Lyrical Sentiment and Word Importance Across Music Genres

A UVA Data Science Case Study by Nathan Patton, 2023

Calling all music fans! Yes, that's you! Did you hear about the new Lebron James album that just dropped?

Now that your attention is here, you must help other music fans out.

Context: Song lyrics have a strong influence on one's state of mind. Words can portray a variety of connotations, such as positive and negative, as well as many different emotions when put together. The ability to recognize these emotions is a universal skill that does not always depend on previous listening to the musical style. Such emotions may include anger, anticipation, disgust, fear, joy, sadness, and trust.

Why does this matter?

If song lyrics play a large role on the listeners state of mind, wouldn't you be concerned if your family member or friend was consistently listening to very negative songs conveying emotions of sadness, disgust, fear, or anger? The answer is yes. Therefore, your goal is to figure out how these lyrical sentiments (emotions) and their word importance (count) vary across genres, so that other music listeners, as well as yourself, can limit their exposure to the genres that correlate most strongly to the negative connotations and emotions.

Deliverables: Produce a machine learning classification model that is able to identify a specific genre of a song based on its lyrical content. More in depth instructions about the final Github repository can be found in the rubric.