NLP Project

Nada Bakeer 49-0360, Nour Shehab 49-0463

March 14, 2024

1 Motivation

This paper investigates the research domains of natural language processing and information retrieval in the context of lyrics production, depending on artist names. By analyzing a vast dataset of lyrics associated with certain artists, we expect to uncover patterns and insights influenced by their styles and personalities. This technique applies NLP and IR research to artistic expression, potentially impacting music recommendation systems and tailored content production.

2 Introduction

In this paper we are going to be discussing some of the previous work, followed by data preparation and analysis to prepare our dataset for the training process.

3 Literature Review

In a paper by Gill et al.[Gom22], a method for generating lyrics using deep learning was explored. They create a word dictionary from lyrics from various musical genres and use a neural network to anticipate the next word, newline, or punctuation in the lyrics. The network design consists of an embedding layer, an LSTM layer with dropout for regularization and a linear layer to generate a probability vector based on the vocabulary. Due to computational constraints, training was limited to six genres. Results showed that a specific input size (k = 16) effectively captured overarching structural dependencies in lyrics, such as verse-chorus structures, whereas smaller values caused complications such as repetition. The study's limitations were constraints in computational resources that restricted genre selection for training and limited the amount of the training datasets. Furthermore, the task of adapting pre-existing models such as GPT-2 from prose to lyrical creation highlighted the complexity of transferring models across different writing styles. Future research suggestions include exploring alternative model types such as Markov Models and GRUs Overall, the study shows potential in creating lyrics for various genres utilizing modern deep learning algorithms.

In another study, Saeed et al.[SIZ19] presented Creative-GAN, a modified Generative Adversarial Network (GAN) designed for the development of creative literature such as poems, lyrics, and metaphors. Their approach involved training a language model using backpropagation over time, with a focus on maximum likelihood estimation (MLE) during the pre-training phases and including discriminator creative rewards during GAN training. The discriminator gives useful feedback to the generative model via a cost function that encourages the use of inventive tokens. Creative-GAN consists of a model trained using the AWD-LSTM and TransformerXL architectures, as well as a discriminator with an encoder similar to the generator encoder but with additional decoding layers. Based on the input sequence's hidden states, the discriminator decoder generates a numerical output between 0 and 1. Creative-GAN was tested on a variety of creative datasets, including classical and contemporary English poems, metaphor sentences, and song lyrics. It outperformed baseline language models and GumbelGAN models in terms of perplexity scores, indicating improved performance in generating creative text. Nonetheless, the non-differentiable nature of text presents issues, as does the previously described GPT-2 technique[Gom22]. To address this, the generator model parameters are changed using policy gradients. Although Creative-GAN produces promising results in creative text production, additional optimisation of the architecture and training methods may be required for various creative writing jobs.

Rodrigues et al.[RdPOMdAP22]present a novel way for creating song lyrics by refining a pre-trained GPT-2 model with English and Portuguese lyric datasets. Using the multitasking capabilities of the GPT-2 model, which is built on the Transformer architecture and offered by OpenAI, the project intends to develop musically coherent and grammatically sound lyrical content using a model that was not specifically designed for this purpose.

A study of the created lyrics is offered, with emphasis on spelling, grammar, and semantics. The research digs into efforts to identify underlying similarities in the generated texts and compares them to human-authored song lyrics. The findings show promise in using pre-trained models such as GPT-2 to generate poetic content, highlighting the benefits of transfer learning in optimizing computational resources.

However, significant limitations are highlighted, such as potential issues in dealing with metaphorical language which was also highlighted in the previously mentioned papers, the use of synonyms, and paraphrasing in the created lyrics. Furthermore, the study may not fully capture all aspects of lyrical creation and may require additional modification to improve the quality and diversity of the created texts.

In this paper, Madhumani et al.[MYH+20] use automatic neural lyrics and melody composition (AutoNLMC). The AutoNLMC comprises two models, the first being the lyrics generator, which utilizes a recurrent neural network (RNN) that generates the upcoming token. In other words, the RNN contains a hidden state that is used to build the base upon which the model will predict the next token which in this case is a sequence of syllables. The objective of this model is to learn the syllable distribution and to predict it later on. The RNN uses the input tokens to adjust the hidden state, for it to be used in the upcoming token prediction, as the essence of RNN is that it generates the new token based on the previous tokens as well as the current token (input). This is executed using a non-linear function that could be a sigmoid function or long short-term memory (LSTM).
Moving on to the melody composer which operates using an RNN as well. The model uses a sequential encoder-decoder, where the melody composer model takes lyrics tokens as input to process and outputs hidden states. The hidden state is then passed on to the decoder to generate a melody. Lastly, the encoder generates a dynamic context vector that encodes parts that should be emphasized for the melody composer RNN model. The proposed model was able to generate 10 full songs that were constrained to 100 syllables, the length of the seed lyrics used for the model input was variable. Multiple forms of lyrics encoding were used, thus models were trained separately for each representation. The forms used were syllable embedding (SE), addition of syllable and word vector (ASW), syllable and word embedding concatenation (SWC), and concatenated syllable, word, and syllable-projected word vector (CSWP). Conclusively, both embedding forms ASW and CSWP proved to be of better performance in identifying the correlation between lyrics and melodies.

4 Data Preparation and Analysis

4.1 Dataset

The dataset chosen in this project, is the 57,650 spotify songs dataset. It was chosen as it has a wide variety of artists and each artist has multiple song, which helps in not overfitting the model.

4.2 Data Cleaning

In this section we will be discussing the data cleaning processes done to prepare the data for usage.

First the data was checked for null values, however no null values were found, so the next step was renaming the column "text" to "lyrics" for a clearer more expressive title. The next step was checking for full row duplicates and removing them if any, however the dataset was clean of that too. Next the lyrics characters and the artist names were all switched to lower case to ease searching and information retrieval.

Following this we removed any non regular words characters from the lyrics to prepare it for tokenization, hence the next step was tokenizing each song lyrics.

When preparing the data we decided against stop words removal and lemmatization to avoid changing the sentence structure which is a key part of the song rhyme and over all coherence. Spelling check was also ignored to keep words such as "sleepin'"form changing the rhyme to "sleeping".

By this we conclude the data cleaning and prepare for the data analysis.

4.3 Data Analysis

To be able to study the dataset data analysis is required. The first technique we used was finding out the most frequent artist and how many songs do they have. This helps in estimating the most songs the model would need to analyze at a time. The results showed that the top artist is "Donna Summer" with 191 songs in the dataset.

The second analysis method we used was figuring out the length of each song as shown in Figure 1, which would later help in identifying the expected output length of the model.

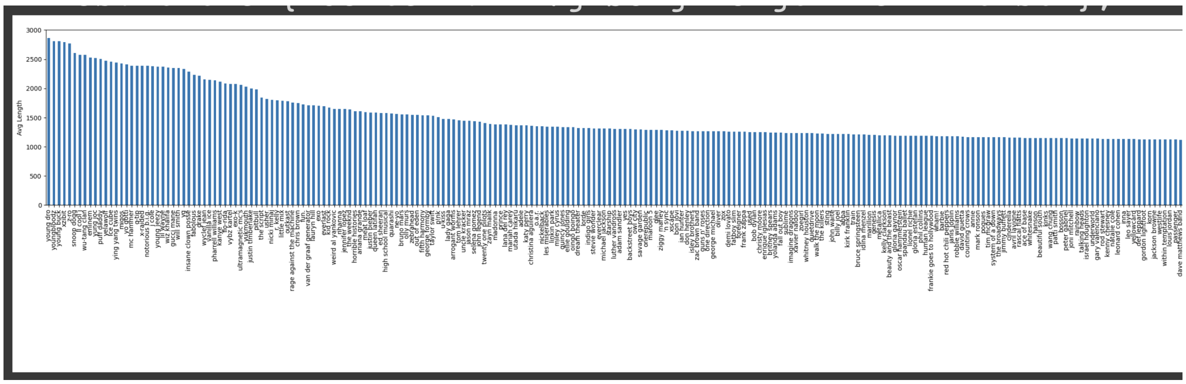

Figure 1: Partial snapshot of the song length analysis

The third data analysis method we implemented was figuring out the most frequent word per each artist songs. This created a table of artists and the most frequent word they use. This may be helpful in analysing the model output later on.

Using the most frequent word per artist we then calculated the tf-idf for each which might be helpful later on for keyword extraction and information retrieval.

The results of the most frequent words and the tf-idf weren't as relative as expected so stop words removal was implemented before applying them.

References

[Gom22] Walid Gomaa. Lyrics analysis of the arab singer abdel elhalim hafez. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 22:1 – 27, 2022.

[MYH+20] Gurunath Reddy Madhumani, Yi Yu, Florian Harscoët, Simon Canales, and Suhua Tang. Automatic neural lyrics and melody composition. *arXiv preprint arXiv:2011.06380*, 2020.

[RdPOMdAP22] Matheus Augusto Gonzaga Rodrigues, Alcione de Paiva Oliveira, Alexandra Moreira, and Maurílio de Araújo Possi. Lyrics generation supported by pre-trained models. *The International FLAIRS Conference Proceedings*, 2022.

[SIZ19] Asir Saeed, Suzana Ilić, and Eva Zangerle. Creative gans for generating poems, lyrics, and metaphors. 09 2019.