dibandingkan Hierarchical Clustering. Sistem rekomendasi yang dikembangkan dapat memberikan saran lagu yang relevan berdasarkan kemiripan karakteristik audio. Hal ini didukung oleh studi kasus dan visualisasi radar chart. Oleh karena itu, K-Means dianggap lebih sesuai sebagai metode utama untuk sistem rekomendasi musik berdasarkan fitur audio dalam dataset yang digunakan dalam penelitian ini.

REFERENSI

- Al Aini, N. A., Endynda, R. C., & Al Rosyid, H. (2026). Penerapan algoritma K-Means untuk klusterisasi lagu terpopuler 2025 versi Spotify. *Jurnal Sains dan Teknologi*, 2(3), 147–157. Retrieved from <https://jurnal.globalscients.com/index.php/jsit/article/view/939>
- Branchris, Yech, K. A., & Wijaya, A. (2026). Klusterisasi lagu pada dataset Spotify berdasarkan fitur audio menggunakan algoritma K-Means. *Sinergi: Jurnal Ilmiah Multidisiplin*, 2(1). Retrieved from <https://publikasi.aahlkamal.com/index.php/sinergi>
- Chinta, C. A., & Fibriani, C. (2025). Analyzing Spotify user experience using UEQ and fuzzy K-Means clustering. *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, 14(2), 1174–1182. doi:10.35889/jutisi.v14i2.2928
- Damayanti, S. E., Fajriana, M. I., Meilani, D., & Fatimah, S. S. (2024). Menjelajahi wawasan industri musik: Klusterisasi lagu terpopuler di Spotify 2024 dengan metode K-Means clustering. In *Seminar Nasional CORISINDO* (p. 562).
- Dharmasena, K. B., & Pramatha, C. (2024). Segmentasi pengguna Spotify berdasarkan preferensi musik dengan algoritma K-Means clustering. *Jurnal Nasional Teknologi Informasi dan Aplikasinya*, 3(1), 37–42. doi:10.24843/JNATIA.2024.v03.i01.p05
- Farhani, F. S., & Qoiriah, A. (2024). Klusterisasi musik menggunakan K-Means. *JINACS: Journal of Informatics and Computer Science*, 6(2). Retrieved from <https://ejournal.unesa.ac.id/>
- Firmansyah, M. I., Rohman, R. S., & Marsusanti, E. (2023). Penerapan algoritma klustering K-Means untuk fitur atribut pada layanan streaming musik Spotify. *Indonesian Journal on Computer Science*, 2(2), 58–66. doi:10.31294/ijcs.v2i2.2465
- Fitradhi, N. R., Hidayat, M. F., Saputro, T. W., Alifiant, M. G., & Sari, A. P. (2023). Rekomendasi musik Spotify menggunakan metode K-Means. *Prosiding Seminar Nasional Informatika Bela Negara (SANTIKA)*, 3, 86–91. Retrieved from <https://santika.upnjatim.ac.id/submissions/index.php/santika/article/view/201>
- Hamid, A., Abdullah, S., Pradana, G. A., & Purnama, D. G. (2025). K-Means analysis for music segmentation in playlist recommendation. *International Conference on Digital, Social, and Science*, 2.
- Marlia, S., Setiawan, K., & Juliane, C. (2024). Analisis fitur musik dan tren popularitas lagu di Spotify menggunakan K-Means dan CRISP-DM. *SISTEMASI: Jurnal Sistem Informasi*, 13(2), 595–607. doi:10.32520/stmsi.v13i2.3757
- Mukhopadhyay, S., Kumar, A., Parashar, D., & Singh, M. (2024). Enhanced music recommendation systems: A comparative study of content-based filtering and K-Means clustering approaches. *Revue d'Intelligence Artificielle*, 38(1), 365–376. doi:10.18280/ria.380138
- Munajad, M. A., Ridwan, A., & Pratama, T. G. (2025). Pengembangan sistem rekomendasi musik dengan K-Means dan KNN berbasis cosine similarity. *Sainteks*, 22(2), 153–165. doi:10.30595/sainteks.v22i2.27815
- Nuriska, D., Irawan, B., Bahtiar, A., & Dikananda, A. R. (2023). Klusterisasi data lagu terpopuler Spotify 2023 berdasarkan suasana hati menggunakan algoritma K-Means. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(6). Retrieved from <https://ejournal.stmik-ikmi-cirebon.ac.id/>
- Rachman, M. D., & Voutama, A. (2024). Implementasi algoritma K-Means dalam sistem rekomendasi musik menggunakan Python. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(3). Retrieved from <https://ejournal.stmik-ikmi-cirebon.ac.id/>
- Rohman, N., & Wibowo, A. (2024). Clustering of popular Spotify songs in 2023 using K-Means method and silhouette coefficient. *Pilar Nusa Mandiri: Journal of Computing and Information System*, 20(1), 18–24. doi:10.33480/pilar.v20i1.4937
- Safitri, R. I., & Ningsih, S. (2025). Klusterisasi lagu pada platform Spotify berdasarkan fitur audio menggunakan algoritma K-Means dan K-Means++. *Jurnal Sistem Informasi Bisnis (JUNSIBI)*, 6(2), 247–257. doi:10.55122/junsibi.v6i2.1684

Shulhiyana, S., & Voutama, A. (2025). Penerapan metode K-Means clustering untuk pengelompokan musik berdasarkan karakteristik audio. *Jurnal Nasional Ilmu Komputer*, 6(2), 59–70. doi:10.47747/jurnalnik.v6i2.2751

Sudirwo, Hadi, A., Judijanto, L., Purwandari, N., Zain, N. N. L. E., Rambe, K. H., ... Yusufi, A. (2024). *Artificial intelligence: Teori, konsep, dan implementasi di berbagai bidang*. Jambi, Indonesia: PT. Sonpedia Publishing Indonesia.

Tundjungsari, V. (2024). *Dasar machine learning*. Yogyakarta, Indonesia: Deepublish.

Zaidah, A. R., Septiarani, C. I., Nisa, M. S., Yusuf, A., & Wahyudi, N. (2021). Komparasi algoritma K-Means, K-Medoid, agglomerative clustering terhadap genre Spotify. *Jurnal Ilmiah Ilmu Komputer*, 7(1). Retrieved from <http://ejournal.fikom-unasman.ac.id>