

Sistem Rekomendasi Lagu Berbasis Knna Menggunakan Fitur Audio Dari Spotify Api

^{1*}**Nur Kholidh Yoga Witanto, ²Daniel Swanjaya, ³Danar Pamungkas**

^{1,2,3} Teknik Informatika, Universitas Nusantara PGRI Kediri

E-mail: ^{*}nur.kholiqyw@gmail.com, ²daniel@unpkediri.ac.id, ³danar@unpkediri.ac.id

Penulis Korespondens : Nur Kholidh Yoga Witanto

Abstrak—Sistem rekomendasi lagu berbasis fitur audio menjadi solusi untuk membantu pengguna menemukan lagu baru sesuai preferensi mereka. Penelitian ini bertujuan mengembangkan sistem rekomendasi lagu menggunakan algoritma K-Nearest Neighbor (KNN) dengan memanfaatkan fitur audio dari Spotify API seperti tempo, energy, valence, acousticness, dan danceability. Proses dimulai dengan mengumpulkan data lagu beserta fitur audionya, selanjutnya lakukan normalisasi pada data, sesudah data dinormalisasi dilakukan perhitungan menggunakan algoritma KNN untuk mendapatkan rekomendasi berdasarkan lagu yang dipilih oleh pengguna dan dilakukan evaluasi menggunakan Top-K Accuracy. Hasil pengujian menunjukkan bahwa sistem mampu memberikan rekomendasi dengan tingkat akurasi sebesar 68%, yang berarti 3-4 lagu dari 5 lagu yang direkomendasikan relevan dengan preferensi pengguna. Penelitian ini membuktikan bahwa KNN dapat diterapkan secara efektif dalam sistem rekomendasi musik berbasis fitur audio, sehingga meningkatkan pengalaman pengguna dalam menikmati musik.

Kata Kunci— fitur audio, K-Nearest Neighbor, rekomendasi lagu, Spotify API, Top K-Accuracy

Abstract— Audio feature-based song recommendation systems provide a solution to help users discover new songs tailored to their preferences. This study aims to develop a song recommendation system using the K-Nearest Neighbor (KNN) algorithm, leveraging audio features from the Spotify API such as tempo, energy, valence, acousticness, and danceability. The process begins with collecting song data along with their audio features, followed by normalizing the data. After normalization, the KNN algorithm is used to calculate recommendations based on the song selected by the user, and the system is evaluated using Top-K Accuracy. Test results indicate that the system achieves a recommendation accuracy of 68%, meaning 3-4 out of 5 recommended songs are relevant to user preferences. This research demonstrates that KNN can be effectively applied in audio feature-based music recommendation systems, thereby enhancing user experience in enjoying music.

Keywords— audio features, K-Nearest Neighbor, song recommendation, Spotify API, Top K-Accuracy

This is an open access article under the CC BY-SA License.



I. PENDAHULUAN

API (Application Programming Interface) adalah mekanisme yang memungkinkan beberapa aplikasi di platform berbeda untuk saling berkomunikasi[1]. Spotify Web API memungkinkan aplikasi berinteraksi dengan layanan streaming Spotify, seperti mengakses metadata lagu. API ini dapat dimanfaatkan sebagai dasar untuk membangun sistem rekomendasi lagu yang lebih personal [2][3].

Pengguna kerap mengalami tantangan dalam mencari lagu baru yang sesuai dengan preferensi mereka. Oleh karena itu, sistem rekomendasi musik menjadi sangat penting untuk

membantu pengguna dalam menemukan lagu baru yang cocok dengan selera mereka [4]. Dengan memanfaatkan Spotify API, masalah ini dapat diatasi melalui penyediaan data lagu yang kaya, termasuk fitur audio yang mendukung sistem rekomendasi [5].

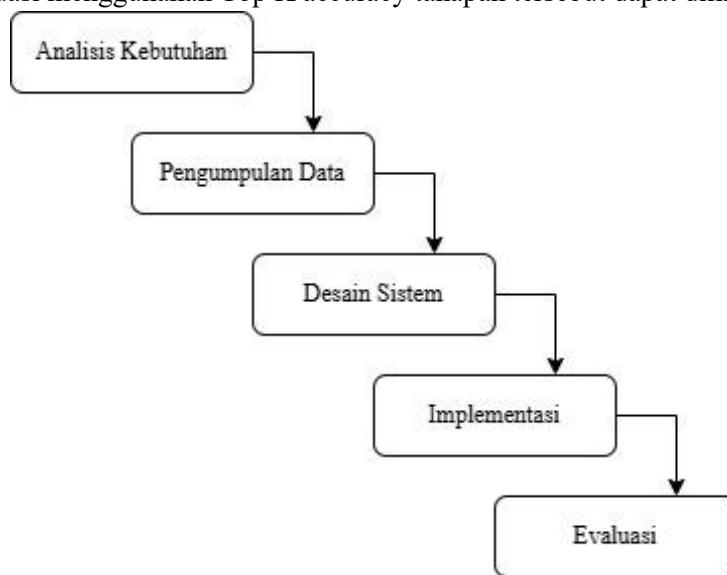
Penelitian mengenai sistem rekomendasi musik telah banyak dikembangkan sebelumnya dengan beragam pendekatan dan algoritma. Salah satunya dilakukan oleh Giri dkk. [6] yang menerapkan metode Case-Based Reasoning (CBR) dengan memanfaatkan data Listening History dan berhasil memperoleh nilai precision rata-rata sebesar 0,79. Sementara itu, penelitian yang akan saya lakukan akan memanfaatkan fitur audio dari lagu, seperti tempo, valence, energy, danceability, dan acousticness, sebagai dasar dalam memberikan rekomendasi musik.

Penelitian ini bertujuan untuk menghasilkan rekomendasi lagu yang disesuaikan dengan preferensi pengguna, dengan memanfaatkan fitur-fitur audio. Algoritma K-Nearest Neighbor (KNN) digunakan dalam proses ini untuk mengklasifikasikan data berdasarkan kedekatan jarak antara satu data dengan data lainnya [7]. Dengan demikian, diharapkan penelitian ini dapat mempermudah pengguna dalam menemukan lagu-lagu baru yang sesuai dengan selera mereka, sehingga dapat meningkatkan kenyamanan dan kepuasan dalam menikmati musik.

II. METODE

2.1 Metode Penelitian

Pada penelitian ini menggunakan metode waterfall, tahap pertama melakukan analisis kebutuhan guna mengetahui yang diperlukan dalam melakukan penelitian dengan cara studi literatur dan observasi dalam pemahaman fitur audio, selanjutnya melakukan pengumpulan data menggunakan dataset dari Kaggle yang berisi fitur audio dengan format numerik, lalu melakukan desain sistem mulai dari preprocessing data dengan cara melakukan normalisasi pada data yang akan digunakan agar fitur audio seperti tempo tidak mendominasi dalam melakukan rekomendasi, tahap berikutnya melakukan implementasi pada data yang sudah siap digunakan, pada tahap akhir melakukan evaluasi menggunakan Top K accuracy tahapan tersebut dapat dilihat di Gambar 1.



Gambar 1. Metode Penelitian

2.2 Spotify

Spotify merupakan platform streaming musik yang berasal dari Stockholm, Swedia, dengan kantor pusat yang berlokasi di London. Platform ini didirikan pada tahun 2006 dan resmi

diluncurkan di Eropa pada tahun 2008. Spotify menyediakan dua jenis layanan, yaitu layanan gratis dan layanan premium. Pengguna yang memilih layanan premium dapat menikmati pemutaran lagu tanpa batas dan bebas dari iklan, baik sebelum maupun sesudah lagu diputar [8].

2.3 Sistem Rekomendasi

Sistem rekomendasi adalah salah satu komponen dalam sistem pencarian atau pengambilan informasi (*Information Retrieval*) dan termasuk dalam ranah kecerdasan buatan (*Artificial Intelligence*) [9]. Sistem ini berfungsi sebagai perangkat lunak yang menyajikan prediksi berupa saran item kepada pengguna berdasarkan karakteristik dan preferensi individu, dengan tujuan membantu pengambilan keputusan agar pengguna lebih cenderung menyukai item yang direkomendasikan. Dalam proses pemberian rekomendasi, sistem tidak hanya bergantung pada satu aspek, seperti penilaian pengguna, melainkan juga dapat menerapkan pendekatan multi-kriteria. Pendekatan ini mempertimbangkan berbagai faktor penting dari suatu item. Sebagai contoh, dalam sistem rekomendasi musik, kriteria yang dapat dianalisis meliputi lirik, elemen visual, karakteristik audio, suara, ritme, genre, dan lainnya. Dengan memperhitungkan berbagai atribut tersebut, sistem rekomendasi dapat memahami preferensi pengguna secara lebih mendalam dan meningkatkan ketepatan dalam menyajikan saran [10].

2.4 K-Nearest Neighbour

Algoritma KNN merupakan salah satu metode klasifikasi yang bekerja dengan membandingkan jarak terdekat antara data yang ingin diklasifikasikan dengan data lainnya. Penentuan nilai K yang paling sesuai ditentukan berdasarkan karakteristik data yang dimiliki. Jika nilai K terlalu besar, pengaruh data yang menyimpang (noise) dapat diminimalkan, namun hal ini juga dapat menyebabkan batas antara kelas-kelas menjadi kurang jelas [7]. Berikut adalah rumus yang digunakan untuk mengukur tingkat kesamaan data :

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

d = jarak antara lagu x dan y

N = Jumlah total fitur yang dibandingkan

x = Data fitur audio dari lagu pertama

y = Data fitur audio dari lagu kedua

i = Indeks dari setiap fitur

2.5 Top K- Accuracy

Top K-Accuracy merupakan metrik yang bersifat fleksibel dan dapat digunakan baik dalam model klasifikasi berbasis K-Nearest Neighbor (KNN) maupun untuk mengevaluasi sistem rekomendasi yang didasarkan pada urutan relevansi. Metrik ini bertujuan untuk membantu dalam menentukan nilai K yang paling sesuai atau dalam menilai kinerja sistem dalam menyajikan rekomendasi teratas. Dalam konteks sistem rekomendasi, Top K-Accuracy biasanya disebut juga evaluasi Top-K, yang mengukur sejauh mana sistem mampu memberikan rekomendasi relevan di antara K item teratas. Metrik ini tidak hanya menilai ketepatan prediksi, tetapi juga memperhatikan relevansi rekomendasi bagi pengguna. Sebagai ilustrasi, pada platform e-commerce, Top K-Accuracy digunakan untuk mengetahui seberapa sering produk yang benar-benar diminati pengguna muncul dalam daftar rekomendasi teratas [11]. Berikut adalah rumus yang digunakan untuk mengukur tingkat akurasi.

$$\text{Top K Accuracy} = \frac{\text{Jumlah Prediksi Benar di K teratas}}{\text{Jumlah total Sampel}} \quad (2)$$

III. HASIL DAN PEMBAHASAN

3.1 Implementasi K-Nearest Neighbor

Penelitian ini menerapkan metode K-Nearest Neighbor dalam memberikan rekomendasi lagu berdasarkan kemiripan fitur audio dari lagu yang dipilih oleh pengguna dengan lagu yang akan direkomendasikan. Berikut adalah bentuk datasetnya pada Tabel 1.

Tabel 1. Dataset

spotify_id	name	artist	tempo	energy	valence	acousticness	Danceability
0kO3njY9N1Rxgv27Ha1Lh	Numb	Linkin Park	110.018	0.863	0.243	0.0046	0.496
01vQXuD FQPlizZb7 XE3BqN	Zombie	The Cranberries	166.821	0.664	0.265	0.0202	0.311
0BWx4W KgUXKHj 8npThsbc H	Dani California	Red Hot Chili Peppers	96.184	0.913	0.73	0.0193	0.556
03hLW9os M6UuNql pwL8XPo	Stairway to Heaven	Led Zeppelin	80.36	0.361	0.212	0.625	0.354
0Aau2Ju1 RoLAhL9 0zRcDx4	Welcome to the Jungle	Guns N' Roses	130.139	0.91	0.154	0.106	0.272

Lalu dataset tersebut dilakukan normalisasi yang bertujuan supaya salah satu fitur audio tidak mendominasi dalam melakukan rekomendasi. Berikut adalah dataset yang sudah dinormalisasi ditampilkan pada Tabel 2.

Tabel 2. Dataset sudah dinormalisasi

spotify_id	name	artist	tempo	energy	valence	acousticness	danceability
0B0sl8U1Acm G4rbNdydN7P	Story of My Life	One Direction	0.506791686724293	0.663	0.28801611278952666	0.22590361445783133	0.6085192697768763
03fqO4EyMiNTf0sA6NM4gw	When it Rains	Eli Young Band	0.5026392348102723	0.636	0.2719033232628399	0.16967871485943778	0.6064908722109533
0BWx4WKgUXKHj8npThsbcH	Dani California	Red Hot Chili Peppers	0.4026203980828396	0.913	0.7351460221550855	0.019377510040160646	0.5638945233265721

03hLW9osM6 UuNqlpwL8X Po	Stairway to Heaven	Led Zeppelin	0.3363 820925 511207	0.36 1	0.21349 446122 86002	0.62751004 01606426	0.35902636 9168357
0Aau2Ju1RoL AhL90zRcDx4	Welcome to the Jungle	Guns N' Roses	0.5447 539714 100337	0.91	0.15508 559919 436052	0.10642570 281124498	0.27586206 89655173

Selanjutnya pengguna memilih judul lagu yang dimana nantinya fitur audio dari lagu tersebut akan digunakan sebagai dasar dari rekomendasi dengan cara mengehitung fitur audio dari lagu yang dipilih oleh pengguna dan semua lagu yang berada pada dataset. Untuk contohnya ada pada Tabel 3 dan Tabel 4 berikut.

Tabel 3. Lagu yang dipilih oleh pengguna

name	tempo	energy	valence	acousticness	danceability
Story of My Life	0.506791686724 293	0.663	0.2880161127895266 6	0.225903614 45783133	0.60851926 97768763

Tabel 4. Fitur audio Lagu lain dari dataset

name	tempo	energy	valence	acousticness	danceability
When it Rains	0.5026392 348102723	0.636	0.2719033232628399 628399	0.1696787148594 3778	0.6064908722109 533

Langkah berikutnya menghitung perbedaan nilai setiap fitur lagu yang dipilih pengguna dengan lagu-lagu yang terdapat dalam dataset. Untuk perhitungannya disajikan pada Tabel 5.

Tabel 5. Perhitungan tiap fitur

Tempo	0.506791686724293	-	0.5026392348102723	=	0,0041524519140207
Energy	0.663	-	0.636	=	0,027
Valence	0.28801611278952666	-	0.2719033232628399	=	0.01611278953
Acousticness	0.22590361445783133	-	0.16967871485943778	=	0.0562248996
Danceability	0.6085192697768763	-	0.6064908722109533	=	0.002028397566

Selanjutnya hasil dari selisih setiap fitur audio di akar kuadratkan seperti pada Tabel 6.

Table 6. Hasil akar kuadrat

Tempo	(0,0041524519140207)2	=	0,000017242858
Energy	(0,027)2	=	0,000729
Valence	(0,01611278953)2	=	0,0002596219608686976209
Acousticness	(0,0562248996)2	=	0,00316123818516608016
Danceability	(0,002028397566)2	=	0,000004114386

Melakukan penjumlahan semua fitur audio hasil dari fitur audio yang sudah di akar kuadratkan:

$$0,000004114386 + 0,000729 + 0,0002596219608686976209 + 0,00316123818516608016 \\ 0,000004114386 = 0,0041712173900347777809$$

Proses akhir melibatkan pengambilan akar kuadrat dari hasil penjumlahan semua fitur audio guna memperoleh jarak antara fitur audio lagu pilihan pengguna dan fitur audio lagu lainnya.

$$\sqrt{0,0041712173900347777809} = 0,064583414$$

3.2 Top K-accuracy

Pengujian akurasi pada sistem rekomendasi ini bertujuan untuk mengevaluasi tingkat relevansi lagu yang disarankan oleh sistem berdasarkan preferensi pengguna. Pada penelitian ini, Top-K Accuracy digunakan sebagai indikator keberhasilan sistem dalam memberikan rekomendasi lagu yang sesuai. Data untuk pengujian diperoleh dari sepuluh pengguna yang memberikan penilaian terhadap lima lagu yang direkomendasikan kepada mereka.

Tabel 7. Top-K Accuracy

Pengguna	Jumlah Relevan	Akurasi
1	4	0,8
2	5	1
3	1	0,2
4	4	0,8
5	3	0,6
6	3	0,6
7	4	0,8
8	3	0,6
9	4	0,8
10	3	0,6

$$\frac{0,8 + 1 + 0,2 + 0,8 + 0,6 + 0,6 + 0,4 + 0,6 + 0,8 + 0,6}{10} = 0,68$$

Tingkat relevansi sistem dalam memberikan rekomendasi lagu kepada pengguna ada 68%. Dari hasil tersebut dapat disimpulkan bahwa sistem dapat memberikan 3 sampai 4 lagu yang relevan. Hasil tersebut menunjukkan bahwa metode K-Nearest Neighbor mampu memberikan rekomendasi yang cukup baik.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa sistem rekomendasi lagu berbasis fitur audio menggunakan algoritma K-Nearest Neighbor (KNN) mampu memberikan rekomendasi yang relevan dengan preferensi pengguna. Fitur audio seperti tempo, energy, valence, acousticness, dan danceability berperan penting dalam menentukan kemiripan antar lagu. Proses normalisasi data juga terbukti efektif untuk menghindari dominasi salah satu fitur dalam perhitungan jarak. Evaluasi menggunakan metrik Top K-Accuracy menunjukkan tingkat akurasi sebesar 68%, yang mengindikasikan bahwa sistem dapat merekomendasikan 3-4 lagu relevan dari 5 lagu yang disajikan. Dengan demikian, penelitian ini berhasil membuktikan bahwa pendekatan KNN dengan memanfaatkan fitur audio dari Spotify API dapat menjadi solusi yang efektif untuk meningkatkan pengalaman pengguna dalam menemukan lagu baru. Untuk pengembangan selanjutnya, dapat dilakukan eksperimen dengan algoritma lain atau penambahan fitur audio untuk meningkatkan akurasi rekomendasi.

DAFTAR PUSTAKA

- [1] A. Wahyudin, A. Anisyah, D. Ahmaddifa, K. Bandung, A. P. Interface, and F. Cloud, “Pengembangan Sistem Notifikasi Real-Time untuk Aplikasi Manajemen Persuratan Multiplatform menggunakan Firebase Cloud Messaging dan Application Programming Interface,” vol. 09, pp. 284–300, 2024.
- [2] F. A. Anderson and M. M. Mutoffar, “Tinjauan Pustaka: Penerapan Teknologi Artificial Intelligence Pada Fitur ‘Made For You’ Aplikasi Spotify,” *INTEK J. Inform. dan Teknol. Inf.*, vol. 7, no. 1, pp. 41–46, 2024.
- [3] R. C. Maringka, A. Khoirunnita, R. Maringka, E. Utami, and K. Kusnawi, “Analisa Perkembangan Musik Pada Spotify Menggunakan Structured Query Language (SQL),” *CogITO Smart J.*, vol. 7, no. 1, pp. 1–14, 2021, doi: 10.31154/cogito.v7i1.287.1-14.
- [4] R. Amanda, “Music streaming dalam industri musik era industri 4.0,” *J. Stud. Komun. (Indonesian J. Commun. Stud.)*, vol. 6, no. 1, pp. 358–382, 2022, doi: 10.25139/jsk.v6i1.3772.
- [5] A. R. Ilhami, A. Putra, A. C. Aurelie, K. Junaidi, and S. Israwana, “Implementasi MapReduce Pada Dataset Spotify Top Music Untuk Mengetahui Artist yang paling banyak Didengar Dalam Kurun Waktu 10 Tahun,” *J. Netw. Comput. Appl.*, vol. 2, no. 2, pp. 37–51, 2023, [Online]. Available: <https://jurnal.netplg.com/>
- [6] G. A. V. M. Giri, M. L. Radhitya, M. A. Raharja, and I. W. Supriana, “SISTEM REKOMENDASI MUSIK BERDASARKAN DATA KONTEKS PADA LISTENING HISTORY MUSIK DAN KETERKAITAN ARTIS INDONESIA,” *J. Resist. (Rekayasa Sist. Komputer)*, vol. 5, no. 1, pp. 86–93, Apr. 2022, doi: 10.31598/jurnalresistor.v5i1.1044.
- [7] S. R. Cholil, T. Handayani, R. Prathivi, and T. Ardianita, “Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Seleksi Penerima Beasiswa,” *IJCIT (Indonesian J. Comput. Inf. Technol.)*, vol. 6, no. 2, pp. 118–127, 2021, doi: 10.31294/ijcit.v6i2.10438.
- [8] J. Wang, “A Critical Research of Spotify’s Business Model—The Case of Discover Weekly,” *BCP Soc. Sci. Humanit.*, vol. 21, 2023, doi: 10.54691/bcpssh.v21i.3873.
- [9] Sitti Aliyah Azzahra, Syafran Nurrahman, and Aep Saefullah, “Integrasi Kecerdasan Buatan Dalam Sistem Rekomendasi Produk Untuk E-Commerce,” *J. Sains dan Teknol.*, vol. 3, no. 1, pp. 21–28, 2024, doi: 10.58169/saintek.v3i1.394.
- [10] F. R. Hariri and L. W. Rochim, “Sistem Rekomendasi Produk Aplikasi Marketplace Berdasarkan Karakteristik Pembeli Menggunakan Metode User Based Collaborative Filtering,” *Teknika*, vol. 11, no. 3, pp. 208–217, Nov. 2022, doi: 10.34148/teknika.v11i3.538.
- [11] D. Dhimas *et al.*, “Sistem Rekomendasi Laptop Menggunakan Metode Collaborative Filtering Dan Weighted Product Pada Toko Online IndoJaya Computer,” vol. 4, pp. 594–604, 2025.