Compression for AGI

Yuxuan Lu

Nov. 3 2023

Northeastern Human-Centered AI Lab

- Generative models are *Lossless* compressors
  - "ChatGPT Is a Blurry JPEG of the Web"? No!
- LLMs are *State-of-the-Art* text compressors
  - comparing to deflate(gzip), Zstd, etc.
- Re-think about the training objective of foundation models

Revisit the Training Crocess of LLMs

Revisit the Training Crocess of LLMs