
Genre, Artist, and Melody Dependent Lyric Generation Using Pretrained Transformer Based Language Models

Abstract

This paper describes a GPT-2-based lyric generation model, fine-tuned to generate stylistic lyrics from existing songs. The work is decomposed into three phases. First, a BERT-based classifier is built to identify the genre of songs to show that the style of the lyrics change depending on their genre. Second, a GPT-2 based generator is built to create lyrics from the song dataset based on their genre and artist. Finally, the generator is extended to create lyrics based on the given music in MIDI format. We evaluate the quality of the generated lyrics with respect to genre, artist, and melody.

1 Introduction

The primary factor in differentiating between songs is identifying the separate artist and genre background. Genres range based on their cultural identities, and songs with a respective background tend to be different from songs of other genres [7]. As a result, it is possible to identify which genre a song is from when given information and examples from each genre.

Our objective is to be able to identify genres and generate lyrics accordingly from data consisting of hundreds of thousands of compositions. We want to be able to verify that lyrics are more respective to their genres than initially thought and subsequently generate lyrics that are independent to an artist or genre. To do this, we use a classifier, to split the lyrics into their genres, and a generator to create new lyrics. Furthermore, the generator generates lyrics based on specific artists and uses the melodies of existing songs to generate lyrics. Because most artists have a relatively unique style of music, it is easier to evaluate the quality of the lyrics generated.

Lyric generation has been an active research area, especially since the wide-spread use of neural network based approaches for language generation [11]. Previous studies mainly utilized recurrent neural networks with long short term memory (LSTM) [4] cells that are trained from a collection of lyrics, and trained models specialized for a genre [6, 8]. Furthermore, stylistic features of artists [12] or artist embeddings learned by a CNN classifier and trained to predict artists based on MEL spectrograms of their song clips [9] were also used as input. This work employs transformers, similar to [2]. However, we start from a large, pre-trained language model, and focus on specialization to genre, artist, and melody. The third phase of our model draws parallels with [13], where a melody is generated from lyrics using an LSTM in an adversarial training setup. Our model aims to do the reverse; it generates lyrics based on a melody and uses the dataset released in their work. [10] has proposed an LSTM based approach to generate lyrics given melody using a small dataset.

While GPT-2 based approaches have been extensively used for various language generation tasks such as story generation [3] or poetry generation [1], to the best of our knowledge it has not yet been employed for lyric generation.

The reason for our interest in this topic is based on our prior knowledge about music and songwriting, as we believe that research into this topic will be beneficial and innovative in creating music. Our novel contributions include (1) fine-tuning pre-trained transformer-based language models for genre- and artist-dependent lyric generation, (2) an analysis of dependency of lyrics on genre, and (3) conditioning lyric generation on melody to generate matching lyrics.

2 Approach

This study is conducted in three phases: In the first phase, we classify genres given lyrics, relying on BERT-based embeddings with a linear classification layer without fine-tuning the model. The data

Data	Ep.	Representative Generated Lyrics
All songs	1	Every life, every light. It is a journey of the brave. Every single night. When the moon is full of moonlight. We will ride..
All songs	6	Don't say your prayers. Don't say you miss me. Cuz I'm missing you much too. I'm missing you too. Don't say your prayers..
Rock	1	I wanna tell you about the things we learned.. In the days when the sun went dry. I just thought I'd give up.. Sometimes my heart..
Rock	6	My buddies tell me I'm crazy. That I've been missing. Oh, I can see that. They believed me. Then they went away..
Eminem	1	My album hits the road. That's right up here! (F***ing up). I'm still here, the one time the sun won't shine in my face..
Eminem	6	Can I borrow your car, my car. There's something in the air that's making me sick. Are you trying to save me, save me..

Table 1: Sample lyrics generated by our model trained using all, only Rock and only Eminem songs.

Actual	Predicted
You always make me smile when I'm feeling down, you He sticks to his guns, he take the road as it comes Hey, hey, bye, bye, oh, oh, I'm doin' this	If you were me an ey would you remember Don't wish you out of mine love is real Hey hey Here we at first we at first we were raised We were raised up!

Table 2: Actual and predicted lyrics of songs with given notes in MIDI files.

used in the first phase comes from a Kaggle competition which consists of thousands of different artists and their respective genres along with over 200,000 songs and their lyrics¹. We clean the data, to filter non-English songs for this study, and use only the artists whose songs belonged to a single genre, as the genre annotations are transferred from artist annotations. We use 9400 songs from the filtered data for testing, 9400 for validation, and 76000 for training. We evaluate its quality based on its ability to classify the genres correctly out of the six genres we used: rock, pop, hip-hop, samba, sertanejo, and funk carioca.

The second phase uses a GPT-2 model fine-tuned on the same lyrics data. We consider three setups: using a specific genre, specific artist, and all the data. Using specific artist we have seen that model tends to overfit quickly, as there is not too much data. During generation, we have employed top-k and nucleus sampling with with beam search to get more creative lyrics [5].

The third phase uses data from [13] in which MIDI files are created from lyrics. However, using the same data, the reverse was done, as lyrics were generated from MIDI files. Each MIDI note is featured as a triplet: the ID of the note (0 to 127), the duration of the note, and the duration of the pause before the note. The GPT-2 model is fine-tuned, with 7500 songs, as if these values are tokens IDs, similar to adaptation to a new language or vocabulary. During testing, melody is given as context and lyrics are generated similar to phase two.

3 Discussion and Future Work

In the first phase, we test the model by classifying songs into distinct genres. We achieve an accuracy of 83.9%, which is significantly higher than the chance accuracy of 61.6% (choosing all as "rock"). This verifies our intuition that each genre's respective lyrics have unique stylistic jargon.

In the second phase, we generate lyrics based on the artist or genre using the same data. Table 1 provides three sets of generated lyrics: using all the data, using a particular artist (in our case, Eminem), using a specific genre (in this case, rock). We use the lyrics generated after the first epoch and the sixth epoch of the fine-tuning. As shown from the examples, there is a clear distinction between lyrics from Eminem, lyrics from the overall genre of rock, and all of our lyrics combined. Eminem's lyrics contain more vulgar language and slang, as most hip-hop songs are, and the rock lyrics contain language that is less commonly seen in other genres. The rock lyrics consist of less modern jargon and themes that are present in pop, hip-hop, and most other genres. Sample lyrics are generated after each epoch, as provided in the Appendix Art Work.

We assess the quality of the generated lyrics for phase three using 74 unseen songs. Table 2 provides three example results from the initial portions of the actual and predicted lyrics. As shown in Table 2, the lyrics are parallel to the genuine lyrics in following the melody. We have provided the corresponding 3 MIDI files under <https://tinyurl.com/creative2020midi>.

All models employ Huggingface Transformers library². Upon acceptance, the code will be made publicly available. Our analysis demonstrates that after the initial epochs of fine-tuning, the model manages to generate meaningful lyrics before eventually over-fitting. As part of future work, we plan to use the MIDI note alignment with the generated words.

¹<https://www.kaggle.com/neisse/scrapped-lyrics-from-6-genres>

²<https://huggingface.co/transformers/>

81 4 Ethics

82 Before designing the model, it is important to consider the ethics of a project similar to this study.
83 Firstly, in order to generate meaningful lyrics, existing lyrics are required for the training, which may
84 not be available for various reasons. Second, since the model may generate words and phrases that
85 are offensive, it is especially important to check the lyrics for inappropriate content, such as cuss
86 words or disagreeable innuendo.

87 References

- 88 [1] B. Bena and J. Kalita. Introducing aspects of creativity in automatic poetry generation. In *Proceedings of*
89 *the International Conference on NLP*, 2019.
- 90 [2] P. S. Castro and M. Attarian. Combining learned lyrical structures and vocabulary for improved lyric
91 generation. *arXiv preprint arXiv:1811.04651*, 2018.
- 92 [3] S. Goldfarb-Tarrant, T. Chakrabarty, R. Weischedel, and N. Peng. Content planning for neural story
93 generation with aristotelian rescoring. In *Proceedings of the EMNLP*, 2020.
- 94 [4] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.
- 95 [5] A. Holtzman, J. Buys, L. Du, M. Forbes, and Y. Choi. The curious case of neural text degeneration. *arXiv*
96 *preprint arXiv:1904.09751*, 2019.
- 97 [6] O. Lampridis, A. Kefalas, and P. Tzallas. Greek lyrics generation. In *IFIP International Conference on*
98 *Artificial Intelligence Applications and Innovations*, pages 445–454. Springer, 2020.
- 99 [7] R. Mayer, R. Neumayer, and A. Rauber. Rhyme and style features for musical genre classification by song
100 lyrics. In *Proceedings of the ISMIR Conference*, 2008.
- 101 [8] P. Potash, A. Romanov, and A. Rumshisky. Ghostwriter: Using an lstm for automatic rap lyric generation.
102 In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages
103 1919–1924, 2015.
- 104 [9] O. Vechtomova, H. Bahuleyan, A. Ghabussi, and V. John. Generating lyrics with variational autoencoder
105 and multi-modal artist embeddings. *arXiv preprint arXiv:1812.08318*, 2018.
- 106 [10] K. Watanabe, Y. Matsubayashi, S. Fukayama, M. Goto, K. Inui, and T. Nakano. A melody-conditioned
107 lyrics language model. In *Proceedings of the NAACL/HLT Conference*, 2018.
- 108 [11] T.H. Wen, M. Gasic, D. Kim, N. Mrksic, P. Su, D. Vandyke, and S. Young. Stochastic language generation
109 in dialogue using recurrent neural networks with convolutional sentence reranking. In *Proceedings of*
110 *SigDial*, 2015.
- 111 [12] X. Wu, Z. Du, Y. Guo, and H. Fujita. Hierarchical attention based long short-term memory for chinese
112 lyric generation. *Applied Intelligence*, 49(1):44–52, 2019.
- 113 [13] Y. Yu and S. Canales. Conditional lstm-gan for melody generation from lyrics. *arXiv preprint*
114 *arXiv:1908.05551*, 2019.

5 Appendix: Phase II Generated Lyrics

Data	Ep.	Representative Generated Lyrics
All songs	1	"Yeah..we made our way..we made our way over to the next day's place.... we made our way over to the next day's place..."
All songs	1	"The Last Temptation. Juliette Lewis and The Licks. Well my sweet mama said "son, I'm bound to lose my mind". My mama said "son"
All songs	1	"Kerry Livgren. Every life, every light. It is a journey of the brave. Every single night. When the moon is full of moonlight. We will ride. A stormy"
All songs	2	"There's no turning back. Even once we know what it means. There's no turning back. Even once we know what it means. There's no turning back. Ah, don't"
All songs	2	"Mop, yo. I got my glass full of all that s**t. Don't ask the DJ, ask the DJ. Play my song, cause I'm your girl. I got my glass"
All songs	2	"Spoken (spoken) : This is an apology. For everybody who's been through this,. I apologize that I make you cry. It's my mistake but I'll be ok. I'll"
All songs	3	"Will.i.am joint. Hit Boy. Britney. Waka. Weezy. And they call me baby baby b, baby. Won't you come out tonight. And they call"
All songs	3	"Ronnie VanZant – Gary Rossington. There's an old man sitting out on a corner. And he don't understand what to say. He walks along the railroad track. And his"
All songs	3	"Spoken: (Reanimation) :46. (Warriors! Come together, chain all day, I know what to do...). The way that you shake it loose. Gotta"
All songs	4	"Everytime I see ya ni***z is selling pu**y, talkin 'bout hoes and cars snortin money, you ain't even s**t ni****"
All songs	4	"Mannie Fresh:. This is truly an eventful day. When she walks by I am thinking. And when I'm sad and blue. I don't cry (I am not talking)."
All songs	4	"The l means a new measure and the / means the bass note on the current chorded /. note of the / Prodigy's note on the current chorded /.. ll'm"
All songs	5	"Intro: DJ Khaled. We the Best Music!. Another one. DJ Khaled. I'ma have one. You're already know, it's DJ Khaled. I'ma"
All songs	5	"Who's that girl coming through the door?. I heard her name was Linda Lu. I heard her name was Linda Lu. I heard her name was Linda Lu. She'd help me to"
All songs	6	"Heaven ain't so easy to get out of the cup. How I wish I knew, he ain't so easy to get out of the cup. I try my best to make you see"
All songs	6	"[R. Kelly]. Uhh, yeah. (Uhh). Yeah. (Bow Wow). Uh, uh. (Bow Wow). You know, I know, you know. And all my"
All songs	6	"Chorus:. Don't say your prayers. Don't say you miss me. Cuz I'm missing you much too. I'm missing you too. Don't say your prayers. Don"
All songs	7	"feat. Theze Man, Sheryl Crow, Jay Lyriq and Fabolous.. [Intro: xzibit]. Yeah! Yeah, man! Yeah!. Big Rob"
All songs	7	"[1st Verse]. Girl, you've been touched before. Restless in de dreams. Restless in dis beautiful dreams. Diggin' graves instead of planting alibis. For you and"
All songs	7	"Love on a subway train, I got love on a moonlight boat. In the summer all of us in my mind, in the heart of the storm. In the heart of the night."
All songs	8	"Eins, zwei, drei vier. Steve: Yeah....eh, ah, ah. Rain on down in London town. Dreamin' of a place that's lovely and
All songs	8	"The l means a new measure and the / means the bass note on the current chord gauge the l means the bass note on the current chord. lEb lCb l. lE"
All songs	8	"Hook:. Don't say you love me. Don't say you care. Don't make me cry. Don't make me give up. Don't make me give up on you."
All songs	8:	"M. Jagger/K. Richards). What does a good girl know. She's not satisfied. With a good girl's touch. She don't know how to love a man. With"

Table 3: Representative lyrics generated by the model using all songs.

Rock	1	"Coverdale. I wanna tell you about the things we learned.. In the days when the sun went dry.. I just thought I'd give up.. Sometimes my heart is beating..Sometimes i think"
Rock	1	"Verse. (Hey yeah). The best is a price. (Hey yeah). I can't find the right words to describe. (Hey yeah). The worst is a time. It never"
Rock	1	"Joan Jett/Kenny Laguna/Marty Jones. I'd wait like that for you to come and talk to me. I could walk you through the window, I could steal"
Rock	2	"Jon Anderson/Chris Squire.. I was born and raised on the seven seas. Now I have got to find where my homeland is. How will I live when the great Thames down. ..."
Rock	2	"Written by Robert Johnson. He told them all that he'd be. It's a morning to brighten his life, and a song to write. He told them all that his eyes would be ..."
Rock	2	"Ronnie Lane, Ron Wood. I went down to San Antone. Asked Tom. Did you ever feel the same. But never one day it's gonna be that old. Cause when you ..."
Rock	3	"Words & music by J. Thomas. Great Britian. Some things may change and some rivers may turn. But I will still love you more. Yes, I will, yes I will, ..."
Rock	3	"words & music by Ivory Joe Hunter. Doo doo doo doo doo doo. Doo doo, doo. Doo doo, doo doo do ..."
Rock	3	"The sun is rising. The moon is on fire. So many people. They're not in need. We are the chosen. You will see us through the night. And when I have nothing ..."
Rock	4	"Verse 1:. I'm going to pack my bags, I'm going to pack my aunt and me. I'm gonna jump the jump the jump the jump the jump the jump the jump ..."
Rock	4	"(Reed). Yeah, this is gonna be the real life. 'Cause we're running at a speed limit. No, we're running at a speed limit. Well, I can't find ..."
Rock	4	"Words & Music by Rick Nielsen. Hey man, I'm gonna sit right down and write this song down. 'Cause it will all go on. I know I should not have called it that ..."
Rock	5	"Kerry Livgren. If I hadn't for one minute of my life. You know it would blow my life away. But then how'd it have been that we would make it through. ..."
Rock	5	"Hear the sirens, hear the sirens. Hear the sirens, hear the circus. For the very first time. There was magic in the air. In the air that night, ..."
Rock	5	"Music: Hexum; Lyrics: Hexum/Martinez. You can see it in my eyes. You can hear it in the street. You can taste it in my nails. You ..."
Rock	6	"M. Jagger/K. Richards). I been walkin' down the street, baby. Just wanna talk to you. Ain't nothin' gonna knock me down. Just a matter ..."
Rock	6	"words & music by Gene Autry. My buddies tell me I'm crazy. That I've been missing. Oh, I can see that. They believed me. Then they went away. All ..."
Rock	6	"J. Rotten/G. Rossington/J. Wayne. I know what itâ™s like to be a star. Itâ™s just a matter of time. ..."

Table 4: Representative lyrics generated by the model using Rock songs.

Eminem	1	"Intro:. (Eminem). Aye... (Eminem)... (Eminem). (Eminem). [Eminem]. [Eminem]. The whole world is gonna be here ..."
Eminem	1	"Chorus:. When you pull your hand. That's when they pull your hand. And when you pull your hand. That's when they pull your hand. That's when they pull your hand. That's when they pull your ..."
Eminem	1	"Hook, what! It's you! It's me! Take it back! You can smell it! I need a lighter man! Tell yo my name (f***in' b***h). I ..."
Eminem	2	"The Green Lantern and the Green Lantern, Green Lantern:! Yeah Green! Yeah Green! Green! Let me go back to that.. when I'm dead, huh...yeah. Guess who's" ..."
Eminem	2	"Chorus:. I'm going to jail, I'm going to jail, I'm going to jail. I'm going to jail, I'm going to jail, I'm going to jail ..."
Eminem	2	"Lickety Split. I'm getting to snare you guys,. The beef youve got in your cup lunch, that's beef you can make your own soup,. Pick a fight, eat"
Eminem	3	"Bizarre. A.L.O.V. Crazy Crazy Crazy. Yeah... What do you guys know about?. You know you got freaky things you don't know. You got ..."
Eminem	3	"R.O.V. [Obie Trice]. What?. This.... this?. [Verse 1]. Paper folding/unchucking Hey Em, we are rolling over ..."
Eminem	3	"Can-i-b***h. Can-i-b***h. Can-i-b***h-b***h. Can-ij-b***h-b***h-b***h. Can- ..."
Eminem	4	"Rihanna. Eminem. Jealous, b***h on the case, beating the s***t to death. (Eminem -4X). (Eminem). (Eminem). Yeah ..."
Eminem	4	"Eminem. Some people ask me. Where the hell ."""
Eminem	4	"were the wee in the bust. I don't know. But I don't know. Where'd the army go?. Hell yeah ..."
Eminem	5	"Lunch in the Thurston Howell, southwest *Bizarre. Got to have it.. uh, I'm gettin' here.. I'll have it right now. ..."
Eminem	5	"Lotto: Paul, s**t.. Hey, what's up man?. Ballin' up Paul, s**t.. Hey, what's goin' to get Paul?. I didn't know ..."
Eminem	5	"Lotto: Skylar Montgomery – (G-Unit). This is it all. Let's say it ain't so bad a thing. We the fans say that airplanes in the night sky are ..."
Eminem	6	"Chorus:. (4X). Eminem [echoes]. "Bust it rugged, shine like a gold nugget" (yeah man). [Verse 1]. It's worth ..."
Eminem	6	"Eminem - "Shady!". (scratched) (3X):. Hey Em, It's Cashis. Meet Dre, the guy's dead. I'm on top b***h ..."
Eminem	6	"Eminem. Better watch out sucker, sucker that sucker. We came here to set this party off. Right let's bounce tonight. And if they don't let us in through. So ..."

Table 5: Representative lyrics generated by the model using Eminem songs.

116 6 Appendix: Phase III Original MIDI Files

117 See attached MIDI files corresponding to the order in Table 2:

118 <https://tinyurl.com/creative2020midi>

119 test0.mid

120 test1.mid

121 test2.mid