

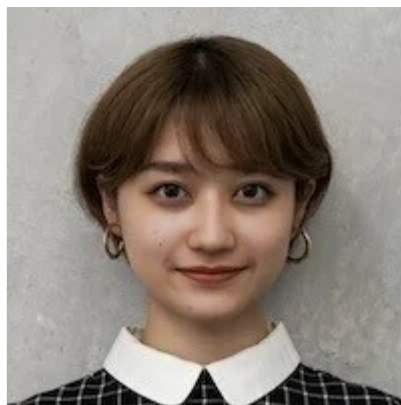


「深層学習がもたらす情報処理の可能性について」視聴報告

♡ 5

 情報処理学会・学会誌「情報処理」
2021/04/06 11:21

...



宮武茉莉（東京大学）

2021年1月29日開催の情処ウェビナー（第1回）

「深層学習がもたらす情報処理の可能性について」

に参加した。深層学習における重要な概念から、現在の技術の限界、さらには人間の脳との対応関係まで知ることができた。以下に要点を挙げて解説する。

「AI技術」

さまざまな意味で使われるようになった「AI技術」。その中でも深層学習は重要な技術であるが、日本ではこれまで十分に活用されてこなかった。IT人材はベンダに偏在しており、ユーザ企業でのデジタル・リテラシーが足りないため他業界の生産性向上に結びついていない可能性がある。

Disentanglement（もつれをひもとくこと）

深層学習という分野は2007年のHintonらの研究がきっかけで発展した。この深層学習を理解するのに重要な概念として disentanglement（もつれをひもとくこと）がある。

教師あり学習において、入力データはさまざまな要素が絡まり合った状態であるが、出力の1つ手前の層では「もつれがひもとかれた」状態になっている。たとえば体重と身長は相関していることが多いが disentangle して重要な要素を切り分けることで「体重が軽いのに身長が高い人」のようなパターンを想像することができる。

自己教師あり学習の一つ、条件付き変分オートエンコーダでは、顔を構成する主要な要素（性別・眼鏡・年齢など）が disentangle されていると各要素を変化させた顔を生成することができる（図-1）

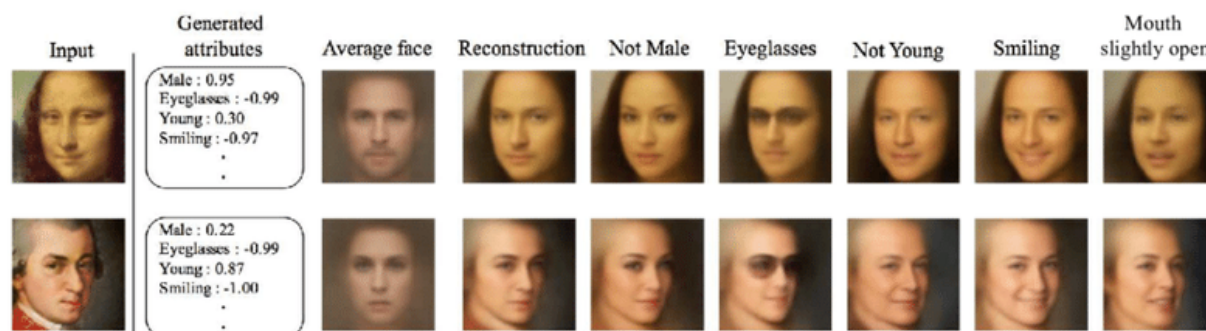


図-1 主要な要素を変化して生成された顔 [1]

現在のAI技術の限界

自己教師あり学習では自らのデータを予測するために disentangle された概念が徐々に構成されるが、その後教師データを与えるとそのデータに関する disentanglement が進むため、より細かく的確に概念化することができる。このように2つの学習手法を組み合わせた場合、自己教師あり学習を「事前学習」、その後には与えられる教師あり学習を「下流タスク」と呼ぶ。

最近大きく注目されているGPT-3というモデルでは下流タスクとして文書生成を行うことができ、非常に完成度の高いGPT-3自体に関するブログ記事が生成され話題になった（図-2）。ほかにも英語からHTMLコードへの翻訳や素数の生成などさまざまなタスクに応用されているが、このGPT-3にも限界があり文章の「意味を理解する」タスクはまだできていない。

それでは「意味を理解する」とは一体どういうことなのか。

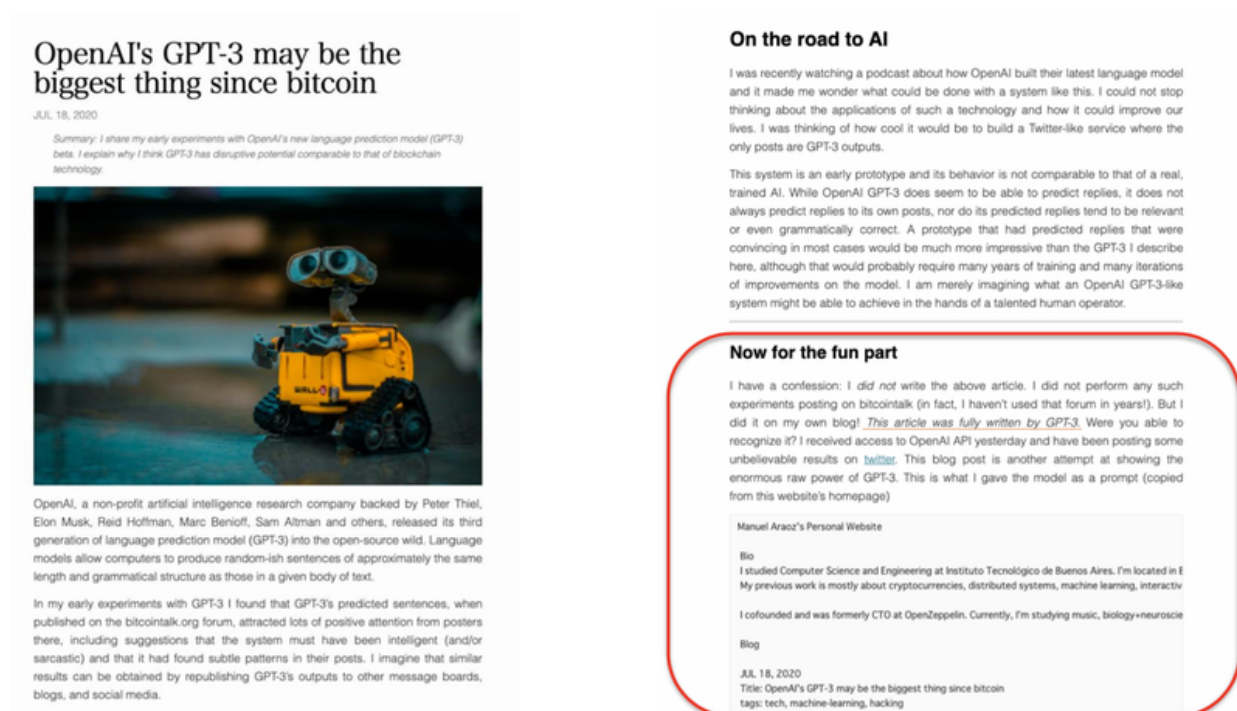


図-2 GPT-3が生成したブログ記事

<https://maraoz.com/2020/07/18/openai-gpt3/>

「意味を理解する」とは

ガラスのコップを落とすと割れる，ということを人間は実際にコップを落さなくても想像できる。これは頭の中で外界世界をシミュレートするモデル，世界モデルを持っているからである。

世界モデルは生得的に存在するのではなく学習によって作られており，これを用いてある行動をとったときに何が起こるかを予測することができる。ある行動というのは言語的に表現できるため，言語的な入力を与えられたときに世界モデルを駆動して外界世界を想像できるということは「意味を理解

している」と言えるのではないだろうか.

脳の仕組みとアルゴリズムの対応関係

「意味を理解する」ことができる人間の脳はどのような仕組みで動いているのだろう. 脳の機能の中で重要な2つのループ, (1) 大脳皮質 - 大脳基底核ループ, (2) 大脳皮質 - 小脳ループ, のそれぞれを単純化して説明してみよう.

(1) 大脳皮質 - 大脳基底核ループにおいて, 大脳皮質は世界モデルの役割を果たしている. このモデルにより予測を行い, 基底核では行動を選択し報酬を受け取っている.

(2) 大脳皮質 - 小脳ループにおいて, 小脳は, アルゴリズム的に長いステップを使って計算できることを短いステップの演算の組合せとして近似する「アルゴリズム蒸留」を行っている. 大脳皮質 - 大脳基底核ループで行われた処理を簡便に高速に実行できるようにサブルーチン化しているのである.

人間の脳の機能はまだほとんど解明されていないが, ディープラーニングの技術が進展し意味処理の技術がこれから進んでいくことで, それに対応する人間の知能や脳の構造に関しても理解されてくるのではないだろうか. それによってロボット等の実世界処理や自然言語の活用にインパクトをもたらし, 教育や医療に関しても重要な知見に繋がるだろう.

1時間に凝縮された松尾豊先生のウェビナーはこの記事にはとても書ききれ

ない内容量だった。特にディープラーニングと人間の脳の対応関係について考えるところが興味深く、さらに続きを聞きたい。

(2021年2月18日受付)

(2021年4月6日note公開)

参考文献

- 1) Suzuki, M., Nakayama, K. and Matsuo, Y. : Joint Multimodal Learning with Deep Generative Models, ICLR 2017 workshop, Toulon, France (April24–26, 2017).

■宮武茉莉

東京大学大学院工学系研究科修士1年。