Ikhlas Attarwala
05-23-19 / Week #8
Text Generation Using NNLMs / Paper 2

Sequence to Sequence Learning with Neural Networks
by Sutskever, Ilya et al.

Google's Brain team proposed an end-to-end approach for sequence learning. This method uses two deep LSTMs: one for (encoding) mapping input sequences to vectors, and the other for (decoding) mapping the vectors to output sequences. In comparison to traditional LSTMs, as this model uses two, there are more model parameters with very little additional computing cost. This model also uses a 4-layer deep LSTM. This model also does something a little strange.. it reverses the order of the input in order to enhance short/long-term predictions of the LSTM. It's important to note that, Tensorflow and PyTorch weren't around or promising at this time. The task Sutskever et al propose is to translate a sequence of sentences from English to French.

Beam Search seemed to help in BLEU scores (results), but it was unclear just how computationally expensive this change was from a beam width of 1 (greedy) to a beam width of 12. The best suggestion was to ensemble 5 randomized, reversed LSTMs. I can't read French, but it appears that longer sentences showed fairly good results. This paper highlighted the benefits of LSTMs in Sequence to Sequence Learning.