Ikhlas Attarwala
05-29-19 / Week #8
Applications of NNLMs / Paper 2

Achieving Open Vocabulary Neural Machine Translation with Hybrid Word-Character Models
by Luong, Minh-Thang and Manning, Christopher D.

Here we are introduced to a fascinating new neural machine translation for an open vocabulary system (can purely pay attention to either words or characters, but I believe a hybrid model was developed too). This system is quite fast to train and doesn't produce unknown words in translations. In order to do so, it uses an unk replacement technique, that replaces an unknown word with an 'unk' token, which is then finally replaced with a word based on alignment. While an advantage of this system is that it is much faster and easier to train, it's difficult to know whether the system really understands anything because the words are treated as independent entities.

The architecture of the model uses a deep LSTM encoder-decoder for word and character translation, and apart from unk, is fairly standard. The dataset used was the WMT '15 English to Czech translation. Their hybrid model that combined the strengths of both word- and character-models outperformed all other systems. The character model (achieved the goal of open vocabulary NMT) heavily relied on attention to improve performance, and longer backprop helped as well. Pure character NMT models also outperform past pure word NMT models.