



Neural Machine Translation: Explaining the Meaning Behind Lyrics

Taite Sandefer

/github.com/tsandefer
in /in/tsandefer
✉ taite.sandefer@gmail.com

Background

As a rhetoric student, I wanted to further explore how neural networks could pick up on the more subtle semantic patterns in natural language.

One of the main challenges in NLP is a lack of labeled data, especially in regard to pragmatics (the study of meaning with respect to goals and intentions).

Data

Genius.com is an online community where users browse, rate, and create annotations for lyrics to help explain their meaning and context.

Example from Childish Gambino's 3005:

Edit Lyrics
Now the thrill is gone, got no patience, 'cause I'm not a doctor
Bino plays on the "patients/patience" homophone to say that he doesn't have the tolerance to help other people with their problems because he has enough to deal with on his own. (Patients are the people that go to the doctor to get treated)
Girl why is you lying, girl why you Mufasa

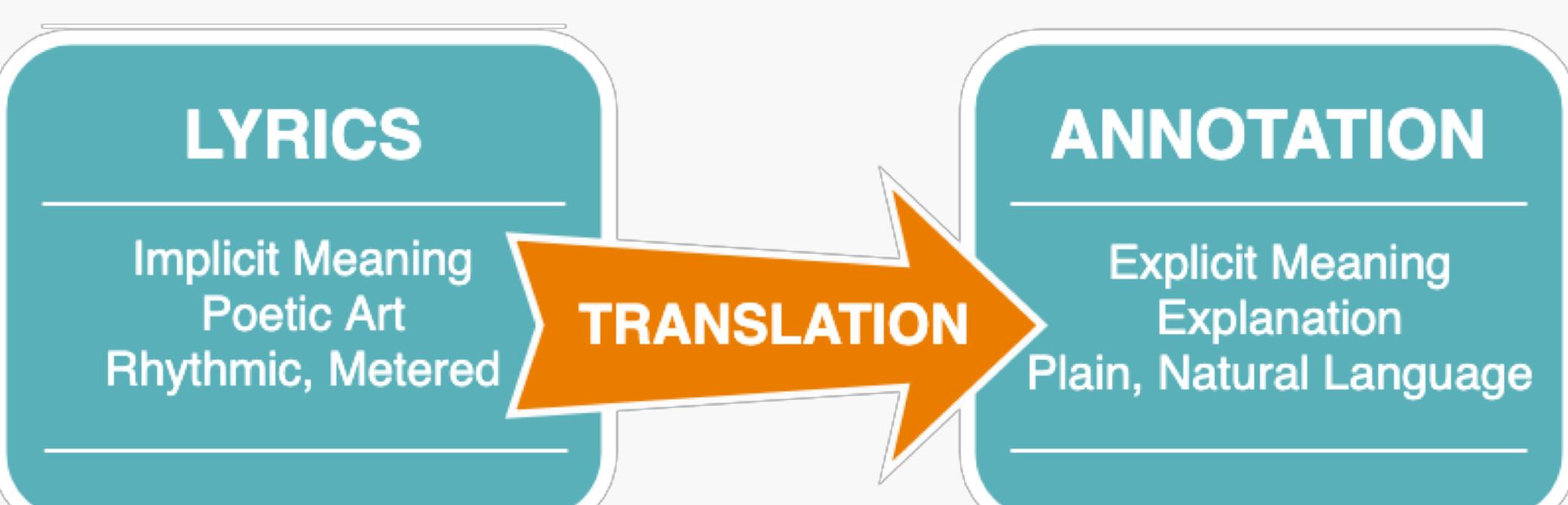
- 9,828 lyric-annotation pairs & metadata from API
- Most popular songs from top 20 artists on Genius



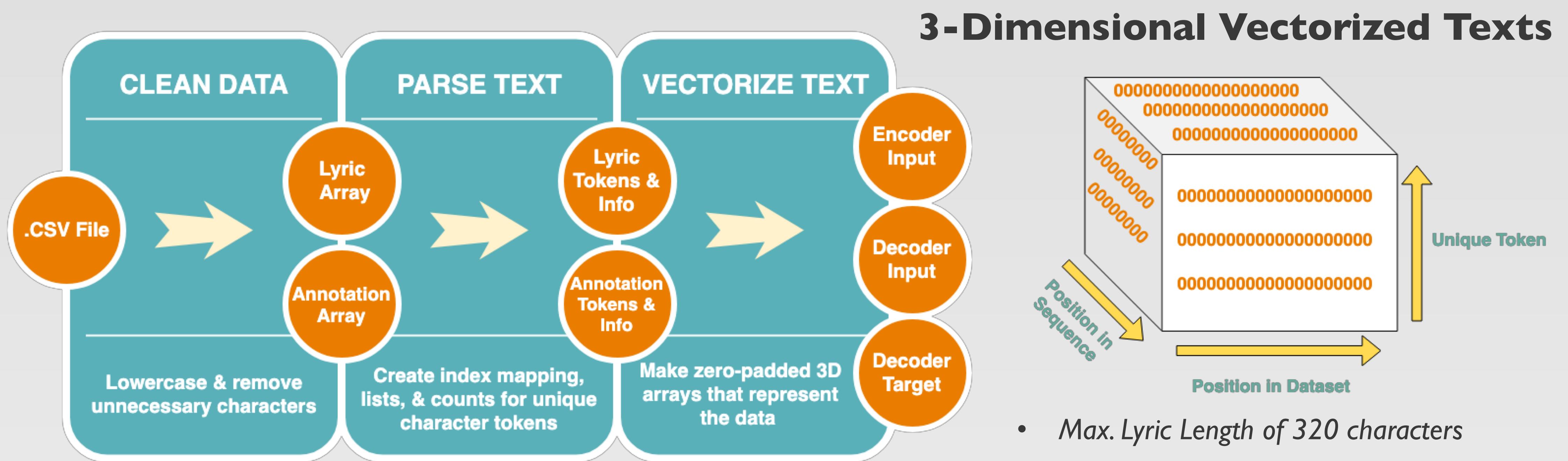
Objectives

The semantic and contextual similarity between these lyrics and their annotations allows us to view this as a translation problem.

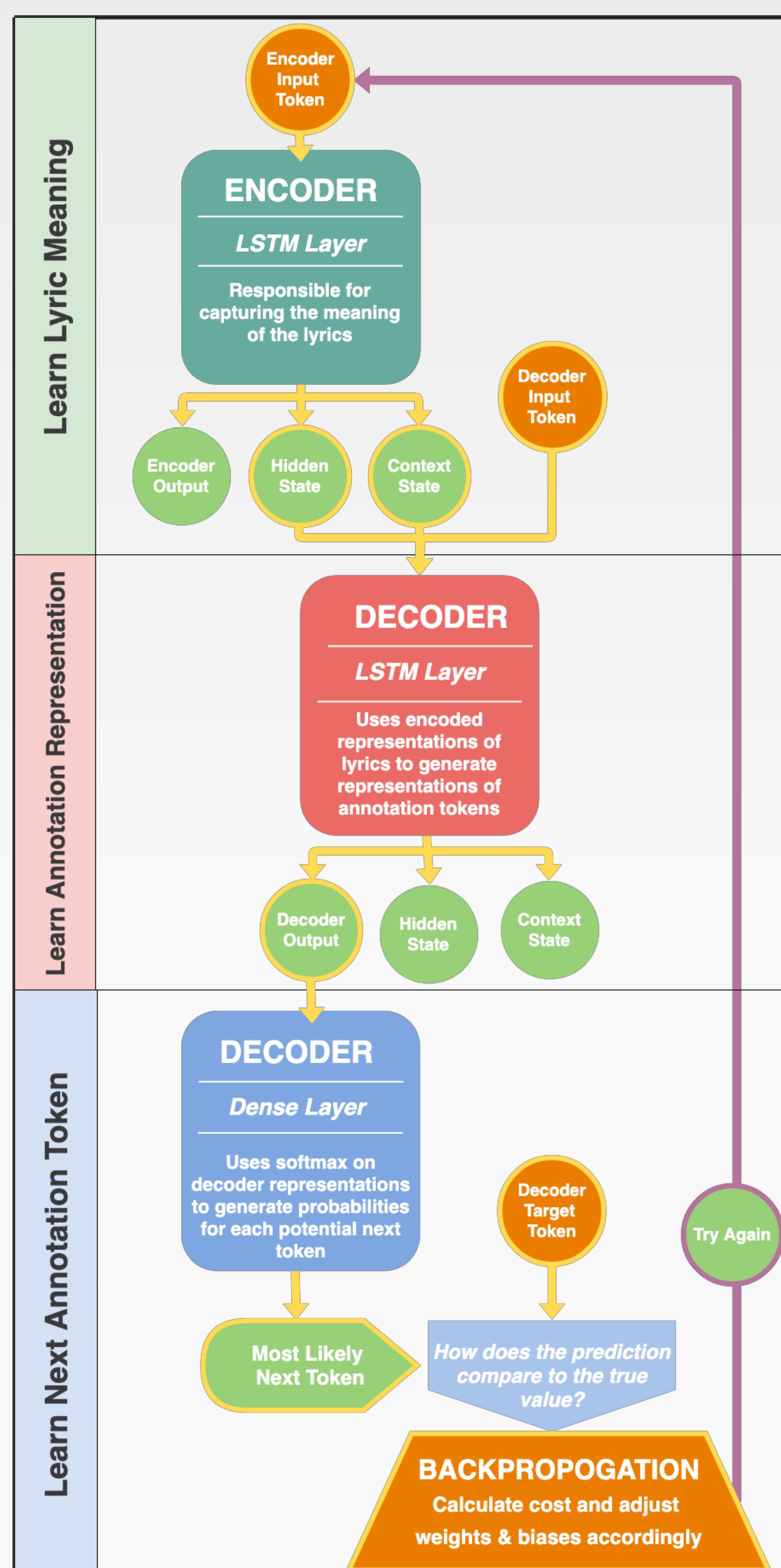
Can a Seq2Seq LSTM model use the linguistic patterns in the data to learn how to explain lyrics?



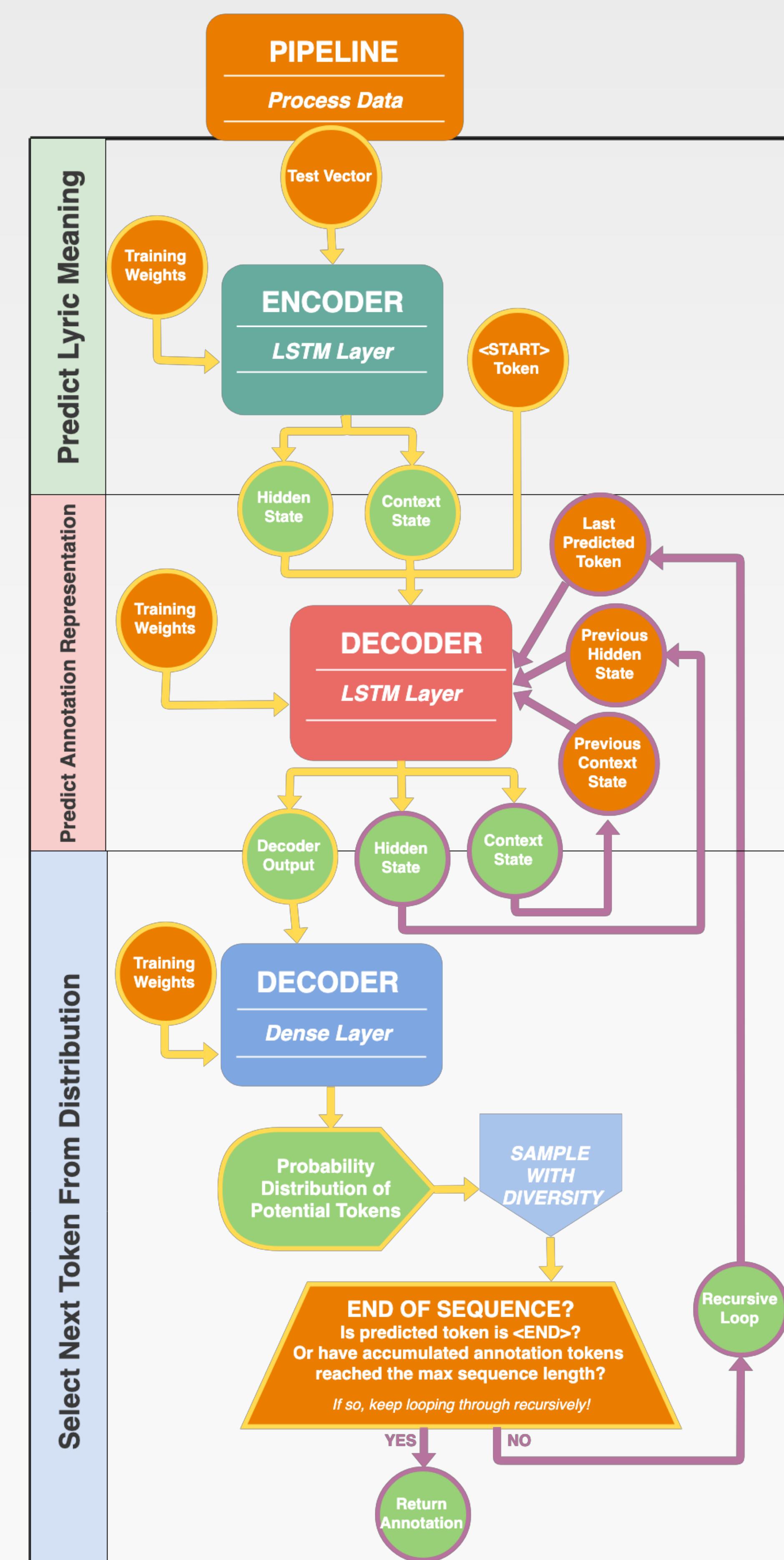
Data Preparation



Training Process



Inference & Generation



Chosen Model Specs

Model Architecture

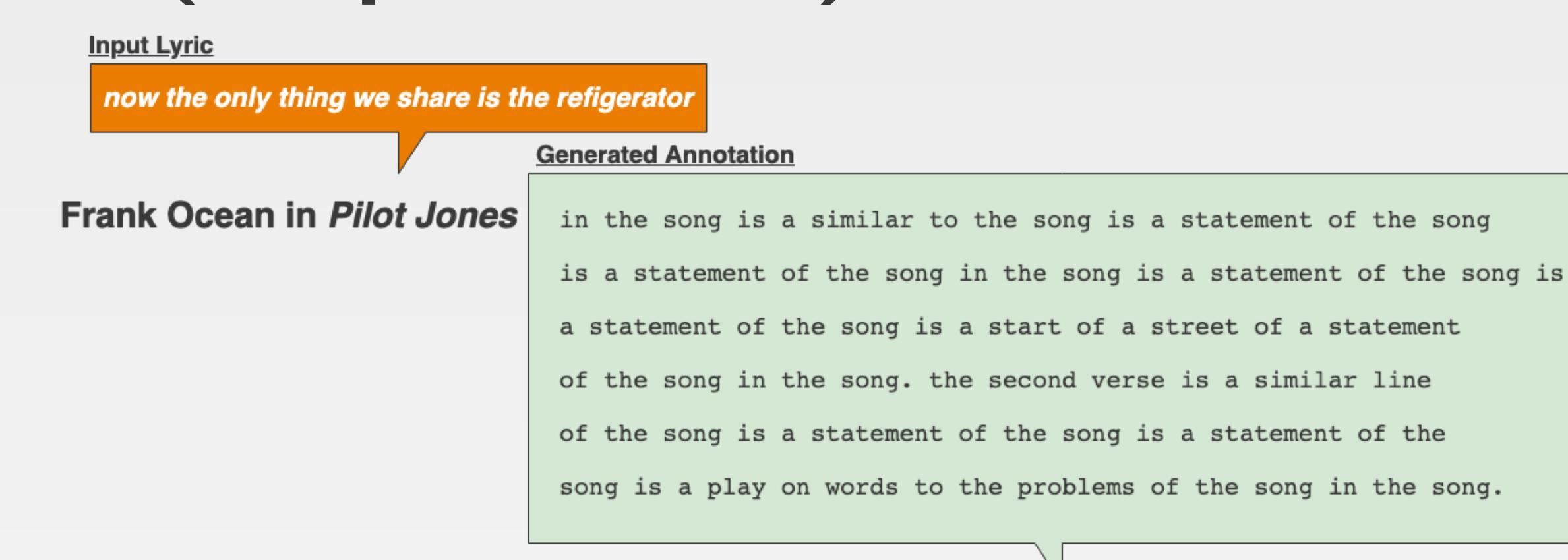
- Character-Based Vanilla Seq2Seq LSTM model
- 512 Latent Dimensions
- Adam optimizer
- Categorical cross-entropy (loss)

Training

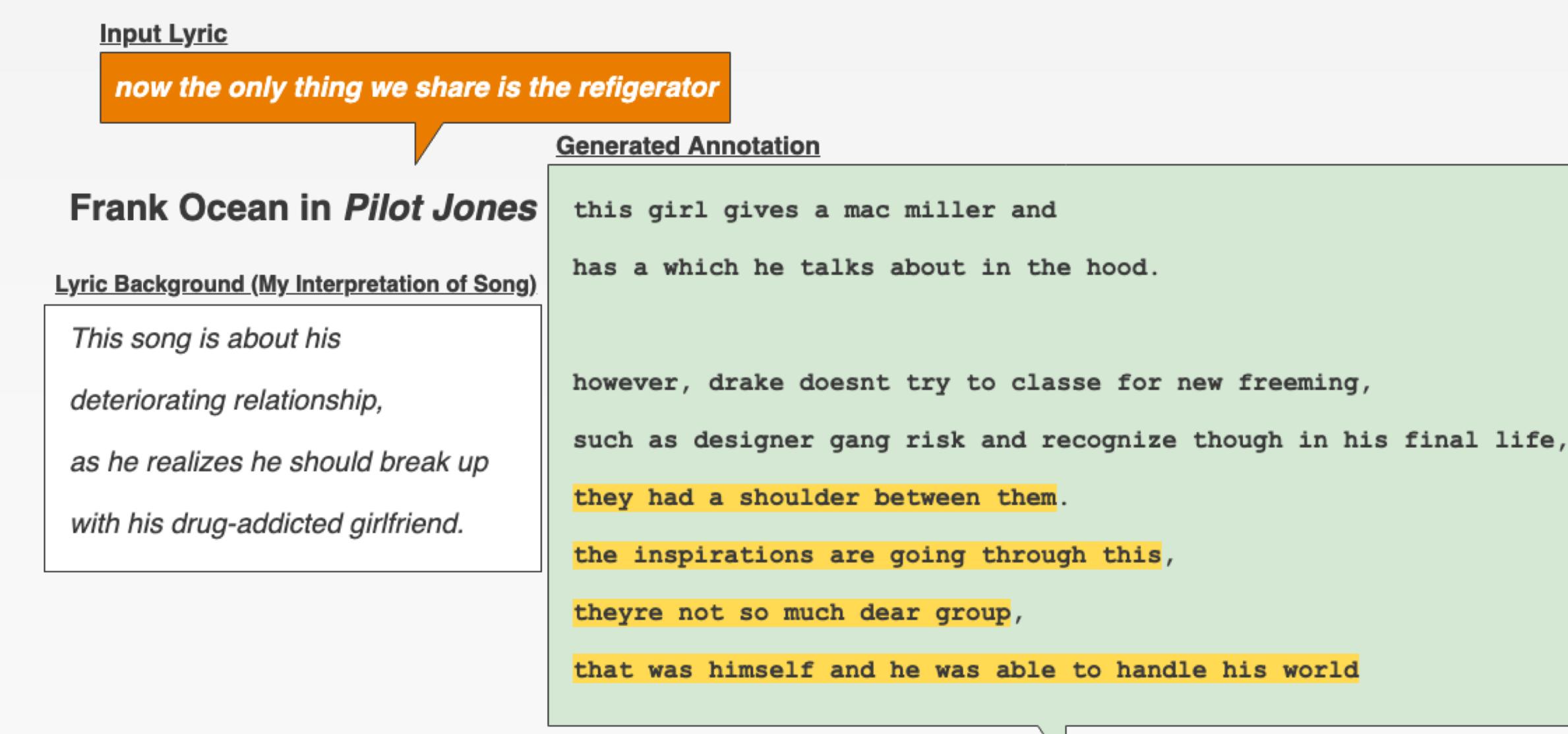
- 100 Epochs
- Batch size of 64
- Trained on 7,889 pairs
- Validated on 876 pairs

Results

Sampling without Diversity (Temperature = 1)



Sampling with Diversity (Temperature = 0.71)



Future Work

Relative Self-Attention & Transformer Models

- Better able to model self-similarity
- Explicit connections to entire history, which are weighted based on positional and contextual importance to current token