Deep*Script*: Handwriting Generation Network

Chris Chute*, Max Lam**, Justin Myles***, Mansheej Paul****

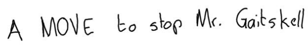
* chute@stanford.edu
** maxlam@stanford.edu
*** jmyles@stanford.edu
**** mansheej@stanford.edu

Introduction

We generate realistic handwriting by combining ***stacked LSTMs****, an* ***Attention mechanism, Mixture Density Networks (MDNs)***, and ***Professor Forcing***—a recent technique for training RNNs as sequential generative models.

Dataset

We used the ***IAM online handwriting database*** to train our model [3]. It contains sequences of pen tip positions and end of stroke tokens annotated with ASCII characters. Total of 86,272 words in 13,049 lines from 221 writers. Sample line below.



Professor Forcing [2]

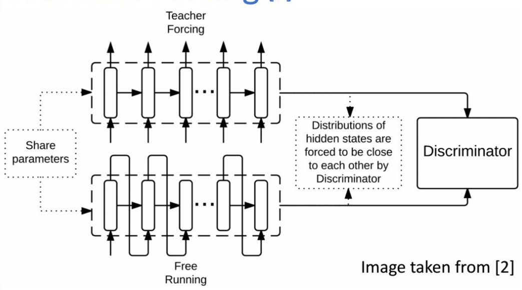


Image taken from [2]

Professor Forcing applies a GAN like framework to training RNNs. The similarity in the dynamics of the two sampling modes leads to longer and more robust sequences.

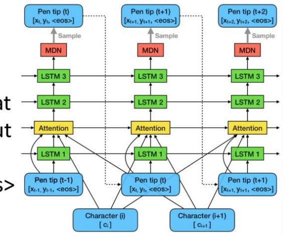
Handwriting Generator [1]

Network models p(**x**|**c**).

c : ASCII character seq. Input to network.

p : mixture of Gaussians at every time step. Output parameters of p.

x : pen tip position + <eos> token seq. Sampled.



Loss: [2] train to predict **x** and fool discriminator.

$$NLL(\theta_{gen}) = E_{(\mathbf{c},\mathbf{x})\sim data}\left[-\log p(\mathbf{x}|\mathbf{c})\right]$$

$$C_{free}(\theta_{gen}|\theta_{disc}) = E_{\mathbf{c}\sim data, \mathbf{x}\sim P_{\theta_{gen}}(\mathbf{x}|\mathbf{c})}\left[-\log D(g(\mathbf{c},\mathbf{x},\theta_{gen}),\theta_{disc})\right]$$

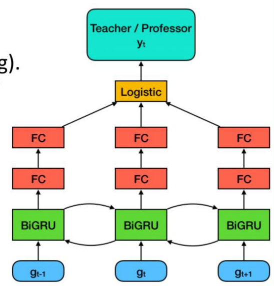
$$C_{gen} = NLL + C_{free}$$

Discriminator

Network predicts if generator is sampling in teacher forcing mode – D(g).

g : seq. of hidden states of generator. Here, parameters of p. Input to network.

FC weights shared across time.



Loss: [2]

$$C_{disc}(\theta_{disc}|\theta_{gen}) = E_{(\mathbf{c},\mathbf{x})\sim data}\left[-\log D(g(\mathbf{c},\mathbf{x},\theta_{gen}),\theta_{disc})\right] + E_{\mathbf{x}\sim P_{\theta_{gen}}(\mathbf{x}|\mathbf{c})}\left[-\log(1 - D(g(\mathbf{c},\mathbf{x},\theta_{gen}),\theta_{disc})\right]$$

Results and Discussion

Goal: Generate longer, more legible sequences.

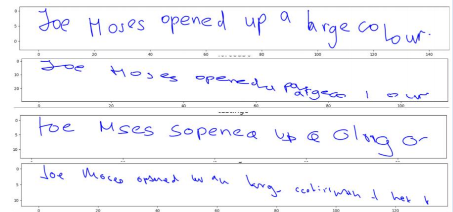
c = "Joe Moses opened up a large coloured"

Teacher t = 0

Teacher t = 50

Free t = 0

Free t = 50



Able to generate longer sequences.

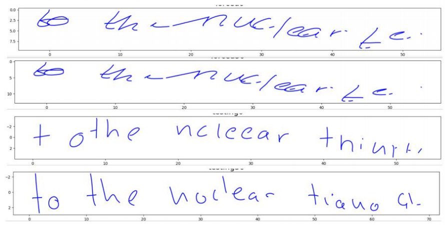
c = "to the nuclear tests"

Teacher t = 0

Teacher t = 30

Free t = 0

Free t = 30



Prof. Forcing loses style. Future work: regularize generators loss: $C_{gen} = NLL + \lambda C_{free}$

References

[1] A. Graves, "Generating Sequences With Recurrent Neural Networks," *ArXiv e-prints*, June 2014.

[2] A. Lamb, A. Goyal, Y. Zhang, S. Zhang, A. Courville, and Y. Bengio, "Professor Forcing: A New Algorithm for Training Recurrent Networks," *ArXiv e-prints*, Oct. 2016. NIPS 2016 Accepted Paper.

[3] "IAM On-Line Handwriting Database." *FKI: Research Group on Computer Vision and Artificial Intelligence, INF, University of Bern.* http://www.fki.inf.unibe.ch/databases/iam-on-line-handwriting-database.

[4] S. Greydanus, "Scribe: realistic handwriting with TensorFlow," *Github*, 21-Aug-2016. https://greydanus.github.io/2016/08/21/handwriting/.