Cook your own Language Model

- Use LSTMs or GRUs, and gradient clipping
  https://colah.github.io/posts/2015-08-Understanding-LSTMs/

- Start with one layer, then stack 3-4, use skip connections

- Use dropout for regularization:

    Zaremba, Sutskever, Vinyals. Recurrent Neural Network Regularization, 2014.

- Have a look into TF tutorial for a working model:
  https://www.tensorflow.org/tutorials/recurrent

- Tune learning rate schedule in SGD or use Adam

- Explore state-of-the-art improvements:

  - **July 2017:** On the State of the Art of Evaluation in Neural Language Models.

  - **August 2017:** Regularizing and Optimizing LSTM Language Models.