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.

To develop Korean lyrics reflecting the distinctive features of K-pop, we innovatively utilized reversed K-pop lyrics data as training data. This approach enabled the model to focus on specific elements of sentences, such as predicates and conjunctions, to initiate string generation. Our proposed method for generating song lyrics also incorporates contextual considerations between lyrics. During lyric generation, the model undergoes randomized adjustments based on a blank, ensuring diverse outputs. Comparative analysis confirmed that lyrics generated from the reversed data exhibit a more natural contextual flow compared to those generated from forward data. Additionally, our method facilitates the creation of new lyrics with structures akin to existing ones.

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.

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.

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.

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.

```
Avg. song length:  1179.504891587164
Avg. song length - no stopwords:  670.6573633998265
Max. song length:  3997
Max. song length - no stopwords:  2807
Min. song length:  400
Min. song length - no stopwords:  102
```

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

	values	counts
0	Donna Summer	191
1	Gordon Lightfoot	189
2	Bob Dylan	188
3	George Strait	188
4	Loretta Lynn	187

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.