Non-NLP models

Description

As an NLP neophyte I thought the mechanisms of automated natural language processing, such as ChatGPT, esoteric and arbitrary. At best, the general pre-training transformer (GPT) model is structured as a series of mathematical functions that apparently work in simulating human language, but bear little relation to the processes by which language is actually spoken, written, read and understood by human beings.

To challenge this thought, I want to recast automated NLP processes in terms of *non*-automated natural language processes. How does language work outside of the mechanics of automated NLP?

How language works

1. Patterns

We favour developing and refining our ability to use language in a learning context that is flexible, fluid, dynamic, distributed, and positioned within a community of other language users. I say this by way of contrast to a rule-based, and lexical approach to language acquisition, a route through grammars and vocabularies. Language acquisition and use is better described as a process of absorbing, recalling, and synthesising patterns in speech and writing. The deployment of neural network architectures attempts to address this pattern processing aspect of language. See post: [Just one neuron.](#)

2. Corpus

Language patterns are acquired through exposure to many examples of spoken and written text in the context of life experiences. NLPs are â??trainedâ?• on very large corpora of texts to incorporate and deliver language patterns. See post: [Nothing beyond the text.](#)

3. Tokens

A competent language learner develops the ability to incorporate words they have never seen or heard, due in part to the way words are made of smaller units, standard prefixes (un-, re-, pre-, dis-, con-, sub-,

etc), word endings (-s, -es, -ed, -ious, -ology, etc), and syllables (arch, tec, bio, etc). Words also form assemblies, recognisable as meaningful blocks (â??down town,â?• â??matter of fact,â?• â??living on the edgeâ?•). Sometimes they are phrases, idioms, cliches that have recognised meanings without the speaker or reader having to assemble and analyse them from scratch. In fact, itâ??s fair to say that words are not the main units of language, but their subunits and combinations. NLP models typically deal with such units as *tokens*, generated via statistical methods from a corpus of texts. See post: [By the same token.](#)

4. Semantic relationships

Competent language users exercise the ability to fill in the blanks from context, or at least to hypothesise whatâ??s missing, e.g. â??That photo of a sunset is very <blank>.â?• The missing <blank> could be â??beautiful,â?• â??dramatic,â?• etc. If the speaker uses a non-existent word, â??That photo of a sunset is highly *evocacious*,â?• then the listener has recourse to the patterns in the malopropism to infer that the speaker meant â??evocative.â?• Context and word units and subunits help establish meaning. Importantly, we accomplish this by incorporating the relationships between words and word units to one another. Language competence involves the ability to substitute one word or token for another, deal with synonyms, and therefore analogy and metaphor. NLPs are remarkably good at this, in part due to **semantic encoding**, that positions tokens in multi-dimensional semantic spaces. See post: [Architecture in multidimensional feature space.](#)

5. Context

In formulating what they are about to say, language users incorporate what they just said. We can observe this capability in sentences. Speech would likely be ungrammatical and incomprehensible without that ability to position each word or subword in an ordered sequence of previous words. That stream constitutes a kind of context window that extends over time through the speakerâ??s own utterances including those of a companion interlocutor, and beyond, though with diminishing influence on what a person is about to say. Evidence for this capability comes from the use of pronouns. â??He wanted him to give it to himâ?• has a different meaning when it follows â??John liked Paulâ??s coat,â?• than â??John saw that Paulâ??s brother needed the coat.â?• Confusion will follow if the referents to the pronouns are too far back in the context window. The size of the context window is a major parameter in NLP models: the larger the context window the greater the capability of the model to contribute to the flow of a conversation, though at substantial computational cost. See post: [Training your AI.](#)

6. Position

As well as this extent of the stream of utterance within this context window, the order in which people say things really matters. â??I **only** eat vegetablesâ?• carries a different meaning to, â??**Only** I eat vegetables.â?• So the position of a word or phrase within a context window is important in establishing meaning. *Positional encoding* in NLPs attempts to operationalise this capability. See post: [Words in order.](#)

7. Attention

Not all words, parts of word and phrases carry equal status as we speak or write. We direct attention to whatâ??s important in a sentence. Speakers can indicate these priorities by giving vocal emphasis to certain words in a sentence, though not all languages support this. In written texts itâ??s common to

italicise words that require particular emphasis. But generally we rely on context to inform the reader where to focus attention. â??I didnâ??t throw the potato across the room. I only *eat* vegetablesâ?• places different emphasis than â??Please remove the meat from my plate. I only eat *vegetables*.â?• This focus on particular parts of an utterance is dealt with as *attentional encoding* in NLPs and mechanisms that deal with multiple attentions. See post: [Attending to more than one thing at a time.](#)

8. Tuning

Finally, language users need the capacity to reflect on and adjust their language usage, including by learning from others. This includes not only corrections to grammar and vocabulary usage, but what we may regard as content: facts, inferences, reasoning, interpretations, appropriateness, etc. This is accommodated by NLP developers recruiting human participants to adjust neural network parameters, through evaluating NLP outputs. Predictions and outputs are adjusted to integrate specialised language usages, and to correct or deprecate unwanted language outcomes. See post: [Fine tune your AI.](#)

Note

Featured image generated by Dall-E3 via Bing prompted by â??photorealistic AI urban apocalypse at high resolution include a billboard that says â??Non-NLP MODEL â?? Richard Coyne.'â?•

Category

1. Artificial Intelligence

Tags

1. chatgpt
2. NLP

Date Created
October 7, 2023
Author

rcoyne99

default watermark

Footer Tagline