Mr. Speice

Independent Study Mentorship 3A

20 October 2020

Neural Networks for Natural Language Processing

Assessment 5 - Research Assessment

Date: 20 October 2020

Subject: Recurrent Neural Networks and LSTM Networks

Works Cited:

Olah, Christopher. “Understanding LSTM Networks.” Understanding LSTM Networks -- Colah's Blog, Github Pages, 27 Aug. 2015, https://colah.github.io/posts/2015-08-Understanding-LSTMs/.

Assessment:

Neural Networks are some of the most powerful technologies created in machine learning and artificial intelligence. They are able to do tasks a wide array of tasks including but not limited to classification and regression. Unfortunately, traditional neural networks lack one important aspect: memory. In a field like natural language processing, it is necessary to be able to have some sort of memory structure in order to be able to understand not only when it is given as a standalone object but also in the broader context of a passage or document. Recurrent Neural Networks (RNN) and Long Short-Term Memory Networks (LSTM) have been developed as a result of these limitations.

In synthesis, a Recurrent Neural Network is a network where “chained information” is passed through each subsequent layer, allowing for the neural network to take into account prior information in the piece of data. This serves to be valuable because now not only will the network understand the data in terms of the physical characters and words but also with some

sort of context attached. These context-based inferences will prove most beneficial in applications such as chatbots. It may prove helpful to read a chain of messages or read on the context that was presented much earlier in the conversation. Through this, the neural network can deduce the correct course of action and reply with a more appropriate response since it is able to look at the broader context. Another instance where this may prove useful is in sarcasm analysis to understand whether the earlier part of the Tweet indicates sarcasm for a latter part of the Tweet. This would not be possible without the loop-like nature of a Recurrent Neural Network, making it very powerful in comparison to traditional neural networks. Collectively speaking, Recurrent Neural Networks are a leap in the field of natural language processing as they remove the necessity of needing to use more traditional methods such as N-Gram models and allow for a pure machine learning approach to solving the problems in natural language processing.

The Long Short-Term Memory network is a special type of neural network that helps combat one of the Recurrent Neural Networks' biggest limitations: transferring localized information to a broader scale. Long Short-Term Memory Networks help combat this by adding four different layers that help keep track of and update important information. An LSTM Network probably would be unnecessary in classifying Tweets as depressive or not depressive because tweets are limited to 280 characters and therefore there is generally one central piece of information in a Tweet. However, LSTMs will be extremely beneficial in the chatbot portion of this project because direct messages on Twitter can be much longer than 280 characters. By applying an LSTM Network on the chatbot, it will ensure that the chatbot has the correct context in mind and is able to refer to context much earlier in messages. In synthesis, LSTM Networks will enable a more stable chatbot feature, allowing for a more refined product.

By understanding Recurrent Neural Networks and LSTM Networks, one of the largest challenges previously discovered has now lessened: context. Through leveraging the power of these networks, there is almost a guarantee that the original work will produce better results and have a higher performance. Collectively speaking, through developing knowledge over these networks, the barriers presented by natural language processing have dissipated. In future research, it will be important to gain insight into how these machine learning architectures can be incorporated into popular frameworks such as Tensorflow or PyTorch.

Annotations