CS395T Project Proposal

Lucas Kabela and Bill Yang

September 24th 2020

1 Goal Of the Project

This project aims to make progress on conditional text generation through exploring improvements in conditional lyric generation. In particular, we aim to improve on previous attempts at conditioning on artist, while most prior efforts on enforcing theme and rhyme scheme in lyric generation.

Conditioning generation on persona has seen success in improving consistency and engagement in dialog response agents (Liu et al. 2020 [[5]]) as well as improvements in text style transfer. As such, we hypothesize conditioning on artist / persona will provide improvements to the quality of lyrics generated in conditional lyric generation as well.

2 Motivation

Early models for long-text generation were able to produce syntactical and fluent text, but their output lacks direction and is not natural. Conditional text generation has recently become a popular technique for controlling long text to provide a tune-able direction of output. We however identify 2 key areas that differentiate lyric generation from other long text generation tasks: songs have more structure than summaries, stories, and other long-texts and lyrics follow rhyme schemes.

Prior approaches to lyric and text generation have explored incorporating attributes to steer generation for the issues mentioned. However, these approaches lack consistency with respect to artists. There are many factors that differentiate the style of artists. Previous literature has largely overlooked how personas can be incorporated with attributes to further influence text generation.

3 Related Work

First, we aim to explore works focused on lyric generation. These texts are a mixture of unconditional and conditional generation attempts (Potash et al. 2015 [10]), (Malmi et al. 2016 [6]). (Nikolov et al. 2020 [8]), (Wu et al. 2018 [15])

For our specific problem, (Vechtomova et al. 2018 [14]) attempts to perform genre conditioning, which is a broader version of our work. Rhyming is explored in (Hopkins et al. 2017 [3]).

This is a recent paper on conditioned poem generation which we hope to use for direction in enforcing structure (Yang et al. 2018 [16]).

Lastly we have a variety of papers that deal with authorship of text, personas, and style transfer: (Manjavacas et al. 2017 [7]), (Tikhonov et al. 2018a [11]), (Oliveira et al. 2015 [9]), (Tikhonov et al. 2018b [12]), (Li et al. 2016 [4]), (Fu et al. 2017 [2]), (Liu et al. 2020 [5]).

4   Dataset and Proposed Architecture

For our dataset we will use Genius, which has a large collection of annotated songs and artist profiles.

We will use an LSTM based approach (Potash et al. 2015 [10]) as our baseline, but may consider using the VAE based model(Yang et al. 2018 [16]), or Transformers (Vaswani et al. 2017 [13].

For our own approach, we have a handful of approaches to potentially explore using LSTMs (Potash et al. 2015 [10]), CVAE (Yang et al. 2018 [16]), artist attribute models, (Dathathri et al. 2020 [1]), personas (Li et al. 2016 [4]), and/or discriminators.

5   Evaluation

The main things we want to measure is if our generated text seem like song lyrics, if the lyrics are unique, and if the style is similar to the given artist. Rhyme density (RD) is a common metric used in lyric evaluation to see if fluent rhyme scheme has been achieved. We plan to use distinct N-gram/self-BLEU matching to determine the uniqueness of the lyrics generated. (Potash et al. 2015 [10]) presents an automatic metric to measure an artist similarity score, and uniqueness of the lyrics generated. Sample cross entropy (Tikhonov et al. 2018b [12]) can also be used to determine the similarity between texts based on their artists.

We will also perform some kind of human evaluation as well. Details TBD.

References

[1] Sumanth Dathathri, Andrea Madotto, Janice Lan, Jane Hung, Eric Frank, Piero Molino, Jason Yosinski, and Rosanne Liu. Plug and play language models: A simple approach to controlled text generation, 2019.

[2] Zhenxin Fu, Xiaoye Tan, Nanyun Peng, Dongyan Zhao, and Rui Yan. Style transfer in text: Exploration and evaluation. *CoRR*, abs/1711.06861, 2017.

[3] Jack Hopkins and Douwe Kiela. Automatically generating rhythmic verse with neural networks. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 168–178, Vancouver, Canada, July 2017. Association for Computational Linguistics.

[4] Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. A persona-based neural conversation model. *CoRR*, abs/1603.06155, 2016.

[5] Qian Liu, Yihong Chen, Bei Chen, Jian-Guang Lou, Zixuan Chen, Bin Zhou, and Dongmei Zhang. You impress me: Dialogue generation via mutual persona perception. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 1417–1427, Online, July 2020. Association for Computational Linguistics.

[6] Eric Malmi, Pyry Takala, Hannu Toivonen, Tapani Raiko, and Aristides Gionis. Dopelearning: A computational approach to rap lyrics generation. *CoRR*, abs/1505.04771, 2015.

[7] Enrique Manjavacas, Jeroen De Gussem, Walter Daelemans, and Mike Kestemont. Assessing the stylistic properties of neurally generated text in authorship attribution. In *Proceedings of the Workshop on Stylistic Variation*, pages 116–125, Copenhagen, Denmark, September 2017. Association for Computational Linguistics.

[8] Nikola I. Nikolov, Eric Malmi, Curtis G. Northcutt, and Loreto Parisi. Conditional rap lyrics generation with denoising autoencoders, 2020.

[9] Hugo Gonçalo Oliveira. Tra-la-lyrics 2.0: Automatic generation of song lyrics on a semantic domain. *J. Artif. Gen. Intell.*, 6(1):87–110, 2015.

[10] Peter Potash, Alexey Romanov, and Anna Rumshisky. GhostWriter: Using an LSTM for automatic rap lyric generation. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1919–1924, Lisbon, Portugal, September 2015. Association for Computational Linguistics.

[11] Aleksey Tikhonov and Ivan Yamshchikov. Sounds wilde. phonetically extended embeddings for author-stylized poetry generation. In *Proceedings of the Fifteenth Workshop on Computational Research in Phonetics, Phonology, and Morphology*, pages 117–124, Brussels, Belgium, October 2018. Association for Computational Linguistics.

[12] Alexey Tikhonov and Ivan P. Yamshchikov. Guess who? multilingual approach for the automated generation of author-stylized poetry. *CoRR*, abs/1807.07147, 2018.

[13] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *CoRR*, abs/1706.03762, 2017.

[14] Olga Vechtomova, Hareesh Bahuleyan, Amirpasha Ghabussi, and Vineet John. Generating lyrics with variational autoencoder and multi-modal artist embeddings. *CoRR*, abs/1812.08318, 2018.

[15] Xing Wu, Zhikang Du, Y. Guo, and H. Fujita. Hierarchical attention based long short-term memory for chinese lyric generation. *Applied Intelligence*, 49:44–52, 2018.

[16] Xiaopeng Yang, Xiaowen Lin, Shunda Suo, and Ming Li. Generating thematic chinese poetry with conditional variational autoencoder. *CoRR*, abs/1711.07632, 2017.