

Rap Lyrics generation using LSTM

COL 772 : NLP project

Deepak - 2017MCS2071
Harish - 2017MCS2074

Problem Statement

- To create a model for the task of ghost-writing rap lyrics.
- Ghost writing: Generating rap lyrics for an artist that are -
 - similar in style to the artist.
 - but different than the existing lyrics.



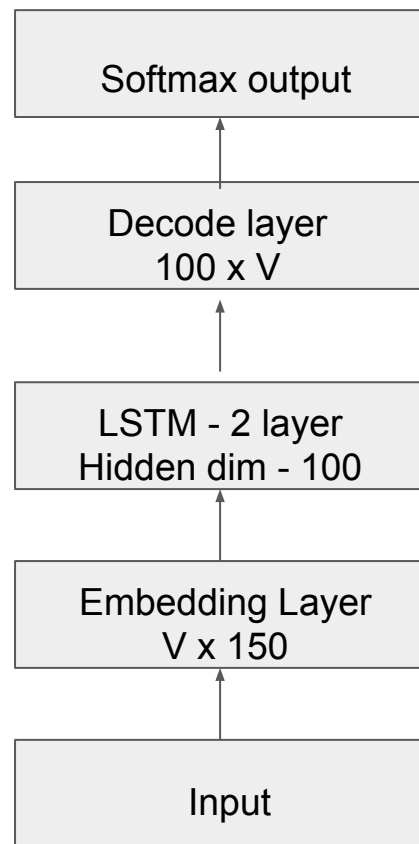
Original Data set

- To compare with the baseline model, we will use lyrics of ‘Fabolous’ artist.
- The data set contains lyrics of only **191!** songs.
- Downloaded from *ohhla.com*.



Baseline model

- The baseline model is from EMNLP paper (2015): *Ghostwriter* - by Peter Potash, Alexey Romanov, Anna Rumshisky.
- The model uses word level LSTM for language modeling.
- We changed the embedding dimensions from $V \times V$ to $V \times 150$, $V \sim 6K$.



Problems and our solution

- Problems

- a. Data set too small!
- b. Naive implementation of model (For e.g. size of embedding = vocab_size = 6k)
- c. Too many parameters to learn compared to data set, model close to overfitting.
 - As a result, the generated verse contained lines directly taken from the training corpus.

- Solution

- a. Smaller number of parameters using character embeddings.
- b. Using transfer learning by pre-training the model on a relatively large lyrics corpus to learn better embeddings.

Related work

- DopeLearning: A Computational Approach to Rap Lyrics Generation (by Malmi et al. (ACM,2015).).
 - Generate fixed 16 line verses.
 - verse generated line-by-line using full lines from existing rap songs.
 - RankSVM algorithm is used to the most relevant next line .
 - Human engineered features as well as semantic features generated by a Deep Neural Network are used.
 - Demo can be found at <http://deepbeat.org/>

Character Level Model

Input

Embedding
Layer 100×128

LSTM
 128×128

Decode Layer
 128×100

Output

Gold Output

n	f	i	n	i	t	y		W	a	r
↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑
I	n	f	i	n	i	t	y		W	a

Input

- Dictionary of all english characters (lower and upper) + some special symbols. (100 different characters).
- Given current letter and all previous letters, try to predict the next character.
- During training for each sequence use all its characters except the last one as an input and the same sequence starting from the second character as ground truth.

Pretraining Model on lyrics corpus

Kaggle Lyrics Dataset

- 57650 songs
- 640 different artists.

Train model

Train and save the character level model on the kaggle dataset to learn the character embeddings.

Train on original dataset

Load the model trained on the kaggle dataset.

Retrain the model parameters except the character embeddings on the small original dataset.

Softmax over the output gives the multinomial distribution over all possible characters. The next character is generated by taking a sample from this distribution.

Evaluation method

- The two metrics used for evaluating the generated lyrics -
 - Rhyme density -
 - Rhyme feature : Defined as the total number of rhymed syllables divided by the total number of syllables.
 - **Good model should produce lyrics that have rhyme density similar to that of the artist's.**
 - We used an external tool by Hirjee and Brown, 2010.
 - Similarity -
 - **Good model should produce lyrics that have minimum similarity to existing lyrics of artist.**
 - TF-IDF matrix is created for the data set and similarity is measured using cosine similarity.
- So aim is to
 - Generate lyrics having rhyme density close to rapper's original rhyme density.
 - Minimize the maximum similarity of the generated lyrics to that of the data set.

Continued...

- To evaluate our models -
 - We generate about 10 sentences from the model every 10 epochs.
 - Calculate the average rhyme density.
 - Save the model where the average rhyme density turns out to be closest to that of the artist.
 - Calculate the maximum similarity of the lyrics generated by this model to that of the lyrics in the dataset

Results

Method	Maximum similarity at rhyme density 0.34 (artists rhyme density)
Baseline LSTM model(original)	0.59
Baseline LSTM model (modified)	0.515
Char level LSTM	0.127
Pretrained character level LSTM	0.140

Note: The similarity score obtained by the char level as compared to the pretrained one is low because the former had relatively higher number of spelling errors which lead to low similarity score.

Qualitative analysis

Sample rap verse generated from char-LSTM model

(verse start)

*yo all have to guns is for the glins
i cinaymna i smell myself and one off your rollino
with a re0 scrape sundy cross the yeah
they be flam baggy just real nigga forthlats
wit your pumpin battle to robe that's it muttalt
my came
i have turn on the baby get hit the few mac
so how my gon nuggon atteations
see aquain i just disaming in the minus
at it like they get man sauny ya*

(verse end)

Sample rap verse generated from pre trained char-LSTM model

(verse start)

*all they in the hand and do it like all dub with by this pill
and friends in the problems are on the rason just like me
i'm there for that cash hadded i jusned a right
i ain't come out bikin as i won't see on a slack by a pring
dough
thatk these feeling in the aught so i don t care for the yan
i'm standin a little bit on a chrome when i go and i ain't
wonderfly*

(verse end)

Demo

Let's Generate some raps!!