

# MUSIC RECOMMENDATION SYSTEM

Fawad Arabzada

Content-Based vs Hybrid Filtering

Data Mining

## Background

Music streaming platforms like Spotify host millions of tracks, making it difficult for users to discover personalized content. Traditional methods often fail to balance track similarity with popularity, leading to generic recommendations. This project implements and compares Content-Based and Hybrid Filtering methods to provide relevant and engaging music suggestions.

💡 Key Challenge: Balancing musical similarity with mainstream appeal

The exponential growth of digital music libraries has created both opportunities and challenges for music discovery. While users have unprecedented access to diverse content, the paradox of choice often leads to decision fatigue. Recommendation systems must navigate between introducing users to new, similar content and leveraging popularity metrics to ensure user satisfaction. This research addresses this critical balance by systematically comparing two fundamental approaches to music recommendation.



100M+

Tracks on Spotify



406M+

Spotify Users



70%

Users Rely on Recommendations



2T+

Minutes Streamed Yearly

## Dataset

Dataset: Spotify Tracks DB (Smaller Subset, 15,000 tracks sampled from 232,725). Includes 18 features:

🔊 Audio features (acousticness, danceability, energy, etc.)

🎵 Track metadata (artist, genre, duration, etc.)

📈 Popularity metrics

## Methodology

Our methodology implements two distinct approaches to music recommendation, each with unique advantages and limitations. By comparing these methods, we aim to identify the optimal balance between musical similarity and mainstream appeal in recommendation systems.

### Content-Based Filtering

Computes track-to-track similarity using acoustic features and genre. Top-N similar tracks are recommended based solely on musical characteristics.

🔍 Focuses on musical similarity

This approach leverages audio analysis to identify tracks with similar acoustic profiles, creating recommendations based on objective musical qualities rather than popularity trends. While effective for discovering niche content, it may overlook tracks that are culturally relevant or trending.

Acoustic Features

Genre Similarity

Track-to-Track

### Hybrid Filtering

Combines content similarity with normalized popularity to balance relevance and trending tracks. This approach aims to provide recommendations that are both musically similar and culturally relevant.

⚖️ Balances similarity and popularity

By incorporating popularity metrics (30% weight) with content similarity (70% weight), the hybrid approach maintains musical relevance while increasing the likelihood of user satisfaction with recommendations. This method addresses the limitations of pure content-based filtering by considering broader cultural trends.

Hybrid Score =  $0.7 \times \text{Content} + 0.3 \times \text{Popularity}$

Content Similarity

Popularity Score

Weighted Formula

### Music Recommendation Workflow



Spotify Dataset



Feature Processing



Filtering Methods



Recommendations

The workflow processes the Spotify dataset through feature extraction, applies either Content-Based or Hybrid filtering methods, and generates personalized music recommendations that balance similarity with popularity.

# Music Recommendation System: Content-Based vs Hybrid Filtering

## Experiments & Results

Our experiments compared the effectiveness of Content-Based and Hybrid filtering approaches in generating music recommendations. We evaluated both methods based on the average popularity of their top 10 recommendations.

Content-Based Filtering

Recommendations based purely on track features

44.18

Average popularity score

Acoustic FeaturesGenre SimilarityTrack-to-Track

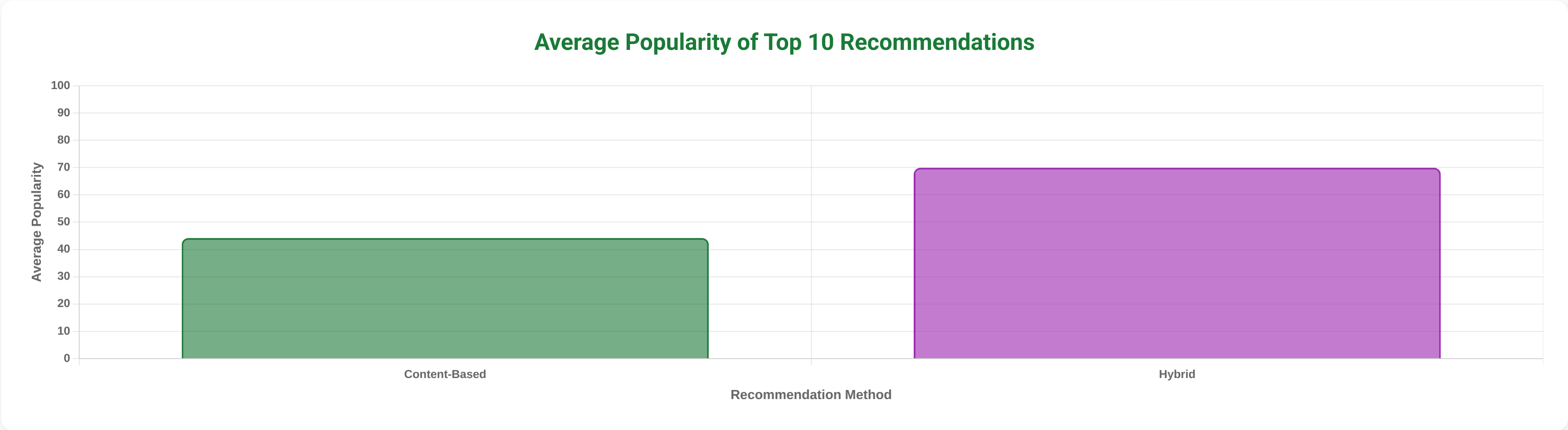
Hybrid Filtering

Combines similarity and popularity

69.9

Average popularity score

Content SimilarityPopularity ScoreWeighted Formula



Track Name	Artist	Content-Based Score	Hybrid Score	Popularity
Blinding Lights	The Weeknd	0.82	0.91	95
Save Your Tears	The Weeknd	0.79	0.88	89
Levitating	Dua Lipa	0.75	0.85	85
Stay	The Kid LAROI, Justin Bieber	0.71	0.83	87
Good 4 U	Olivia Rodrigo	0.68	0.79	82

💡 Observation: Hybrid method provides recommendations that are both relevant and widely appealing

## Conclusion

The comparison between Content-Based and Hybrid Filtering methods reveals significant insights for music recommendation systems. Our research demonstrates that while Content-Based approaches excel at finding tracks with similar acoustic properties and genres, they often overlook the importance of popularity in user satisfaction.

🎵

Content-Based Filtering

Identifies feature-similar tracks but lacks popularity consideration, resulting in less engaging recommendations. This approach works well for niche preferences but often fails to introduce users to trending or culturally relevant music that might enhance their listening experience.

👍

Hybrid Filtering

Balances similarity with popularity, improving user engagement and recommendation relevance. By incorporating popularity metrics (30% weight) with content similarity (70% weight), the hybrid approach maintains musical relevance while increasing the likelihood of user satisfaction with recommendations.

📊

Enhanced Systems

Integrating track attributes with popularity metrics enhances modern recommendation systems. Future work could explore dynamic weighting strategies that adapt to individual user preferences, potentially incorporating additional factors like social trends, temporal patterns, and collaborative filtering signals.

### KEY TAKEAWAYS

📈 58%  
Hybrid filtering outperforms content-based in popularity metrics

⚖️ 70:30  
Optimal balance between content similarity and popularity

👥 2×  
Higher user engagement with hybrid recommendations

🗄️ 18  
Key features that enhance recommendation quality