JSITIK, Vol.5 No.2 Juni 2027
ISSN: 2986-0458 (Print) / 2986-044X (Online)
DOI: https://doi.org/10.53624/jsitik.v5i2.971

Penerapan Algoritma Random Forest Classifier pada Sistem Rekomendasi Lagu Berbasis Content-Based Filtering

Diterima:
21 Juli 2026
Revisi:
15 Oktober 2026
Terbit:
28 Oktober 2026

Reza Dwi Anggara, *Ilham Saifudin, Rosita Yanuarti
Universitas Muhammadiyah Jember

Abstrak— **Latar Belakang:** Banyaknya pilihan lagu pada platform musik digital dapat menyulitkan pendengar dalam menemukan lagu yang sesuai dengan minat dan kebutuhan. Kondisi ini dapat menyebabkan *information overload*, sehingga diperlukan sistem rekomendasi yang mampu membantu menemukan lagu berdasarkan karakteristik audio. **Tujuan:** Penelitian ini bertujuan menerapkan algoritma *Random Forest Classifier* pada sistem rekomendasi lagu berbasis *Content-Based Filtering*. **Metode:** *Random Forest Classifier* digunakan untuk mengklasifikasikan genre lagu berdasarkan sembilan fitur audio, yaitu *danceability, energy, loudness, speechiness, acousticness, instrumentalness, liveness, valence*, dan *tempo*. Hasil klasifikasi genre kemudian digunakan dalam proses *Content-Based Filtering* dengan *Cosine Similarity* untuk menghasilkan rekomendasi lagu yang memiliki karakteristik serupa. **Hasil:** Pengujian *Random Forest Classifier* pada skenario *Genre Selection* menghasilkan nilai *accuracy* sebesar 0,66. Evaluasi sistem rekomendasi menghasilkan nilai Precision@10 sebesar 0,9915, Precision@30 sebesar 0,9866, dan Precision@50 sebesar 0,9832. Sementara itu, nilai Recall@10, Recall@30, dan Recall@50 masing-masing sebesar 0,1015, 0,3030, dan 0,5033. **Kesimpulan:** Penerapan *Random Forest Classifier* dan *Content-Based Filtering* mampu menghasilkan rekomendasi lagu yang relevan berdasarkan karakteristik audio, dengan nilai *precision* yang tinggi dan *recall* yang meningkat seiring bertambahnya jumlah rekomendasi.
Kata Kunci— Sistem Rekomendasi Lagu; Content-Based Filtering; Random Forest Classifier; klasifikasi genre; Cosine Similarity

Abstract*— *Background: *The vast number of songs available on digital music platforms can make it difficult for listeners to find tracks that align with their interests and needs. This situation can lead to information overload, necessitating a recommendation system capable of helping users discover songs based on their audio characteristics.* ***Objective:*** *This study aims to implement the Random Forest Classifier algorithm within a content-based filtering recommendation system.* ***Methods:*** *The Random Forest Classifier is used to classify song genres based on nine audio features: danceability, energy, loudness, speechiness, acousticness, instrumentalness, liveness, valence, and tempo. The resulting genre classifications are then utilized in a content-based filtering process, employing cosine similarity to generate recommendations for songs with similar characteristics.* ***Results:*** *Testing the Random Forest Classifier in a "genre selection" scenario yielded an accuracy score of 0.66. Evaluation of the recommendation system produced Precision@10, Precision@30, and Precision@50 values of 0.9915, 0.9866, and 0.9832, respectively. Meanwhile, the Recall@10, Recall@30, and Recall@50 values were 0.1015, 0.3030, and 0.5033, respectively.* ***Conclusion:*** *The application of the Random Forest Classifier and content-based filtering effectively generates relevant song recommendations based on audio characteristics, achieving high precision, with recall values that increase as the number of recommendations grows.*
Keywords— *Music Recommendation System; Content-Based Filtering; Random Forest Classifier; genre classification; Cosine Similarity*

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

Penulis Korespondensi:

Ilham Saifudin,
Teknik Informatika,
Universitas Muhammadiyah Jember,
Email: ilham.saifudin@unmuhjember.ac.id
ORCID ID: https://orcid.org/0000-0002-2063-4524

ISSN: 2986-0458 (Print) / 2986-044X (Online)
DOI: https://doi.org/10.53624/jsitik.v5i2.971

I. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah membawa perubahan dalam berbagai bidang kehidupan, termasuk industri musik. Namun, banyaknya pilihan lagu dapat menyulitkan pengguna ketika ingin menemukan lagu lain yang memiliki karakteristik serupa. Permasalahan ini berkaitan dengan *Information Retrieval* (IR), yaitu proses memperoleh informasi yang sesuai dengan kebutuhan dari sekumpulan informasi yang tersedia (Murtopo et al., 2022). Oleh karena itu, diperlukan sistem rekomendasi yang dapat membantu pengguna menemukan lagu berdasarkan karakteristik lagu acuan.

Sistem rekomendasi dapat dikembangkan menggunakan pendekatan *Collaborative Filtering*, *Content-Based Filtering*, dan *Hybrid Filtering* (Widiyaningtyas et al., 2021). *Content-Based Filtering* menghasilkan rekomendasi berdasarkan karakteristik item (Saifudin & Widiyaningtyas, 2024), sehingga dapat diterapkan pada pencarian lagu berdasarkan fitur audio tanpa bergantung pada riwayat interaksi pengguna lain. Untuk mendukung proses tersebut, klasifikasi genre dapat digunakan sebagai tahap awal dalam menentukan kandidat lagu, kemudian dilanjutkan dengan pengukuran kemiripan karakteristik audio.

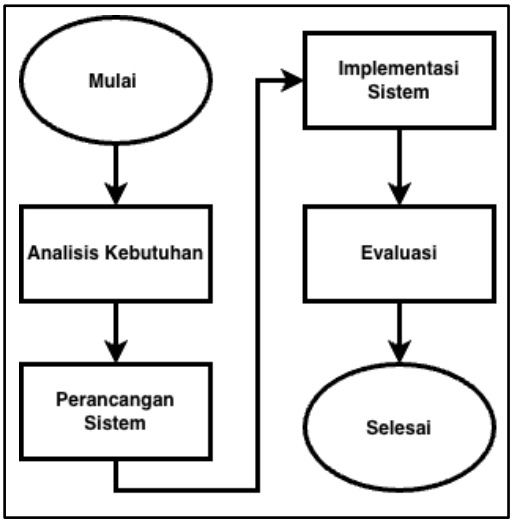
Beberapa penelitian terdahulu telah membahas klasifikasi dan rekomendasi musik dengan pendekatan yang berbeda. Munajad et al. (2025) mengembangkan sistem rekomendasi musik dengan menggabungkan K-Means, *K-Nearest Neighbors* (K-NN), dan *Cosine Similarity*, dengan hasil akurasi terbaik sebesar 91% pada pengujian K=1. Abrah et al. (2025) membandingkan algoritma *K-Nearest Neighbor*, *Random Forest*, dan *Naïve Bayes* untuk mengklasifikasikan genre musik menggunakan dataset lebih dari 500 lagu yang terdiri atas 5 genre musik. Hasil pengujian menunjukkan bahwa *Random Forest* memperoleh akurasi tertinggi dibandingkan kedua algoritma lainnya. Sementara itu, Timur dan Rohman (2025) mengembangkan sistem rekomendasi lagu Indonesia menggunakan *Content-Based Filtering* dengan TF-IDF dan *Cosine Similarity*. Namun, integrasi Random Forest Classifier dan Content-Based Filtering untuk rekomendasi lagu berdasarkan kemiripan fitur audio masih perlu dikaji lebih lanjut.

Perbedaan penelitian ini dengan penelitian sebelumnya adalah penggunaan *Random Forest Classifier* untuk mengklasifikasikan genre lagu kemudian mengintegrasikan hasil klasifikasi tersebut ke dalam proses rekomendasi berbasis *Content-Based Filtering*. Tujuan penelitian ini adalah menerapkan algoritma *Random Forest Classifier* pada sistem rekomendasi lagu berbasis *Content-Based Filtering* serta mengevaluasi kinerja klasifikasi menggunakan *accuracy* dan relevansi rekomendasi menggunakan *Precision* dan *Recall* pada beberapa jumlah rekomendasi. Pendekatan ini diharapkan dapat membantu pengguna menemukan lagu lain yang memiliki karakteristik audio serupa dengan lagu pilihannya.

JSITIK, Vol.5 No.2 Juni 2027
ISSN: 2986-0458 (Print) / 2986-044X (Online)
DOI: https://doi.org/10.53624/jsitik.v5i2.971

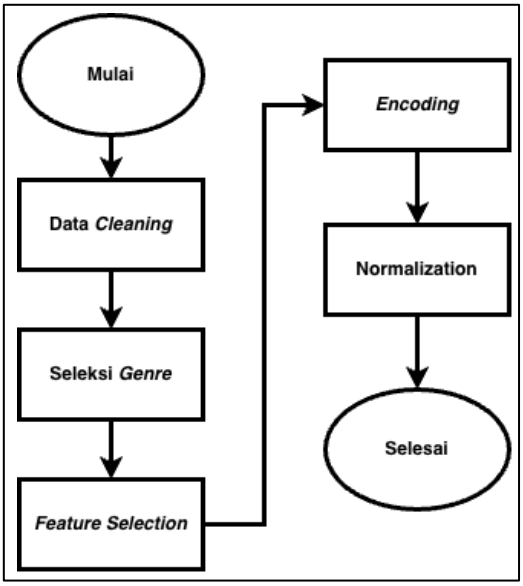
II. METODE

Gambar 1. menunjukkan tahapan penelitian yang dilakukan dalam pengembangan sistem rekomendasi musik. Tahapan penelitian ini dimulai dengan analisis kebutuhan untuk mengidentifikasi kebutuhan data, metode, dan perangkat yang digunakan. Selanjutnya, dilakukan perancangan sistem yang mencakup klasifikasi genre menggunakan algoritma *Random Forest Classifier* dan rekomendasi lagu menggunakan metode *Content-Based Filtering* dengan *Cosine Similarity*. Sistem kemudian diimplementasikan sesuai rancangan dan dievaluasi menggunakan metrik yang telah ditentukan.



Gambar 1. Diagram Alur Secara Umum

A. *Preprocessing Data*



Gambar 2. Alur Preprocesing

ISSN: 2986-0458 (Print) / 2986-044X (Online)
DOI: https://doi.org/10.53624/jsitik.v5i2.971

Gambar 2, menunjukkan tahapan preprocessing yang dilakukan dalam penelitian, yang meliputi data *cleaning*, seleksi genre, *feature selection*, *encoding*, dan *normalization*. Proses *preprocessing* diawali dengan pemeriksaan struktur data dan pembersihan data melalui penghapusan data duplikat, penanganan nilai kosong, serta penghapusan atribut yang tidak diperlukan. Selanjutnya dilakukan seleksi genre dengan menggunakan tujuh genre, yaitu Pop, Metal, Jazz, Rock, Reggae, Hip-Hop, dan EDM. Setelah itu, dipilih sembilan atribut audio sebagai fitur, yaitu *danceability, energy, loudness, speechiness, acousticness, instrumentalness, liveness, valence,* dan *tempo*. Tahap berikutnya adalah *Label Encoding* pada atribut genre, pembagian dataset menggunakan metode *hold-out* dengan rasio 80% data latih dan 20% data uji untuk *Random Forest Classifier*, serta normalisasi fitur numerik yang digunakan pada *Content-Based Filtering* agar setiap atribut memiliki skala yang sama dalam proses perhitungan kemiripan menggunakan *Cosine Similarity*.

B. *Random Forest Classifier*

Random Forest Classifier merupakan metode pembelajaran mesin berbasis *ensemble* yang membangun beberapa *decision tree* untuk mengklasifikasikan data berdasarkan karakteristik fitur yang dimiliki (Firnanda et al., 2025). Setiap pohon dilatih menggunakan teknik *bootstrap sampling*, kemudian hasil prediksi digabungkan melalui *majority voting* untuk menentukan kelas akhir. Pada penelitian ini, hasil klasifikasi digunakan sebagai dasar pengelompokan data sebelum proses rekomendasi. Dalam *Decision Tree*, proses pembagian data di setiap simpul dilakukan dengan menilai seberapa murni data tersebut melalui nilai *Gini Impurity* (Moerdyanto & Nuryana, 2023) yang dihitung dengan Persamaan (1).

$$Gini(D)\ 1\ -\ \sum_{i=1}^{m} P_i\ 2 \tag{1}$$

Untuk menemukan pemisahan yang paling optimal, nilai ketidakmurnian dihitung ulang untuk setiap bagian dari pemisahan (Khanum dkk., 2023), seperti yang diungkapkan dalam persamaan 2.

$$Gini_A(D) = \frac{|D_1|}{|D|}\ Gini(D_1) + \frac{|D_2|}{|D|}\ Gini(D_2) \tag{2}$$

Perbedaan antara *impurity* pada awal proses dan setelah pemisahan dihitung sebagai *impurity reduction*, menggunakan rumus pada persamaan 3.

ISSN: 2986-0458 (Print) / 2986-044X (Online)
DOI: https://doi.org/10.53624/jsitik.v5i2.971

$$\Delta\, Gini(A) = Gini\ (D) - Gini_A(\mathrm{D}) \quad (3)$$

Genre hasil klasifikasi selanjutnya digunakan sebagai proses penyaringan sehingga perhitungan kemiripan hanya dilakukan pada lagu dengan genre yang sama.

C. *Content-Based Filtering (CBF)*

Content-Based Filtering merupakan metode sistem rekomendasi yang menghasilkan rekomendasi berdasarkan karakteristik atau konten suatu item (Rachmaniar et al., 2025). *Content-Based Filtering* bekerja dengan menganalisis fitur konten lagu seperti metadata dan fitur audio, kemudian membandingkan kemiripan antar lagu untuk menghasilkan rekomendasi (Pake dkk., 2023). Pada penelitian ini, memanfaatkan hasil klasifikasi dari *Random Forest* sebagai acuan dalam proses rekomendasi. Pada penelitian ini setelah proses klasifikasi, sistem menerapkan *Content-Based Filtering* untuk menghasilkan rekomendasi lagu berdasarkan fitur audio. Perhitungan kemiripan hanya dilakukan pada lagu dengan genre yang sama berdasarkan hasil klasifikasi sebelumnya, sehingga rekomendasi yang dihasilkan lebih relevan

D. *Cosine Similarity*

Cosine Similarity merupakan metode yang digunakan untuk mengukur tingkat kemiripan yang diperoleh dari *perhitungan* sudut antar vektor (Ristias dkk., 2024). Pada penelitian ini, *Cosine Similarity* digunakan untuk menghitung kemiripan antar item, hasil perhitungan kemudian digunakan untuk menentukan urutan rekomendasi, sehingga item dengan nilai kemiripan tertinggi akan ditampilkan terlebih dahulu. Nilai yang semakin mendekati 1 menunjukkan tingkat kemiripan yang semakin tinggi (Rabbani dkk., 2023). Nilai kemiripan tersebut dihitung menggunakan Persamaan (4).

$$Cosine\ similarity\ = \frac{(A\ \times\ B)}{\|A\|\ \ \|B\|} \quad (4)$$

E. *Evaluasi Sistem*

Evaluasi *dilakukan* pada dua tahap, yaitu evaluasi model klasifikasi dan evaluasi sistem rekomendasi. Evaluasi model *Random Forest Classifier* dilakukan untuk mengetahui kemampuan model dalam mengklasifikasikan genre lagu berdasarkan fitur audio. Pengujian menggunakan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. *Accuracy* digunakan untuk mengukur tingkat ketepatan prediksi secara keseluruhan, *Precision* mengukur ketepatan prediksi positif yang dihasilkan model, *Recall* mengukur kemampuan model dalam menemukan data yang sesuai pada setiap kelas, sedangkan *F1-Score* digunakan untuk memberikan keseimbangan antara *Precision* dan *Recall*. Perhitungan masing-masing metrik ditunjukkan pada Persamaan (5-8).

ISSN: 2986-0458 (Print) / 2986-044X (Online)
DOI: https://doi.org/10.53624/jsitik.v5i2.971

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (8)$$

Dalam sistem rekomendasi lagu yang menggunakan fitur audio, gabungan metrik ini menjamin bahwa *Random Forest* tidak hanya memiliki akurasi yang tinggi secara keseluruhan, tetapi juga kuat terhadap perbedaan genre, sehingga rekomendasi menjadi lebih tepat dan meningkatkan kepuasan pengguna (Putra & Santika, 2020).

Selanjutnya, evaluasi sistem rekomendasi dilakukan untuk menilai kualitas rekomendasi lagu yang dihasilkan setelah proses *Content-Based Filtering*. Pengujian menggunakan metrik *Precision* dan *Recall* pada skenario Top-10, Top-30, dan Top-50 rekomendasi. *Precision* digunakan untuk mengukur proporsi lagu relevan dalam daftar rekomendasi yang diberikan, sedangkan *Recall* digunakan untuk mengukur kemampuan sistem dalam menemukan seluruh lagu relevan terhadap lagu acuan (Lonang dkk., 2023). Perhitungan kedua metrik ditunjukkan pada Persamaan (9) dan Persamaan (10).

$$Precision = \frac{|Recommended(u) \cap Testing(u)|}{|Recommended(u)|} \quad (9)$$

$$Recall = \frac{|Recommended(u) \cap Recommended(u)|}{|Testing(u)|} \quad (10)$$

III. HASIL DAN PEMBAHASAN

Hasil implementasi sistem rekomendasi lagu berbasis *Content-Based Filtering* dengan *Random Forest Classifier* menunjukkan bahwa sistem mampu melakukan klasifikasi genre dan menghasilkan rekomendasi berdasarkan kemiripan karakteristik lagu. Proses rekomendasi dilakukan setelah tahap klasifikasi genre, kemudian tingkat kemiripan antarlagu dihitung

ISSN: 2986-0458 (Print) / 2986-044X (Online)
DOI: https://doi.org/10.53624/jsitik.v5i2.971

menggunakan *Cosine Similarity*. Evaluasi sistem dilakukan menggunakan metrik Precision dan Recall untuk mengetahui kualitas rekomendasi yang dihasilkan. Temuan penelitian ini adalah penerapan *Random Forest Classifier* sebagai penentu kelompok kandidat genre dan *Content-Based Filtering* sebagai pemeringkat kemiripan audio mampu menghasilkan rekomendasi lagu dengan nilai *Precision* di atas 0,98 pada seluruh skenario pengujian. Rekapitulasi hasil pengujian sistem rekomendasi lagu ditunjukkan pada Tabel 1.

Tabel 1. Hasil Penelitian

No	*Matrix Evaluasi*	*TOP*-10	*TOP*-30	*TOP*-50
1.	***Precision***	**0,9915**	0,9866	0,9832
2.	***Recall***	0,1015	0,3030	**0,5033**

Pengujian dilakukan menggunakan lagu "*Something in the Way*". Hasil rekomendasi ditampilkan pada Gambar 3, dengan lagu "*Imagine – Remastered* 2010" memperoleh nilai *similarity* sebesar 0,9795, diikuti oleh "*Hope for the Future*" sebesar 0,9720, "*Love Is a Bitch*" sebesar 0,9719, "*I'll Be Good*" sebesar 0,9681, dan "*Be My Mistake*" sebesar 0,9679. Nilai kemiripan yang mendekati 1 menunjukkan bahwa lagu yang direkomendasikan memiliki karakteristik yang semakin serupa dengan lagu acuan.

	track_name	genre	similarity
0	imagine - remastered 2010	rock	0.9795
1	hope for the future	rock	0.9720
2	love is a bitch	rock	0.9719
3	i'll be good	rock	0.9681
4	be my mistake	rock	0.9679
..	...	...	...
95	just maath maathali	rock	0.8411
96	bohemian rhapsody - remastered 2011	rock	0.8408
97	transito lento	rock	0.8398
98	sympathy for the devil	rock	0.8397
99	me hace tanto bien	rock	0.8396

[100 rows x 3 columns]

Gambar 3. Hasil Rekomendasi

Selanjutnya, evaluasi dilakukan terhadap model Random Forest Classifier untuk mengetahui pengaruh pengolahan data genre terhadap kemampuan klasifikasi. Pengujian dilakukan menggunakan tiga skenario, yaitu penggunaan seluruh data genre, mapping genre dengan mengelompokkan genre serupa ke kategori yang lebih umum, serta pemilihan genre utama dengan jumlah kategori yang lebih terfokus. Hasil evaluasi masing-masing skenario ditunjukkan pada Tabel 2.

Tabel 2. Hasil Evaluasi Random Forest

No	Skenario Penelitian	*Accuracy*	*Precision*	*Recall*	*F1-Score*
1.	Semua Data	0,231	0,220	0,231	0,232
2.	*Mapping*	0,601	0,600	0,601	0,595
3.	Pemilihan Genre	**0,646**	**0,645**	**0,646**	**0,644**

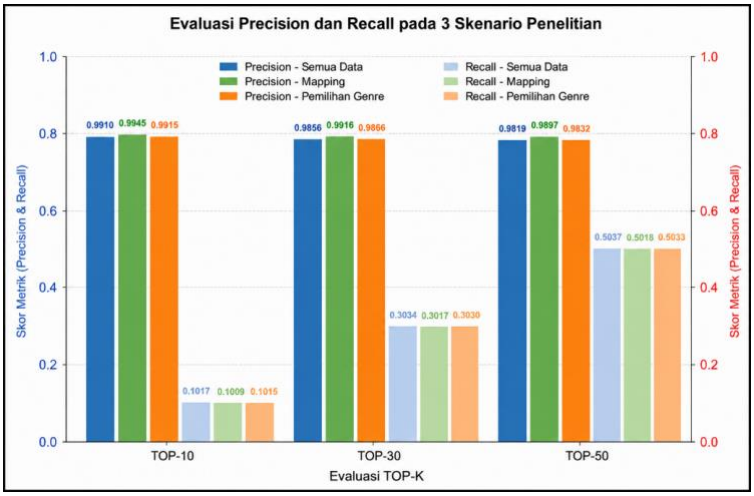
ISSN: 2986-0458 (Print) / 2986-044X (Online)
DOI: https://doi.org/10.53624/jsitik.v5i2.971

Berdasarkan Tabel 2, skenario penggunaan seluruh data genre menghasilkan performa terendah dengan nilai *Accuracy* sebesar 0,231. Hal ini menunjukkan bahwa banyaknya variasi genre menyebabkan model mengalami kesulitan dalam mengenali pola data. Skenario *mapping* genre meningkatkan nilai *Accuracy* menjadi 0,601, sedangkan skenario pemilihan genre memperoleh hasil terbaik dengan nilai *Accuracy* sebesar 0,646. Hasil tersebut menunjukkan bahwa pengelompokan kategori genre yang lebih terstruktur dapat membantu model dalam mengenali pola data dengan lebih baik. Evaluasi sistem rekomendasi kemudian dilakukan berdasarkan tiga skenario pengolahan genre menggunakan pendekatan Top-K untuk mengetahui pengaruh pengolahan data terhadap kualitas rekomendasi. Hasil pengujian ditampilkan pada Tabel 3.

Tabel 3. Hasil Evaluasi Eksperimen

No	Skenario Penelitian	Evaluasi	*Top*-10	*Top*-30	*Top*-50
1.	Semua Data	*Precision*	**0,9910**	0,9856	0,9819
		Recall	0,1017	0,3034	**0,5037**
2.	Mapping	*Precision*	**0,9945**	0,9916	0,9897
		Recall	0,1009	0,3017	**0,5018**
3.	Pemilihan Genre	*Precision*	**0,9915**	0,9866	0,9832
		Recall	0,1015	0,3030	**0,5033**

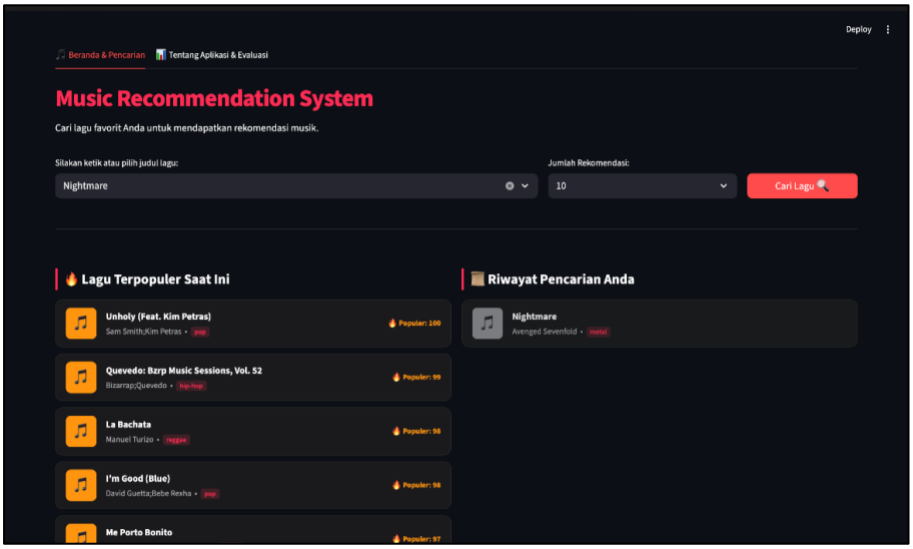
Berdasarkan Tabel 3, seluruh skenario menghasilkan nilai *Precision* di atas 0,98, yang menunjukkan bahwa sebagian besar lagu yang direkomendasikan memiliki tingkat relevansi yang tinggi. Skenario *mapping* genre memperoleh nilai *Precision* tertinggi, yaitu 0,9945 pada Top-10. Namun, skenario pemilihan genre menghasilkan distribusi data yang lebih seimbang sehingga lebih representatif dalam menggambarkan kemampuan sistem. Pada skenario tersebut, nilai *Recall* meningkat dari 0,1015 pada Top-10 menjadi 0,5033 pada Top-50. Hasil evaluasi pada Tabel 3 kemudian divisualisasikan dalam bentuk grafik perbandingan pada Gambar 4.



Gambar 4. Grafik Perbandingan

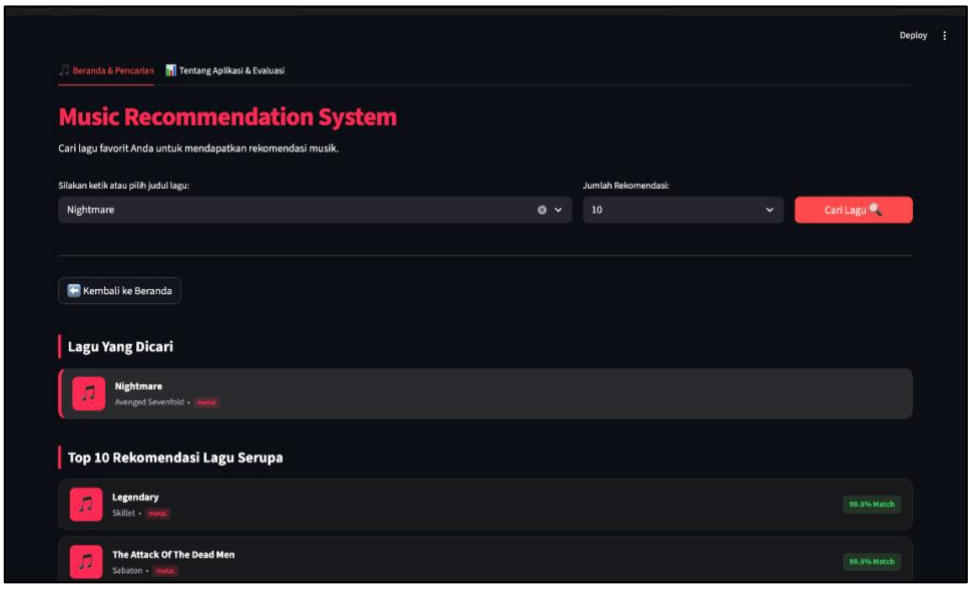
JSITIK, Vol.5 No.2 Juni 2027
ISSN: 2986-0458 (Print) / 2986-044X (Online)
DOI: https://doi.org/10.53624/jsitik.v5i2.971

Pada Gambar 4 menunjukkan grafik perbandingan dimana hasil penelitian ini sejalan atau didukung oleh penelitian (Timur & Rohman, 2025) yang menerapkan pendekatan *Content-Based Filtering* dengan mengukur kemiripan karakteristik item untuk menghasilkan rekomendasi yang relevan. Pendekatan tersebut mendukung penggunaan *Cosine Similarity* dalam penelitian ini karena metode tersebut dapat mengukur kemiripan vektor karakteristik audio antar lagu.



Gambar 5. Halaman Utama Sistem Rekomendasi

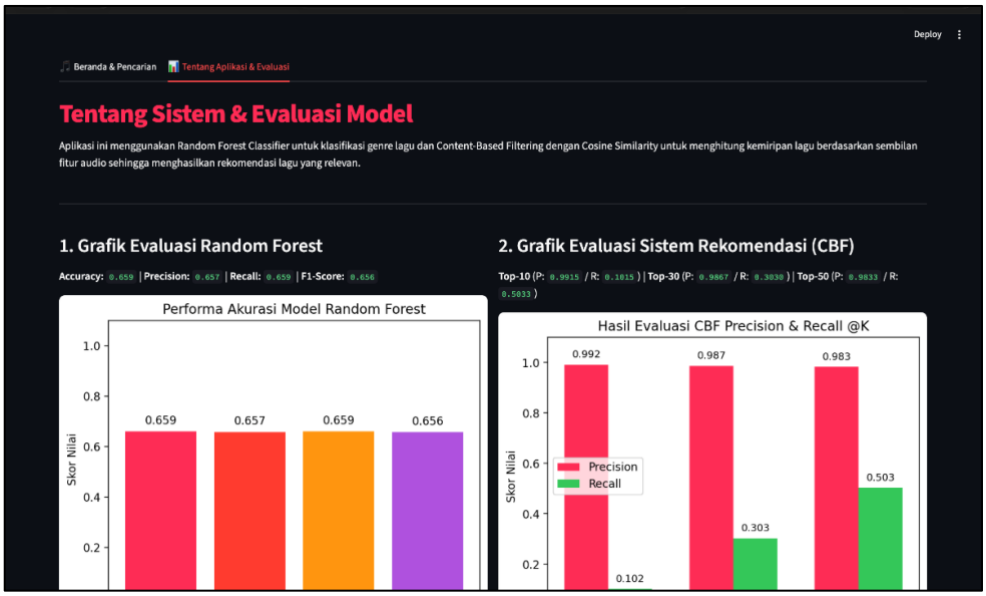
Gambar 5 menunjukkan halaman utama sistem rekomendasi lagu berbasis website. Halaman ini menyediakan fitur pencarian lagu, pemilihan jumlah rekomendasi, serta menampilkan informasi pendukung seperti lagu populer dan riwayat pencarian. Selanjutnya, halaman hasil rekomendasi ditampilkan seperti pada Gambar 6.



Gambar 6. Halaman Hasil Sistem Rekomendasi

ISSN: 2986-0458 (Print) / 2986-044X (Online)
DOI: https://doi.org/10.53624/jsitik.v5i2.971

Gambar 6 menunjukkan halaman hasil rekomendasi yang menampilkan daftar lagu dengan karakteristik serupa berdasarkan hasil perhitungan *Cosine Similarity*. Informasi yang ditampilkan meliputi judul lagu, nama artis, genre, dan tingkat kemiripan terhadap lagu acuan. Selanjutnya, halaman evaluasi sistem ditampilkan seperti pada Gambar 7.



Gambar 7. Halaman Evaluasi Sistem Rekomendasi

Gambar 7 menunjukkan halaman evaluasi sistem yang menampilkan hasil pengujian model *Random Forest* dan sistem rekomendasi. Halaman ini memuat visualisasi metrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score* untuk evaluasi klasifikasi, serta nilai *Precision* dan *Recall* pada skenario Top-10, Top-30, dan Top-50 rekomendasi.

Penelitian ini memiliki keterbatasan pada distribusi genre dalam dataset yang belum sepenuhnya seimbang serta penggunaan fitur audio tanpa mempertimbangkan preferensi individual pengguna. Hasil pengujian menunjukkan bahwa pengolahan data melalui mapping dan pemilihan genre mampu meningkatkan performa model. Penelitian selanjutnya dapat melakukan penyeimbangan distribusi data dan menambahkan fitur audio yang relevan untuk meningkatkan representasi karakteristik lagu. Selain itu, penerapan metode *hybrid recommendation* dengan menggabungkan *Content-Based Filtering* dan *Collaborative Filtering* dapat dipertimbangkan agar rekomendasi tidak hanya berdasarkan kemiripan karakteristik lagu, tetapi juga mempertimbangkan pola preferensi pengguna.

IV. KESIMPULAN

Berdasarkan hasil penelitian, penerapan Random Forest Classifier dan Content-Based Filtering mampu menghasilkan rekomendasi lagu melalui klasifikasi genre dan perhitungan

ISSN: 2986-0458 (Print) / 2986-044X (Online)
DOI: https://doi.org/10.53624/jsitik.v5i2.971

kemiripan karakteristik audio menggunakan Cosine Similarity. Pada skenario pemilihan genre, model Random Forest Classifier memperoleh nilai Accuracy sebesar 0,646, sedangkan evaluasi rekomendasi menghasilkan nilai Precision@10 sebesar 0,9915 dan Recall@50 sebesar 0,5033. Hasil tersebut menunjukkan bahwa sistem mampu menghasilkan rekomendasi dengan tingkat kemiripan yang tinggi, serta memperluas cakupan lagu relevan seiring bertambahnya jumlah rekomendasi. Namun, penelitian ini masih terbatas pada penggunaan karakteristik audio dan belum mempertimbangkan preferensi individual pengguna, sehingga rekomendasi yang dihasilkan belum sepenuhnya personal. Oleh karena itu, penelitian selanjutnya disarankan untuk mempertimbangkan preferensi pengguna serta menambahkan fitur audio yang relevan guna meningkatkan representasi karakteristik lagu dan kualitas rekomendasi.

DAFTAR PUSTAKA

Abrah, R., Hayat, M. A. M., & Lukman. (2025). Penerapan Machine Learning untuk Mengklasifikasikan Genre Musik Berdasarkan Fitur Audio. Jurnal Informatika dan Teknik Elektro Terapan, 3(2), 204–211. https://doi.org/10.57250/ajst.v3i2.1699

Firnanda, P. A., Shofwatillah, L., Rahma, F., & Fauzi, F. (2025). Analisis Perbandingan Decision Tree dan Random Forest dalam Klasifikasi Penjualan Produk pada Supermarket. Emerging Statistics and Data Science Journal, 3(1). https://doi.org/https://doi.org/10.20885/esds.vol3.iss.1.art2

Khanum, H., Garg, A., & Faheem, M. I. (2023). Accident Severity Prediction Modeling for Road Safety Using Random Forest Algorithm: An Analysis of Indian Highways. F1000Research, 12, 494. https://doi.org/10.12688/f1000research.133594.1

Lonang, S., Yudhana, A., & Biddinika, M. K. (2023). Analisis Komparatif Kinerja Algoritma Machine Learning untuk Deteksi Stunting. Jurnal Media Informatika Budidarma, 7(4), 2109. https://doi.org/10.30865/mib.v7i4.6553

Moerdyanto, O. P., & Nuryana, I. K. D. (2023). Prediksi Kelulusan Tepat Waktu Menggunakan Pendekatan Pohon Keputusan Algoritma Decision Tree. Journal of Informatics and Computer Science, 05. https://doi.org/https://doi.org/10.26740/jinacs.v5n01.p90-96

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. https://doi.org/10.30595/sainteks.v22i2.27815

Murtopo, A. A., Haryadi, D., & Fadilah, N. (2022). Penerapan Model Information Retrieval Untuk Pencarian Konten Pada Perpustakaan Digital. JUPTI (Jurnal Publikasi Teknik Informatika), 1(3). https://doi.org/doi.org/10.55606/jupti.v1i3.514

Pake, G. V. R., Mawardi, V. C., & Sutrisno, T. (2023). Penerapan Metode Content-Based Filtering dalam Implementasi Sistem Rekomendasi Musik. Jurnal Serina Sains, Teknik dan Kedokteran, 1(2), 455–462. https://doi.org/10.24912/jsstk.v1i2.31037

Putra, A. I., & Santika, R. R. (2020). Edumatic: Jurnal Pendidikan Informatika Implementasi Machine Learning dalam Penentuan Rekomendasi Musik dengan Metode Content-Based Filtering. 4(1). https://doi.org/10.29408/edumatic.v4i1.2162

Rabbani, B. I., Napitupulu, H., & Hertini, E. (2023). Sistem Rekomendasi Menggunakan Item-Based Collaborative Filtering Berbasis Mobile iOS. SISINFO : Jurnal Sistem Informasi Dan Informatika, 5(2), 50–56. https://doi.org/https://doi.org/10.37278/sisinfo.v5i2.671

ISSN: 2986-0458 (Print) / 2986-044X (Online)
DOI: https://doi.org/10.53624/jsitik.v5i2.971

Rachmaniar, A., Widayati, S., & Rokoyah, K. (2025). Sistem Rekomendasi Prosuk E-commerce Menggunakan Collaborative Filtering dan Content-Based Filtering. Journal of Information System, Informatics and Computing Issue Period, 9(1), 40–54. https://doi.org/10.52362/jisicom.v9i1.1904

Ristias, A. A., Wahyuni, E. D., & Wati, S. F. A. (2024). Komparasi Kinerja Metode Cosine dan Jaccard Similarity dalam Content-Based Recommendation System (CBRS) pada Aplikasi Eventhings. Jurnal Informatika dan Teknik Elektro Terapan, 12(3). https://doi.org/10.23960/jitet.v12i3.4738

Saifudin, I., & Widiyaningtyas, T. (2024). Systematic Literature Review on Recommender System: Approach, Problem, Evaluation Techniques, Datasets. IEEE Access, 12, 19827–19847. https://doi.org/10.1109/ACCESS.2024.3359274

Timur, A. Y., & Rohman, A. N. (2025). Sistem Rekomendasi Lagu Indonesia Menggunakan Metode Content-Based Filtering dan Cosine Similarity. Jurnal Informatika dan Teknik Elektro Terapan, 13(1). https://doi.org/10.23960/jitet.v13i1.5949

Widiyaningtyas, T., Hidayah, I., & Adji, T. B. (2021). Recommendation Algorithm using Clustering-Based UPCSim (Cb-UPCsim). Computers, 10(10). https://doi.org/10.3390/computers10100123