Project Ideas

Jonah Philion
jonahphilion@college.harvard.edu
github

March 13, 2018

I have two ideas for a final project. The first is relevant to the talk by Marc'Aurelio Ranzato which got me very excited about unsupervised translation. The second connects what I know about computer vision with the NLP we've learned in 287.

1 Unsupervised Translation

In Facebook's Unsupervised Translation paper, training occurs in two steps. First, a parallel dictionary is learned in an unsupervised way. Second, a single encoder-decoder is learned using the parallel dictionary for word vectors.
I want to know if Unsupervised Translation can improve if

1. Word-pieces are used instead of word-vectors (as in Google's Zero-Shot)

2. Training occurs in one step instead of two

To do so, I'm proposing that the translation loss function include a language modeling loss. Not only should should the next word be predicted by the previous words, the next word should also be predicted by the translation of the previous words.
Another way to solve the two points mentioned might be to create a universal intermediate language: learn a lanuage L such that every human language l can be encoded into L and decoded out of L. The same loss function used in the Facebook paper (de-noising and back-translation) is used. To translate, we always pass through the intermediate language L.

$$l_1 \xrightarrow[l_1^{enc}]{} \xrightarrow[L_{l_1}^{dec}]{} L \xrightarrow[L_{l_2}^{enc}]{} \xrightarrow[l_2^{dec}]{} l_2$$

I argue this model is more expressive than just adding more layers to a decoder. Apart from that, I think the structure of the language L is interesting since the computer is making it's own language. Finally, this model has the benefit that we can train unsupervised translation of n distinct languages at once instead of just 2 as is done by Facebook. A system more similar to zero-shot would bring model complexity from linear to constant in the number of languages.

2   ResNet Does NLP

In HW1, I showed that ResNet can do sentiment classification on screenshots of text to about the same accuracy that logistic regression can do on word vectors. Link is here. How well can ResNet do on other NLP tasks? i.e. language modeling, encoding. Does ResNet improve performance in tasks involving character-based languages?