

Text Generation and Neural Style Transfer

S. Singhal, K. Siddarth, P. Agarwal, A. Garg
Mentor: N. Asnani

Department of Computer Science and Engineering
IIT Kanpur

22nd November 2017

Introduction

- Text generation is a foundational task in Natural Language Processing
- The aim is to produce a natural language text in order to meet specified communicative goals.
- Takes non-linguistic representation of information as input and outputs text, documents, reports, etc.
- Has a diverse set of applications ranging from image captioning to text summarization.

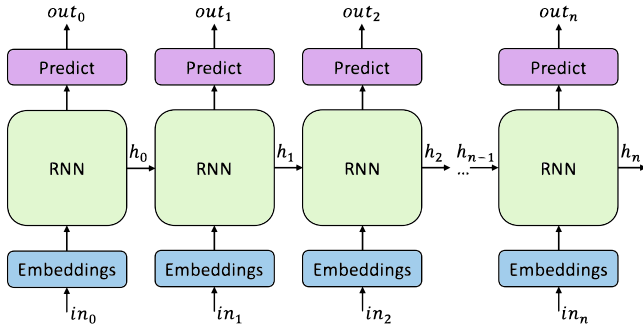
Goals

- Attempt to generate coherent text in the style of an author
- Experiment with different models to see which works best
- Design a model that takes text in the style of one author and convert it to that of another author

Previous work

- Our work is inspired by Andrej Karpathy's use of character level RNN's to generate text
- At every time-step it feeds in a character, and the RNN predicts the next character.

Previous work



- w_i - input tokens of source article
- h_i - Encoder hidden states
- $P_{vocab} = \text{softmax}(Vh_i + b)$ is the distribution over vocabulary from which we sample out_i

Previous work

- Our work is inspired by Andrej Karpathy's use of character level RNN's to generate text
- At every time-step we feed in a character, and the RNN predicts the next character.
- One very basic problem with this model is that character RNN's can conjure up words on their own.
- A very easy fix is to use word level models instead of character level models.

Character vs Word

- Both have size 512 and 3 stacked layers
- Character level

KINGequeses, wifely A mighty **vanagy** died, and is it **sotis** being
note but by flatter, which, I rather be! Hear over-blown swifled by;
The king was timely followed.

- Word level

King VI: First Citizen: And will will tell you, I have not I is to be
content; it are not that is a more than all the writing. DUKE OF
YORK: My lord, I am a bond, and we is the writing. DUKE OF
YORK: What is the writing.

2 vs 3 layers

- While testing, we found that having more layers with a vanilla RNN leads to nonsensical outputs

2 layers

KING RICHARD III: Ay, if you know the general is not so far with me. QUEEN ELIZABETH: My lord, I will not not a man of such good Than not to see him in the Duke of York. KING RICHARD III: Ay, but you will not be a traitor to the people, And yet thou art a soldier, and that is not so much with me for his eye

3 layers

[illegible]

RNN vs LSTM

- Both have size of 1024 and 3 stacked layers

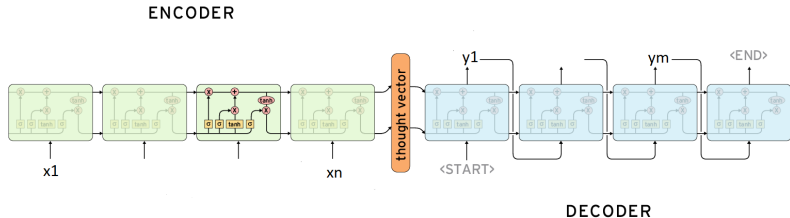
RNN

[illegible]

LSTM

King VI: First Citizen: And will will tell you, I have not I is to be content; it are not that is a more than all the writing. DUKE OF YORK: My lord, I am a bond, and we is the writing. DUKE OF YORK: What is the writing. DUKE OF YORK: What is the writing.

Sequence to Sequence models

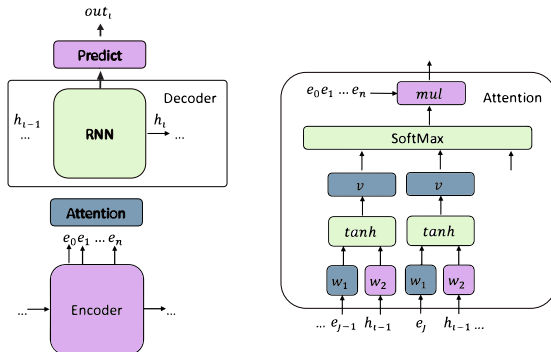


- 1
 - It consists of an Encoder(Bidirectional LSTM) and a Decoder LSTM network.
 - The final hidden state from the Encoder(thought vector) is passed into the Decoder.

¹Image from colah.github.io

Attention

- $importance_{i,t} = V * \tanh(e_i W_1 + h_t W_2 + b_{attn})$.
- Attention Distribution $a^t = \text{softmax}(importance_{i,t})$
- Context vector $h_t^* = \sum_i e_i * a_i^t$



Our Novel model

- Inspired by the work of in vision.
- They manage to separate the style of the image and the content of the image by passing the image through a CNN, and then reconstructing the image from the representation.
- This works in a very similar way to an autoencoder model.

Style Transfer

- Here we aim to take a corpus of text from one author and generate text with the same meaning in the style of another author.
- There has not been much work on transfer of style from one author to another.
- In the paper by Gatys et. al. [GEB15] the authors find that content and style in a Convolutional Neural Net(CNN) are separable, and therefore can be manipulated separately.

Our Novel model

We propose a very simple seq2seq model for style transfer.

Step 1

- We make a seq2seq encoder-decoder work as an auto-encoder first. That is given an input sentence, we train it output the same sentence.
- We train this for **Author 1**. We did this for Agatha Christie and Shakespeare
- As these models can't handle multiple sentences well, we only train these on single sentence to single sentence

Step 2

- Once the seq2seq auto-encoder is trained, we input the sentence of **Author 2**, in our case Sir Arthur Conan Doyle.

Why should it work?

- We think that while training on the first author, the network would first learn a good encoding of that sentence. And then using that encoding it needs to learn regenerate the sentence.
- So it makes sense for the model to encode only the content part of the sentence in the encoding because style is same for the author and that can be learned by the decoder.
- We use different weights for encoder and decoder.
- So when, we feed in the sentence of second author it's content gets encoded by the encoder.
- Then the decoder styles that content in the style of the first author

Parameters

- LSTM
- Size = 1024
- Depth = 2
- Embedding size = 500
- beam width = 5
- max decode step = 300

How good is our Auto-Encoder

- We use the BLEU metric to test how well our model does self encoding
- We got a BLEU score of 55.13, meaning it does the autoencoding pretty well

Results

Sherlock Holmes (Original)

Was there a secret marriage ? Absolutely none . None . No sign of it ? Come in ! & said Holmes . Seven ! & I answered . She will not sell . And I. My own seal . We have tried and failed . Stolen , then . I was mad - insane . To ruin me . We were both in the photograph .

Generated

Absolutely . None . ; No sign of it ? Come in ! & said . Lord ! ; I answered . She will not see . And My mother . We have come and rushed . Welcome , then . I was mad - . To me me . We both were in the photograph .

Original

How many ? I don't know . Holmes laughed . It is quite a pretty little problem , said he . My photograph . Stolen . What do you make of that ? asked Holmes . I am about to be married . I think that I had better go , Holmes . My private note-paper . No legal papers or certificates ? I promise , said Holmes . I carefully examined the writing , and the paper upon which it was written .

Generated

What do you make of that ? asked asked . I am going to be married . I think that I had really go , I had My private private . No girl or or two ? I dare , said gruffly . I carefully the man , and the paper paper which it was written .

Ongoing Work

- We plan to train two auto-encoder (A1-D1) and (A2-D2) for author 1 and 2 respectively.
- Then combine the A1 and D2 to get a style transfer model to convert text from author 1 to author 2.

References I



A. Karpathy

The Unreasonable Effectiveness of Recurrent Neural Networks.
Andrej Karpathy blog 2015.



Bahdanau, Dzmitry and Cho, Kyunghyun and Bengio, Yoshua
Neural machine translation by jointly learning to align and
translate.
arXiv preprint arXiv:1409.0473, 2014.



Leon A. Gatys, Alexander S. Ecker, Matthias Bethge
A Neural Algorithm of Artistic Style
arXiv preprint arXiv:1508.06576, 2015

References II



Zhiting Hu, Zichao Yang, Xiaodan Liang, Ruslan Salakhutdinov, Eric P. Xing

Toward Controlled Generation of Text

arXiv preprint arXiv:1703.00955, 2017