

CS410 FA22 Project Progress Report

Team Name: GroupOf4

Team Member

Ying Sun(yingsun3) - Captain

Chang li(changli8)

Yuteng Zhuang(yutengz2)

Qinhan Xia(qinhanx2)

1) Which tasks have been completed?

- Loaded and cleaned all lyrics data to dataframe by using pandas.
- Finished sentiment analysis for all lyrics and put the grade of analysis in the data set.
- Convert all lyrics into BagOfWords list with delimiters deleted, and then map theses data with the frequency of each term for each doc into a dictionary
- Compute the term frequency for each doc
- Compute the compute inverse data frequency for each doc
- Compute TF-IDF for each doc
- Constrcuted fron-end framework
- Designed some front-end components, including two search pages and detail page:

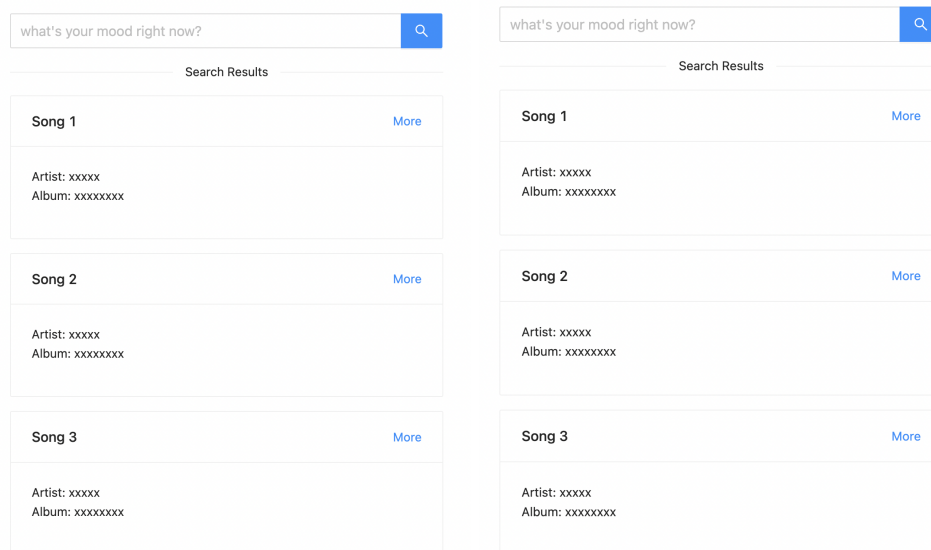


Figure1&2 Two search pages

Song Info

Artist: Lady Gaga

Album: XXXXX

Released: Oct. 22, 2012

Genre: Pop, soft rock

Lyrics

Love game intuition, play the cards with spades to start
And after he's been hooked, I'll play the one that's on his heart
Oh, whoa, oh, oh
Oh, oh-oh
I'll get him hot, show him what I got
Oh, whoa, oh, oh
Oh, oh-oh
I'll get him hot, show him what I got
Can't read my, can't read my
No, he can't read my poker face
(She's got me like nobody)
Can't read my, can't read my
No, he can't read my poker face
(She's got me like nobody)

Figure 3 Detail page

2) Which tasks are pending?

- Frontend web in local host
- We are working on implementing sentiment based search algorithms, 60% done and 40% remaining.
- Haven't add the algorithm using TF-IDF to rank the result of the query
- Search based on sentiment of lyrics.
- Testing & Debugging

3) Are you facing any challenges?

- We are still exploring how to build the frontend local host by using host.
- Combine the preprocessing indexing data with the search algorithm