Generation Model to Generate Song Lyrics by Artist Name and Genre

Marwan Ashraf Ali

49-5622

Introduction and Motivation

This report explores the utilization of Natural Language Processing (NLP) techniques to generate artist-specific song lyrics, driven by a fusion of artistic curiosity, technological innovation, and the desire to understand the intricate nuances of musical expression. By delving into the linguistic patterns, themes, and stylistic flourishes of artists' lyrical repertoire, we aim to unravel the mysteries of musical identity, while also shedding light on the evolution of genres, cultural trends, and socio-cultural influences. Beyond academic inquiry, this endeavor holds potential for assisting aspiring songwriters, enhancing music recommendation systems, and fostering collaborative innovation within the music industry, thus enriching our understanding of the profound connections between language, music, and human creativity.

Literature Review

A novel method for generating rap lyrics is introduced, emphasizing both creativity and lyrical complexity. This approach incorporates a prediction model utilizing RankSVM and a unique deep neural network structure to identify the next line of lyrics from a pool of candidates. The model demonstrates an accuracy of 17% in predicting the true next line, significantly outperforming random selection. By leveraging this prediction model, existing song lines are combined to create lyrics with meaningful content and rhyme. Evaluation indicates that the generated lyrics surpass human rappers in quantitative rhyme density by 21%. The resulting rap lyrics generator, named DeepBeat, is available online and its performance is validated through usage logs, revealing alignment between machine-generated rankings and user preferences. [1]

The field of automatic lyrics generation has witnessed significant interest from both the music and AI communities over the years. Traditional rule-based methods have largely given way to deep learning-based systems, thanks to advancements in computational power and data-driven models. However, many existing approaches either heavily rely on prior musical and lyrical knowledge or oversimplify the task by neglecting melodic information. In this study, we propose an end-to-end melody-conditioned lyrics generation system using Sequence Generative Adversarial Networks (SeqGAN). This system generates lines of lyrics based on given melodies, with an additional investigation into the impact of incorporating thematic input conditions. Our findings reveal that incorporating input conditions such as melody and theme does not negatively affect evaluation metrics, while enhancing the network's ability to produce more meaningful results. [2]

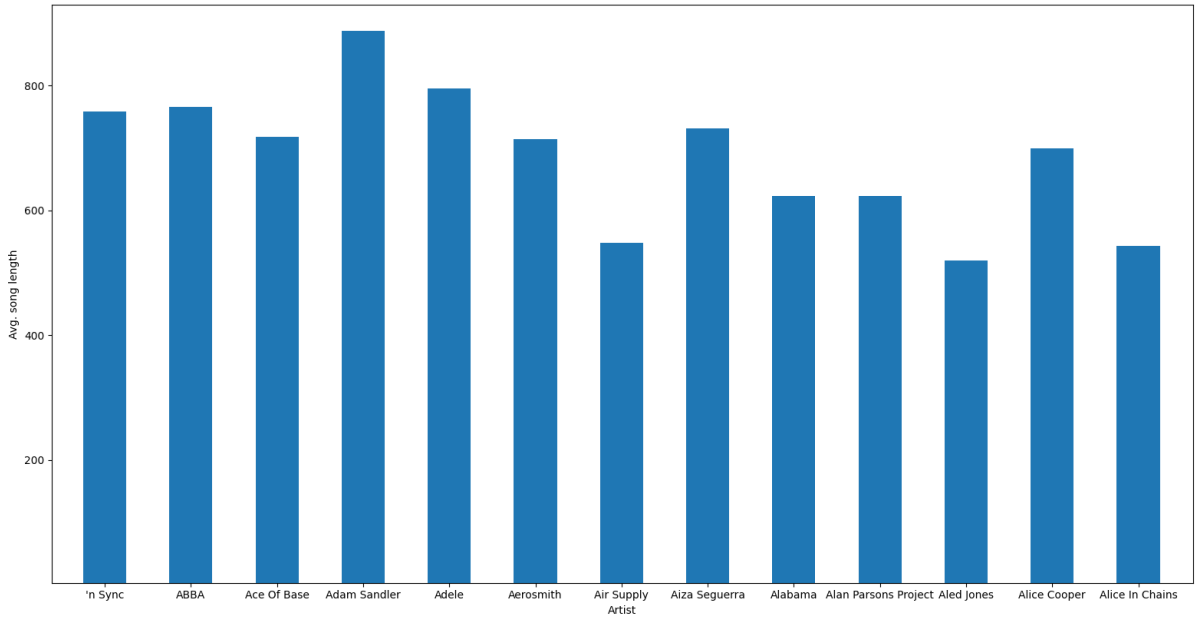
This study investigates the potential of deep learning techniques in generating lyrics tailored to a particular musical genre. While prior research in computational linguistics has primarily focused on lyric generation within specific genres using Recurrent Neural Networks (RNN) or Gated Recurrent Units (GRU), we introduce the use of a Long Short Term Memory (LSTM) network for this purpose. Our approach involves generating lyrics for a designated genre based on a given sample lyric. Furthermore, we conduct an evaluation of the generated lyrics using various linguistic metrics, comparing them with those of other genres and the training dataset to analyze linguistic similarities, differences, and the efficacy of our network in producing semantically aligned lyrics. Our findings demonstrate that the LSTM model performs well in generating rap and pop lyrics, closely replicating average line length and word variation both within songs and across genres compared to the training data. [3]

We introduce AI-Lyricist, a system designed to generate original and meaningful lyrics based on a provided vocabulary and MIDI file. This task involves addressing several challenges, such as automatically identifying the melody, extracting a syllable template from multi-channel music, generating creative lyrics in line with the input music's style and syllable alignment, and adhering to vocabulary constraints. To tackle these challenges, we propose an automatic lyrics generation system comprising four modules: (1) A music structure analyzer to derive the musical structure and syllable template from the MIDI file, employing the concept of expected syllable number for improved melody identification. (2) A SeqGAN-based lyrics generator optimized through multi-adversarial training with twin discriminators for text quality and syllable alignment. (3) A deep coupled music-lyrics embedding model to project music and lyrics into a joint space, facilitating fair comparison of melody and lyric constraints. (4) A Polisher module to ensure vocabulary constraints are met by applying a mask to the generator and substituting words to be learned. We trained our model on a dataset containing over 7,000 music-lyrics pairs, enriched with manually annotated labels for theme, sentiment, and genre. Both objective and subjective evaluations demonstrate AI-Lyricist's superior performance compared to state-of-the-art approaches for the specified tasks.[4]

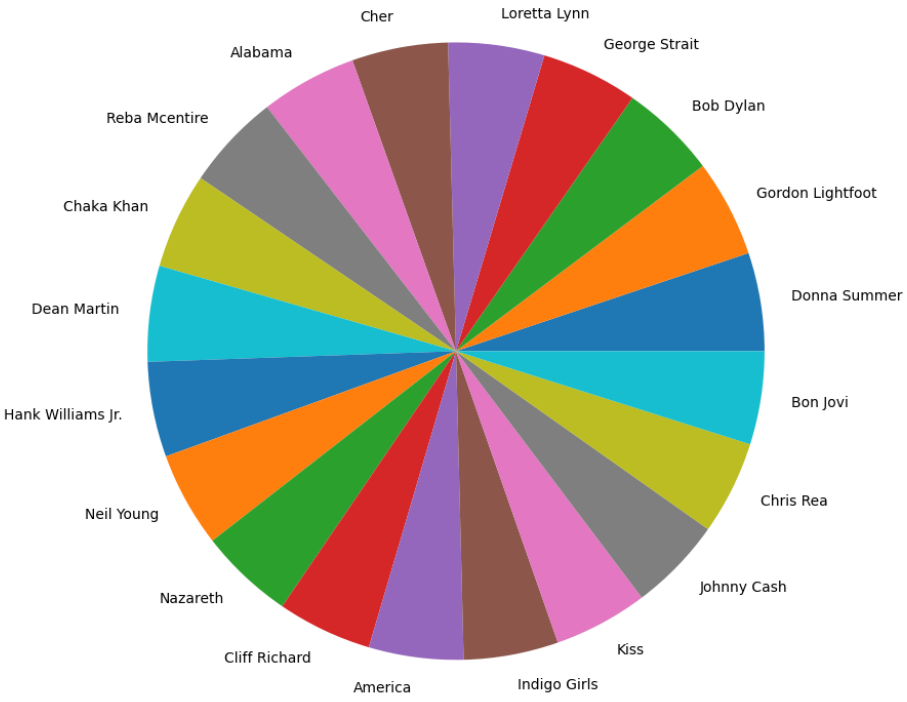
Over time, both the music and AI industries have shown keen interest in automating lyric creation, transitioning from early rule-based methods to deep learning systems with advancements in computing power and data-driven models. This study investigates the potential of deep learning models to generate lyrics tailored to specific musical genres and artists, shifting focus from previous reliance on Recurrent Neural Networks (RNN) or Gated Recurrent Units (GRU) to LSTM networks. It explores LSTM's capacity to craft lyrics based on designated genres or artists using provided seed lyrics as input, while also examining how modern deep learning techniques can enhance the songwriting process by learning from established artists and their significant works. The proposed LSTM model adeptly generates rap and pop lyrics, closely resembling characteristics such as average line length and word variation within songs and across genres. Through parameter adjustments, the neural network is pretrained on lyrics from various composers and musicians, employing a multilayer LSTM-based training model with bidirectional neurons and BERT integration. Input lyrics are segmented into word and rhyme indices to generate a comprehensive set of lyrics, with each trial producing a unique outcome and achieving a loss of less than 0.06 when parameters are optimized. The study analyzes the impact of different dropout positions on results and identifies some model-generated lyrics as of suitable quality, ultimately suggesting that deep learning techniques hold promise in supporting the creative process of songwriting. [5]

Data Analysis

Analysis of the data has shown some general statistics about the data regarding song lengths. This is before and after cleaning the data, where the strings cleaned were the actual songs themselves. The statistics are shown in the below figure.



Beyond that, the data told us who were the artists most listened to shown by the below figure.



Data Limitations

The data does not provide any information regarding how many times a single song was listened to, which could have given valuable insights into what makes a song more attractive and more probable to be played, and hence better from a listener's perspective.

Methodology

In order to generate song lyrics based on our given dataset, the data given is first tokenized using the Keras tokenizer. Afterwards, input sequences are created based on those tokens then padded to ensure they are the same length, before we finally define and train a Long-Short Term Memory (LSTM) model, which is trained and used to generate the lyrics based on seed text. The model is a sequential model consisting of an embedding layer, an LSTM layer and a dense layer for a total of three layers. The data is fed into the model in batches to reduce memory usage. On fitting and training the model to the training data, despite the implementation of batching mechanisms, the code still crashed due to using up all the memory. After all other solutions were exhausted, the size of the training data was reduced all the way

down to two hundred rows, not only to solve the error being raised due to insufficient memory, but also to reduce training time, admittedly at the expense of accuracy.

Results

The model successfully generates words based on the training data, although the output data is just gibberish. This is attributed to the small training data given to the model. Various seed texts were tested to see the output of the model to ascertain the fact that this indeed the general case, and is not something specific to certain inputs. Below is an example seed text and its corresponding output.

Seed text: "take me on a whim"

Output: "take me on a whim side dont walk away dont walk away dont want say dont stop dont stop dont stop dont stop dont stop"

In the future, it is advised to try adjusting the batch size in order to accommodate a bigger, more sensible training dataset size. Moreover, it is recommended to add a spell checker to handle spelling mistakes in the text as part of cleaning the data for better quality of output.

Resources

[1] Malmi, E., Takala, P., Toivonen, H., Raiko, T., & Gionis, A. (2016, August 1). *A Computational Approach to Rap Lyrics Generation*.

[2] Chen, Y., & Lerch, A. (2021, January 22). *Melody-Conditioned Lyrics Generation with SeqGANs*.

[3] Gill, H., Lee, D., & Marwell, N. (2021, August 21). *Deep Learning in Musical Lyric Generation: An LSTM-Based Approach*.

[4] Ma, X., Wang, Y., Kan, M.-Y., & Lee, W. S. (2021, October 17). *AI-Lyricist: Generating Music and Vocabulary Constrained Lyrics*.

[5] Dhandapani, A., Ilakiyaselvan, N., Mandal, S., Bhadra, S., & Viswanathan, V. (2023, June 11). *Lyrics Generation Using LSTM and RNN*.