# Text Mining Project Writeup and Reflection Lyrics Generator

## **Daniel Daugherty**

#### **Project Overview**

To create a new song using the lyrics of a specific artist, I used <u>www.metrolyrics.com</u> to gather all the data for song lyrics. By searching the album page of a specific artist, I was able to get the links to the lyrics of all the songs in the album. After compiling the lyrics from all songs within the album into a list, I utilized Markov Analysis to detect patterns in the syntax structure of a particular album to produce a new song for the album as if it were written by the original artist.

## **Implementation**

In order to get all the lyrics from an album, I pickled the album webpage and used Python's BeautifulSoup package to parse through and collect the links to each song within the album. For each song, I parsed through its webpage to get the lyrics and compiled all of them into a list. With this list of words, I used markov analysis to detect patterns in the order of words. For my markov analysis, I created a dictionary matching groups of 3 words with corresponding groups of words that tend to follow. Finally, to generate a new song, I randomly selected keys within the dictionary and printed out the prefixes with the corresponding suffixes.

One of the biggest design decisions I faced was choosing how many songs from a specific artist to grab. At first I tried to gather all of the songs on the artist's page; however, I felt like I was gathering a lot of data and it was difficult to collect only the song links using BeautifulSoup. I found that a more elegant solution would be using the artist's album page and collecting the songs with the name of the album within each link code. This allowed me to get a smaller selection of songs and also a more focused group of lyrics by being in the same album.

#### Results

My code was able to successfully generate random lyrics based on the lyrics in any album. Typically songs contain somewhere between 159 and 300 words. To account for this, each song I produce contains a random amount of words somewhere between 159 and 300. Unfortunately the generated lyrics don't make much sense, but I still think it was really cool to write lyrics like several various artists did.

### **Examples of generated lyrics**

Artist: Adele Album: 25

for there is something in your loving that tears down my walls i wasn't ready i'm ready now i'm heading straight for you you will only be eternally the one that i belong to the sweetest devotion i've known sweetest it's the sweetest sweetest it's the sweetest it's the sweetest it's the sweetest it's the sweetest it's the sweetest sweetest it's the sweetest sweetest it's the sweetest it's the sweetest sweetest it's the sweetest it's

Artist: Queen Album: Greatest Hits

gonna be a big man of me get oh you gonna take me home tonight down beside that red fire light oh you gonna let it all hang out fat bottomed girls fat bottomed all right ride 'em cowboy fat bottomed girls you make the rocking world go round hey i was just a skinny lad never knew no good from bad but i knew love before i left my nursery left alone with big fat fanny she was such a naughty nanny heap big you made a bad boy out of me hey i've been singing with my band across the across the land i've seen every blue eyed floozy on the way but their beauty and their style went kind of smooth after a while take me to them naughty ladies every time won't you take me home tonight oh down beside that red fire light you gonna let it all hang out fat bottomed girls you make the rocking world go round get on your bikes and ride them fat bottomed girls we are not in a position to display these lyrics due to licensing sorry for the i want to be free baby oh how i want to break yea i i i want i want to break yea i i want to break free i've got to break free i've fallen in love i've fallen in love it's strange but it's hey yea i can't get over the way

Artist: Justin Bieber Album: My Worlds Acoustic

you needed a little time for my mistakes it's funny how you use that time to have me replaced but did you think that i wouldn't see you out at the movies what you doin' to me you're takin' him where we used to and now we're and where do we when there's is no road to get to your lets start over so it's up to and its up to and its up to that we meet in the on our way back down to our way back down to back down to down to earth x8 you were always and i live out of so tell me how i can make a i close my eyes and i can see a better day i close my eyes and pray i close my eyes and i can see a better day i close my eyes and pray i close my eyes and pray i close my eyes and i can see a better day i clos

Artist: Taylor Swift Album: 1989

Artist: The Beetles Album: Abbey Road

me cry because the sky is it makes me cry because the sky is ah you're gonna carry that weight a long time here come old flat top he come groovin' up slowly he got joo joo eyeballs he one holy rollers he got hair down to his knee got to be a joker he just do what he please he wear no shoeshine he got toe jam football he got monkey finger he shoot coca cola he say i know you know me one thing i can tell you is you got to be a joker he just do what he please he wear no shoeshine he got toe jam football he got monkey finger he shoot coca cola he say i know you know me one thing i can tell you is you got to be a joker he just do what he please he wear no shoeshine he got toe jam football he got monkey finger he shoot coca cola he say i know you know me one thing i can tell you is you got to be good looking cause he's so hard to see come together right now over me i want to tell you i feel hung up and i don't know something in the way she woos me i don't want to leave her now

#### Reflection

I feel like I learned a lot throughout this project. Coming into this project, I had no previous experience working with parsing through web page source data and now I feel like I have no problem teaching it to others. One of the things I think could be improved in future projects is unit testing. I found that I would run into type errors or other kind of simple errors late into the project which could have easily been identified with doctests. If I had more time to work on this project I would try to get the lyrics to make more sense and have more structure. Instead of just a block of text, I think it would be interesting to work on producing randomly generated lines of a song that would have a rhyme scheme.