**Analysis of song lyrics using NLP and Toxicity analysis models**

**By: Shahaf Malka & Ofek biton**

**As part of a course on advanced topics in machine learning**

**Lecturer: Chen Hajj**

[**https://github.com/Ofekbiton4/Final-project---advanced-topics-in-machine-learning**](https://github.com/Ofekbiton4/Final-project---advanced-topics-in-machine-learning)



**Table of Contents**

**Abstract………………………………………………………………..3**

**Introduction…………………………………………………………...3**

**pulling the data……………………………..……………….………3-4**

**Data Structure and Features…………………………………………4**

**Preprocessing……………………………………..…………………..5**

**Methodology…………………………………………...……………5-6**

**Results and conclusions……………………………………….……6-7**

**A few words about teamwork**

Teamwork was essential to the success of the project. The pronunciation of the idea was shared, as well as the construction of the entire project and report.

Despite what was stated, responsibility for fixing bugs and complete success was divided according to the following division:

Ofek is responsible for pulling the data and performing the preliminary preparation

Shahaf is responsible on the statistical analysis and visualizations.

The selection of the models and their implementation were done in a completely joint manner.

**Abstract**

Our project presents a solution for analyzing existing songs and playlists on Spotify.

The analysis is based on existing models trained in Python that perform sentiment analysis (NLP model) and in addition models that perform toxicity analysis for song lyrics.

The sentiment analysis model is a structured model whose main focus is specialization in social networks, but it was found to be suitable for performing the analysis in our case and by means of words, connecting words and punctuation marks determines the sentiment of the song.

The toxicity analysis model locates offensive, sexist and toxic words in the lyrics and scores the songs with a grade on their level of toxicity.

The following analysis will enable accurate choices for customers and listeners as to what type of music and singer they want to listen to according to their mood and will also enable a decision on the level of exposure to offensive words while listening to music.

**Introduction**

Music is a significant part of many people's lives and has a great influence on their mood. Depressed people will use music to lift their mood, on the other hand, some songs may cause damage and further decline in mood.

Children and teenagers are also big consumers of music. Many times and sometimes without noticing, they consume music with negative effects, offensive and sexist language that can be added to the vocabulary of the children and in more serious cases affect their perception of the world.

In the next report we will detail an analysis of songs taken from Spotify. The analysis will check the lyrics of the songs using NLP and Sentiment Analysis models. We will find offensive words and classify the songs as those with rude or offensive content and thus we can recommend not playing these songs to certain populations

**pulling the data**

In this project, we have interfaced with Spotify's API, an Application Programming Interface, to access their extensive music database. An API is effectively a mediator that allows our software to request and receive data from Spotify's servers in a structured, programmatic way.

During development, we encountered certain limitations associated with the Spotify API. Specifically, we had to obtain proper authentication credentials—Client ID and Client Secret—by registering our application on the Spotify Developer Dashboard. Handling Spotify's rate limits was also pivotal, as these are in place to prevent abuse and ensure service reliability. We made sure to design our data retrieval to stay within these limits to avoid any interruptions to API .

**Data Structure and Features**

The Spotify data we received came in the form of an Excel file, which organized the information into a structured format with various columns, each representing a different attribute of the music tracks. The structure and features included in our data file are as follows:

* Name of the Song: This column holds the title of the track as a string. It's the primary descriptor of the song that users would recognize.
* AlbumType: Indicating the album type as a string, values include 'Single', 'Album', or 'Compilation', providing insight into the release format of the track.
* Artist Name: This column specifies the name of the artist or band performing the song, given as a string and is crucial for identifying the creator.
* Album: Contained within this column is the name of the album to which the song belongs, also provided as a string.
* Year: An integer representation of the release year of the song, which lets us understand the temporal context of the music.
* Length: This column records the duration of the song in milliseconds, allowing for analyses based on song length.
* Spotify ID: The unique identifier for each song within Spotify's system, listed as an object. This ID is vital for retrieving additional data through Spotify's API
* Popularity: An integer scoring the song's popularity on Spotify, ranging from 1 to 100. This metric helps gauge the song's reception and reach.
* Playlist ID: Similar to the Spotify ID for songs, this is the unique identifier for playlists in which the song appears, also noted as an object and essential for playlist-related queries.
* Lyrics: The full lyrics of the song provided as a string. This unstructured text data allows for linguistic and sentiment analysis.

This structured dataset enabled us to conduct comprehensive analyses, from understanding music trends over time to developing correlation studies between song popularity and various attributes

**Preprocecing**

Before we could feed the collected data into our analytical models, it was essential to perform several preliminary data preparation steps to ensure the integrity and quality of our analysis. This preparation included:

* Data Cleaning: We began by ensuring the data was free from inconsistencies or any irrelevant information. This involved removing any duplicate records, correcting any aberrant data entries, and handling missing values appropriately to preserve the dataset's reliability.
* Text Processing: For the 'Lyrics' field, which contains textual data, we conducted tokenization to break down the text into individual words or phrases, removed common stop words that do not offer substantial information, and applied stemming or lemmatization to reduce words to their root forms.

There are many other preparations that can be done such as normalization, changing year to date and time type, creating category columns and more.

Since our model relies on the column of words only, there is no need for this preliminary preparation.

**Methodology**

For a deeper understanding of the sentiment and toxicity analyses within our project, it’s important to comprehend the methodologies of the VADER and Detoxify models and how they process text data.

VADER Sentiment Analysis Model

VADER is a lexicon and rule-based sentiment analysis tool specifically attuned to sentiments expressed in social media. VADER uses a combination of a sentiment lexicon—a list of lexical features (e.g., words) which are generally labeled according to their semantic orientation as either positive or negative. The model is unique because it is sensitive to both the polarity and the intensity of emotions.

The VADER model doesn't only tally up positive and negative markers but also considers intensifiers, such as adverbs, and grammatical constructs that can alter the sentiment intensity of a phrase. For example, VADER understands that "love" is positive and that "don’t love" is a negation that flips the sentiment. Additionally, it can interpret the differing sentiment in phrases like "love" (positive) and "totally love" (more positive). It assigns each piece of text a compound score that ranges from -1 (most negative) to +1 (most positive), as well as categorical scores that classify the sentiment as either positive, negative, or neutral.

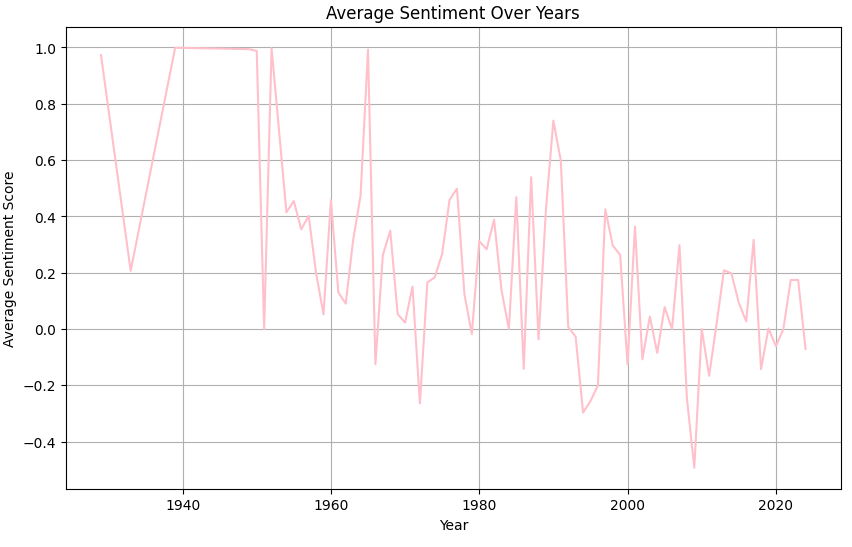
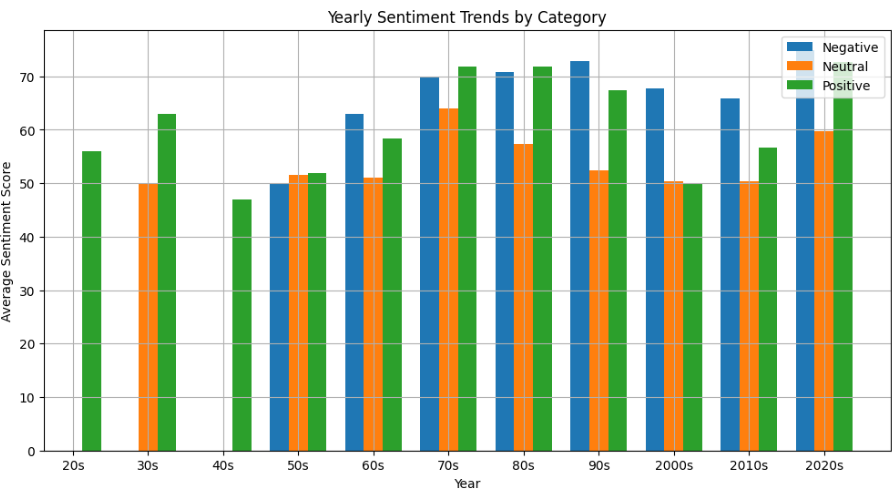
Detoxify Toxicity Classification Model

Detoxify operates on a different principle. It originated from machine learning models trained on large datasets with human-annotated examples of toxic behavior. Through this training, Detoxify can predict text toxicity, identifying various forms of negative interactions, including insults, threats, and hate speech. It does so by understanding textual patterns and context that typically suggest offensive content.

The model represents text using numerical vectors—a format conducive for machine learning algorithms. It then runs these vectors through a neural network to assess how closely they match with learned patterns of toxic speech. Based on this evaluation, Detoxify assigns a probability score indicating the likelihood that a given text is toxic.

**Results and conclusions**

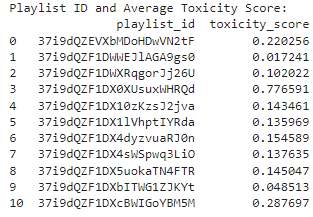
The average sentiment score over the years:

Despite the fluctuations in the graph, you can clearly see the decrease in the sentiment score of the songs from the 20s until today. there was a noticeable shift towards more negative sentiment in songs from the 1980s and 1990s, possibly reflecting societal changes or cultural influences during that time period.

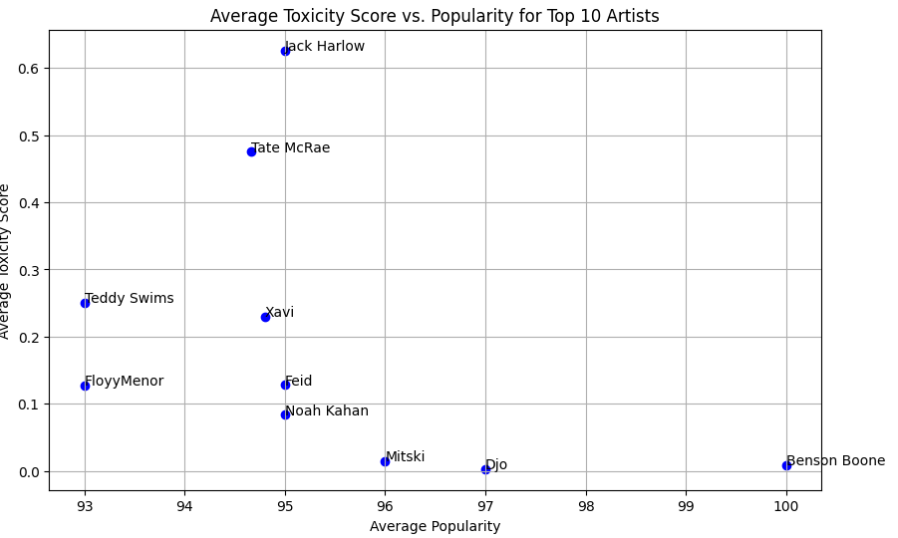
Impact of Lyrics on Popularity:

* There seemed to be a correlation between sentiment category and song popularity. Songs categorized as having a positive sentiment tended to be more popular, while those with negative sentiments were less popular on average.
* This suggests that audiences may gravitate towards music with uplifting themes and messages.

Playlist Restrictions:

By examining the proportion of offensive songs in each playlist, we were able to identify playlists that may not be suitable for certain audiences, such as those under 18 years old. Playlists with a high percentage of offensive songs could be flagged for further review or restricted access. For example Playlist ID: 37i9dQZF1DX0XUsuxWHRQd is Prohibited Playlists for Under 18. In addition, you can use the average toxicity score of each playlist to determine what the toxicity threshold is that we want to consume and thus choose the playlists we want to listen to.

The toxicity of music versus its popularity

When we examined the 10 most popular singers and their average toxicity score, we discovered that 8 out of 10 singers are below a toxicity score of 0.3. This trend indicates a strong correlation between artist messaging and audience values, potentially reflecting a broader societal shift towards more uplifting and socially conscious music consumption patterns. Such alignment likely enhances the appeal and resonance of these artists' work, contributing substantially to their widespread recognition and success.

These findings have implications for content curation and recommendation algorithms in music streaming platforms. By considering sentiment analysis and offensive language detection, platforms can better tailor music recommendations to individual preferences and sensitivities.

Follow-up research:

expand the analysis to include a larger dataset spanning more genres, languages, and time periods. Additionally, refining the sentiment analysis and offensive language detection models could lead to more accurate insights and recommendations.