

Large Language Models: A Bird's Eye View

Graham Neubig



Carnegie Mellon University
Language Technologies Institute



<https://phontron.com/slides/neubig231lms.pdf>

NLP Before Large Language Models (Train/Test)



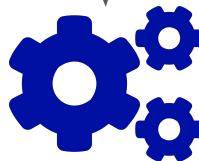
Who is the president of the US?

Joe Biden

Where is Carnegie Mellon located?

Pittsburgh

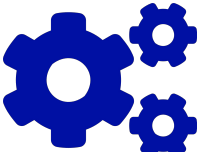
What is the capital of Pennsylvania?



Harrisburg

NLP With Large Language Models ([Prompting](#))

Q: What is the capital of Pennsylvania?

A: —  → Harrisburg

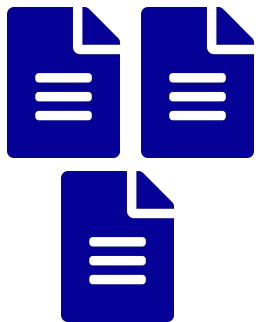
It's for real this time. After months of legal drama, bad memes and will-they-or-won't-they-chaos to put your favorite rom-com to shame, Elon Musk has closed his \$44 billion acquisition of Twitter. Musk sealed the deal Thursday night, taking Twitter private and ousting a handful of top executives — CEO Parag Agrawal included — in the process.

TL;DR: —  → Elon Musk has bought Twitter.

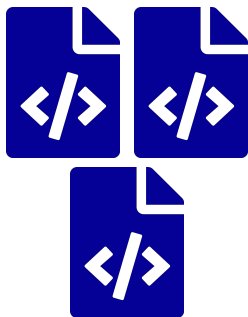
- Mostly through APIs such as [GPT](#), [Cohere](#), [PaLM](#)

From Zero to ChatGPT

Lots of web text



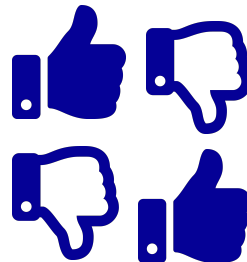
Lots of GitHub code



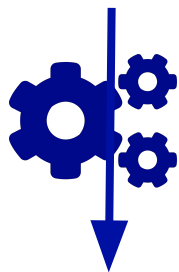
Lots of annotated data



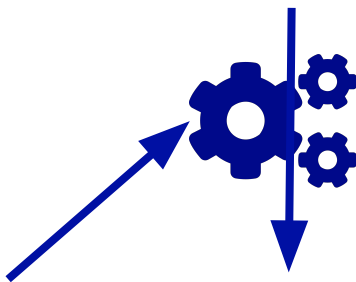
Human judgements of response quality



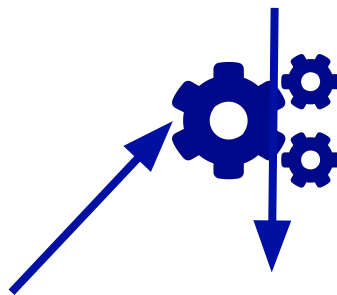
Chat-oriented data



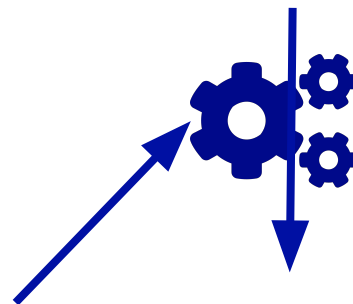
davinci



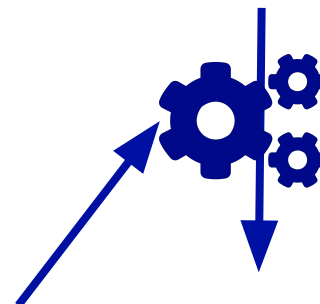
code-
davinci-002



text-
davinci-002



text-
davinci-003



davinci-
3.5-turbo
(ChatGPT)



What building is the CMU School of Computer Science Located in?



The Carnegie Mellon University School of Computer Science is located in the Gates and Hillman Centers on the university's campus in Pittsburgh, Pennsylvania.



🔄 Regenerate response






What building is the CMU School of Computer Science Located in?



The Carnegie Mellon University School of Computer Science is located in the Gates and Hillman Centers on the university's campus in Pittsburgh, Pennsylvania.

 Regenerate response



Which presidents were born in New York City?



Only one President of the United States was born in New York City: Theodore Roosevelt. He served as the 26th President.

Stop generating



What is Easier w/ LLMs?

- Achieving **good accuracy** on existing NLP tasks
- Quickly **tackling a new and creative task**
- **Coherent generation** of medium-length text (one to several pages)
- **Code** generation and code completion

4 Big Questions about LLMs

What to do about the “black box”?

- Current LLMs are proprietary and little information is released about them, what to do?
- **Solutions?**
 - Build competitive open and well-documented models (e.g. Meta’s [OPT/LLaMa](#), Stanford’s [Alpaca](#))
 - Probe models for properties (e.g. [BIGBench](#), adversarial “[red-teaming](#)”)
 - Urge/force providers to be more transparent (e.g. [petitions](#), regulation such as the [EU AI Act](#))

Lack of factuality, reliability of results

- Model outputs are often fluent but wrong, toxic, or otherwise undesirable.
- **Solutions?**
 - “Citing your sources” through retrieval-based models (e.g. [Bing search](#), [perplexity.ai](#))
 - Getting models to “know what they know” through [calibration](#)
 - Post-hoc verification of outputs (e.g. Inspired [Critique](#))

Lack of robustness

- Models are less effective when tried on new applications, domains, and languages.
- **Solutions?**
 - Methods for engineering/prompting on individual tasks
 - Training special-purpose models for domains (e.g. Meta [Galactica](#), Google [MedPaLM](#), [BloombergGPT](#))
 - Fine-tuning on or upweighting on low-resource languages (e.g. [BLOOM](#))

Societal Impacts

- LLMs could cause large-scale societal disruption:
 - **Job loss** or realignment, expansion of **disparities**
(Solutions? Forward-looking projections and policy)
 - Increase in targeted **phishing, fraud, or manipulation**
(Solutions? Stricter verification of information sources and detection of machine-generated text)
 - Increased use of **algorithmic decisions** exacerbating disparities
(Solutions? Self- and imposed regulation for fairness)

Final Words



A good way to end a presentation is to leave your audience with a clear and concise message that they can remember and act on. You can do this by summarizing your key points, providing a call to action, or sharing a memorable quote. You should also thank your audience for their time and attention.

Here are some tips for ending your presentation effectively:

- **Summarize your key points.** This is a great way to ensure that your audience remembers the most important information from your presentation. You can do this by simply recapping your main points or by creating a visual summary on a slide.
- **Provide a call to action.** This is a great way to get your audience to take the next step after your presentation. You can do this by asking them to sign up for a newsletter, visit your website, or take some other action.
- **Share a memorable quote.** This is a great way to leave your audience with something to think about after your presentation. You can use a quote from a famous person, a personal anecdote, or something else that will resonate with your audience.
- **Thank your audience.** This is a simple but important way to show your appreciation for your audience's time and attention. You can do this by saying something like, "Thank you for your time today. I hope you found my presentation helpful."