SOMETHING SOMETHING SOMETHING...

Lukas Nabergall

April 3, 2017

1 Initial Idea Overview

The basic idea is to use a "massively hierarchical" transfer learning approach to learn better (general) language models with faster training during adaptation to new domains. We use a deep neural network (of some type) as the model. Each layer (or adjacent group of layers) of the network corresponds to a different level in the language hierarchy: the first layer (or group of layers) is learned on a very large dataset containing a large number of samples from every available language type, of every available language, of every available linguistic style, etc.; the second layer is specific to a given language type (and trained on the corresponding restricted data set); the third layer is specific to a given language; the fourth layer is specific to a given linguistic style; etc. The learning task is (something like) to predict the next character (and/or word) in a sequence.

Three reasons to think transfering from such different languages and, in particular, from such different language types, will be (at least a little) beneficial:

- At a high level, even seemingly vastly different language types, e.g. natural languages and programming languages, do have some common features—separation into "words", similarly structured whitespace and punctuation, some intersections in vocabulary, etc. In particular, a programming language and a natural language are more similar than a programming/natural language and a randomly sampled sequence of characters.

- We can typically describe a given program using natural language. Although this often assumes a significant amount of background knowledge, it already points to large overlaps in the expressible meanings between the two language types, indicating the existence of some shared features.

- It is likely (although currently unverified) that when a human reads and writes, say, natural language and programming language, they are using some common parts of the brain. Although the intersection between the parts of the brain used for each of the tasks may be small, this also strongly suggests the existence of common features between the two language types (in particular, that a neural network should likely be able to learn high-level representations useful to modeling/processing both language types).

2 Problem

Some Rough Ideas:

1. Recognizing/generating "comprehensible" natural language

2. Recognizing/generating/classifying any "structured" language (generally, English, Chinese, HTML, C, LaTeX, Morse, etc.)

3. Converting between linguistic styles (e.g. academic $\leftrightarrow$ Trump $\leftrightarrow$ Shakespeare $\leftrightarrow$ Twitter)

4. Correcting informal/sloppy/incorrect natural language (or, generally, correcting any "structured" language) — more specifically, domain-specific natural language correction (e.g. correcting text from/for Twitter, Wikipedia, a Trump speech, a Senate bill; more generally, HTML, C, LaTeX, Morse, Chinese, etc.)

3 Related Work

- Plank, What to do about non-standard (or non-canonical) language in NLP, (2016)
- Jozefowicz, Vinyals, Schuster, Shazeer, Wu, Exploring the Limits of Language Modeling, (2016)
- Zhang, Liu, Wang, Zhu, Neural Personalized Response Generation as Domain Adaptation, (2017)
- Wang and Zheng, Domain Adaptation of Recurrent Neural Networks for Natural Language Understanding, (2015)
- Wang and Zheng, Transfer Learning for Speech and Language Processing, (2015)
- Yoon, Yun, Kim, Park, Jung, Efficient Transfer Learning Schemes for Personalized Language Modeling using Recurrent Neural Network, (2017)
- Yosinski, Clune, Bengio, Lipson, How transferable are features in deep neural networks, (2014)
- Shazeer, Mirgoseini, Maziarz, Davis, Le, Hinton, Dean, Outrageously large neural networks: The Sparsely-Gated Mixture-of-Experts Layer, (2017)
-
-
-

4 Model

Model

Data

Training

Testing/Evaluation