Lyric Generation using Recurrent Neural Network

SMAI Course Project

Introduction

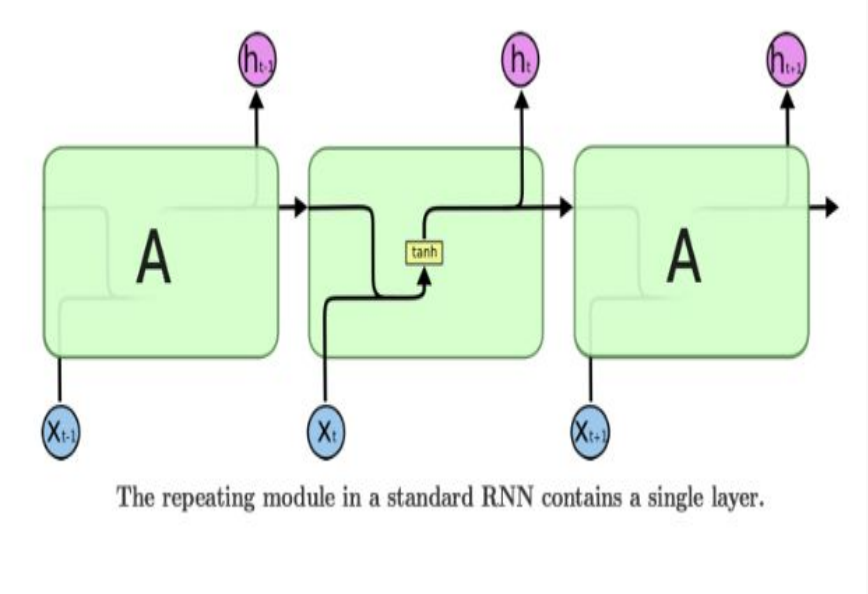
- This project demonstrates the effectiveness of a LSTM in our initial efforts to generate lyrics.
- The goal of this model is to generate lyrics that are to that of a given song.
- We’ve studied various architecture while doing the same.

MODELS

“RNN is able to process sequences of inputs by utilizing the internal state. Hence, it is regarded as a very promising candidate to solve NLP tasks.”

Why RNN?

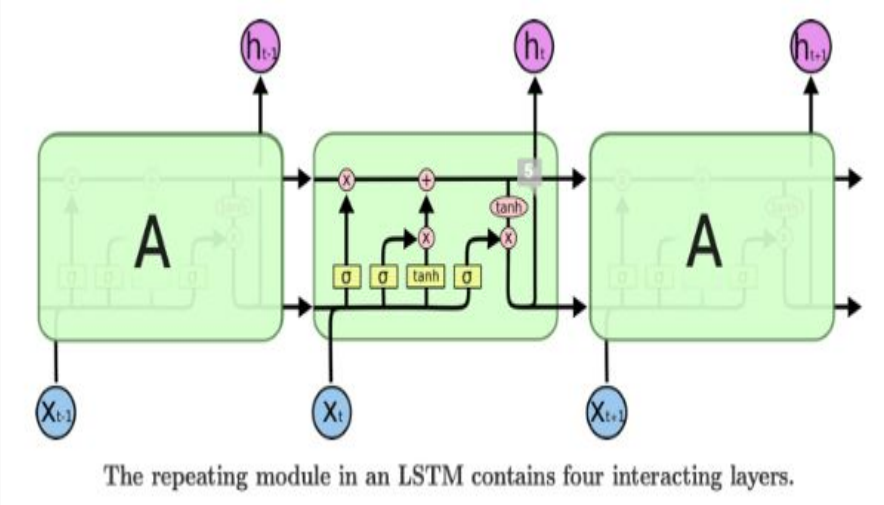
- Because lyrics generation can be modeled as a time series data.
- RNNs are able to process sequences of inputs (such as words and sentences) by utilizing the internal state (memory state).
- Lyrics generation can be thought as sequence of characters/words given and predicting next character/word.



The repeating module in a standard RNN contains a single layer.

Why LSTM?

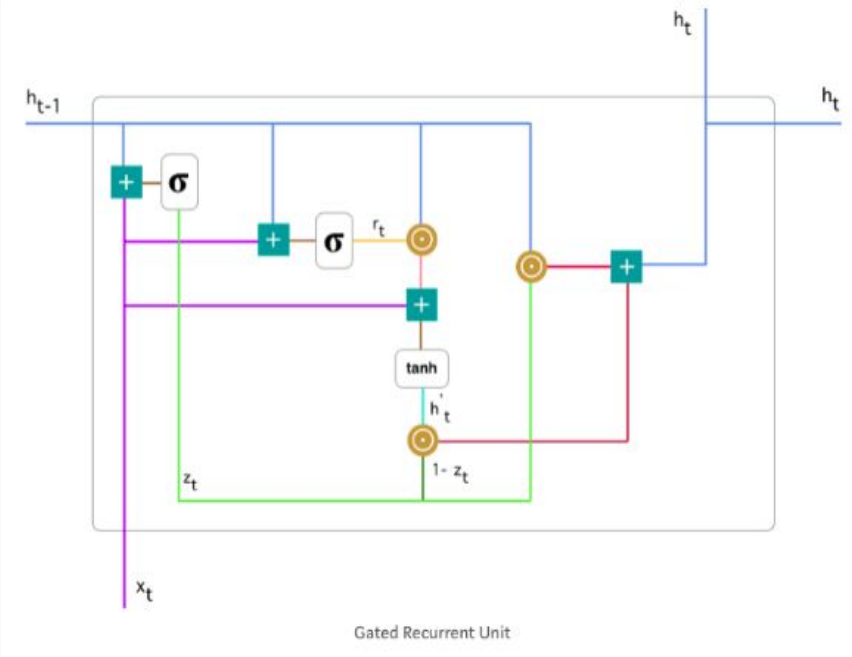
- LSTM are a special kind of RNN, capable of learning long-term dependencies. They are explicitly designed remember information for long periods of time .
- LSTMs chain-like structure, consisting of four layers, interacting in a special way.
- While experimenting, we've also tried Bidirectional LSTM.



The repeating module in an LSTM contains four interacting layers.

Why GRU?

- GRU, a LSTM variant, aims to solve the vanishing gradient problem by using update gate and reset gate.
- These are two vectors which decide what information should be passed to the output.
- They can be trained to keep information from long ago, without washing it through time or remove information which is irrelevant to the prediction.



Gated Recurrent Unit

APPROACHES

Character based approach

- The idea here is to train the model with many sequences of characters and the target next_character.
- The idea is that after many epochs the model will learn “the style” of how the corpus is written.
- Trying to adjust weights of the network such that the model will be able to predict the next character given a sequence of Maxlen previous characters.

```
In [18]: sentences[0]
Out[18]: ['h', 'e', 'y', ' ', 't', 'h', 'e', 'r', 'e', ' ']

In [19]: next_char[0]
Out[19]: 'd'

In [20]: sentences[1]
Out[20]: ['e', 'y', ' ', 't', 'h', 'e', 'r', 'e', ' ', 'd']

In [21]: next_char[1]
Out[21]: 'e'
```

Word based approach

- The idea here is to train the model with many sequences of words and the target next_word.
- We don’t actually send the strings, but a vectorized representation of the word inside a dictionary of possible words
- After many epochs the model will learn “the style” of how the corpus is written, trying to adjust the weights of the network to predict the next word given a sequence of the N previous words.

```
In [4]: sentences[0]
Out[4]: ['hey', 'there', 'delilah', 'whats', 'it', 'like', 'in', 'new', 'york', 'city']

In [5]: next_words[0]
Out[5]: 'I'

In [6]: sentences[1]
Out[6]: ['there', 'delilah', 'whats', 'it', 'like', 'in', 'new', 'york', 'city', 'I']

In [7]: next_words[1]
Out[7]: 'am'
```

RESULTS

Char based result (English Dataset)

Input (seed)	Generated lyrics
“ work this treacherous road for all the “	work this treacherous road for all the side is that some could i don't know on was and but of your down and when i thought the light of i a
“the chance on you chorus the first noel “	the chance on you chorus the first noel the sayin' for a little boy and i can't follow baby i was some to my search that i say i was long an
“ the mother mother touched, and dude ain”	the mother mother touched, and dude ain't down to hear a look is so bad with you so dead to help and we know that i ride my read or on our

Input (seed)	Generated lyrics
"Sand walking with someone"	sand walking with someone and holding his hand i should walk with me it seems i stood and forget all sometimes lovers don't do all they could sometimes lovers hide in the cover of innocence and blame i know that we act just the same your heart's
"stop the tears how"	stop the tears how i hate to spend our dust in the stars eyes for the end of the road so, look out for me, darling on the power of love a man don't crawl that we need to walk around here tonight gotta go down

Word based result English Dataset

Char based result (Hindi Translated Dataset)

Input (seed)	Generated lyrics
"hout you chorus i wont live i will die w"	hout you chorus i wont live i will die wanc in as in and beauth a my heart in beant like me the what no le the san and in sore and in and in
"eloved didnt come beloved didnt come tha"	eloved didnt come beloved didnt come that beloved didnt come beaut the beade god on shall and shall god about you are and seare you are the

Word based result
Hindi Translated Dataset

Input (seed)	Generated lyrics
"oh eyes deceivers "	oh eyes deceivers your eyes are big deceivers you are the days of the night you are in my eyes ill take care of you in my breaths ill keep you in my eyes ill take care of you in my breaths ill keep
" shake that "	 shake that youll he became full of joy since you are the bride of badri baby see the girls you burns of live without you dont talk about the world now the world of stars now since i have seen thy face i have

OBSERVATIONS

Both Models are working good. However , both have their pros and cons.

- Character Based Model
  - Can be trained with less data.
  - Having semantic level understanding is difficult.
  - Sometimes, it is having tough times to even predict right spellings.

- Word Based Model
  - Require large dataset to train for having enough data for each words.
  - Learns semantic easily compared to Character based models.
  - Can’t generate results for never seen words.

Thanks!

Praveen Balireddy
Arpit Gupta
Aman Joshi
Shubham Pokhriyal