

Final Capstone

Proposal

By Karen McGee

Problem Statement

- Perform text analysis of song lyrics to classify differences between songs and sentiment analysis.
 - Based on the lyrics can the model determine the type of mood the song?
 - Based on the mood of the song can the model recommend other songs?

How is the solution valuable?

- Understanding how text can have an impact on how you feel is valuable information. This type of information can be utilized by marketing companies to help drive consumers to buy certain products.
- In addition, having a model that can recommend additional songs based on the current mood of a song adds value to the individual person. The individual gets a sense of pleasure of hearing recommended songs that they haven't heard in a while or wouldn't have thought of at the time to play.

What is your data source and how will you access it?

- 100 songs from various artist.
- The lyrics will be obtained from Genius.com or Lyrics.com via scrapy or an API.

What techniques from the course do you anticipate using?

- Web scraping.
- Natural language processing.
- Unsupervised learning models and evaluations

What do you anticipate to be the biggest challenge you'll face?

- Generating enough features that will adequately help the model determine differences between songs.
- Classifying the mood of the song based on the lyrics will be challenging as the lyrics may contain a series of negative words but end up on an uplifting and positive note (the opposite of this idea is also true).
- Generating the file using scrapy API, every site is different and as a result understanding how to scrape a site to pull the information I am looking for maybe challenging.