

Kelly Cosgrove

10/24/2021

Pop Punk Song Search Project Proposal

1. What are the names and NetIDs of all your team members? Who is the captain? The captain will have more administrative duties than team members.
 - a. Kelly Cosgrove - kellyc4 - Captain
2. What is your free topic? Please give a detailed description. What is the task? Why is it important or interesting? What is your planned approach? What tools, systems or datasets are involved? What is the expected outcome? How are you going to evaluate your work?
 - a. Pop Punk is a genre that many people listened to in their angsty teen years. Often times it's hard for us to recollect a song; rather, we remember a lyric, the theme, or a few key words from it. The goal of this project is to facilitate this rediscovery by building a vector space model-based search engine that allows a user to type these lyrics/themes/key words and receive a list of possible song matches.
 - b. This project will be interesting because rediscovering the songs one listened to the most in High School can be a particularly nostalgic experience.
 - c. I plan to build a corpus consisting of song lyrics from the most popular Pop-Punk bands from the last 30 years, use this corpus to code a vector space model, and then build a GUI for users to interact with the model, containing a search bar and model output.
 - d. I will be using the Gensim Python library to build the vector space model and the [lyrics-corpora](#) Python package to build the corpus. I will be using the [Dash](#) library to build the GUI.
 - e. I expect the user will be able to type key words into a search bar which will then be run through the model and the closest song matches will be returned on a GUI.
 - f. I will evaluate my work by creating a set of queries with relevance judgements and computing the F measure.
3. Which programming language do you plan to use?
 - a. Python
4. Please justify that the workload of your topic is at least $20 \times N$ hours, N being the total number of students in your team. You may list the main tasks to be completed, and the estimated time cost for each task.
 - a. 1 hour - Research most popular Pop-Punk artists of last 20 years
 - b. 4 hours - Research lyrics-corpora library and build corpus of Pop-Punk songs
 - c. 8 hours - Research Gensim library and create vector space model
 - d. 8 hours - Research Dash and create Dash application
 - e. 2 hours - Evaluate using set of pre-judged queries