# LYRICS BY GENRE: Classification and Generation using BERT and LSTM

NLP FINAL PROJECT PRESENTATION

Chloe Circenis, Gavin Hanville, Ima Mervin, Mia Ray, Mariana Vadas-Arendt

# PROBLEM STATEMENT

How can we automatically classify and generate song lyrics by genre using NLP techniques?

# **MOTIVATION**

Music is a culturally rich medium, and genre classification has applications in recommendation systems, music analysis, and creative AI. Generating genre-consistent lyrics is a challenging yet meaningful task in creative AI research

# **DATA COLLECTION**

API Authentication	Input Methods	Data Collection	Rate Limiting	Data Processing & Storage
	Direct Playlist ID	Spotify Track Data	200	
SpotifyClientCredentials	- playlist argument  Search Term	get_playlist_tracks(playlist_id)     extract metadata (name, artist, album)     optional : get_track_features(track_id)	ylist	pd.Dataframe
	- search argument	Genius Lyric Data	20 ecc	
lyricgenius.Genius(token)	Reccomendations	get_lyrics(artist_name, track_name)     genius.search_song(track_name,	time.sleep() every tracks for r	CSV Export
	-get_recommendations() method	artist)  3. return song.lyrics if found	time.sle	

# TASK FORMULATION

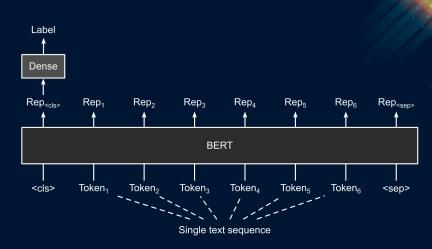
Using BERT for Sequence Classification

Both models are trained using song lyrics and the song's corresponding genre

Inputs are song lyrics

- Block Text as a string
- ❖ BERT Tokenizer (30,500 Tokens)

Outputs are the predicted song genre



# **MODEL SELECTION**

#### **BERT**

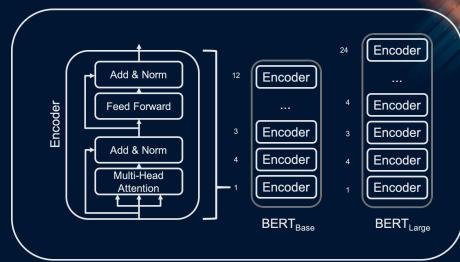
- Uncased
- Base
- For Sequence Classification

**Provides Bi-Directional Context** 

Baseline Model Performance:

- Accuracy
- Precision
- **❖** F-1

# **BERT Encoder**

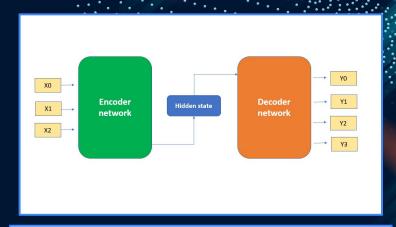


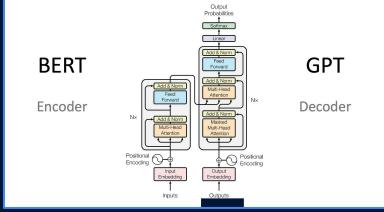
# **MODEL SELECTION**

Encoder-Decoder

# What and Why Encoder-Decoder?

- Originally popularized for machine translation
- Well suited for generating structured sequences – LYRICS!!





# **TRAINING DETAILS**

## Consistently Updating Library

- Currently ~10,000

### Hyperparameters:

- Batch size: 32
- Learning rate: 2e-5
- Epochs: 8

## Optimizer:

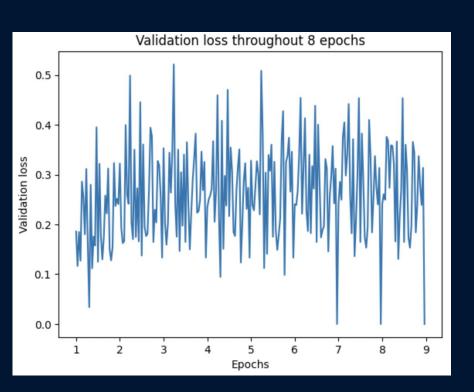
- AdamW with epsilon of 1e-8

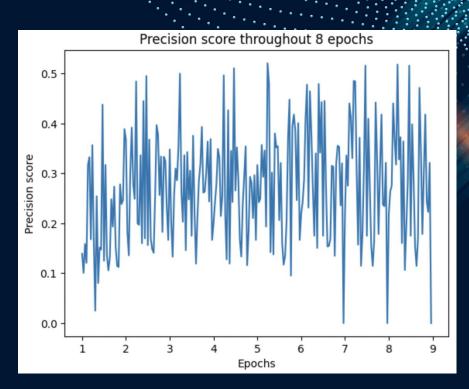
### Linear Scheduler

- Essential for BERT

Trained using Google Colab







# TASK FORMULATION

Generate lyrics based on genre, without sound data

# **MODEL SELECTION**

#### **LSTM**

- Prior promising results in "Deep learning in musical lyric generation: An Istm-based approach." (Gill et. al. 2020)
- Sequences, long term dependencies, light weight!

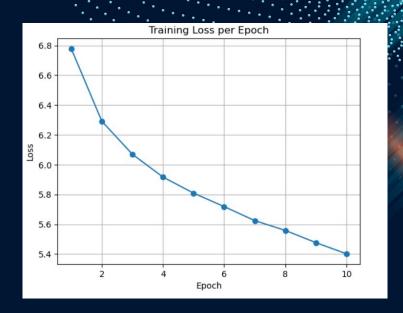
# **TRAINING DETAILS**

Smaller data set for initial generations

- 4 genres : rock, pop, hiphop, R&B
- Songs: 3468
- Breakdown: 22% rock, 46% R&B, 8% Hiphop, 24% pop

Hyperparameters:
Batch size: 16
Epochs: 10
1 LSTM layer
Teacher Forcing

Optimizer: Adam with epsilon of 1e-3 for fast convergence on small dataset



## **GENERATION**

```
Zero-init, data #1
hip hop jazz rock =
<UNK><UNK><UNK><UNK><UNK><UNK><
```

Genre-init, data #2 hip hop jazz rock = you 're me,,i,you,i,you,i,you,i,you,i,you,i,you,i,you,i,you,i,you,i,

# **GENERATION**

Genre-init Data #3 Top-P

HipHop = isnt that leke people sorry that youre talking to paid these I girls is mile girls and tryna the day just told me run on now it is the lick the so i i i bottegas and you now all me when i hold on and im fun at my best tryna leave no door do did youre four i and she they can hear it up when im gon the done youre nobody on weed bitch boaw can em make you snow me when i went to done drunk in if you was think no he but nobody else

Encoder-Decoder

**®** Baseline evaluation metrics

- Gill et. al. use an LSTM to generate genre-specific lyrics
- Lyrics are not uniform genres vary stylistically and structurally
- They propose 5 linguistic metrics to evaluate genre adherence

Deep Learning in Musical Lyric Generation: An LSTM-Based Approach

Harrison Gill<sup>1</sup>, Daniel (Taesoo) Lee<sup>1</sup>, Nick Marwell<sup>1</sup>

<sup>1</sup>Department of Linguistics, Yale University

Figure 2: Computed Linguistic Characteristics of Training
Data by Genre

Genre	Average Line Length	Song Word Variation	Genre Word Variation	I vs. You	Word Repetition
Metal	8.98	0.54	0.55	1.35	1.91
Rock	8.48	0.48	0.42	1.48	2.03
Rap	8.17	0.42	0.36	4.71	10.96
Pop	7.33	0.41	0.39	2.25	5.38
Country	8.18	0.48	0.62	1.43	1.52
Jazz	7.17	0.48	0.36	1.23	1.69

Figure 2: Computed Linguistic Characteristics of Training Data by Genre

## Encoder-Decoder

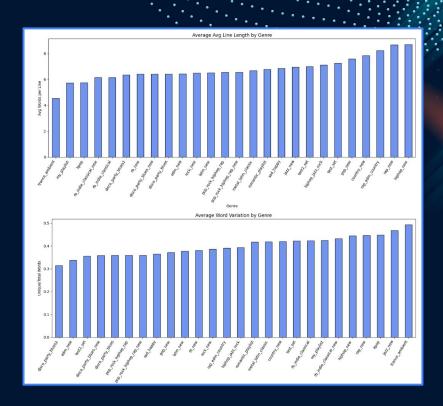
#### (\*Our scraped data\*)

country_new: 187 songs							
disco_party_blues: 625 songs							
disco_party_blues2: 78 songs							
disco_party_blues_new: 628 songs							
edm_new: 57 songs							
hiphop_jazz_rock: 65 songs							
hiphop_new: 105 songs							
jazz_new: 125 songs							
kpop: 187 songs							
latin_new: 13 songs							
metal_latin_classic: 595 songs							
my_playlist: 18 songs							
pop_new: 114 songs							
pop_rock_hiphop_rap: 190 songs							
pop_rock_hiphop_rap_new: 190 songs							
rap_edm_country: 84 songs							
rap_new: 104 songs							
rb_indie_classical: 542 songs							
rb_indie_classical_new: 840 songs							
rb_new: 229 songs							
rock_new: 578 songs							
romantic_playlist: 105 songs							
sad_happy: 74 songs							
test2_set: 328 songs							
test_set: 44 songs							
trance_ambient: 8 songs							

Organized data

Top 5	Songs in G	enre:	HIPHOP_JA	ZZ_N	OCK					
	ge	enre	avg_line_le	ngth	word_varia	tion	i_vs_y	rou	repetiti	ons
1575	hiphop_jazz_	rock	9.84	8837	0.203	070		-1		78
1576	hiphop_jazz_	rock	7.11	3208	0.251	989		15		C
1577	hiphop_jazz_	rock	7.08	1081	0.396	947		17		9
1578	hiphop_jazz_	rock	5.60	6061	0.443	243		15		3
1579	hiphop_jazz_	rock	8.65	0000	0.497	110		3		1
Top 5 Songs in Genre: HIPHOP NEW										
	genre	avg	line_length	wo	rd_variation	i_vs	you	rep	etitions	
1640	hiphop_new		11.803922		0.433555		9		16	
1641	hiphop_new		10.952381		0.363768		-9		7	
1642	hiphop_new		8.329412		0.539548		0		10	
1643	hiphop_new		7.693878		0.267905		10		- 11	
1644	hiphop new		7.222222		0.364103				24	
							2			
		lenre	: HIPHOP_I	NEW						on
Top 5	Songs in G	enre avç	: HIPHOP_I	NEW th w		on				
Top 5	Songs in G	enre avç	: HIPHOP_I	NEW th w	vord_variati	on 55		ou		1
Top 5	Songs in G genre hiphop_new	lenre avç	: HIPHOP_I g_line_lengt 11.80392	NEW th w	ord_variati	on 55		ou 9		10
Top 5	Songs in G genre hiphop_new	ienre avç	: HIPHOP_! g_line_lengt 11.80392 10.95238	NEW th w 22 31	ord_variati 0.4335 0.3637	on 55 68 48	i_vs_y	ou 9 -9		10
Top 5 1640 1641 1642	Songs in d genre hiphop_new hiphop_new hiphop_new hiphop_new	enre avç	: HIPHOP_I _line_lengt 11.80392 10.95238 8.32941	NEW th v 22 31 12	ord_variati 0.4335 0.3637 0.5395	on 55 68 48	i_vs_y	9 -9		10
Top 5 1640 1641 1642 1643	Songs in d genre hiphop_new hiphop_new hiphop_new hiphop_new	enre avç	: HIPHOP_1 J. line_lengt 11.80392 10.95238 8.32941 7.69387 7.22222	NEW 222 31 12 78	ord_variati 0.4335 0.3637 0.5395 0.2679	on 555 68 48 05	i_vs_y	9 -9 0		10
Top 5 1640 1641 1642 1643	Songs in d genre hiphop_new hiphop_new hiphop_new hiphop_new	avg	: HIPHOP_I I_line_lengt 11.80392 10.95238 8.32941 7.69387 7.22222	NEW tth w 222 331 122 78	0.4335 0.3637 0.5395 0.2679	on 555 68 48 05	i_vs_y	9 -9 0		10
Top 5 1640 1641 1642 1643	Songs in G genre hiphop_new hiphop_new hiphop_new hiphop_new hiphop_new	denre avg	: HIPHOP_! J. line_lengt 11.80392 10.95238 8.32941 7.69387 7.22222	NEW tth w 222 231 12 222 2	0.4335 0.3637 0.5395 0.2679	on   555   668   448   005   003	i_vs_y	9 -9 0 10 2	repetiti	10
Top 5 1640 1641 1642 1643 1644 Top 5	Songs in G genre hiphop_new hiphop_new hiphop_new hiphop_new hiphop_new	denre avg	: HIPHOP_! J. line_lengt 11.80392 10.95238 8.32941 7.69387 7.22222	NEW tth w 222 231 12 222 2	0.4335 0.3637 0.5395 0.2679 0.3641	on   555   668   448   005   003	i_vs_y	9 -9 0 10 2	repetiti	10 11 24
Top 5 1640 1641 1642 1643 1644 Top 5	Songs in G genre hiphop_new hiphop_new hiphop_new hiphop_new hiphop_new Songs in G genre a	denre avg	: HIPHOP_! J. line_lengt 11.80392 10.95238 8.32941 7.69387 7.22222 : JAZZ_NEL	NEW tth w 222 231 12 222 2	0.4335 0.3637 0.5395 0.2679 0.3641	on   555   668   448   005   003	i_vs_you	9 -9 0 10 2	repetiti	11 11 2-
Top 5 1640 1641 1642 1643 1644 Top 5	Songs in @ genre hiphop_new hiphop_new hiphop_new hiphop_new hiphop_new Songs in @ genre a jazz_new	denre avg	: HIPHOP_I J. line Jengt 11.80392 10.95238 8.32941 7.69387 7.22222 : JAZZ_NEL se_Jength 3.862069	NEW tth w 222 231 12 222 2	0.4335 0.3637 0.5395 0.2679 0.3641 d_variation 0.437500	on   555   668   448   005   003	i_vs_you	9 -9 0 10 2	repetition	10 11 11 20
Top 5 1640 1641 1642 1643 1644 Top 5 1745 1746	Songs in 6 genre hiphop_new hiphop_new hiphop_new hiphop_new Songs in 6 genre a jazz_new jazz_new	denre avg	: HIPHOP_I J.line_lengt 11.80392 10.95238 8.32941 7.69387 7.22222 : JAZZ_NEW 12.862069 5.823529	NEW tth w 222 231 12 222 2	0.4335 0.3637 0.5395 0.2679 0.3641 4_variation 0.437500 0.646465	on   555   668   448   005   003	i_vs_you 7	9 -9 0 10 2	repetitions	1 1 2

Pandas Dataframes



Pyplot!

Encoder-Decoder

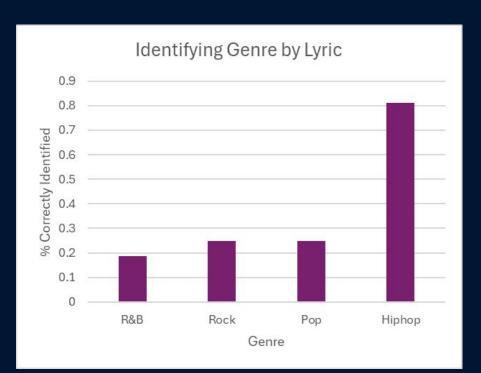
- © Comparison of real vs. generated lyrics
- Gill et al.'s cosine similarity test
- Quantitatively measure style similarity

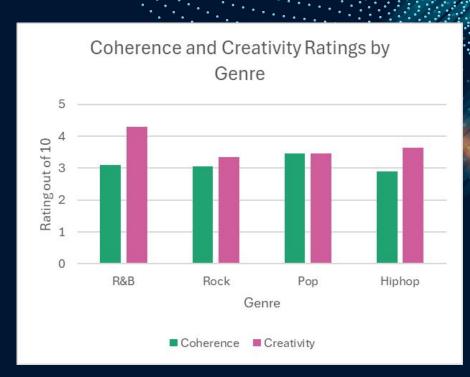
# Figure 5: Percentage Changes from original to generated lyrics for each metric and genre

Genre	Average Line Length	Song Word Variation	Genre Word Variation	I vs. You	Word Repetition
Metal	-54.3	-25.1	-13.5	82.7	88.2
Rock	-38.4	-5.0	-5.9	77.0	34.9
Rap	-52.3	-64.6	-99.0	75.5	-49.6
Pop	-28.0	-16.7	-44.6	94.6	36.5
Country	-74.1	-36.6	-23.0	-229.6	75.2
Jazz	-24.1	-38.7	-78.5	73.7	94.6

$$\cos( heta) = rac{A \cdot B}{\|A\| \|B\|}$$

# **Human Evaluation of LSTM Lyrics**





# **NEXT STEPS**

## **LSTM**

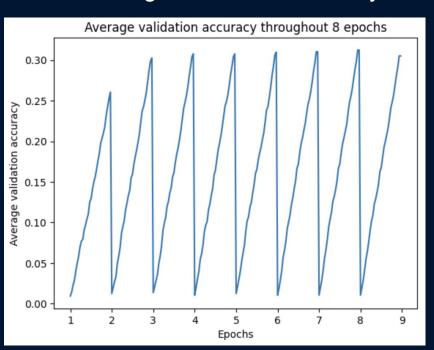
- ❖ MORE DATA
- More than 1 LSTM layer, introduce dropout
- Introduce temperature to top-p sampling
- More genres (limited)
- learned embedding layer from the fine-tuned BERT model as init hidden state?

## **BERT**

- Simplify Genres
- Language Correction
- Introduce Warmup
- Identify Task-Inappropriate Songs
- More Computation Power
  - Larger Batch Sizes
  - > CURC

# **BERT RESULTS...so far**

# Average validation accuracy



# Validation loss

