

IMPLEMENTASI ALGORITMA K-MEANS DALAM SISTEM REKOMENDASI MUSIK MENGGUNAKAN PYTHON

Muhammda Daffa Rachman, Apriade Voutama

Sistem Informasi, Universitas Singaperbangsa Karawang
Jl. HS.Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang, Indonesia
2110631250052@student.unsika.ac.id

ABSTRAK

Musik telah menjadi bagian integral kehidupan, dengan teknologi memfasilitasi akses mudah ke berbagai jenis musik. Layanan streaming menggunakan algoritma untuk merekomendasikan musik kepada pengguna. Artikel ini mengevaluasi efektivitas algoritma k-means dalam mengelompokkan preferensi musik berdasarkan fitur-fitur seperti valensi, akustik, danceability, dan lainnya. Data dari platform Kaggle diproses dan dinormalisasi sebelum dilatih dengan algoritma k-means. Metode Elbow digunakan untuk menentukan jumlah cluster optimal yang kemudian dievaluasi menggunakan Silhouette Score. Hasilnya menunjukkan bahwa clustering menggunakan algoritma k-means belum optimal, dengan potensi peningkatan melalui evaluasi lebih lanjut terhadap fitur-fitur dan jumlah cluster yang digunakan. Langkah-langkah ini diharapkan meningkatkan kualitas rekomendasi musik.

Kata kunci : Rekomendasi musik, algoritma K-means, sistem rekomendasi, Python, Silhouette Score

1. PENDAHULUAN

Musik adalah bagian dari kehidupan sehari-hari banyak individu di seluruh dunia. Dengan perkembangan teknologi informasi dan internet, akses terhadap berbagai jenis musik semakin mudah. Musik merupakan salah satu bentuk seni yang paling populer di dunia. Musik dapat memberikan hiburan, meningkatkan mood, dan bahkan membantu orang untuk bersantai. Dengan banyaknya pilihan musik yang tersedia, tidak heran jika banyak orang yang mencari rekomendasi musik yang sesuai dengan selera. Musik merupakan bagian integral dari kehidupan masyarakat di seluruh dunia, berfungsi sebagai seni yang memberikan rangsangan kepada jiwa dan mengungkapkan beragam emosi melalui lagu-lagu. Sebagai salah satu bentuk hiburan yang populer, musik memiliki berbagai genre yang beragam. Musik dapat ditemui di berbagai tempat, seperti dalam film, iklan, permainan video, dan bahkan sebagai nada sambung panggilan yang menggantikan nada panggilan standar. [1]

Kemajuan teknologi saat ini telah mengakibatkan perubahan yang signifikan, salah satunya adalah transformasi menuju bentuk digital dalam banyak hal. Era digital secara umum telah mengubah gaya hidup manusia menjadi lebih tergantung pada perangkat elektronik. Sistem teknologi digital membuat kehidupan manusia menjadi lebih efisien dan membantu, yang menandai transisi menuju era digital.[2]

Layanan seperti Spotify dan Apple Music memiliki puluhan juta pengguna dan menyediakan puluhan juta lagu. Untuk mengelola sistem sebesar ini, layanan streaming menggunakan berbagai teknologi, termasuk algoritma. Pendengar kontemporer tidak hanya mengandalkan antarmuka depan layanan untuk menemukan musik; semakin banyak, mereka juga mengandalkan saran dari layanan itu sendiri -

bergantung pada daftar putar yang dibuat oleh layanan atau stasiun radio untuk menemukan lagu-lagu individu. Penurunan dalam pentingnya album bukanlah hal baru: proses 'pemisahan' album menjadi trek tunggal dimulai dengan peluncuran iTunes. Yang baru adalah fakta bahwa lagu-lagu individu sekarang 'dipromosikan' kepada pendengar melalui daftar putar, daripada dicari secara aktif. Ini adalah bagian dari tren lebih luas, di mana algoritma telah menjadi 'logika kunci yang mengatur aliran informasi yang kita andalkan'. [3]

Dalam era digital, pertumbuhan bisnis telah mengalami kemajuan pesat. Kemajuan teknologi dan internet telah membuka peluang baru bagi bisnis untuk meraih pasar yang lebih luas, meningkatkan efisiensi operasional, dan mempercepat pertumbuhan bisnis secara keseluruhan. [4]Perkembangan AI dalam industri digital telah memberikan dampak yang signifikan pada berbagai bidang, termasuk dalam pengembangan sistem rekomendasi musik. Kecerdasan Buatan merupakan salah satu bidang penelitian yang paling mencuat dalam beberapa tahun terakhir. Dalam satu dekade, AI telah bergerak dari topik penelitian akademis yang sebagian besar dipelajari di departemen ilmu komputer, matematika, dan psikologi, menjadi insentif ilmiah dan bisnis global. AI menjadi alat yang meresap digunakan untuk memajukan pengetahuan dalam banyak bidang seperti fisika, ekonomi, genetika, atau ilmu sosial, di antara lain. Dan teknologi ini diterapkan di inti dari berbagai aplikasi yang digunakan oleh jutaan orang setiap hari. [5]Rekomendasi musik dapat diberikan dengan menggunakan berbagai metode, salah satunya adalah dengan menggunakan algoritma clustering. Algoritma clustering adalah metode data mining yang digunakan untuk mengelompokkan data menjadi beberapa kelompok yang memiliki kemiripan. Dalam konteks rekomendasi musik, algoritma k-means dikenal

sebagai salah satu pendekatan yang relatif sederhana untuk diterapkan dalam analisis data. Cara konsumen musik mengakses dan membeli musik dipengaruhi oleh pengenalan layanan musik online dan peningkatan kecepatan broadband.[6]

Penggunaan algoritma k-means pada sistem rekomendasi musik terletak pada kemampuannya dalam mengelompokkan data yang memiliki jumlah dimensi yang besar. Ini berarti algoritma ini dapat digunakan untuk menganalisis dataset yang kompleks dan besar, memungkinkan pengelompokan data dalam skala yang luas tanpa mengorbankan kinerja secara signifikan. K-means memperhitungkan variasi yang besar dalam data, sehingga cocok digunakan pada dataset dengan banyak atribut atau dimensi.

2. TINJAUAN PUSTAKA

Kecerdasan buatan umum mengacu pada kemampuan mesin untuk berkomunikasi, berpikir, dan beroperasi secara mandiri dalam situasi yang dikenal maupun baru dengan cara yang mirip dengan manusia. Hal ini masih jauh di luar jangkauan metode saat ini dan bukanlah yang dimaksudkan ketika istilah 'AI' umumnya digunakan. Sebagian besar referensi terhadap AI saat ini sering digunakan sebagai istilah yang dapat dipertukarkan dengan 'machine learning' atau 'deep learning'. [7]

Unsupervised learning adalah metode pembelajaran mesin dimana algoritma menggunakan data tanpa label atau data yang belum ditata untuk mencari hubungan, kaitan, atau kelompok yang tidak diketahui di antara data tersebut. Algoritma ini tidak menggunakan informasi yang diberikan oleh manusia untuk mengidentifikasi hubungan atau kelompok dalam data. Hasil akhir *unsupervised learning* bisa berupa label kelas (dalam klasifikasi) atau angka riil (dalam regresi). [8]

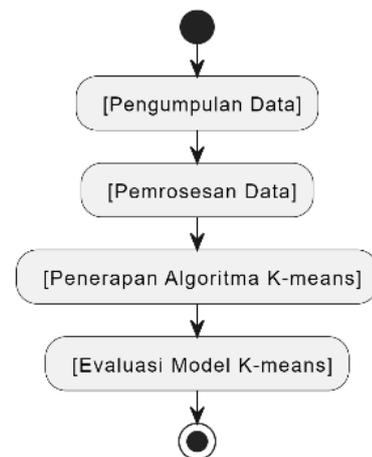
Algoritma K-Means adalah salah satu algoritma partisional karena bergantung pada penentuan jumlah awal kelompok dengan menetapkan nilai centroid awalnya. Ini menggunakan proses iteratif untuk menghasilkan cluster data. Jumlah cluster awal yang diinginkan diberikan sebagai masukan, dan jumlah cluster akhir dihasilkan sebagai keluaran. Ketika algoritma diminta untuk menghasilkan cluster K , akan ada K titik awal dan K titik akhir. K-Means memilih K pola awal centroid secara acak. Jumlah iterasi untuk mencapai centroid cluster dipengaruhi oleh posisi awal centroid yang dipilih secara acak; jika posisi baru centroid tidak berubah, nilai K yang dipilih dianggap sebagai pusat akhir. Ini dihitung menggunakan Euclidean Distance untuk menentukan jarak terdekat antara titik centroid dan data/objek. Data dengan jarak terdekat ke centroid membentuk sebuah cluster.[9]

Metode Elbow adalah teknik yang digunakan untuk menentukan jumlah cluster yang optimal dengan cara memeriksa perbandingan hasil varians antara jumlah cluster yang dibentuk. Metode ini memberikan ide dengan memilih jumlah cluster dan menambahnya secara bertahap untuk membentuk model data dalam

menentukan cluster terbaik. Persentase varians yang dihasilkan menjadi pembandingan antara jumlah cluster yang ditambah. Perbedaan persentase varians dari setiap jumlah cluster dapat ditampilkan melalui grafik sebagai sumber informasi. Jika perubahan varians antara jumlah cluster pertama dan kedua menunjukkan penurunan yang signifikan dalam grafik atau memiliki siku yang terbesar, maka jumlah cluster tersebut dianggap sebagai yang optimal.[9]

Python adalah salah satu dari banyak bahasa pemrograman. Seperti bahasa manusia, ada banyak bahasa komputer lainnya, seperti Java, LISP, PHP, dan Perl, dan... jangan lupa C atau bahasa lainnya, serta hal-hal berguna seperti scripting UNIX. Kebanyakan bahasa memiliki kelebihan masing-masing - sebagai contoh, menulis program yang mudah dipindahkan adalah keunggulan Java dan mengakses database serta memasukkannya ke dalam halaman web adalah keahlian khusus PHP. Tetapi pada dasarnya, semua bahasa ini sangat mirip pada tingkat konsep inti - kebanyakan memiliki data dalam variabel dan fungsi (prosedur, metode) untuk melakukan operasi pada data tersebut. Beberapa bahasa bahkan menggabungkan data dan fungsi menjadi paket yang disebut objek, dan yang lain seperti LISP memungkinkan Anda memperlakukan fungsi seperti variabel, dan sebaliknya. Python adalah bahasa pemrograman yang kuat dan elegan yang mudah dibaca dan dipahami. Ini menunjukkan sebagian besar fitur-fitur ini yang umum di banyak bahasa lain dan berguna untuk aplikasi dunia nyata.[10]

3. METODE PENELITIAN



Gambar 1. Flowchart penelitian

- Pengumpulan data : Tahap ini melibatkan proses pencarian dan pengumpulan data dari platform Kaggle.
- Pemrosesan data : Tahap ini melibatkan proses penting yaitu pembersihan dan transformasi data, di mana data-data yang telah dikumpulkan akan diolah dengan mengidentifikasi dan mengatasi

ketidakesesuaian, serta mengubahnya ke dalam format yang lebih sesuai untuk penelitian.

c. Penerapan algoritma k-means : Pada langkah ini, algoritma K-means diterapkan pada data preferensi musik yang telah diproses sebelumnya. Algoritma ini digunakan untuk mengelompokkan data ke dalam kluster berdasarkan kesamaan preferensi musik.

d. Evaluasi model K-means : Langkah ini bertujuan untuk mengevaluasi model K-means yang telah diimplementasikan. Evaluasi dilakukan untuk memastikan bahwa pengelompokan yang dihasilkan relevan dan berguna.

Berikut adalah definisi dari atribut-atribut yang ada dalam dataset :

Tabel 1. Definisi atribut dataset

Atribut	Penjelasan
valence	Menggambarkan tingkat positivitas atau kebahagiaan yang terkandung dalam lagu. Nilai valensi yang tinggi menunjukkan lagu yang lebih positif atau bahagia, sedangkan nilai valensi yang rendah menunjukkan lagu yang lebih sedih atau negatif.
year	Tahun rilis lagu.
acousticness	Seberapa besar proporsi suara dalam lagu yang dihasilkan dari instrumen akustik dibandingkan instrumen elektronik. Nilai mendekati 1 menunjukkan lagu yang sangat akustik, sementara nilai mendekati 0 menunjukkan lagu yang sangat elektronik.
artists	Nama-nama artis yang terlibat dalam pembuatan lagu.
danceability	Menggambarkan seberapa cocok lagu untuk menari berdasarkan tempo, ritme, stabilitas ritme, dan reguleritas umum lainnya. Lagu dengan danceability yang tinggi biasanya lebih mudah untuk digunakan sebagai musik dansa.
duration_ms	Durasi lagu dalam milidetik (ms).
energy	Tingkat energi dalam lagu. Lagu dengan energi yang tinggi cenderung terasa cepat, keras, dan intens.
explicit	Menunjukkan apakah lagu tersebut mengandung konten eksplisit yang mungkin tidak cocok untuk semua pendengar. Nilai 1 menunjukkan ada konten eksplisit, sedangkan nilai 0 menunjukkan tidak ada konten eksplisit.
id	ID unik untuk lagu dalam dataset.
instrumentalness	Menggambarkan seberapa besar lagu tersebut tidak mengandung vokal. Nilai mendekati 1 menunjukkan lagu yang hampir sepenuhnya instrumental, sedangkan nilai mendekati 0 menunjukkan lagu yang memiliki vokal yang dominan.
key	kunci musik yang digunakan dalam lagu tersebut, misalnya C major, D minor, dan sebagainya. Dalam dataset, mungkin direpresentasikan dengan nilai numerik atau string yang menunjukkan kunci musik.
loudness	Volume relatif dari lagu, diukur dalam desibel (dB).
mode	Menunjukkan apakah lagu tersebut dalam mode mayor atau minor. Nilai 1 menunjukkan mayor, sedangkan nilai 0 menunjukkan minor.
name	Nama lagu.
popularity	Tingkat popularitas lagu, mungkin berdasarkan berbagai faktor seperti jumlah streaming, unduhan, dll.
release_date	Tanggal rilis lagu.
speechiness	Menggambarkan seberapa banyak kata-kata atau "bicara" terdapat dalam lagu. Nilai mendekati 1 menunjukkan lagu yang lebih cenderung bersifat "bicara", seperti pidato atau rap, sedangkan nilai mendekati 0 menunjukkan lagu yang lebih cenderung bersifat musik.
tempo	Kecepatan ritme musik, diukur dalam ketukan per menit (bpm).
genres	kategori atau jenis tertentu dari musik yang memiliki karakteristik yang serupa. Genre musik digunakan untuk mengelompokkan musik berdasarkan berbagai elemen seperti ritme, melodi, instrumen yang digunakan, gaya vokal, dan tema lirik.

4. HASIL DAN PEMBAHASAN

Sample yang dipakai untuk penelitian ini terdapat 202.409 data. Dataset tersebut dibagi menjadi 4, data_semua, data_by_artist, data_by_year, dan data_by_genre. Pada dataset data_semua terdapat 19

atribut, pada data_by_artist terdapat 15 atribut, pada data_by_genre terdapat 14 atribut. Berikut adalah contoh visualisasi dari setiap dataset:

Tabel 2. Visualisasi data_semua

no	1	2
valence	594	963
year	1921	1921
acousticness	982	732
artists	['Sergei Rachmaninoff', 'James Levine', 'Berliner Philharmoniker']	['Dennis Day']
danceability	279	8,19E+15
duration_ms	831667	180533
energy	211	341

explicit	0	0
id	4BJqT0PrAfrxzMOxytFOIz	7xPhfUan2yNtyFG0cUWkt8
instrumentalness	878	0
key	10	7
liveness	665	16
loudness	-20096	-12441
mode	1	1
name	Piano Concerto No. 3 in D Minor, Op. 30: III. Finale. Alla breve	Clancy Lowered the Boom
popularity	4	5
release_date	1921	1921
speechiness	366	415
tempo	80954	6,0936E+15

Tabel 3. Visualisasi data_by_genre

no	1	2
mode	1	1
genres	21st century classical	432hz
acousticness	9,79333E+15	49478
danceability	1,62883E+16	2,99333E+15
duration_ms	1,60298E+16	1,04889E+15
energy	7,13167E+15	4,50678E+15
instrumentalness	60683367	4,77762E+15
liveness	3616	131
loudness	-3,15143E+16	-16854
speechiness	4,05667E+15	7,68167E+15
tempo	753365	1,20286E+16
valence	1,03783E+16	22175
popularity	2,78333E+15	525
key	6	5

Tabel 4. Visualisasi data_by_artist

no	1	2
mode	1	1
count	9	26
acousticness	5,90111E+15	8,62538E+15
artists	"Cats" 1981 Original London Cast	"Cats" 1983 Broadway Cast
danceability	4,67222E+15	4,41731E+15
duration_ms	2,50319E+15	2872800
energy	3,94003E+15	4,06808E+15
instrumentalness	1,13999E+16	8,11583E+15
liveness	2,90833E+15	3,15215E+15
loudness	-14448	-1069
speechiness	2,10389E+16	1,76212E+16
tempo	1,17518E+16	1,03044E+16
valence	3895	2,68865E+15
popularity	3,83333E+16	3,05769E+15
key	5	5

Pada tahap ini, fokus utama adalah membersihkan dataset dari data yang hilang dan data yang terduplikasi untuk memastikan keakuratan dan kebersihan data. Data yang hilang atau tidak lengkap dapat mengurangi keandalan analisis dan membuat hasil yang diperoleh menjadi tidak representatif. Sementara itu, keberadaan data duplikat dapat mengganggu analisis statistik dan menghasilkan hasil yang tidak akurat. Oleh karena itu, tahap pemrosesan data ini menjadi kunci dalam menyediakan dataset yang bersih dan siap untuk tahap analisis lebih lanjut. Dengan menghilangkan data yang hilang dan duplikat,

kita dapat meningkatkan kualitas dataset dan meningkatkan kepercayaan dalam hasil analisis yang dihasilkan.

Setelah data dibersihkan, selanjutnya menggabungkan ketiga dataset, yaitu dataset data_semua, data_by_genre, dan data_by_artist menjadi 'x'

Selanjutnya adalah mentekun atribut yang dipakai. Atribut tersebut yaitu, valence, acousticness, danceability, energy, instrumentalness, liveness, loudness, speechiness, tempo, popularity, key, dan mode.

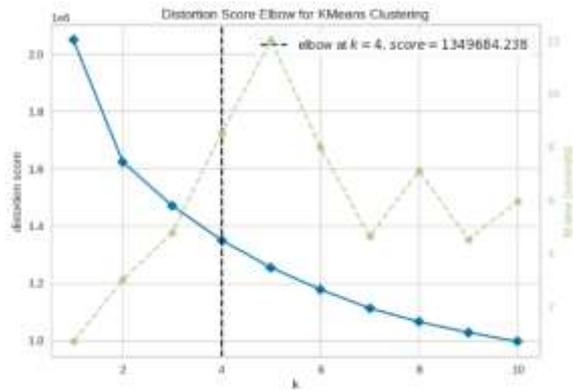
Setelah atribut ditentukan tahap selanjutnya adalah melakukan transformasi data. Transformasi yang dilakukan di sini adalah normalisasi data menggunakan metode standarisasi, di mana setiap fitur dalam data akan diubah sedemikian rupa sehingga memiliki mean nol dan variansi satu. Hal ini dapat meningkatkan performa model machine learning yang sensitif terhadap skala fitur. Dengan melakukan normalisasi, kita memastikan bahwa setiap fitur memiliki skala yang seragam, sehingga tidak ada satu fitur pun yang mendominasi pengaruhnya terhadap hasil analisis atau pemodelan.

```

from sklearn.preprocessing import StandardScaler
X = StandardScaler().fit_transform(X)
X
array([[ -1.7828247 ,  1.27618658, -1.46701268, ..., -1.25680847,
         1.36558797,  0.64391197],
       [ 1.65068832,  0.61134711,  1.59877887, ..., -1.21099271,
         0.5121232 ,  0.64391197],
       [-1.858821 ,  1.22034007, -1.18882049, ..., -1.21099271,
        -0.62582982,  0.64391197],
       ...,
       [ 0.41194856, -1.06670771,  0.5484614 , ...,  2.04192615,
        -0.34134156, -1.55300732],
       [-1.26756976, -1.30876246,  0.75852489, ...,  1.7670316 ,
        -0.91031807,  0.64391197],
       [ 0.43094764, -0.98426761,  1.80884237, ...,  1.95029463,
         0.5121232 ,  0.64391197]])
    
```

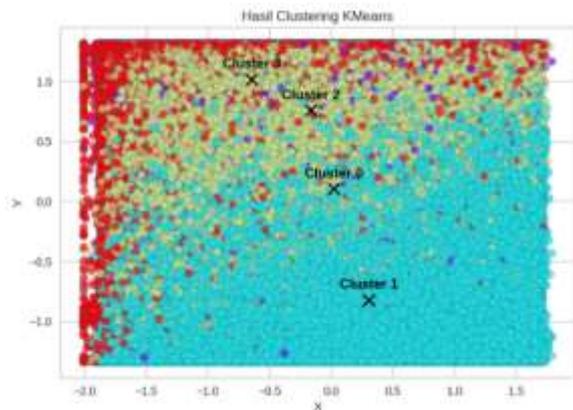
Gambar 2. Transformasi data dengan python

Setelah data ditransformasi, selanjutnya adalah menentukan nilai K menggunakan metode Elbow. Metode Elbow digunakan untuk mengevaluasi jumlah cluster yang optimal dalam model KMeans menggunakan metode elbow visualizer.



Gambar 3. Hasil metode *elbow*

Dari gambar 3 diatas, didapat nilai K untuk pelatihan model dengan nilai K adalah 4 dengan score 1349684,238. Dari nilai K yang didapat, selanjutnya adalah tahapan penerapan algoritma KMeans. Dari hasil model KMeans yang dilatih, hasil clustering dan centroid divisualisasikan menggunakan *scatter plot*.



Gambar 4. Visualisasi *cluster*

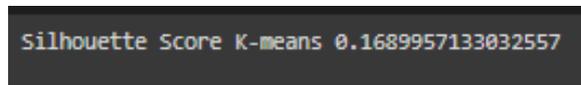
Dari visualisasi cluster tersebut menunjukkan 4 warna, 4 warna tersebut didapat dari nilai K yang sebelumnya sudah didapat melalui metode *Elbow*. Tanda silang menunjukkan posisi centroid pada setiap cluster. Pada centroid cluster 0, data dalam cluster ini cenderung memiliki nilai positif rendah pada kedua fitur. Ini menunjukkan bahwa kelompok data ini memiliki karakteristik yang mendekati titik asal (0,0) dalam ruang fitur. Pada centroid cluster 1, data dalam cluster ini memiliki nilai fitur pertama yang positif dan nilai fitur kedua yang negatif. Ini mungkin menunjukkan pola atau kondisi tertentu yang memisahkan data ini dari cluster lainnya, seperti hubungan negatif antara kedua fitur. Pada centroid cluster 2, data dalam cluster ini cenderung memiliki nilai fitur pertama yang negatif dan fitur kedua yang positif.

Cluster ini menunjukkan kelompok data dengan karakteristik yang berkebalikan pada kedua fitur dibandingkan dengan Cluster 1. Pada centroid cluster 3, data dalam cluster ini memiliki nilai fitur pertama yang lebih negatif dan fitur kedua yang positif, menunjukkan data dengan karakteristik yang lebih

ekstrem pada dimensi pertama dibandingkan dengan Cluster 2.

Kesimpulan yang dapat diambil dari visualisasi tersebut adalah setiap cluster terletak di kuadran yang berbeda atau memiliki kombinasi nilai fitur yang unik, menunjukkan bahwa KMeans telah berhasil mengelompokkan data berdasarkan pola dalam dua fitur yang dianalisis. Setiap data yang serupa cenderung memiliki titik data yang berdekatan satu sama lain.

Tahapan selanjutnya adalah mengevaluasi hasil model KMeans yang dilatih menggunakan *Silhouette Score* dengan python. Fungsi ini digunakan untuk menghitung *Silhouette Score*, yang merupakan metrik evaluasi yang umum digunakan untuk mengevaluasi kualitas kluster dalam analisis kluster. *Silhouette Score* berkisar antara -1 hingga 1, dengan nilai yang lebih tinggi menunjukkan kluster yang lebih baik. Hasil *Silhouette Score* dapat dilihat pada gambar berikut.



Gambar 5. Hasil *Silhouette Score*

Dari hasil evaluasi *Silhouette Score* tersebut, algoritma k-means berada pada nilai 0.1689957133032557, menunjukkan bahwa hasil clustering menggunakan algoritma K-Means tidak optimal. Nilai 0.1689957133032557 mengindikasikan bahwa ada beberapa overlap atau campuran antara cluster, dan bahwa ada ruang untuk peningkatan dalam pemisahan antara cluster-cluster yang terbentuk.

5. KESIMPULAN DAN SARAN

Dengan *Silhouette Score* sebesar 0.1689957133032557, hasil clustering yang diperoleh dalam sistem rekomendasi musik tampaknya tidak optimal. Meskipun demikian, ada beberapa langkah yang dapat diambil untuk meningkatkan kualitas clustering. Pertama, perlu dilakukan evaluasi lebih lanjut terhadap fitur-fitur yang digunakan dalam analisis. Kemudian, eksplorasi lebih lanjut terhadap jumlah kluster dan penggunaan metode inisialisasi yang berbeda juga dapat membantu. Selain itu, penting untuk mempertimbangkan teknik evaluasi lainnya seperti Davies-Bouldin Index atau Calinski-Harabasz Index untuk mendapatkan pemahaman yang lebih holistik tentang kualitas clustering. Dengan melakukan langkah-langkah ini, diharapkan sistem rekomendasi musik dapat memberikan rekomendasi yang lebih relevan dan bermanfaat bagi pengguna.

DAFTAR PUSTAKA

[1] G. Wahyu Dewatara and S. Monik Agustin, "PEMASARAN MUSIK PADA ERA DIGITAL DIGITALISASI INDUSTRI MUSIK DALAM INDUSTRI 4.0 DI INDONESIA".
 [2] M. E. M. Rumapea, "Tantangan Pembelajaran Musik Pada Era Digital," *Gondang: Jurnal Seni*

- dan Budaya*, vol. 3, no. 2, p. 101, Dec. 2019, doi: 10.24114/gondang.v3i2.13168.
- [3] “Beyond_the_black_box_in_music_streaming-O’Dair-Fry”.
- [4] A. Widya Astuti, A. Muharam, P. Siber Cerdika Internasional, and U. Swadaya Gunung Jati, “PERKEMBANGAN BISNIS DI ERA DIGITAL,” 2023, [Online]. Available: <https://jmi.rivierapublishing.id/index.php/rp>
- [5] “AI in the media and creative industries.”
- [6] V. Chang, Y. Yang, Q. A. Xu, and C. Xiong, “Factors Influencing Consumer Intention to Subscribe to the Premium Music Streaming Services in China,” *Journal of Global Information Management*, vol. 29, no. 6. IGI Global, Nov. 01, 2021. doi: 10.4018/JGIM.20211101.0a17.
- [7] X. Du-Harpur, F. M. Watt, N. M. Luscombe, and M. D. Lynch, “What is AI? Applications of artificial intelligence to dermatology,” *British Journal of Dermatology*, vol. 183, no. 3. Blackwell Publishing Ltd, pp. 423–430, Sep. 01, 2020. doi: 10.1111/bjd.18880.
- [8] Z. Ghahramani, “Unsupervised Learning.” [Online]. Available: <http://www.gatsby.ucl.ac.uk/~zoubin>
- [9] “OPTIMASI JUMLAH CLUSTER K-MEANS DENGAN METODE.”
- [10] H. H. Simon Yuill, “Python,” 2006.