Something Something Creative Writing

Ben Greene
Galen Berger-Fletcher
Tenzin Dophen
Malcom Grossman

October 9, 2018

1 Vision

2 Division of work

We plan on working together as a full group on the most critical parts of the implementation of our RNN. This neural language model will be the conceptual core of our project, so it is important that every group member has the opportunity to develop a high-level understanding of how it works. However, Malcolm and Tenzin will take the lead on figuring out the lower-level details of the RNN and optimizing its performance. Ben will focus on finding and processing our training data, while Galen will begin developing the front-facing elements of our project like our UI and web framework.

3 Individual tasks

Although we will be working as a group on much of the project, we have each set individual goals for the remainder of Fall Term, in alignment with our general directions as discussed above. These goals are as follows:

- Malcolm: have a RNN that compiles and runs to completion in Tensorflow
- Tenzin: find possible optimization strategies to improve the efficiency of the neural network
- Galen: design a basic web interface for the application using Balsamiq, set up a simple webapp with Django for the RNN to slot into
- Ben: Develop data sources via a webscraper and efficiently manage RNN training.

4 Resources

For our initial training and testing of our neural language model on fairly small subsets of our data, we plan to request a dedicated lab machine in CMC 307. This machine will eventually also allow us to serve our webapp until we agree upon a longer-term solution.

As the computational requirements of the training process become clearer to us over the second half of fall term, we will likely need more processing power than a single iMac can offer. There are two possible ways for us to get this increased power. First, we can work with Mike Tie to gain access to the new research cluster (summer18.dmz.carleton.edu). This option is limited due to the cluster being primarily intended for faculty research, but we may be able to get our tasks in at a lower priority. Second, we plan on using the Jetson TX2 Dev Kit for our more advanced training. It's a machine-learning-specific graphics card developed by Nvidia and will allow us to train more advanced neural networks faster and better than our other options. Malcolm has access to this resource through the St. Olaf - Carleton Engineering team, which has two available for use.

To build mockups for our web UI (as well as for any other platforms we may extend onto), we plan on using Balsamiq, a wireframing program. We should be able to get by with free trial options, but if we end up relying on it more than expected during our UI design process, we might request a month-or-two-long subscription to its cloud service (at $9 per month).

References

[1] Y. Bengio, R. Ducharme, P. Vincent, and C. Jauvin. A neural probabilistic language model, 2003.

[2] M. Ghazvininejad, X. Shi, Y. Choi, and K. Knight. Generating topical poetry. In *Proceedings of ACL*, pages 1183–1191, 2016.

[3] Y. Ji, T. Cohn, L. Kong, C. Dyer, and J. Eisenstein. Document context language models. In *Proceedings of the 4th International Conference on Learning Representations (Workshop Track)*, 2016.

[4] A. Karpathy. The unreasonable effectiveness of recurrent neural networks. `http://karpathy.github.io/2015/05/21/rnn-effectiveness/`, 2015.

[5] J. Kim, J. Cheng, and M. S. Bernstein. Ensemble: exploring complementary strengths of leaders and crowds in creative collaboration. In *Proceedings of CSCW*, 2014.

[6] T. Mikolov, M. Karafát, L. Burget, J. Cernockỳ, and S. Khudanpur. Recurrent neural network based language model. In *Proceedings of Interspeech*, 2010.

[7] M. Roemmele and A. S. Gordon. Creative help: A story writing assistant. In *International Conference on Interactive Digital Storytelling*, page 81–92, 2015.

[8] M. Roemmele and A. S. Gordon. An encoder-decoder approach to predicting causal relations in stories. In *Storytelling Workshop at the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2018)*, June 5, 2018.

[9] R. Sloan. Writing with the machine. `https://www.robinsloan.com/notes/writing-with-the-machine/`, 2016.

[10] A. Sordon, M. Galley, M. Auli, C. Brockett, Y. Ji, M. Mitchell, J.-Y. Nie, J. Gao, and B. Dolan. A neural netowrk approach to context-sensitive generation of conversational responses. In *Proceedings of ACL*, page 196–205, 2015.

[11] T. Tran. Creating a text generator using recurrent neural network. `https://chunml.github.io/ChunML.github.io/project/Creating-Text-Generator-Using-Recurrent-Neural-Network/`, 2016.

[12] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan. Show and tell: a neural image caption generator. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3156–3164, 2015.