

**IMPLEMENTASI SISTEM REKOMENDASI LAGU
MENGUNAKAN METODE CONTENT BASED FILTERING
& COSINE SIMILARITY**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

Rizka Amella Sari

22.12.2603

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**IMPLEMENTASI SISTEM REKOMENDASI LAGU
MENGUNAKAN METODE CONTENT BASED FILTERING
& COSINE SIMILARITY**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Sistem Informasi



disusun oleh

RIZKA AMELIA SARI

22.12.2603

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

SKRIPSI

**IMPLEMENTASI SISTEM REKOMENDASI LAGU MENGGUNAKAN
METODE CONTENT BASED FILTERING
& COSINE SIMILARITY**

yang disusun dan diajukan oleh

Rizka Amelia Sari

22.12.2603

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 9 Mei 2026

Dosen Pembimbing,



Ir. Akhmad Dahlan, S.Kom., M.Kom
NIK. 190302174

HALAMAN PENGESAHAN

SKRIPSI

IMPLEMENTASI SISTEM REKOMENDASI LAGU MENGGUNAKAN METODE CONTENT BASED FILTERING & COSINE SIMILARITY

Yang disusun dan diajukan oleh

Rizka Amelia Sari

22.12.2603

Telah dipertahankan di depan Dewan Penguji
pada tanggal 18 Mei 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Bety Wulan Sari, S.Kom., M.Kom
NIK. 190302254

Ikmal, S.Kom., M.Kom
NIK. 190302282

Akhmad Dahlan, S.Kom, M
NIK. 190302174

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 18 Mei 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom

NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama Mahasiswa : Rizka Amelia Sari

NIM : 22.12.2603

Menyatakan bahwa Skripsi dengan judul berikut:

IMPLEMENTASI SISTEM REKOMENDASI LAGU MENGGUNAKAN METODE CONTENT BASED FILTERING & COSINE SIMILARITY

Dosen Pembimbing : Ir. Akhmad Dahlan, S.Kom, M.Kom

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian **SAYA** sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab **SAYA**, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini **SAYA** buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka **SAYA** bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, **serta sanksi lainnya** sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 18 Mei 2026

Yang Menyatakan,



Rizka Amelia Sari

HALAMAN PERSEMBAHAN

Dengan rasa Syukur yang mendalam dan hikmat atas kesempatan ini, Terima kasih kepada Allah SWT yang telah memberikan kepercayaan kepada hamba-Nya untuk menyelesaikan skripsi ini, dengan telah diselesaikan nya skripsi dengan judul "Implementasi Sistem Rekomendasi Lagu Menggunakan Metode *Content Based Filtering* dan *Cosine Similarity*". Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta. Topik ini dipilih karena banyaknya pilihan lagu pada platform musik digital membuat pengguna sering kesulitan menemukan lagu yang sesuai dengan selera mereka. Oleh karena itu, penulis tertarik mengembangkan sistem rekomendasi musik berdasarkan kemiripan karakteristik lagu.

Selesaiannya skripsi ini penulis akui tidak terlepas dari bantuan beberapa pihak. Untuk itu penulis mengucapkan terima kasih kepada :

1. Teruntuk kedua orang tua saya, Bapak (Sunarso) dan Ibu (Kriswanti), yang selalu memberikan kasih sayang, doa, dukungan moral maupun materi, serta menjadi alasan terbesar bagi penulis untuk tetap berjuang menyelesaikan pendidikan ini.
2. Mas Naufal yang membantu, membimbing, serta memberikan dukungan selama proses pengerjaan skripsi ini hingga dapat diselesaikan dengan baik.
3. Ivel, Hesti, Ulvi, Najwa, Anggi serta seluruh teman-teman yang tidak dapat disebutkan satu per satu, yang telah menjadi bagian dari perjalanan perkuliahan penulis, memberikan bantuan, semangat, dan kenangan selama proses penyusunan skripsi ini.
4. Seluruh coffee shop di Yogyakarta yang secara tidak langsung telah menjadi tempat penulis mencari suasana baru, mengerjakan revisi, menyusun penelitian, dan menghabiskan banyak waktu selama proses penyelesaian skripsi ini.

KATA PENGANTAR

Alhamdulillah segala puji dan syukur penulis panjatkan kehadiran Allah Subhuanallahu wata'ala atas berkah dan Rahmat Nya yang telah dilimpahkan kepada kita semua khususnya nikmat sehat yang tak ternilai harganya, sehingga penulis dapat menyelesaikan Skripsi dengan baik, tepat waktu yang telah ditentukan. Dalam penulisan Skripsi ini penulis banyak mendapatkan bimbingan dan dukungan serta saran yang begitu besar manfaatnya, baik dari dosen maupun rekan mahasiswa. Oleh karena itu penulis ingin menyampaikan terimakasih kepada:

1. Pak Prof. Dr. M. Suyanto, M.M selaku Rector Universitas Amikom Yogyakarta, yang telah memberikan kesempatan kepada penulis untuk menempuh pendidikan dan mendapatkan pengalaman belajar di Universitas Amikom Yogyakarta.
2. Bu Prof.Dr Kusrini, M.Kom selaku Dekan Fakultas Ilmu Komputer, yang telah memberikan dukungan dalam proses akademik selama perkuliahan.
3. Pak Anggit Dwi Hartanto, M.Kom selaku Ketua Prodi Sistem Informasi, yang telah memberikan arahan dan dukungan kepada mahasiswa selama menjalani proses perkuliahan.
4. Pak Ir.Akhmad Dahlan, S.Kom., M.Kom selaku Dosen Pembimbing, yang dengan sabar telah memberikan arahan, masukan, ilmu, serta motivasi kepada penulis selama proses penyusunan skripsi ini hingga dapat diselesaikan dengan baik.
5. *Last but not least, I wanna thank me, for believing in me, for doing all this hardwork , for having no days off never quilting, for just being me at all times.*

Yogyakarta, 18 Mei 2026

Rizka Amelia Sari

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	x
DAFTAR LAMPIRAN.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xiv
INTISARI	xvi
<i>ABSTRACT</i>	xvii
BAB I PENDAHULUAN.....	1
1. Latar Belakang	1
2. Rumusan Masalah	2
3. Batasan Masalah	2
4. Tujuan Penelitian	3
5. Manfaat Penelitian	3
6. Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5

2.2	Dasar Teori.....	21
BAB III METODE PENELITIAN		27
3.1	Objek Penelitian.....	27
3.2	Alur Penelitian	27
3.3	Alat dan Bahan.....	32
BAB IV HASIL DAN PEMBAHASAN		34
4.1.	Dataset.....	34
4.2.	Data <i>Preprocessing</i>	35
4.3.	Representasi Data.....	38
4.4.	Implementasi Sistem Rekomendasi	44
4.5.	Hasil Dan Pengujian Sistem.....	54
4.6.	Deployment.....	63
BAB V PENUTUP		81
5.1	Kesimpulan	81
5.2	Saran	81
REFERENSI		82
LAMPIRAN		88
1.	Surat Penunjukan Dosen Pembimbing.....	88
2.	Publish Web	88

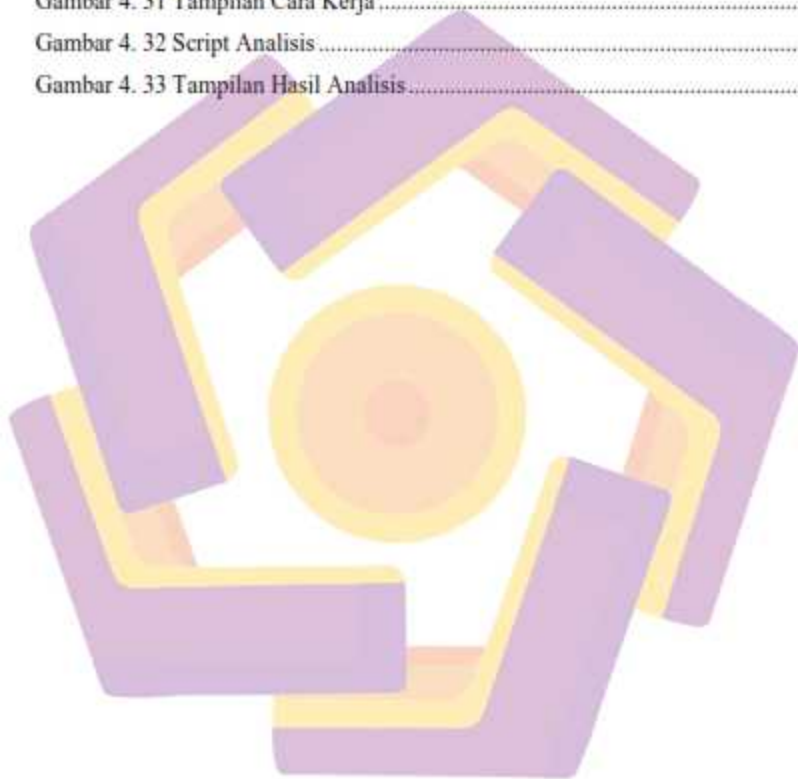
DAFTAR TABEL

Table 2. 1 Keaslian Penelitian	11
Table 4. 1 Data Lagu	38
Table 4. 2 TF	39
Table 4. 3 TF IDF	41
Table 4. 4 Script TF IDF	42
Table 4. 5 Hasil representasi vektor	42
Table 4. 6 script representasi vektor	44
Table 4. 7 Cosine similarity (menggunakan data tf idf)	44
Table 4. 8 perkalian titik	45
Table 4. 9 Representasi TF IDF	49
Table 4. 10 Matriks Similarity	51
Table 4. 11 data mentah evaluasi	56
Table 4. 12 perhitungan manual precision	57
Table 4. 13 Perhitungan manual recall	58
Table 4. 14 Perhitungan manual fl-score	60

DAFTAR GAMBAR

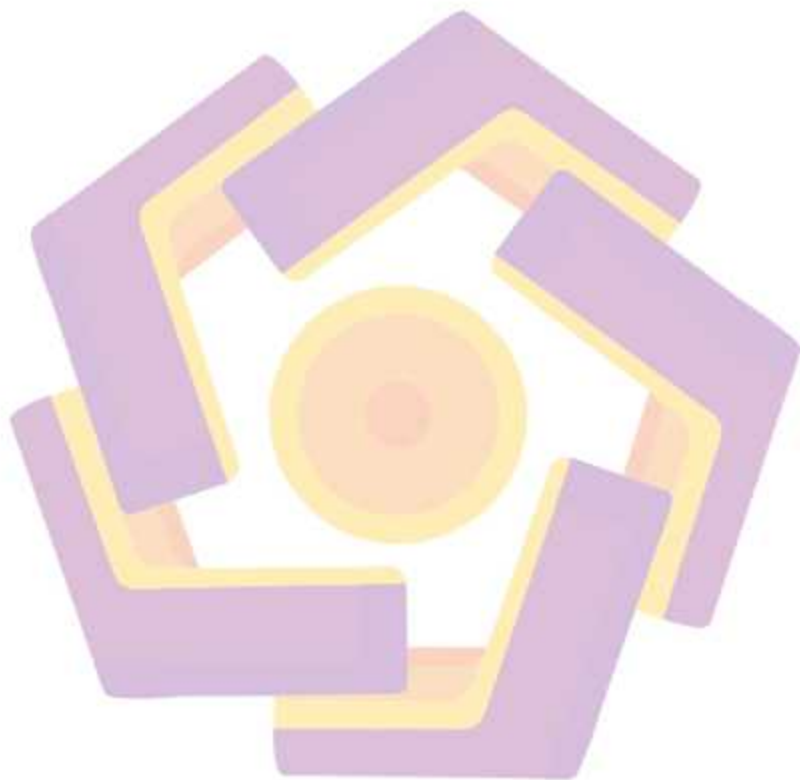
Gambar 2. 1 Rumus Cosine Similarity	22
Gambar 2. 2 Rumus Precision.....	25
Gambar 2. 3 Rumus Recall	26
Gambar 2. 4 Rumus f1-score	26
Gambar 3. 1 Alur Penelitian	28
Gambar 4. 1 Dataset Musik	34
Gambar 4. 2 Script Data Cleaning	35
Gambar 4. 3 Hasil Data Cleaning	36
Gambar 4. 4 Script Case Folding & Tokenization.....	36
Gambar 4. 5 Hasil Tokenisasi & Case Folding.....	37
Gambar 4. 6 Script Normalisasi.....	37
Gambar 4. 7 Hasil Normalisasi.....	38
Gambar 4. 8 script similarity	49
Gambar 4. 9 Script matriks similarity.....	52
Gambar 4. 10 Script pengambilan rekomendasi.....	53
Gambar 4. 11 Script hasil rekomendasi	54
Gambar 4. 12 Gambar Hasil Rekomendasi.....	55
Gambar 4. 13 Script precision.....	58
Gambar 4. 14 Script recall	59
Gambar 4. 15 script f1-score.....	61
Gambar 4. 16 Script evaluasi full	62
Gambar 4. 17 Hasil fungsi evaluasi	63
Gambar 4. 18 Script home	64
Gambar 4. 19 Tampilan Home.....	65
Gambar 4. 20 Script menu rekomendasi 1	66
Gambar 4. 21 Script menu rekomendasi 2	67
Gambar 4. 22 Tampilan Rekomendasi.....	68
Gambar 4. 23 Tampilan Rekomendasi 2.....	69
Gambar 4. 24 Script evaluasi sistem 1	70
Gambar 4. 25 Script evaluasi sistem 2.....	71

Gambar 4. 26 Tampilan Evaluasi Perhitungan	72
Gambar 4. 27 Tampilan Detail Evaluasi	73
Gambar 4. 28 Script eksplorasi data	74
Gambar 4. 29 Tampil eksplorasi data	75
Gambar 4. 30 Script cara kerja.....	76
Gambar 4. 31 Tampilan Cara Kerja	77
Gambar 4. 32 Script Analisis	78
Gambar 4. 33 Tampilan Hasil Analisis	79

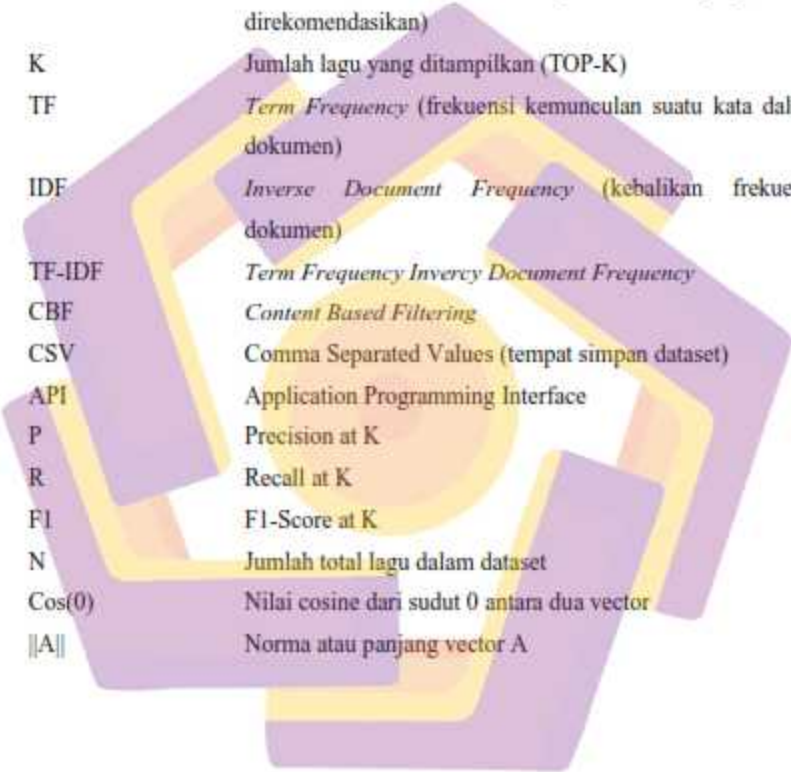


DAFTAR LAMPIRAN

Lampiran 1 surat penunjukkan dosen pembimbing	88
Lampiran 2 publish web.....	88



DAFTAR LAMBANG DAN SINGKATAN



TP	<i>True Positive</i> (jumlah rekomendasi yang relevan)
FP	<i>False Positive</i> (jumlah rekomendasi yang tidak relevan)
FN	<i>False Negative</i> (jumlah lagu relevan yang tidak direkomendasikan)
K	Jumlah lagu yang ditampilkan (TOP-K)
TF	<i>Term Frequency</i> (frekuensi kemunculan suatu kata dalam dokumen)
IDF	<i>Inverse Document Frequency</i> (kebalikan frekuensi dokumen)
TF-IDF	<i>Term Frequency Invercy Document Frequency</i>
CBF	<i>Content Based Filtering</i>
CSV	Comma Separated Values (tempat simpan dataset)
API	Application Programming Interface
P	Precision at K
R	Recall at K
F1	F1-Score at K
N	Jumlah total lagu dalam dataset
$\text{Cos}(0)$	Nilai cosine dari sudut 0 antara dua vector
$\ A\ $	Norma atau panjang vector A

DAFTAR ISTILAH



<i>Content Based Filtering</i>	Metode sistem rekomendasi yang memberikan saran berdasarkan kemiripan konten antar item, tanpa data interaksi pengguna lain.
Cosine Similarity	Untuk menghitung tingkat kemiripan (similarity) antar dua buah objek yang berbobot.
<i>Collaborative Filtering Metode</i>	Sistem rekomendasi berdasarkan kemiripan perilaku pengguna
Dataset	Kumpulan data yang disimpan dalam format csv, untuk digunakan sebagai bahan utama dalam proses penelitian dan pelatihan model.
<i>Pre processing</i>	Proses pembersihan data mentah sebelum digunakan pada proses pemodelan.
<i>Data cleaning</i>	Proses menghapus data yang tidak sesuai, duplikat, kosong atau tidak konsisten dari dataset.
<i>Case folding</i>	Proses mengubah seluruh teks menjadi huruf kecil untuk menyamakan penulisan sebelum di proses lebih lanjut.
Tokenisasi	Proses memecah teks menjadi satuan kata atau unit lebih kecil yang disebut token.
Normalisasi	Proses mengubah nilai data ke rentang 0 hingga 1 menggunakan metode Min-Max Scaling.
Vektor	Representasi data dalam bentuk angka untuk menggambarkan karakteristik suatu item.
Matriks Similarity	Tabel yang berisi nilai kemiripan antar lagu pada dataset.

Deployment

Tahap penerapan model yang sudah diuat dalam bentuk web yang dapat diakses oleh pengguna.

Scrapping

Teknik pengumpulan data secara otomatis dari spotify menggunakan perangkat lunak.

Streamlit

Framework Pyhton yang digunakan untuk menjalankan web berbasis data secara cepat.



INTISARI

Lagu merupakan Kumpulan kata yang telah di susun menjadi sebuah text dan dinyanyikan dengan iringan *Music*. Tetapi, dengan semakin banyak nya lagu yang tersedia di *platform*, pengguna mengalami kesulitan untuk menemukan lagu yang sesuai dengan selera. Sistem rekomendasi adalah suatu teknologi yang digunakan untuk memberikan saran atau pilihan kepada pengguna berdasarkan data dan pola tertentu yang diperoleh dari perilaku pengguna tersebut. Pada penelitian ini menggunakan metode *content-based filtering*. Untuk tahapan penelitian meliputi pengumpulan *dataset* lagu. Selanjutnya di lakukan *Preprocessing* untuk menghapus data yang tidak sesuai. Serta penerapan algoritma *Cosine similarity* untuk menghitung tingkat kemiripan antar lagu berdasarkan nilai fitur yang telah diekstraksi. Lalu hasil perhitungan tersebut digunakan untuk menampilkan beberapa rekomendasi lagu yang memiliki karakteristik paling mirip dengan lagu yang telah di inputkan oleh pengguna. Hasil pengujian menunjukkan sistem menghasilkan rekomendasi yang cukup akurat. Sistem ini membantu menemukan lagu dengan suasana atau pun karakteristik yang serupa tanpa harus mencari secara manual.

Kata kunci: Sistem rekomendasi, lagu, *Content based filtering*, *dataset*, rekomendasi musik

ABSTRACT

A song is a collection of words arranged into lyrics and sung to musical accompaniment. However, with the increasing number of songs available on platforms, users find it difficult to find songs that match their tastes. A recommendation system is a technology used to provide suggestions or options to users based on specific data and patterns derived from their behavior. This study employs the content-based filtering method. The research stages include the collection of a song dataset. Next, preprocessing is performed to remove irrelevant data. Additionally, the Cosine similarity algorithm is applied to calculate the similarity between songs based on the extracted feature values. The results of these calculations are then used to display several song recommendations that share the most similar characteristics with the song inputted by the user. Testing results indicate that the system generates fairly accurate recommendations. This system helps users find songs with a similar mood or characteristics without having to search manually.

Keyword: Recommendation system, songs, content-based filtering, dataset, Music recommendations