KV Multimedia Search and Retrieval

Exercise 1 Group E

Ali Ayadi
Johannes Kepler University
K12042692@students.jku.at

Luca Della Mura
Johannes Kepler University
k12241884@students.jku.at

Ruo Li
Johannes Kepler University
k12247854@students.jku.at

Sara Scheucher
Johannes Kepler University
k1648069@students.jku.at

ABSTRACT

In this paper, a rudimentary text-based retrieval system is being developed. The goal of this paper is to test the effect of different combinations of similarity measures and word embeddings in a retrieval system on its results. For calculating this similarity measure only text features are used. This paper compares the results of tf-idf, word2vec as well as the transformer-based BERT. The similarity measure is in all cases the cos-sim-similarity.

1 Introduction

The increasing availability of digital libraries has paved the way for a new generation of music recommender systems. Music recommendation systems play a pivotal role in helping users discover new tracks, artists, and genres and thus drive customer satisfaction in a significant way.
The purpose of this paper is to evaluate and compare 4 different approaches to music retrieval systems. Within the scope of this paper all features are going to be text-based.
Concretely, the cos-sim-similarity is used as a similarity measure and tf-idf, word2vec and BERT as word embeddings.
The resulting recommendations will be evaluated qualitatively according to the similarity to the queried song.
The data set used for testing the 4 retrieval systems is a subset of the Music4All-Onion dataset which was kindly provided by the university.

2 Methodology

For implementing the assignment, the programming language Python is used as it is most suitable for data analysis and data science purposes. The coding environment used is Jupyter Notebook as it supports the programming language Python. We also used the libraries NumPy and pandas as well as Scikit-learn which provide us with different similarity functions. The code repository is hosted on the platform GitHub. The coordination and integration of code contributions of each team member is therefore ensured using Git.

To ensure that new functionality as well as new algorithms can easily be added the text-based music recommender system in the future a large focus is set on making the code modular.

The input of the system is the string which contains the name and the artist of the query track. The recommender system should output a list of songs with the title and the artist-

To be able to better analyze the results of the recommender system this output of the script is saved in a Python dictionary.

To keep the code modular and make it reusable we first implemented some basic functions in a separate Python file. First, we defined a function to get information about the artist and song name from the IDs. The function takes a list of ids as input and returns a list of tuples of strings containing the name and artist of tracks as the output. Then we implemented a function, which can be used for all text-based analysis. The function itself is also called text based. It takes as input parameters “id”, the id of the query song, “repr” the string which represents the word embedding used in the data, for example, the tf-idf, “N”, the number of tracks retrieved as well as the similarity function for example cosine similarity or Euclidean similarity. These measures help increase the flexibility and reusability of the code. Then we imported the functions into the main file where they can be called with different input parameters.

2.1 The dataset

Music4All-Onion is a large-scale, multi-modal music data set, which expands the Music4All-dataset with additional features and meta-data. For this task, only text-based features are considered (i.e.: title, lyrics, artist, album name and ID).

The provided data is presented in four TSV files (Tab-Separated Values file) this type of file is similar to the famous CSV (Comma-Separated Values) data format, but it uses tabs to separate the parts where each line represents a row of data. To read these files we use the Pandas library that returns the data as dataframes.

The four files are:

- id_information_mmsr.tsv: contains the track IDs in the first column, the artist, song, and album name are contained in the remaining columns.
- id_lyrics_bert_mmsr.tsv: presents the data using the BERT feature vectors with one column for the IDs.
- id_lyrics_tf-idf_mmsr.tsv: presents the data using the Term Frequency-Inverse Document Frequency feature vectors with one column for the IDs.

- id_lyrics_word2vec_mmsr.tsv: presents the data using the word2vec feature vectors with one column containing the IDs.

BERT: is a text-based model trained on a collection of words and is capable of capturing rich contextual information for words in a sentence in a word embedding manner that allows BERT to understand the meanings of words in a given phrase for example. This method is widely used in NLP tasks.

TF-IDF: this method is the mix of two components TF (Term Frequency) and IDF (Inverse Document Frequency), this combination yields a vector capturing the relevance of words in a document relative to a corpus.
TF calculates the occurrence of a word for a given document on the other hand IDF calculates a word's importance for a given collection of documents the product of IDF by TF yields the relative importance of a word to a document within the overall corpus.

Word2vec: is an embedding technique that uses vector space. Words with similar context or meaning have similar vector representations (close vectors in the vector space representation). This ensures capturing semantic relationships between words in the vector space. The method is based on training a neural network on a bunch of words to learn their continuous vector space and present each word as a feature vector.

2.2 Random baseline

For the random baseline, we first shuffled the songs according to a randomly generated order, so we get a different result each time the function is called. Then we excluded the query song from the data frame, so it does not appear in the result list. Afterwards, we retrieve the top N random songs and store them in the result list.

2.3 Cos-sim based on tf-idf

To calculate the cosine-similarity of the tf-idf representation of the lyrics, we created a wrapper function called "cos_sim" that takes two Numpy-arrays as input and reshapes them to 2d arrays so they can be used in the cosine similarity function which is provided by the scikit-learn library. The result of the cos_sim function is the similarity score of the two arrays. The similarity function cos_sim is then passed to the "text_based" function, as well as the query song-id and the dataset containing the tf-idf values. The first thing the test_based function does is search for the query song in the tf-idf dataset and extract the row vector representing the queried song. Then we create an array which is called similarities to store the similarity scores.
Afterwards, the "text_based" function iterates through all the rows in the tf-idf dataset. The similarities between the query-vector and track-vector are then calculated using the cos_sim function. The song-id as well as the similarity-score are then saved in the similarities array. The similarity scores can range from -1 to 1, 1 corresponds to perfect similarity. Afterwards, we sort the list in decreasing order of the similarity score and retrieve the ids of the 10 most similar songs.

2.4 Cos-sim based on word2vec

The next retrieval system uses the word2vec-embedding with the cos-sim-similarity measure. Word2vec represents each word as a high-dimensional vector and uses a neural network to learn the optimal distance of these vectors in the embedding space. The flow of the code is the same as tf-idf. So only one parameter has to be changed when calling the function.

2.5 Cos-sim based on BERT

As the last retriever that we have implemented for this project, we have chosen cosine similarity as its similarity function and lyric representation generated by BERT. As mentioned above, we have adopted a modularized scheme when implementing the text-based function. Thus, for this step, we just set the repr parameter of the text_base function to "bert", and the sim_func parameter to "cos-sim." The flow of execution is the same as when using the other two representations.

3 Qualitative Analysis

For the qualitative analysis, we selected 3 tracks for each retrieval system and retrieved 10 tracks for each query track:

3.1 Random baseline

The random baseline function will generate a random set of tracks every time it is called, irrespective of the query track. It returns often songs of different genres that are unrelated to the query track.

3.2 RS<Cos-sim, tf-idf>

Query Track 1: Love me by the 1975:
Result list:

Song	Oh Yeah	Artist	Big Time Rush
Song	The Gospel	Artist	Alicia Keys
Song	Fire Starter	Artist	Demi Lovato
Song	Rat Fink	Artist	Misfits
Song	How Bad Do You Want It (Oh Yeah)	Artist	Sevyn Streeter
Song	Yeah! (feat. Lil Jon & Ludacris	Artist	Usher
Song	Regarde-moi	Artist	Céline Dion
Song	Miss Independent	Artist	Ne-Yo

Song	Euphoria	Artist	BTS
Song	Let There Be Love	Artist	Simple Minds

In the result, there are songs of different genres like R&B, K-Pop, Pop, Indie-Rock, and Punk. The lyrics of the query song "Love me" contain many occurrences of the tokens "yeah" and "love". By examining the lyrics of the retrieved songs, we also noticed a high number of occurrences of the token "Yeah". The tracks "Let There Be Love" and "Miss Independent" also exhibit a high occurrence of the word "Love" respectively. Other than that, the query song and the retrieved songs do not have so much in common and come from different genres.

Query Track 2: One by U2
Result list:

Song	One	Artist	Mary J. Blige
Song	One Love (feat. Estelle)	Artist	David Guetta
Song	Love the One You're With	Artist	Stephen Stills
Song	One	Artist	Alanis Morissette
Song	No One	Artist	Alicia Keys
Song	One Tribe (Defqon.1 2019 Anthem)	Artist	Phuture Noize
Song	You Can Be the One	Artist	Late Night Alumni
Song	Rape Me	Artist	Nirvana
Song	Palavras No Corpo	Artist	Gal Costa
Song	No One in the World	Artist	Anita Baker

For the second query track, we picked "One" by U2. The data also includes a cover version of it by Mary J. Blige, which appears first in the result set because of the identical lyrics. The result also displays a diversity in genre. The genre of the query track is best described as a rock ballad, whereas the genre of the retrieved songs ranges from the genres R&B, Grunge, EDM and Pop-Rock. There is also one Spanish song in the result.

Query Track 3: Every Christmas by Kelly Clarkson
Result list:

Song	Christmas Conga	Artist	Cyndi Lauper
Song	Three Ships	Artist	Cyndi Lauper
Song	Hellhound On My Trail	Artist	Robert Johnson
Song	St. Patrick's Day	Artist	John Mayer
Song	Last Christmas	Artist	Carly Rae Jepsen
Song	My Only Wish (This Year)	Artist	Britney Spears
Song	Christmas Vacation	Artist	Descendents
Song	Last Christmas - Studio Version	Artist	Jimmy Eat World
Song	The Christmas Song (Merry Christmas To You)	Artist	Nat King Cole
Song	I Shut Doors and Windows	Artist	September Malevolence

For the third query, we analyzed a Christmas song because Christmas songs usually contain recurring tokens in the lyrics, for example, the tokens "year", "wish" or "mistletoe". As we can see in the result, we retrieved 9 Christmas songs and one song with a different theme. The song "I Shut Doors and Windows" by September Malevolence could not be described as a Christmas song but also contains one occurrence of the token "Christmas" in its lyrics. The result includes two tracks from the same artist "Cyndi Lauper", which also happens to be from the same Christmas-themed album.

3.3 RS<Cos-sim, word2vec>

Query song 1: Love me by the 1975:
Result list:

Song	Miss Independent	Artist	Ne-Yo
Song	If Our Love Is Wrong	Artist	Calum Scott
Song	Looking For Clues	Artist	Robert Palmer
Song	Out on the Tiles	Artist	Led Zeppelin
Song	So Much Love	Artist	The Rocket Summer
Song	Let There Be Love	Artist	Simple Minds
Song	In the Evening	Artist	Led Zeppelin
Song	All You Got	Artist	Tegan and Sara
Song	Rosalyn	Artist	David Bowie

Song	How Bad Do You Want It (Oh Yeah)	Artist	Sevyn Streeter

The result contains two different tracks by the same artist “Led Zeppelin”. It is noticed that the result set contains several tracks by British artists. The artist who made the query track is also a British Band. which leads us to the speculation that the cluster might be attributed to different language use between British English and American English. The genres of the retrieved tracks are Rock, Pop, R&B and Rap, when ranked in decreasing order.

Query Track 2: One by U2
Result list:

Song	One	Artist	Mary J. Blige
Song	One Love (feat. Estelle)	Artist	David Guetta
Song	Quien Eres Tu (Feat. Trey Songz)	Artist	María José
Song	Dance With The One That Brought You	Artist	Shania Twain
Song	Apocalyptic	Artist	Halestorm
Song	I Will Survive - Extended Version	Artist	Gloria Gaynor
Song	King For A Day	Artist	Thompson Twins
Song	Never Let Me Down	Artist	Kanye West
Song	Fica Mais um Pouco Amor	Artist	Emicida
Song	I Will Survive	Artist	Gloria Gaynor

For the query track “One” by “U2”, the cover version by Mary J. Blige ranks the first in the result, just like we did with the tf-idf representation. We also retrieved one Spanish and one Portuguese song. The Genres of the retrieved track are Rock, Pop, Electronica, Country, Hip-Hop, Samba and Pagode, when rank in decreasing order of similarity.

Query Track 3: Every Christmas by Kelly Clarkson
Result list:

Song	St. Patrick's Day	Artist	John Mayer
Song	Junesong Provision	Artist	Coheed and Cambria
Song	My Only Wish (This Year)	Artist	Britney Spears
Song	PERFECT!	Artist	WJSN
Song	Si Tu Novio Te Deja Sola	Artist	J Balvin
Song	So Doggone Lonesome	Artist	Johnny Cash
Song	The Best Day	Artist	Taylor Swift
Song	Wait For You	Artist	Elliott Yamin
Song	Jesus Christ	Artist	Brand New
Song	Alone (feat. Big Sean & Stefflon Don)	Artist	Halsey

For the third track, the result includes one Korean song and one song in Spanish. We retrieved all kinds of different genres like country, Rock, and Pop. The entropy of the genre is high. Not as many Christmas themed songs are included in the result as are compared to the result obtained from other data.

3.4 Cos-sim based on BERT

Query song 1: "Love me" by "The 1975"
Result list:

Song	Dance Gavin Dance	Artist	Thug City
Song	Shine	Artist	Take That
Song	One, Two, Three, GO!	Artist	Belanova
Song	Right There	Artist	Ariana Grande
Song	Bing Bing	Artist	Crayon Pop
Song	Come Get It Bae	Artist	Pharrell Williams
Song	We Made You	Artist	Eminem
Song	Here I Am	Artist	Monica
Song	Wannabe	Artist	why mona
Song	Edge of the World	Artist	Faith No More

For the result generated by the first query track with the BERT data, none of the ten retrieved tracks appears in the result from the other two datasets. In terms of genre, the result shows a similar pattern result to the other two datasets. The genres that appeared in the results including Pop, Indie Rock, R&B, Funk, Hip Hop and K-pop, showed no effects on the result.

Query song 2: "One" by "U2"
Result list:

Song	One	Artist	Mary J. Blige
Song	What About Love	Artist	Austin Mahone
Song	All of Your Glory	Artist	Broods
Song	La Tortura	Artist	Shakira
Song	Love One Another	Artist	Cher
Song	Black Lake	Artist	Björk
Song	El Triste	Artist	José José
Song	Love Makes the World Go Round	Artist	Ashlee Simpson
Song	Keep It Together	Artist	Madonna
Song	U Want Me 2	Artist	Sarah McLachlan

The result generated with the second query track and the BERT data, again, does not show similarity with the results obtained from the other two data with the sole exception being "One" by Mary J.

Blige. As explained in the section above. It is a cover version of the query song. Therefore its lyrics is identical to that of the query song. In terms of genre, most tracks appeared in the result belong to the Pop genre, different from Rock, the genre of the query song

Query song 3: “Every Christmas” by “Kelly Clarkson”
Result list:

Song	My Only Wish (This Year)	Artist	Britney Spears
Song	Christmas Conga	Artist	Cyndi Lauper
Song	Merry Christmas, Kiss My Ass	Artist	All Time Low
Song	St. Patrick's Day	Artist	John Mayer
Song	The Christmas Song (Merry Christmas To You)	Artist	Nat King Col e
Song	Last Christmas	Artist	Carly Rae Jepsen
Song	Next Year	Artist	Foo Fighters
Song	December's Boudoir	Artist	Laura Nyro
Song	Last Xmas	Artist	Allie X
Song	Santa Claus Is Coming To Town	Artist	The Jackson 5

In the result generated by the third query track with the BERT data, again, five of the retrieved tracks appear in the result from other datasets. Also, in terms of genre, most tracks appeared in the result belongs to the Pop genre, which could be attributed to the theme of the song, Christmas. Christmas music is known to be associated with the instrumental, Carol and Pop genre. It is also worth mentioning that the song “St. Patricks Day”, which appears in the results obtained from other datasets, is also included in the result. An examination of the lyrics reveals that, despite the title of the song being St Patricks Day, there are repeated references to words such as "cold", "snow", "December" and other words that might be found in other Christmas songs, as well as the phrase "Christmas times" itself appears three times.

4. REFERENCES

Moscati, M., Parada-Cabaleiro, E., Deldjoo, Y., Zangerle, E., & Schedl, M. (2022). Music4All-Onion (Version v0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.6609677