EVOLUTION OF PUBLIC PERCEPTION OF LAW ENFORCEMENT VIA HIP HOP LYRICS: A MACHINE LEARNING APPROACH

Through an analysis of rap/hip-hop lyrics, this project provided insight into how the sentiment towards law enforcement (police) has changed over time based on how lyricists describe them in their songs.

The results showed a sharp spike in negative sentiment at around the same time as the LA Riots following the events involving Rodney King and law enforcement (LAPD).

I have used the rap lyrics dataset, Machine Learning, Natural Language Processing, and Sentiment Analysis to find if there is a positive or negative association with law enforcement and how it has changed over the last few decades.

The study utilized a large corpus of noisy lyrics data which I had to preprocess and train a word embedding model. Multiple variants were considered including FastText and Word2Vec to find the highest-quality embeddings. Using the embedding vectors, I computed the top-k similar words for "police" and analyzed them across a few performance axes. For example, how well they align with expert insights or manually checking how relevant they are to the topic.

Various properties of the embeddings were analyzed like how they change over time. Next, I obtained a set of relevant sentences using Okapi-BM25 on the list of similar keywords obtained from the word embeddings. An alternative method utilizing sentence embeddings to compute a set of similar lyrics was also considered.

Lastly, the study utilized a large language model trained with a sentiment analysis objective to obtain the sentiment. The results provided broader insight into how the rap and hip-hop communities reflect anti-law sentiments through their lyrics, specifically with promising initial data about a lineup between a sharp spike in negative sentiment at around the same time as the LA Riots following the events involving Rodney King and law enforcement (LAPD).