

Final Project Proposal

CS 470

Grant Lewis

Fall 2023

For my final project, I would like to fine-tune a pre-trained Large Language Model (LLM) to generate song lyrics in the style of specific artists. I have always wanted to train a generative model to write song lyrics. I also feel that such a project will help improve my prospects for careers I am interested in after I graduate, even if the results are mediocre at best. My plan is to fine-tune an LLM, probably LLama 2 7b, (I do have access to resources of large enough to fine-tune this model), to generate lyrics based on a title, genre, and artist name.

In order to accomplish this, I will first collect lyrics by several artists through either publicly available datasets or ethical web scraping. The dataset will then be prepared for use with an LLM, and a training script will be prepared using PyTorch and Huggingface. Post fine-tuning, I will conduct some evaluation. Qualitatively, multiple people will score the generated lyrics based on creativity, coherence, and stylistic similarity to the original artist. The evaluation will involve generating lyrics for songs that did not exist prior to the original training of the LLM and were not part of the test set. If time permits, I will explore how differently engineered prompts can affect the results.