

Natural Language Processing has been overrun by large neural language models! What should we make of that?

Christopher Manning (Stanford University)

In Natural Language Processing, the long dominant way of using the structure of human languages in systems for various downstream tasks was by building context-free grammar or richer parsers from hand-annotated morphosyntactic resources that display linguistic structure, that is, treebanks. However, recent deep learning language models are simply large artificial neural networks trained in a self-supervised fashion to predict a masked word in a given context. Nevertheless, once fine-tuned, these models yield much better task performance seemingly without any structural knowledge. What is a right-thinking (psycho)linguist meant to think of this? I first consider recurrent neural network models and introduce the notion of bounded hierarchical languages, showing that RNNs can generate such languages with optimal memory. I then examine how deep contextual language models like BERT learn knowledge of linguistic structure because it helps them in word prediction. Using a new method for identifying linguistic hierarchical structure emergent in artificial neural networks, I show how components in these models focus on syntactic grammatical relationships and anaphoric coreference, and, moreover, there seems to be significant shared cross-linguistic structure, or a kind of Universal Grammar. These results both help explain why recent neural models have brought such large improvements across many language-understanding tasks and provide intriguing hints about the possibility of learning language from observed evidence alone, as human children appear to do.