

Comparison of Critic and Popular Ratings of Hip-hop Lyrics (or, What Makes Trap Popular?)

Cristobal Sciutto - *csciutto* - 06099425

1 Introduction

The project began with the intention of predicting song popularity of all genres according to Billboard charts, taking into account several song features such as length, lyrics, structure, etc. Nevertheless, the complexity of extracting features such as key signatures and primary octaves, BPM, and played instruments, most likely requiring a MIDI file for songs, led me to hone down on lyrics as the major area of study. While artists like Radiohead or The Beatles are certainly known for their poetic lines, I think it will be more interesting to focus on hip-hop and rap, as the genres are the most cited on Billboard and Spotify charts and of my personal taste.

I'm an avid listener of the lyrically intense rap from the 1990s from artists such as A Tribe Called Quest and Gang Starr. This older style of rap tends to receive critical praise for deep lyrics and complex rhymes, in comparison to the trap music which permeates the current pop music scene through songs with repetitive and catchy hooks. Trap artists, such as Travis Scott and Migos, dominate the Billboard charts despite receiving mediocre scores from leading review magazines. They're considered mere "hype" artists, more interested in the rapper lifestyle than in the art. A specifically interesting example is the artist MF DOOM, often called "your favorite rapper's favorite rapper". The dichotomy between his high regard among both critics and rappers, while not sharing a significant popular following, is exactly the kind of anomaly I am trying to understand. While a social analysis would certainly be warranted, I will attempt to understand this phenomenon from a statistical approach.

2 Approach

The project will consist of extracting lyrical features of songs in the three major areas of semantics, structure, and vocabulary. A model will be built around them, and trained on two distinct outputs: Billboard rankings and critic's reviews from distinguished magazines such as Pitchfork and NME. The intention is to generate two models which can respectively predict how a hip-hop song will perform on popular charts, and the review it will likely get from critics. More interestingly, the process will attempt to reveal what song characteristics tend to correlate to critical acclaim versus popular appeal.

An artificial neural-network shall be used in an attempt to encompass all possible linear combinations of the features described below, giving both a more intricate model, and more insights into what makes a song "good". Nevertheless, it is important to have upper and lower bounds to evaluate performance, so the neural-net will be compared to *baseline* and oracle approaches.

2.1 Input-Output Behaviour

- Input: title and lyrics of a song
 - s , tokenized string of title and lyrics, produced via a library such as *nlTK*
- Output: the Billboard's standing (y_b) and critical acclaim of the song (y_c)
 - $y_b = r_b \cdot t_b$, where r_b is a mapping from the peak rank of a song on the Billboard 100 chart to a number between 0 and 1 via a negative exponential, e.g. $r(x) = \frac{(x-100)^2}{9801}$ and t_b is the time on the charts.

- y_c is the normalized average review the song received on Album of the Year, between 0 and 1.
- Features:
 - Semantics
 - * Top 3 affect categories extracted via the EmoLex outside of positive and negative (indicators).
 - * More positive than negative (indicator).
 - Structure
 - * Length of song (# of words)
 - * Presence of hooks
 - * Number of time hook appears
 - * Number of verses
 - * Average length of verses (# of words)
 - Vocabulary
 - * Number of distinct words (as a percentage of number of total words)
 - * Number of distinct rhyming phonemes
 - * Number of rhyming syllables pairs

2.2 Billboard Metrics

For the Billboard’s rankings, the *baseline* approach taken is calculating the average y_b for previous songs by the artist. On the other hand, an *oracle* approach is the actual value of y_b .

e.g. **Mo Money, Mo Problems** - Life After Death - Notorious BIG

Baseline: $avg(1 \cdot 20, 0.978 \cdot 21, 0.978 \cdot 20, 0.902 \cdot 24) = avg(20, 20.538, 19.56, 21.648) = 20.4365$

Oracle: $r(x) = \frac{(1-100)^2}{9801} \cdot 30 = 30$

Due to the incredible time the single stayed on the Billboard charts, the baseline ended up under-predicting the score.

2.3 Critic Review Metrics

A *baseline* approach will consist of averaging all critic reviews from previous albums by the artist. This offers good predictive value, despite being extremely naive, as most artist tend to have consistent style and their songs receive comparable scores from a given critics. On the other hand, the *oracle* approach is merely the actual review.

e.g. **Mo Money, Mo Problems** - Life After Death - Notorious BIG

Baseline: $avg(\frac{84}{100}, \frac{97}{100}) = 0.91$

Oracle: $\frac{84}{100} = 0.84$

Due to the almost perfect score of Biggie Small’s previous album *Ready to Die*, the baseline ended up over-predicting the score.

3 Challenges

Individual songs are rarely reviewed, but rather they are given a rating in context of the album they are released in. This will inherently give some error as every album has a diversity of songs, usually varying in quality. On the other hand, for the Billboard 100 chart, we can get individual songs as our atomic unit.

Another important caveat is that the critic review and Billboard datasets are not complete. While nearly all Billboard 100 songs are reviewed by leading magazines, the converse is usually not true. Some of the

highest regarded hip-hop albums did not sell enough to get on the charts. It is still to be decided if the critic review predictor will be trained on the entirety of the dataset, or rather just those that also achieved popular levels.

4 Data Sources

The data required for this project will be primarily extracted via API of official sources such as Billboard and Genius. For data regarding critics reviews, web scraping will be necessary.

- Lyrics: Genius API docs.genius.com
- Semantics: NRC Word-Emotion Lexicon (EmoLex) will be used to associated the lyrics to emotions.
- Billboard: extra-official API github.com/guoguo12/billboard-charts
- Critics Reviews: webscraping AOTY.com for a single critics score, composed of all major critic magazines that reviewed the song
- Rhymes: previous CS 221 project which created a freestyle-AI will be used as a basis to extract rhymes from song lyrics (see literature review).

5 Relevant Literature

- EmoLex: arxiv.org/pdf/1308.6297.pdf
- Rhyme and Style Features: http://www.ifs.tuwien.ac.at/mayer/publications/pdf/may_ismir08.pdf
- Extracting Rhymes: <https://worksheets.codalab.org/worksheets/0xa30cf00878ba4a85ab239b40fc9b5818/>