

Ethics in NLP

CMSC 473/673 - NATURAL LANGUAGE PROCESSING

Learning Objectives

Identify ethical issues of LLMs/transformers from various lenses (social, environmental, legal, economic, etc.) by...

- Extracting them from the Stochastic Parrots paper
- Extending them with your own perspectives

Determine how these issues apply to any LM

Review: What is a foundation model?

A model that captures “foundational” or core information about a modality (e.g., text, speech, images)

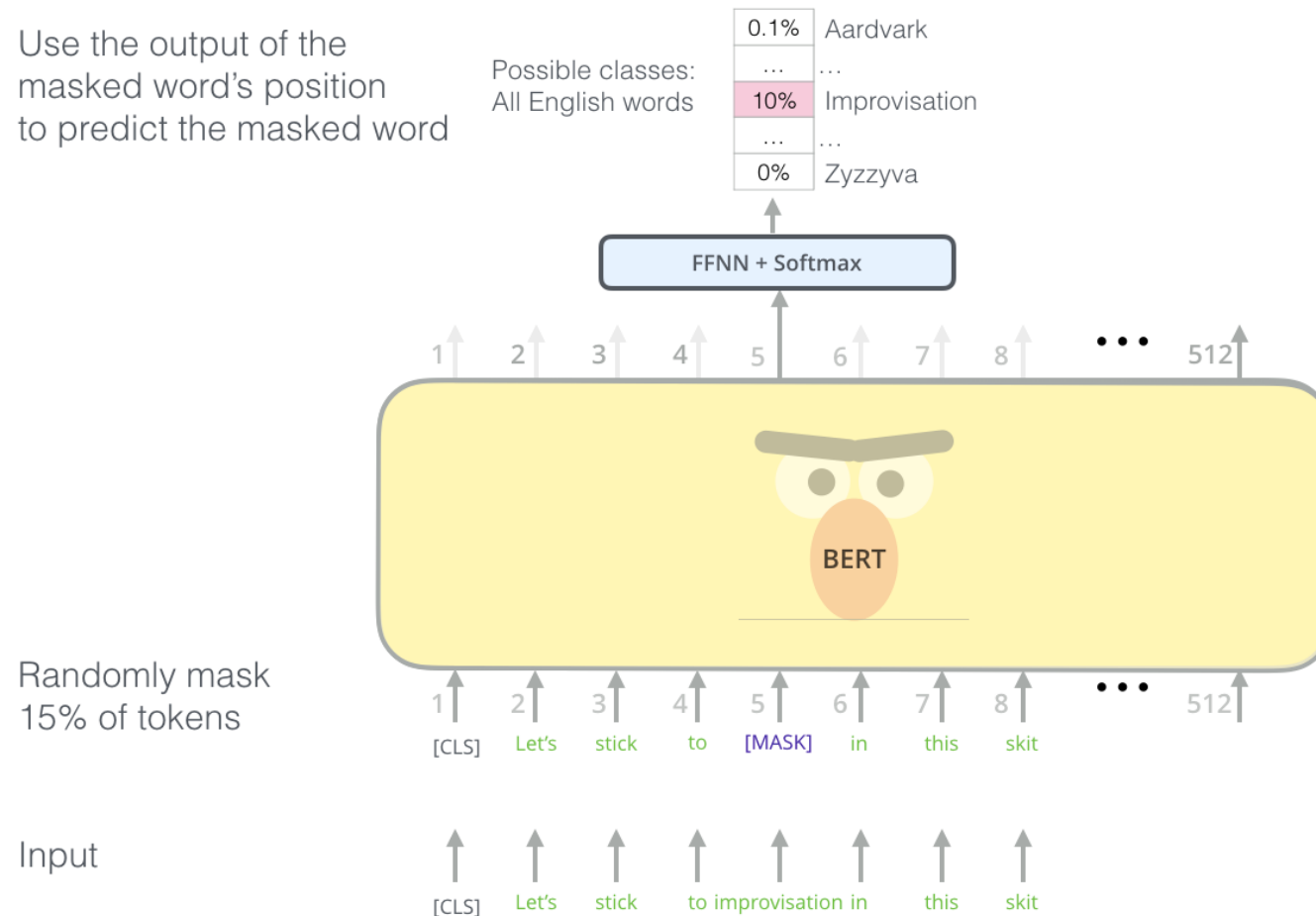
Pretrained on a large amount of data & able to be finetuned on a particular task

Self-supervised

All non-finetuned large language models (LLMs) are foundation models

Review: BERT (Devlin et al. 2019)

Use the output of the masked word's position to predict the masked word



<http://jalammar.github.io/illustrated-bert/>

Review: GPT Family

- Decoder-only
 - Input: Text sequence
 - Goal: Predict the next word given the previous ones
- How to use:
 - Ask GPT* to continue from a prompt.
 - Finetune smaller GPTs for more customized generation tasks.
 - ChatGPT cannot be finetuned since it is already finetuned
 - Use OpenAI's API to get them to fine-tune GPT-3 for you.

Review: T5 Family of Models

- Encoder-decoder
 - Input: Text sequence with random word spans deleted
 - Goal: Generate the deleted word spans
- How to use:
 - Finetune smaller ones for either generation or classification tasks.
 - Prompt tuning (train a sequence of embedding which get prefixed to the input)

Stochastic Parrots

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, 610–623.
<https://doi.org/10.1145/3442188.3445922>

Ethical Issues

1. Environmental
 2. Financial
 3. Diversity
 4. Static Data
 5. Bias
 6. Accountability
 7. Lack of Understanding
 8. Subjective Coherence
 9. Harms
- + 10. Mitigation Strategies

1) Environmental

2) Financial

3) Diversity

4) Static Data

5) Bias

6) Accountability

7) Lack of Understanding

8) Subjective Coherence

9) Harms

10) Mitigation Strategies

Issues of all LMs

How many of these are relevant to smaller LMs?

Reporting Bias

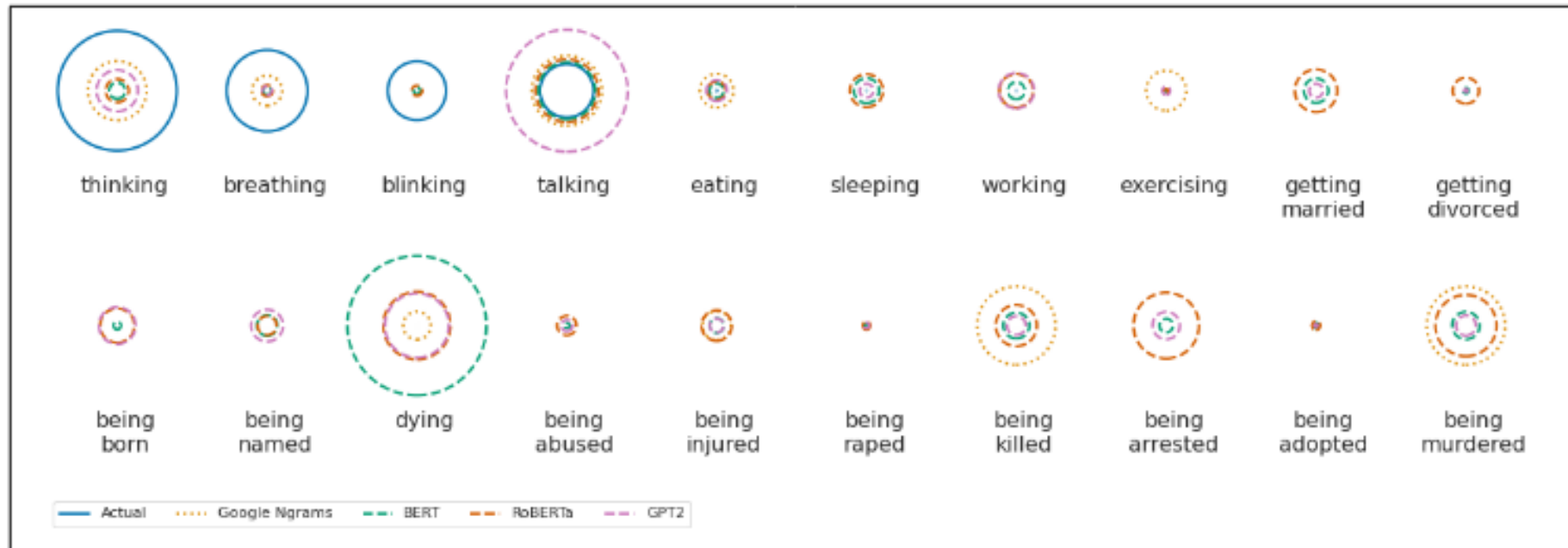


Figure 1: Frequency of actions performed or occurring to people during their lifetime from very frequent (daily), through once in a lifetime events, to very rare (don't happen to most people). Note that actual frequencies of rare events are too small to show. See Appendix A for the exact frequencies.

Near Duplicates in Data

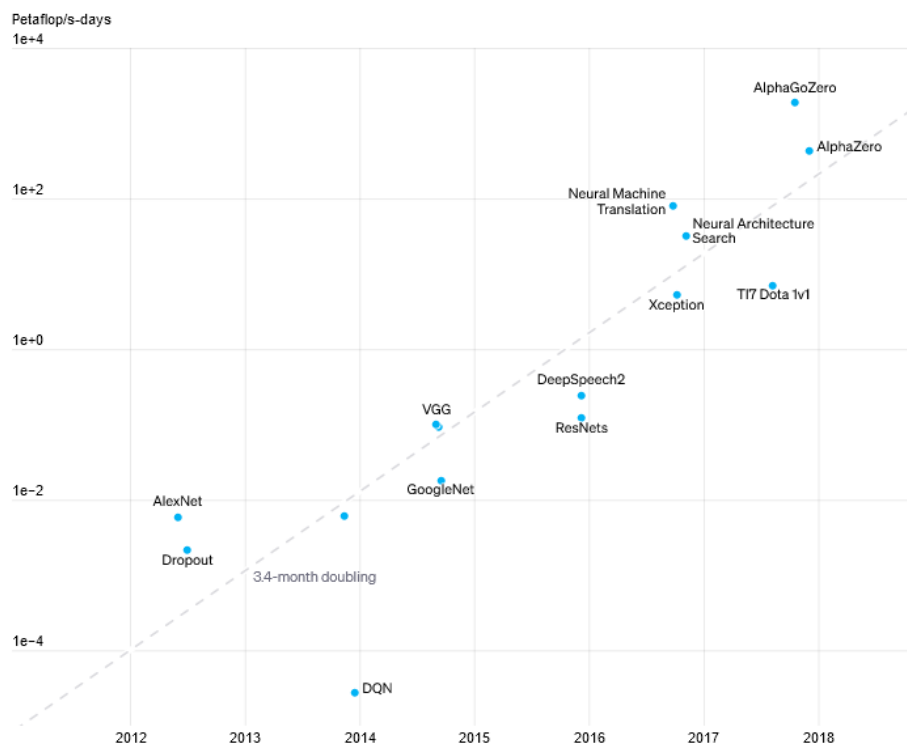
Dataset	Example	Near-Duplicate Example
Wiki-40B	\n_START_ARTICLE_\nHum Award for Most Impactful Character\n_START_SECTION_\nWinners and nominees\n_START_PARAGRAPH_\nIn the list below, winners are listed first in the colored row, followed by the other nominees. [...]	\n_START_ARTICLE_\nHum Award for Best Actor in a Negative Role\n_START_SECTION_\nWinners and nominees\n_START_PARAGRAPH_\nIn the list below, winners are listed first in the colored row, followed by the other nominees. [...]
LM1B	I left for California in 1979 and tracked Cleveland's changes on trips back to visit my sisters .	I left for California in 1979 , and tracked Cleveland's changes on trips back to visit my sisters .
RealNews	KUALA LUMPUR (Reuters) - Roads in Southeast Asia have been getting a little louder lately as motorcycle makers, an aspiring middle class and easy bank credit come together to breed a new genus of motorcyclists – the big-bike rider. [...]	A visitor looks at a Triumph motorcycle on display at the Indonesian International Motor Show in Jakarta September 19, 2014. REUTERS/Darren Whiteside\nKUALA LUMPUR (Reuters) - Roads in Southeast Asia have been getting a little [...] big-bike rider. [...]
C4	Affordable and convenient holiday flights take off from your departure country, "Canada". From May 2019 to October 2019, Condor flights to your dream destination will be roughly 6 a week! Book your Halifax (YHZ) - Basel (BSL) flight now, and look forward to your "Switzerland" destination!	Affordable and convenient holiday flights take off from your departure country, "USA". From April 2019 to October 2019, Condor flights to your dream destination will be roughly 7 a week! Book your Maui Kahului (OGG) - Dubrovnik (DBV) flight now, and look forward to your "Croatia" destination!

Table 1: Qualitative examples of near-duplicates identified by NEARDUP from each dataset. The similarity between documents is highlighted. Note the small interspersed differences that make exact duplicate matching less effective. Examples ending with “[...]” have been truncated for brevity.

Energy of Models

AlexNet to AlphaGo Zero: 300,000x increase in compute

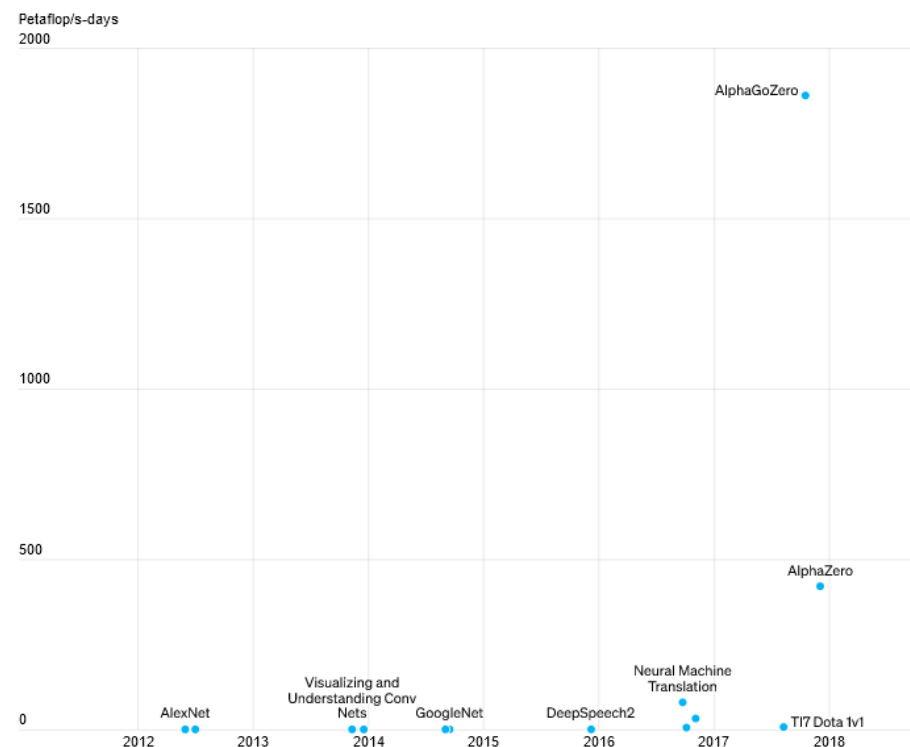
Log scale Linear Scale



The total amount of compute, in petaflop/s-days,^D used to train selected results that are relatively well known, used a lot of compute for their time, and gave enough information to estimate the compute used.

AlexNet to AlphaGo Zero: 300,000x increase in compute

Log scale Linear Scale



The total amount of compute, in petaflop/s-days,^D used to train selected results that are relatively well known, used a lot of compute for their time, and gave enough information to estimate the compute used.

Energy Shift

Two distinct eras of compute usage in training AI systems

Show error bars All ▾

Petaflop/s-days

