

- [Lyrical](#)
 - [Project Overview](#)
 - [Features](#)
 - [Project Organization](#)
 - [Installation](#)
 - [Usage](#)
 - [License](#)

Lyrical

This repository contains an end-to-end Lyrics Generation Model, a machine learning project aimed at generating original song lyrics using deep learning techniques. The model is trained on a dataset of song lyrics obtained from various sources, including Genius Lyrics, and is capable of generating new lyrics that are stylistically similar to the training data.

Project Overview

The Lyrics Generation Model leverages the power of recurrent neural networks (RNNs) and natural language processing to learn the patterns, structure, and vocabulary of song lyrics. By training on a diverse collection of lyrics from different genres and artists, the model can capture the nuances and characteristics of various music styles and intelligently generate lyrics that adhere to specific themes, genres, or even mimic the style of a particular artist.

Features

- **Data Collection:** The model incorporates a web scraping component to collect song lyrics from popular online platforms such as Genius Lyrics. The lyrics, along with associated metadata such as artist name, song title, genre, and album, are gathered to provide additional context for conditioning the model during the generation process.
- **Data Preprocessing:** The collected lyrics are meticulously preprocessed to ensure data quality and consistency. This includes removing irrelevant

information, handling missing values, standardizing text formatting, and performing other data cleaning tasks. The lyrics are also tokenized and converted into numerical representations suitable for training the lyrics generation model.

- **Conditioning and Context:** The model takes advantage of the additional metadata, such as genre, artist name, and album, to condition the generated lyrics. By providing these conditioning factors during the training and generation phases, the model can generate lyrics that align with specific styles, themes, or artists, producing more contextually relevant and coherent results.
- **Model Training:** The lyrics generation model employs advanced deep learning techniques, such as recurrent neural networks (RNNs) or transformer architectures, to learn the underlying patterns and structures within the lyrics dataset. The model is trained using techniques like sequence-to-sequence modeling or language modeling approaches, optimizing it to generate meaningful and creative lyrics.
- **Evaluation and Fine-tuning:** The generated lyrics can be evaluated using metrics such as perplexity, BLEU score, or through human evaluation. This evaluation process helps fine-tune the model by iteratively adjusting hyperparameters, experimenting with different architectures, and training strategies to optimize the results and enhance the quality and creativity of the generated lyrics.
- **Lyrics Generation:** Once trained, the model can generate new lyrics by providing a seed input or prompt along with the desired conditioning information. The model uses the learned patterns and structures from the training dataset to generate lyrics that are coherent, stylistically appropriate, and aligned with the provided conditioning factors.

Project Organization

```
|— LICENSE
|— Makefile          <- Makefile with commands like `make data` or `make train`
|— README.md         <- The top-level README for developers using this project.
(you are here!)
|— data
|   |— external      <- Data from third party sources.
|   |— interim       <- Intermediate data that has been transformed.
|   |— processed     <- The final, canonical data sets for modeling.
|   |— raw           <- The original, immutable data dump.
|— models            <- Trained and serialized models, model predictions, or
model summaries
```

```

├── notebooks          <- Jupyter notebooks. Naming convention is a number (for
ordering),
                        my first name, and a short `` delimited description,
e.g.
                        `1.0-syed-initial-data-exploration`.
├── reports            <- Generated analysis as HTML, PDF, LaTeX, etc.
    └── figures        <- Generated graphics and figures to be used in reporting
├── requirements.txt   <- The requirements file for reproducing the analysis
environment
├── src                <- Source code for use in this project.
    ├── __init__.py    <- Makes src a Python module
    ├── data           <- Scripts to download or generate data
    │   └── make_dataset.py
    ├── features        <- Scripts to turn raw data into features for modeling
    │   └── build_features.py
    ├── models          <- Scripts to train models and then use trained models to
make
    │   │               predictions
    │   ├── predict_model.py
    │   └── train_model.py
    └── visualization   <- Scripts to create exploratory and results oriented
visualizations
        └── visualize.py

```

Installation

To set up the Lyrics Generation Model, follow the instructions below:

1. Clone the repository:

```
git clone https://github.com/syedaddan/lyrical.git
```

2. Navigate to the project directory:

```
cd lyrical
```

3. Create a virtual environment (optional but recommended):

```
python3 -m venv env
```

4. Activate the virtual environment:

```
source env/bin/activate
```

5. Install the required dependencies:

```
pip install -r requirements.txt
```

Usage

The Lyrics Generation Model can be used as follows:

1. Preprocess the dataset:

- Place the [raw data files](#) in the [data/raw](#) directory.
- Run the data preprocessing script:

```
python src/data/make_dataset.py
```

- The preprocessed data will be saved in the [data/processed](#) directory.
- Run the Build features script:

```
python src/features/build_features.py
```

2. Train the model:

- Customize the model configuration in the [src/models/train_model.py](#) script.
- Run the model training script:

```
python src/models/train_model.py
```

- The trained model will be saved in the `models` directory.

3. Generate lyrics:

- Customize the generation settings in the `src/models/predict_model.py` script.
- Run the lyrics generation script:

```
python src/models/predict_model.py
```

- The generated lyrics will be displayed or saved, depending on the configuration.

For more detailed information on dataset preprocessing, model training, and lyrics generation, refer to the [User Guide](#).

License

This project is licensed under the MIT License. See the [LICENSE](#) file for more information.

With the Lyrics Generation Model, you can embark on a creative journey of generating unique song lyrics that capture the essence of different music genres, artists, and themes. Let your imagination flow and dive into the world of AI-assisted lyrics creation!

Happy lyrics generation!