MMSR Task 1 Report: Content-Based Music Retrieval System

Zaunmayr
k01457265

Vlad
k01575100

Delev
k00975000

Mirmanova
k12334743

Suleimenov
k12247291

1 INTRODUCTION

This report presents our work in the "Multimedia Search and Retrieval" course, focusing on the creation of a content-based music retrieval system using a subset of the Music4All-Onion dataset. We explore various retrieval methods, each employing different text-based features and similarity measures to analyze and identify similar songs.

2 METHODOLOGY

2.1 Data Preparation

We used a subset of the Music4All-Onion dataset, which required no special preparation for our analysis.

2.2 Library and Tool Utilization

Our analysis relied on Python libraries including Pandas, NumPy, and scikit-learn, primarily for data manipulation, numerical computations, and implementing machine learning algorithms.

2.3 Implementation of Tasks

2.3.1 Task 1: Random Baseline. We implemented a system that randomly selects tracks from the dataset, serving as a baseline for comparison with other methods.

2.3.2 Task 2: Text-Based Retrieval with TF-IDF and Cosine Similarity. This task involved developing a retrieval system using TF-IDF vectors and cosine similarity to find songs similar to a given query based on lyrical content.

2.3.3 Task 3: Text-Based Retrieval with Alternative Features. We explored retrieval systems using different text-based features such as BERT and Word2Vec, along with various similarity measures.

2.3.4 Task 4: Text-Based Retrieval with New Similarity and Feature Combination. We experimented with a new combination, utilizing Word2Vec and Euclidean distance, to assess its effectiveness in song retrieval compared to traditional methods.

3 RESULTS AND DISCUSSION

3.1 Qualitative Analysis of Retrieved Tracks

For the qualitative analysis, we selected three familiar tracks as queries and used each of the four retrieval systems to retrieve 10 tracks for each query. This resulted in a total of 12 lists (3 tracks × 4 retrieval systems), providing a comprehensive dataset for analysis.

3.1.1 Analysis of Retrieved Tracks. We focused on aspects like the presence of tracks by the same artist, tracks of the same genre, and other noticeable patterns. The relevance of tracks varied across systems. The TF-IDF based system often retrieved tracks with similar lyrical themes, while the BERT-based system surfaced tracks with more nuanced thematic connections.

3.2 Comparative Effectiveness of Retrieval Systems

3.2.1 Random Baseline. This system, as expected, showed the most diversity but the least relevance in its selections. It served as a control group for comparison with the other systems.

3.2.2 TF-IDF with Cosine Similarity. This method effectively retrieved tracks that were not just lyrically similar but often exact matches or very close variants, demonstrating the system's precision in identifying specific phrases or themes within the lyrics.

Analysis with "Jingle Bells" by Frank Sinatra. For the song "Jingle Bells" by Frank Sinatra, the TF-IDF with Cosine Similarity retrieval system found tracks with a high degree of lyrical similarity. The results are detailed in the table below:

Table 1: TF-IDF with Cosine Similarity results for "Jingle Bells"

Cosine Similarity	Artist	Song
0.963528	Gwen Stefani	Jingle Bells
0.601615	V. Carlton	Hear the Bells
0.526900	Lil Xan	Saved by the Bell

Further Analysis with "Happy Birthday" by Stevie Wonder (TF-IDF and Cosine Similarity). The analysis of "Happy Birthday" by Stevie Wonder using the TF-IDF with Cosine Similarity system revealed interesting selections that highlight the system's focus on word frequency. This approach led to the identification of songs with similar lexical elements, even across different languages and themes.

A notable example is the song "Dead" by Korn, which, despite its contrasting musical style to Stevie Wonder's, contains repeated mentions of the word "happy":

> **Korn - Dead Lyrics**
> All I want in life is to be happy, happy
> All I want in life is to be happy, happy

Additionally, the system recommended "Felicità" by Al Bano, an Italian song thematically centered around happiness. The inclusion of this song demonstrates the system's ability to recognize lexical similarities in different languages:

> **Al Bano - Felicità Lyrics**
> Felicità (Happiness)
> È tenersi per mano, andare lontano, la felicità (It's

holding hands, going far, happiness)
È il tuo sguardo innocente in mezzo alla gente, la felicità (It's your innocent look among people, happiness)
È restare vicini come bambini, la felicità (It's staying close like children, happiness)

These examples illustrate how the TF-IDF with Cosine Similarity system, while effective in identifying word-based patterns, may not always align with the broader thematic or emotional context of the music. The focus on specific word frequencies in "Dead" and "Felicità" highlights both the strengths and limitations of this text-based retrieval approach.

Table 2: TF-IDF with Cosine Similarity results for "Happy Birthday"

Cosine Similarity	Artist	Song
0.892669	Square Heads	Happy
0.858457	Al Bano	Felicità
0.857303	Korn	Dead

Analysis of "Wake Me Up When September Ends" by Green Day (TF-IDF and Cosine Similarity). Similar to the analysis conducted in section 3.2.2.2, the TF-IDF with Cosine Similarity system's evaluation of "Wake Me Up When September Ends" by Green Day primarily focused on the statistical occurrence of words. The system identified songs with a high frequency of certain words, similar to those in the Green Day track, without necessarily reflecting thematic or stylistic similarities.

The songs selected for their lexical resemblance included "Wake Up" by Emigrate, "Wake Up (Make a Move)" by Lostprophets, and "Open Your Eyes" by Goldfinger. As with the previous analysis, this demonstrates the system's effectiveness in identifying word-based patterns but also highlights its limitations in capturing the full thematic and emotional scope of the music.

Table 3: TF-IDF with Cosine Similarity results for "Wake Me Up When September Ends"

Cosine Similarity	Artist	Song
0.535697	Emigrate	Wake Up
0.521682	Lostprophets	Wake Up (Make a Move)
0.501407	Goldfinger	Open Your Eyes

3.2.3 Alternative Text-based Features (BERT/Word2Vec). The use of BERT and Word2Vec provided varied results, with BERT being particularly effective in understanding contextual nuances, and Word2Vec focusing more on overall semantic content.

Analysis of "Jingle Bells" by Frank Sinatra (Alternative Features). The alternative feature-based analysis of "Jingle Bells" by Frank Sinatra interestingly brought up songs that, while not sharing direct lyrical similarities (such as the phrase "jingle bells"), resonated thematically around the idea of Christmas.

A notable example is "Hellhound on My Trail" by Robert Johnson. Despite the stark difference in musical style and era, the lyrics of this song evoke the theme of Christmas Eve, as seen in the excerpt below:

Robert Johnson - Hellhound on My Trail Lyrics
If today was Christmas Eve, if today was Christmas Eve
And tomorrow was Christmas Day
If today was Christmas Eve, and tomorrow was Christmas Day
Aw, wouldn't we have a time, baby?

This selection underscores the ability of the alternative feature-based system to capture thematic elements that extend beyond exact word matches, focusing instead on the broader contextual theme of Christmas. The inclusion of this song demonstrates how the system can identify thematic connections even when the explicit lexical similarities are limited.

Table 4: Alternative Text-based Features results for "Jingle Bells"

Cosine Similarity	Artist	Song
0.951122	Gwen Stefani	Jingle Bells
0.662801	Robert Johnson	Hellhound On My Trail
0.623144	Change	The Glow of Love

Further Analysis of "Happy Birthday" by Stevie Wonder (Alternative Features). In the analysis of "Happy Birthday" by Stevie Wonder using alternative features, an interesting outcome was the recommendation of a song in Portuguese, "Meu Aniversário" by Vanessa da Mata, which is also themed around birthday celebrations. This suggests the system's capability to detect thematic resonances across different languages.

The lyrics of "Meu Aniversário" emphasize the theme of birthdays and celebration, aligning thematically with Stevie Wonder's song. The excerpt and its English translation are as follows:

Vanessa da Mata - Meu Aniversário Lyrics
Hoje é meu aniversário (Today is my birthday)
Corpo cheio de esperança (Body full of hope)
Uma eterna criança, meu bem (An eternal child, my dear)
Hoje é meu aniversário (Today is my birthday)

This inclusion highlights the system's ability to transcend language barriers and identify songs that share a common theme, in this case, the celebration of a birthday. The recognition of thematic congruence, despite linguistic differences, showcases a sophisticated aspect of the alternative feature-based retrieval system.

Table 5: Alternative Text-based Features results for "Happy Birthday"

Cosine Similarity	Artist	Song
0.694113	Vanessa da Mata	Meu Aniversário
0.555861	Kool and The Gang	Celebration - Single Version
0.554781	Alicia Keys	New Day

Examination of "Wake Me Up When September Ends" by Green Day (Alternative Features). In the case of "Wake Me Up When September Ends" by Green Day, the alternative feature-based system retrieved songs with similar thematic elements. The top results are presented in the table below:

Table 6: Alternative Text-based Features results for "Wake Me Up When September Ends"

Cosine Similarity	Artist	Song
0.735971	Pink Floyd	Fat Old Sun
0.724557	Summoning	Where Hope and Daylight Die
0.720513	Los Tres	Déjate Caer

3.2.4 New Combination of Similarity Measure and Feature. In this innovative approach, we shifted from BERT representations to Word2Vec and replaced cosine similarity with Euclidean distance for measuring similarity. The Euclidean distance, interpreted as the distance between points in a vector space, was sorted in ascending order, with smaller distances indicating greater similarity. This modification led to distinct outcomes, revealing tracks with unique thematic connections and lexical alignments not evident through traditional methods.

Analysis of "Jingle Bells" by Frank Sinatra (Word2Vec and Euclidean Distance). The Word2Vec representation paired with Euclidean distance for "Jingle Bells" by Frank Sinatra yielded results that align well with the Christmas theme. The songs recommended shared a clear thematic connection with the holiday season.

Table 7: Word2Vec and Euclidean Distance results for "Jingle Bells"

Euclidean Distance	Artist	Song
0.073473	Gwen Stefani	Jingle Bells
0.253046	Nat King Cole	The Christmas Song (Merry Christmas To You)
0.264748	Cyndi Lauper	Christmas Conga

Analysis of "Happy Birthday" by Stevie Wonder (Word2Vec and Euclidean Distance). The analysis of "Happy Birthday" using Word2Vec and Euclidean distance presented surprising results, recommending songs in different languages with small but interesting lyrical similarities related to themes of happiness and celebration.

Table 8: Word2Vec and Euclidean Distance results for "Happy Birthday"

Euclidean Distance	Artist	Song
0.234464	Dead Fish	Bem-Vindo ao Clube
0.239648	Silvio Rodríguez	Pequeña serenata diurna
0.243180	Arvingarna	I Do

Relevant excerpts from the lyrics with translations include:

Dead Fish - Bem-Vindo ao Clube Lyrics
O mundo é que está errado (The world is wrong)
Bem vindo ao clube (Welcome to the club)
Celebrar o fim (Celebrate the end)
Seja feliz (Be happy)

Silvio Rodríguez - Pequeña serenata diurna Lyrics
Soy feliz (I am happy)
Soy un hombre feliz (I am a happy man)
Y quiero que me perdonen (And I want you to forgive me)

Arvingarna - I Do Lyrics
Du gör mig lycklig (You make me happy)
Lycklig (Happy)
I do I do I do (I do I do I do)

These selections, though linguistically diverse, share a focus on themes of happiness and celebration, demonstrating the nuanced capability of the Word2Vec and Euclidean distance approach in capturing subtle thematic connections across languages.

Analysis of "Wake Me Up When September Ends" by Green Day (Word2Vec and Euclidean Distance). The Word2Vec and Euclidean distance analysis for "Wake Me Up When September Ends" produced results that intriguingly revolve around the themes of time passing, endings, and transitions.

Table 9: Word2Vec and Euclidean Distance results for "Wake Me Up When September Ends"

Euclidean Distance	Artist	Song
0.235480	Midlake	Children Of The Grounds
0.239017	Wild Beasts	End Come Too Soon
0.246090	Electric Wizard	Dunwich

Lyric excerpts that emphasize these themes include:

Midlake - Children Of The Grounds Lyrics
It begins to die
So I've come here to wait
For the end of it all
Till I'm gone from here
I'm gone from here

Wild Beasts - End Come Too Soon Lyrics
The end it comes too soon,
Too soon, too soon, too soon
The end it came too soon

Electric Wizard - Dunwich Lyrics
The end has begun
Our time has come
The end has begun

Our time has come

These lyric excerpts, from songs recommended by the system, resonate with the notion of time and endings, reflecting the essence of "Wake Me Up When September Ends." This suggests that the Word2Vec and Euclidean distance approach, while focusing on lexical proximity, can also inadvertently capture thematic undercurrents present in the songs.

4 INSIGHTS AND SUMMARY OF FINDINGS

Throughout our exploration of content-based music retrieval using various text-based features and similarity measures, we have uncovered several key insights and findings:

(1) **Variability Across Methods**: The results varied significantly across the different retrieval methods (TF-IDF with Cosine Similarity, Word2Vec with Euclidean Distance, and others). Each method showed unique strengths in identifying song similarities, whether it be through word frequency, thematic connections, or lexical proximity.
(2) **Thematic vs. Lexical Similarities**: While systems like TF-IDF with Cosine Similarity were effective in identifying songs with similar word usage, they sometimes missed broader thematic connections. In contrast, the Word2Vec and Euclidean distance approach, though primarily focusing on lexical proximity, occasionally revealed thematic undercurrents in its selections.
(3) **Language and Contextual Nuances**: The retrieval systems demonstrated an interesting capability to transcend language barriers, identifying thematic resonances across different languages. This was particularly evident in the selection of songs with similar themes of happiness, celebration, and endings, despite linguistic differences.
(4) **Limitations of Text-Based Analysis**: Our findings also highlighted the limitations of relying solely on text-based analysis for music retrieval. Such systems, while adept at identifying patterns in word usage, may not fully capture the emotional or thematic essence of music.

These insights underscore the complexity of developing effective music retrieval systems and the potential of integrating diverse text-based features and similarity measures. They also highlight the importance of considering both lexical and thematic elements in music analysis.