Making sense of Large Language Models

Theoretical foundations and practical implications for linguists, translators, teachers and other language experts

Piotr Pęzik

LLMs

Large Language Models (LLMs) based on Generative Pretrained Transformers (GPTs) such as OpenAI's chatGPT, Meta's LLama have recently been advertised as a genuine example of Artificial Intelligence, possibly even exhibiting features of Artificial General Intelligence (AGI). LLMs have been expected to largely replace translators, writers, programmers, foreign language teachers, customer support agents, etc. effectively calling into question the need to educate a wide spectrum of professionals. This course puts the advent of the LLM technology into a wider theoretical context of AI and natural language processing research. It introduces the basic principles underlying this technology, emphasizing its key advantages as well as its fundamental limitations, which are not likely to be solved in the immediate future. We will address a number of questions about when and how LLMs can be used in research and professional applications. The author of the course has considerable hands-on experience in fine-tuning Trurl.ai, the first open LLM adapted for a variety of Polish-English tasks, and therefore feels qualified to explain key notions and issues which determine the usability of LLMs, such as prompt engineering, hallucinations, temperature sampling, fine-tuning, the curse of reversal, self-reflection, next-word-prediction and model benchmarking/ evaluation.

Syllabus

- Week 1-2
  - Strong and Weak Artificial Intelligence in science and popular culture
  - Are we there yet?
  - Theoretical foundations
    - Turing test, Turing machine, Von Neumann machine
    - The Computational theory of the mind, The Chinese Room Argument and its refutations
  - Homework: Searle's lectures
  - Listening comprehension tests:
    - http://goo.gl/forms/wgqnNZJHhe
    - http://goo.gl/forms/oX6KRRW1UO
  - https://bard.google.com, https://trurl.ai, http://chat.openai.com

Introducing LLMs

- Week 3-4
  - Training Large Language Models
    - Transformers and Generative Pretrained Transformers
    - Next-word-prediction
    - Finetuning on instructions, chain-of-thought, zero/one/few-shot
    - Reinforcement Learning
  - Inference
    - Prompting, prompt engineering
    - Temperature sampling
    - Hallucinations
    - Self-reflection
  - Evaluation
    - Ad hoc
    - Benchmarks

Tokens vs Words

- https://platform.openai.com/tokenizer