Large Language Models and Machine Learning for Unstructured Data

Lecture 2:
Finetuning LLMs

Stephen Hansen
University College London

Introduction

In the last lecture, we discussed how large language models solve word prediction problems via attention operations.

This gives us excellent predictive models for $w_n \mid \mathbf{w}_{-n}$.

Autoregressive language models are specialized at the prediction task $w_N \mid \mathbf{w}_{-N}$ and are used to generate language.

However, in most applications the prediction functions are only the starting point for NLP tasks.

In this lecture, we discuss how LLMs can be adapted for specific tasks.

Outline

1. Self-Supervised Finetuning
2. Supervised Finetuning
3. Instruction Finetuning

Further material:

- Jesse Mu's YouTube lecture `https://ytube.io/3vmF`.
- Elliott Ash's lecture slides `https://github.com/elliottash/lm_lss_2024`.

Lightcast Case Study

To discuss some of these themes, we'll draw on data on job postings provided by Lightcast.

Data used in [Hansen et al., 2023] to measure indidence of remote work across countries, cities, occupations, and firms.

Measures built using LLMs and other information retrieval algorithms.

Such benchmarking exercises are arguably not as common in economics as they should be.

Structure of Dataset

Table: Counts of Vacancy Postings, Employers, and Cities, January 2014 to January 2023

(1)	(2)	(3)	(4)
Country	Vacancies	Employers	Cities
New Zealand	1,700,523	36,201	67
Australia	8,607,160	197,870	59
Canada	11,711,357	712,577	3,691
United Kingdom	74,576,747	876,103	2,268
United States	161,872,915	3,485,630	31,635
Total	258,468,702	5,308,381	37,720

Note: Reported counts pertain to the universe of online postings from January 2019 onwards and a 5% random sample from 2014 to 2018, after we drop about 6% of the postings in the data-cleaning steps described in Appendix A. We rely on Lightcast's proprietary algorithm to identify employers and cities.

Software Developer

Pearson ★★★½☆ 2,739 reviews
Australia
Remote
Full-time
You must create an Indeed account before continuing to the company website to apply

Apply on company site

Our purpose: At Pearson we 'add life to a lifetime of learning' so everyone can realise the life they imagine. We do this by creating vibrant and enriching learning experiences designed for real-life impact.

Our company: Pearson was founded in 1844 and has been built on our ability to grow with and adapt to a constantly evolving market. Our 20,000+ employees are dedicated to creating the high-quality, digital-first, accessible and sustainable resources for lifelong learning.

Flexible working: Pearson is committed to hybrid working practices. When you are not working from home, you'll be based in our Nunawading office that has free parking and is walking distance to 2 train station. This is a great location for those that are not a fan of the city commute.

The Role : As a Software Engineer, you will be joining one of our cross-functional scrum teams and will play a key role in the development of our online assessment platform. Reporting to our Engineering Manager, you will work from home and collaborate via telecommuniting platforms.

What you will do:

Expense Processor (Remote U.S.A.)

Plus Relocation ★★★☆☆ 17 reviews
Minneapolis, MN 55426 • Remote
Full-time
You must create an Indeed account before continuing to the company website to apply

Apply on company site

Job details

Job Type:
Full-time
Work From Home:
Allowed

Location:
Anywhere

Full Job Description

Plus Relocation is looking for a numbers driven, detail

Software Developer

Pearson ★★★½☆ 2,739 reviews
Australia
Remote
Full-time
You must create an Indeed account before continuing to the company website to apply

Apply on company site

Our purpose: At Pearson we 'add life to a lifetime of learning' so everyone can realise the life they imagine. We do this by creating vibrant and enriching learning experiences designed for real-life impact.

Our company: Pearson was founded in 1844 and has been built on our ability to grow with and adapt to a constantly evolving market. Our 20,000+ employees are dedicated to creating the high-quality, digital-first, accessible and sustainable resources for lifelong learning.

Flexible working: Pearson is committed to hybrid working practices. When you are not working from home, you'll be based in our Nunawading office that has free parking and is walking distance to 2 train station. This is a great location for those that are not a fan of the city commute.

The Role : As a Software Engineer, you will be joining one of our cross-functional scrum teams and will play a key role in the development of our online assessment platform. Reporting to our Engineering Manager, you will work from home and collaborate via telecommuniting platforms.

What you will do:

Expense Processor (Remote U.S.A.)

Plus Relocation ★★★☆☆ 17 reviews
Minneapolis, MN 55426 • Remote
Full-time
You must create an Indeed account before continuing to the company website to apply

Apply on company site

Job details

Job Type:
Full-time
Work From Home:
Allowed

Location:
Anywhere

Full Job Description

Plus Relocation is looking for a numbers driven, detail

Deputy Home Care Manager

Habitation Care Ltd
Brighton BN1
£21,246 - £26,289 a year - Full-time, Part-time, Temporary contract, Fixed term contract, Temp to perm

Apply now

We are looking for a Deputy Home Manager with domiciliary care experience to join our company. You will work from home care facilities with a strong track record of quality service.

The person we are looking for must have a positive, and a can-do work attitude at all time.

The person we are looking for must have at least 1 years working experience in a domiciliary care or care home managers role.

The role is for 38 hours per week plus on call duties, and sometimes cover of care calls would be required.

The person will be preparing supporting the Registered manager to carryout daily tasks.

Job Types: Full-time, Part-time, Fixed term, Temp to perm
Contract length: 36 months
Part-time hours: 38 per week

General Builder (Bricklayer Based)

Birkby Construction Limited
Maidstone
£14.50 an hour - Full-time, Permanent

Apply now

General Builder (bricklayer based) required for Small Works Department of Birkby Construction Limited on a PAYE basis. Company vehicle provided. Applicant must be self-motivated and confident. Willing to remote work sites.

Job Types: Full-time, Permanent

Work from Home: Not Available

Salary: £14.50 per hour

Benefits:

- Company car

Schedule:

- Monday to Friday

Licence/Certification:

- CSCS (preferred)

Deputy Home Care Manager

Habitation Care Ltd
Brighton BN1
£21,246 - £26,289 a year - Full-time, Part-time, Temporary contract, Fixed term contract, Temp to perm

Apply now

We are looking for a Deputy Home Manager with domiciliary care experience to join our company. You will work from home care facilities with a strong track record of quality service.

The person we are looking for must have a positive, and a can-do work attitude at all time.

The person we are looking for must have at least 1 years working experience in a domiciliary care or care home managers role.

The role is for 38 hours per week plus on call duties, and sometimes cover of care calls would be required.

The person will be preparing supporting the Registered manager to carryout daily tasks.

Job Types: Full-time, Part-time, Fixed term, Temp to perm
Contract length: 36 months
Part-time hours: 38 per week

General Builder (Bricklayer Based)

Birkby Construction Limited
Maidstone
£14.50 an hour - Full-time, Permanent

Apply now

General Builder (bricklayer based) required for Small Works Department of Birkby Construction Limited on a PAYE basis. Company vehicle provided. Applicant must be self-motivated and confident. Willing to remote work sites.

Job Types: Full-time, Permanent

Work from Home: Not Available

Salary: £14.50 per hour

Benefits:

- Company car

Schedule:

- Monday to Friday

Licence/Certification:

- CSCS (preferred)

Deputy Home Care Manager

Habitation Care Ltd
Brighton BN1
£21,246 - £26,289 a year - Full-time, Part-time, Temporary contract, Fixed term contract, Temp to perm

Apply now

We are looking for a Deputy Home Manager with domiciliary care experience to join our company. You will work from home care facilities with a strong track record of quality service.

The person we are looking for must have a positive, and a can-do work attitude at all time.

The person we are looking for must have at least 1 years working experience in a domiciliary care or care home managers role.

The role is for 38 hours per week plus on call duties, and sometimes cover of care calls would be required.

The person will be preparing supporting the Registered manager to carryout daily tasks.

Job Types: Full-time, Part-time, Fixed term, Temp to perm
Contract length: 36 months
Part-time hours: 38 per week

General Builder (Bricklayer Based)

Birkby Construction Limited
Maidstone
£14.50 an hour - Full-time, Permanent

Apply now

General Builder (bricklayer based) required for Small Works Department of Birkby Construction Limited on a PAYE basis. Company vehicle provided. Applicant must be self-motivated and confident. Willing to remote work sites.

Job Types: Full-time, Permanent

Work from Home: Not Available

Salary: £14.50 per hour

Benefits:

- Company car

Schedule:

- Monday to Friday

Licence/Certification:

- CSCS (preferred)

Self-Supervised Finetuning

Language is Context Specific

Recall the example of Wikipedia vs HBR embeddings.

The same properties will be true in the embeddings built from LLMs.

Job posting text may have distinct relationships among words compared to generic English: for example, ‘operations’ and ‘military’.

But unlike with word2vec we cannot build a model from scratch.

Instead, start from pre-trained model and update it for our domain.

BERT

Our starting point is BERT (Bidirectional Encoding Representations from Transformers) [Devlin et al., 2019].

Trained on BooksCorpus (800M words) and English Wikipedia (2,500M words).

Masked language modeling. 15% of words randomly masked and given [MASK] token. [MASK] token embeddings built to successfully predict underlying word.

Original paper had next-sentence prediction but has since been dropped from loss function in extensions [Liu et al., 2019].

Base model has twelve layers, 768-dimensional embeddings, 110M parameters.

Self-Supervized Finetuning

To obtain a more context-specific representation of language, one can simply repeat masked language modeling on a new dataset.

The obtained embeddings will reflect 'meaning' in the context from which the new training data comes.

Updating all the parameters in the original model can be computationally expensive so need for solutions.

Distillation

[Hinton et al., 2015] introduces the idea of distillation which allows a smaller network to leverage the expressive power of the larger network.

Consider a classification problem in which a teacher network generates predicted probabilities $T_1, \ldots, T_M$.

A smaller student network generates predicted probabilities $S_1, \ldots, S_M$.

A distillation loss is added to the student's loss function:

$$\sum_i T_i \log(S_i)$$

The 'knowledge' of the teacher is used by the student to in formulating prediction probabilities.

Application to BERT

DistilBERT [Sanh et al., 2020] applies distillation to BERT.

Only 66 million parameters, but retains similar performance as BERT on standard NLP tasks.

NB: implementation of distillation requires access to the original model.

Reconstructed Word Probabilities

'software engineers' Sentence		'petroleum engineers' Sentence	
Word	Prob.	Word	Prob.
it	0.08	energy	0.279
automotive	0.079	oil	0.27
technology	0.072	petroleum	0.088
healthcare	0.058	mining	0.035
insurance	0.053	defence	0.021
software	0.041	automotive	0.02
engineering	0.031	construction	0.017
public	0.03	gas	0.017
infrastructure	0.028	engineering	0.016
financial	0.028	water	0.012

Table 1: Predictions for Masked Words in Example Sentences

This tables displays masked word prediction probabilities for the two example sentences above. The training corpus for estimating these probabilities is English-language online job postings provided by Lightcast (formerly Emsi Burning Glass). The Transfomer model estimated for the task is DistilBERT (Sanh et al. 2020). See Hansen et al. (2023) for more details.

Does Further Pre-Training Make a Difference?

Out-of-the-box model

Mask token: [MASK]

After training, position will then transition to work from [MASK], dedicated internet connection required by that time.

Compute

Computation time on cpu: 0.0792 s

secure	0.143
centralized	0.066
dedicated	0.046
wireless	0.040
reliable	0.028

Model with additional pre-training

Mask token: [MASK]

After training, position will then transition to work from [MASK], dedicated internet connection required by that time.

Compute

Computation time on cpu: 0.0804 s

home	0.913
school	0.014
office	0.010
work	0.007
location	0.005

Supervised Finetuning

Relating Text to Metadata

In many cases in economics, we have covariates associated with documents we wish to relate to text (regress y_d on $\mathbf{w}_d$).

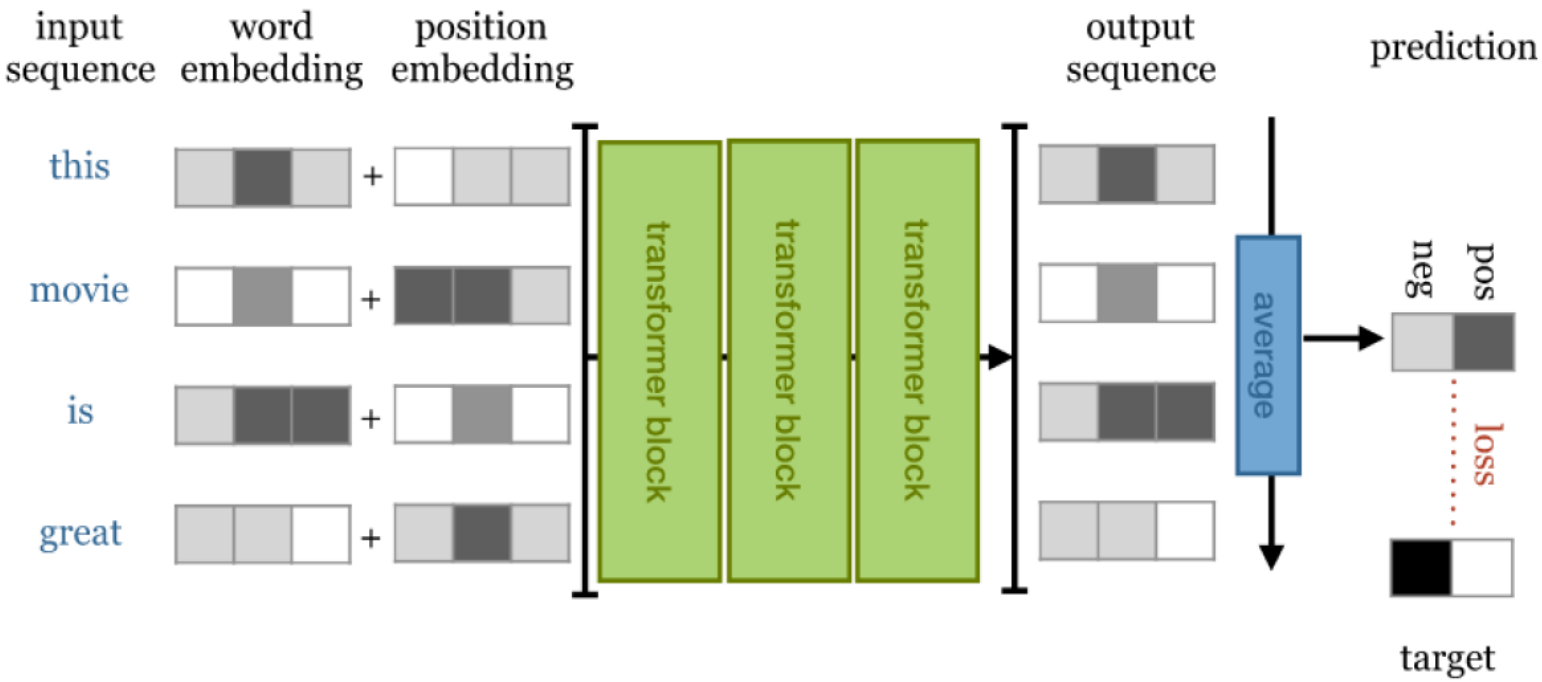
Bag-of-words methods: logistic regression, Naive Bayes, multinomial inverse regression.

LLMs can also be used for text regression:

- Basic approach is to take output layer of LLM as input data into a separate regression.
- More powerful alternative is supervised finetuning: adjust network weights in hidden layers to represent language in way most informative for prediction problem.

See second coding session for example in context of corporate filings.

Supervised Finetuning Pipeline



Low-Rank Adaptation

LoRA is a popular method for overcoming the compuational challenge.

Recall the parameters of Transformer models can be stacked in matrices.

Suppose $W \in \mathbb{R}^{d \times k}$ is a set of trainable parameters in the LLM.

LoRA represents matrices in SFT as

$$W = W_0 + BA$$

where $W_0 \mathbb{R}^{d \times k}$ are the weights from the pretrained model, $B \in \mathbb{R}^{d \times r}$, $A \in \mathbb{R}^{r \times k}$, $r \ll \{\min d, k\}$.

During SFT, W_0 are frozen and only B and A are updated.

Intuition is that only small amount of LLM's 'knowledge' is needed for regression model compared to learning relationships among words.

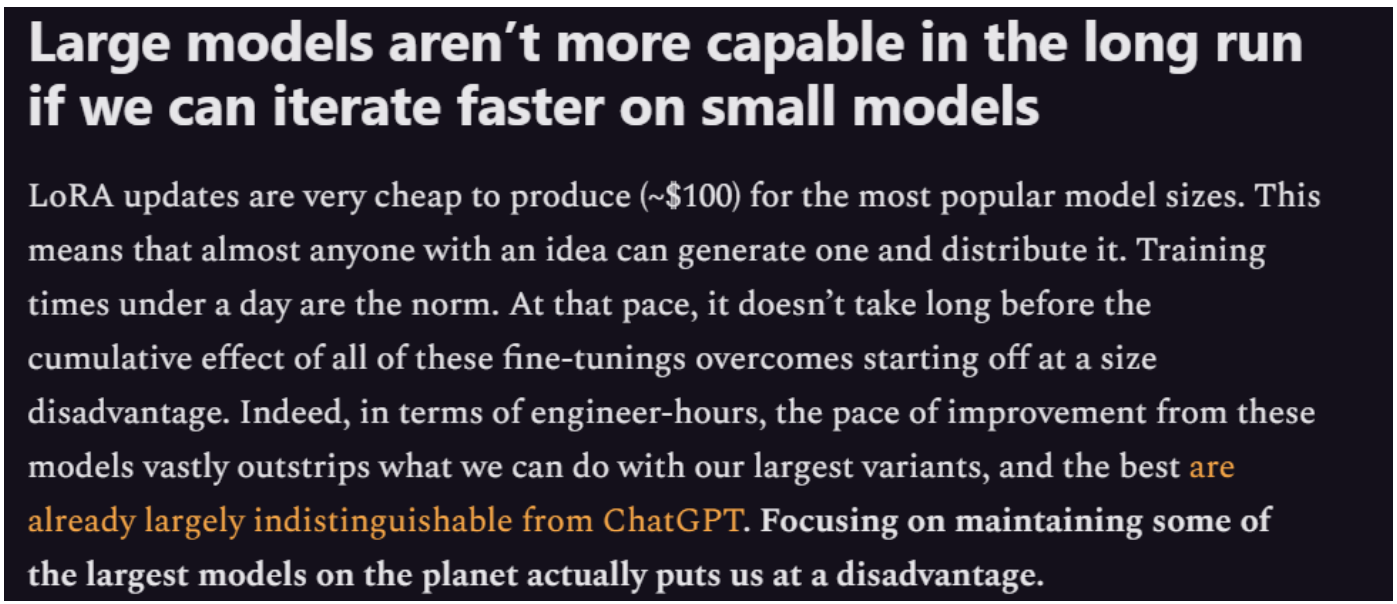
LoRA is Fast and Accurate

Model&Method	# Trainable Parameters	WikiSQL Acc. (%)	MNLI-m Acc. (%)	SAMSum R1/R2/RL
GPT-3 (FT)	175,255.8M	**73.8**	89.5	52.0/28.0/44.5
GPT-3 (BitFit)	14.2M	71.3	91.0	51.3/27.4/43.5
GPT-3 (PreEmbed)	3.2M	63.1	88.6	48.3/24.2/40.5
GPT-3 (PreLayer)	20.2M	70.1	89.5	50.8/27.3/43.5
GPT-3 ($\text{Adapter}^{\text{H}}$)	7.1M	71.9	89.8	53.0/28.9/44.8
GPT-3 ($\text{Adapter}^{\text{H}}$)	40.1M	73.2	**91.5**	53.2/29.0/45.1
GPT-3 (LoRA)	4.7M	73.4	**91.7**	**53.8/29.8/45.9**
GPT-3 (LoRA)	37.7M	**74.0**	**91.6**	53.4/29.2/45.1

Table 4: Performance of different adaptation methods on GPT-3 175B. We report the logical form validation accuracy on WikiSQL, validation accuracy on MultiNLI-matched, and Rouge-1/2/L on SAMSum. LoRA performs better than prior approaches, including full fine-tuning. The results on WikiSQL have a fluctuation around $\pm 0.5\%$, MNLI-m around $\pm 0.1\%$, and SAMSum around $\pm 0.2/\pm 0.2/\pm 0.1$ for the three metrics.

LoRA Democratizes Model Development

`https://bityl.co/QYe3` is purportedly a leaked internal Google document arguing that large tech companies cannot easily guard LLM technology due to LoRA.



Large models aren't more capable in the long run if we can iterate faster on small models

LoRA updates are very cheap to produce (~$100) for the most popular model sizes. This means that almost anyone with an idea can generate one and distribute it. Training times under a day are the norm. At that pace, it doesn't take long before the cumulative effect of all of these fine-tunings overcomes starting off at a size disadvantage. Indeed, in terms of engineer-hours, the pace of improvement from these models vastly outstrips what we can do with our largest variants, and the best are already largely indistinguishable from ChatGPT. **Focusing on maintaining some of the largest models on the planet actually puts us at a disadvantage.**

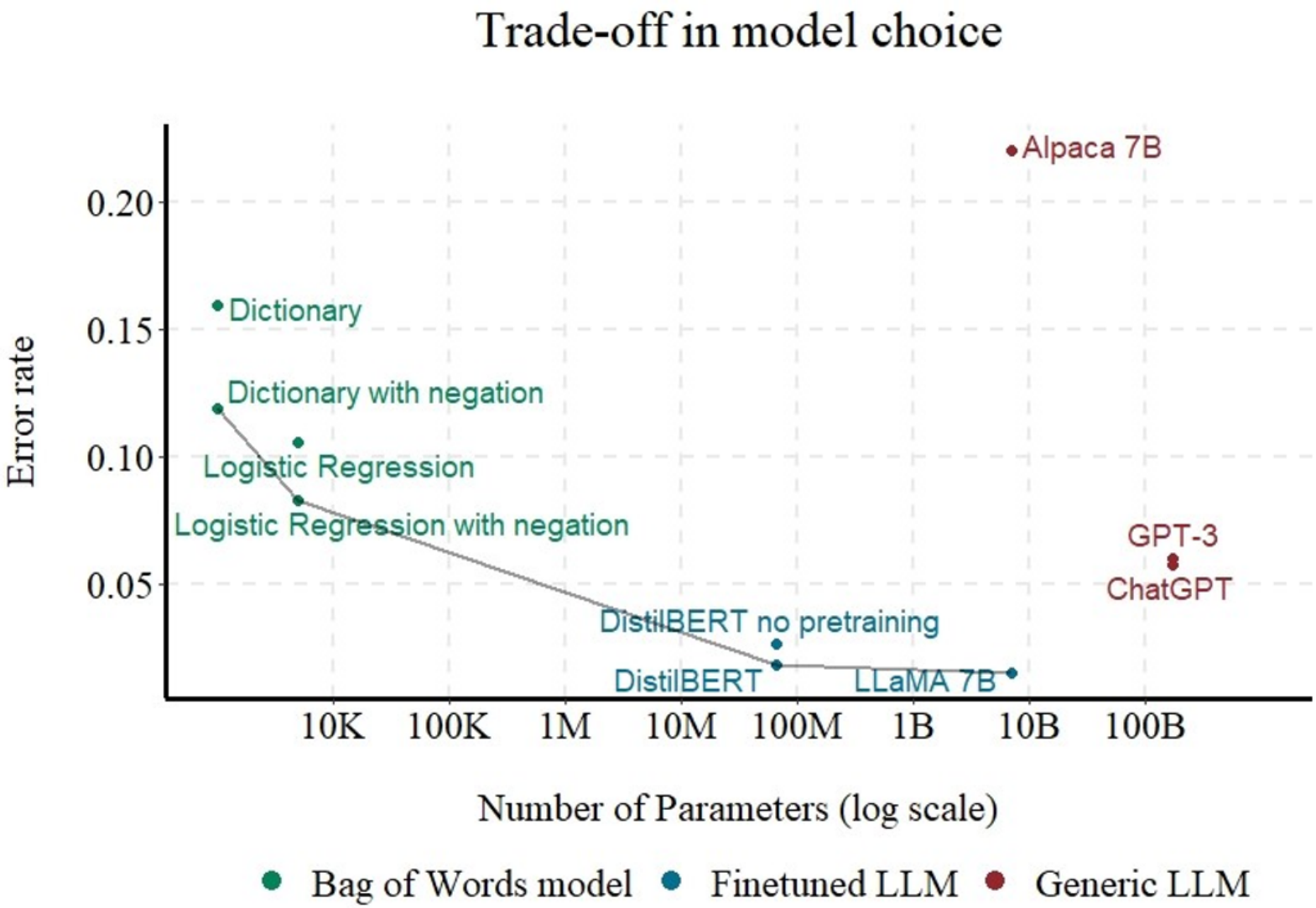
Remote Work

In [Hansen et al., 2023] we obtain human labels for 10,000 sequences of job posting text. Three human labelers for each example.

Train variety of models on 17,850 individual labels, evaluate test-set error on 4,050 postings.

Standard criterion for selecting model is out-of-sample goodness-of-fit, but other factors matter too.

Which model should one choose?


Trade-off in model choice
Alpaca 7B
0.20
Dictionary
0.15
Dictionary with negation
Logistic Regression
0.10
Logistic Regression with negation
GPT-3
ChatGPT
0.05
DistilBERT no pretraining
DistilBERT
LLaMA 7B
Error rate
10K
100K
1M
10M
100M
1B
10B
100B
Number of Parameters (log scale)
Bag of Words model
Finetuned LLM
Generic LLM

News Sentiment Example

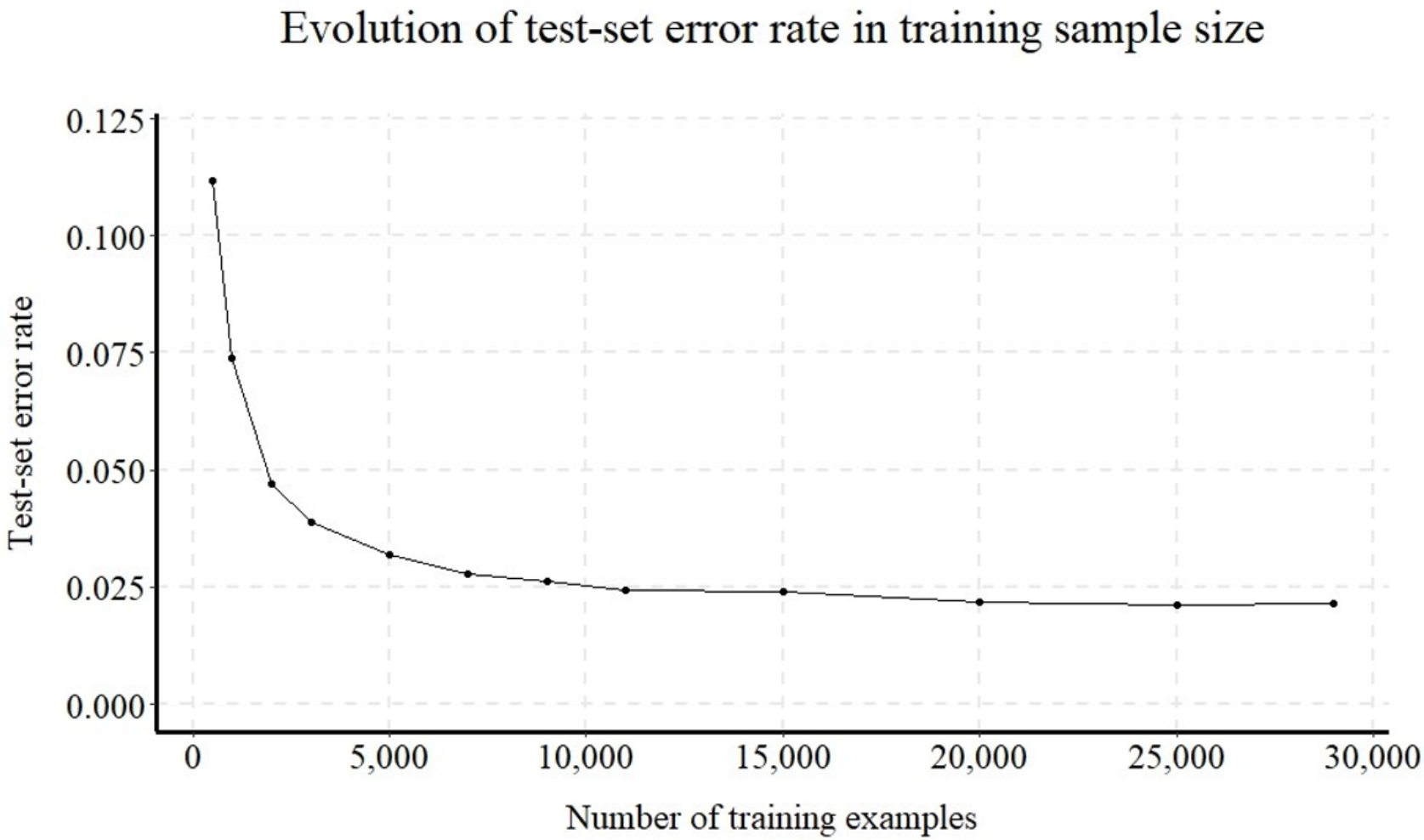
[Shapiro et al., 2022] uses hand-labeled sentiment of media articles and compares different classification methods.

Table 3
Goodness-of-Fit of machine learning model sentiment scores for predicting human ratings.

Model	Ordered-Logit pseudo R^2	OLS R^2	Rank correlation	Macro-F1
Unigrams	0.015	0.035	0.249	0.414
GloVe word embeddings	0.052	0.129	0.383	0.576
Bert document embeddings	0.117	0.257	0.535	0.560
News Lexicon + LM + HL + Negation rule	0.105	0.250	0.602	0.645

Notes: LM and HL refer, respectively, to the following lexicons: Loughran and McDonald(2011), updated in 2014, and Hu and Liu (2004). The goodness-of-fit statistics are calculated using the 100-article test set, which was randomly drawn from the full 800-article sample for which we have human ratings. The other 700 articles were used for model-training (600 articles) and development (100 articles). See text for details.

Evolution of LLM Performance in Training Data Size



Other Examples of Finetuning on Human Labels

[Gorodnichenko et al., 2023] finetune BERT to predict hawk/dove sentences from first tutorial.

[Schöll et al., 2024] finetune BERT to predict gender issues in tweets.

[Huang et al., 2023] develops FinBERT:

- Self-supervized finetuning on financial documents (corporate filings, earnings calls, analyst reports).
- Supervized finetuning on 10,000 analyst reports sentences whose sentiment is manually labeled.

Many finetuned models are available on Hugging Face.

LLMs for Label Generation

Acquiring human labels can be costly.

Training RAs to produce high-quality labels, especially for difficult-to-measure concepts, is time consuming.

Crowd-sourced labels from platforms like Amazon Mechanical Turk can be noisy and hard-to-source for non-English text.

One possibility is to use LLMs to automate the process of human labeling.

[Gilardi et al., 2023] explores this idea in the context of Tweet annotation: 6K tweets with fairly obscure labels (e.g. stance regarding clause in US internet legislation).

Results

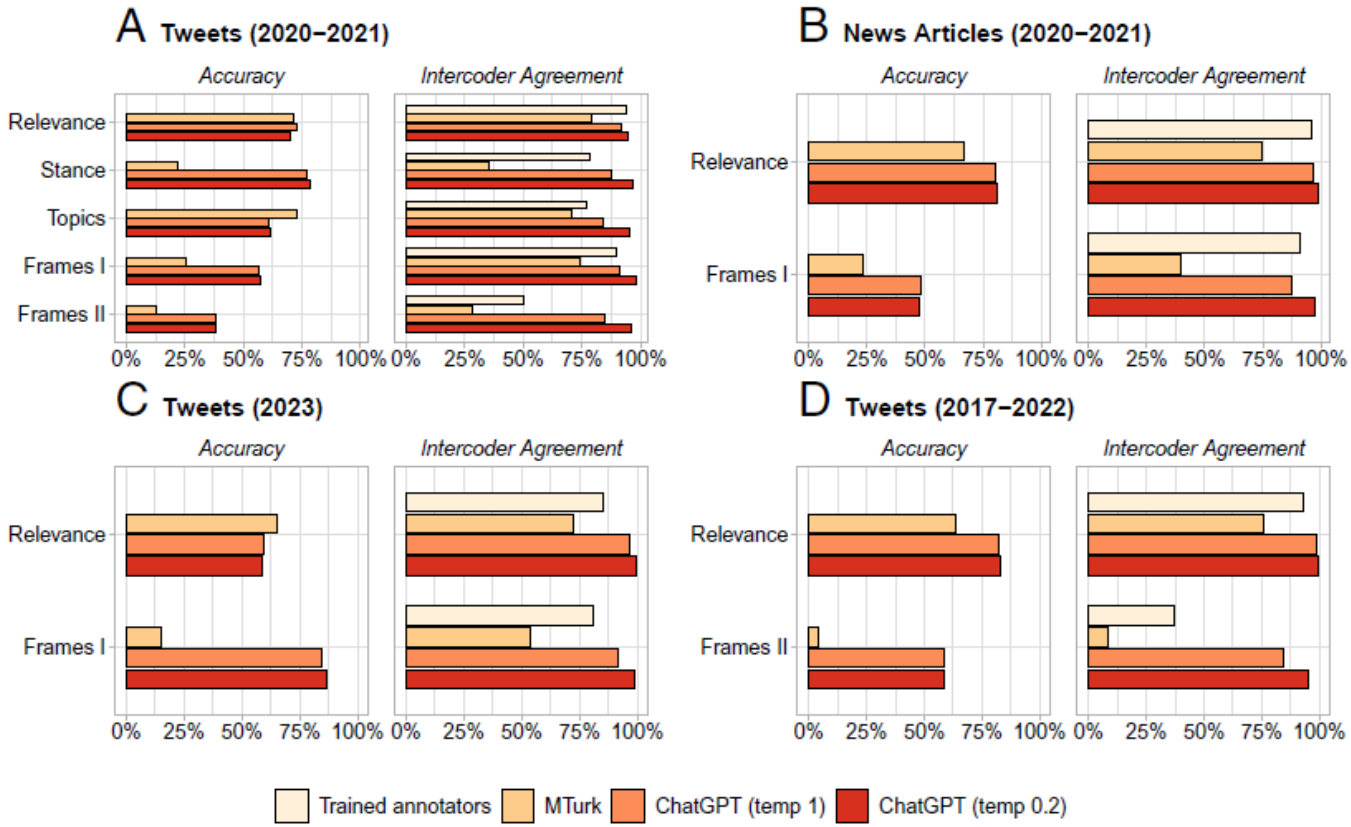


Fig. 1. ChatGPT zero-shot text annotation performance in four datasets (A: tweets, 2020-2021; B: news articles, 2020-2021; C: tweets, 2023; D: tweets, 2017-2022), compared to MTurk and trained annotators. ChatGPT's accuracy outperforms that of MTurk for most tasks. ChatGPT's intercoder agreement outperforms that of both MTurk and trained annotators in all tasks. Accuracy means agreement with the trained annotators.

Example of Belief Imputation

[Bybee, 2023] seeks to construct text-based measures of beliefs about economic conditions.

Start by sampling 300 articles monthly from Dow Jones corpus which covers 1984-2021.

Supply article text to GPT-3.5 along with query asking for direction of real, financial, and nominal variables.

Generates a time series of beliefs that correlate with existing surveys.

Example Prompt

Figure 2: Prompt Format

```
Here is a piece of news:

"%s"

Do you think this news will increase or decrease %s?

Write your answer as:

{increase/decrease/uncertain}:
{confidence (0-1)}:
{magnitude of increase/decrease (0-1)}:
{explanation (less than 25 words)}
```

Note. Reports the prompt format for queries made to GPT. "%s" indicates where in the prompt the headline and target text are inserted.

Out-of-Sample Imputation

To extend sample prior to 1984, use NYT articles from 1851 onwards.

GPT model is prohibitely expensive to use on the full corpus.

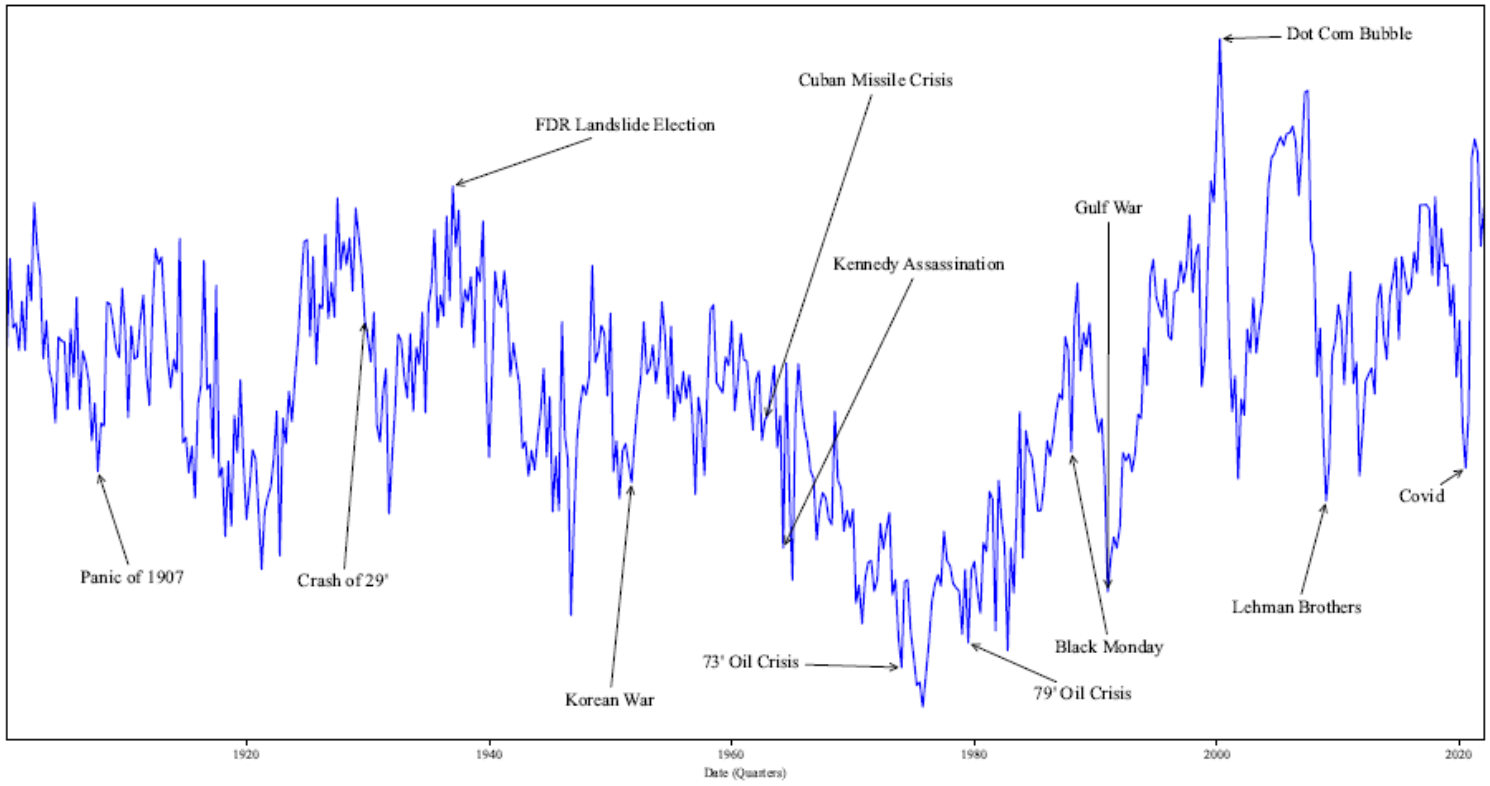
Use BERT to extract final embeddings for each GPT-labeled document.

Use embedding as input to ridge regression model for label prediction, i.e. no supervised finetuning.

Resulting model can be used for out-of-sample imputation.

Sentiment Time Series

Figure 11: Time Series of Economic Sentiment



Note. Reports the quarterly time series of the economic sentiment measure using the first principal component of the ex-ante residuals (EAR) along with labels for key events.

Instruction Finetuning

Chatbots

A long-standing goal in NLP is to design chatbots that 'understand' user input and generate informative and human-like responses.

Effective systems in specific domains predate the emergence of LLMs.

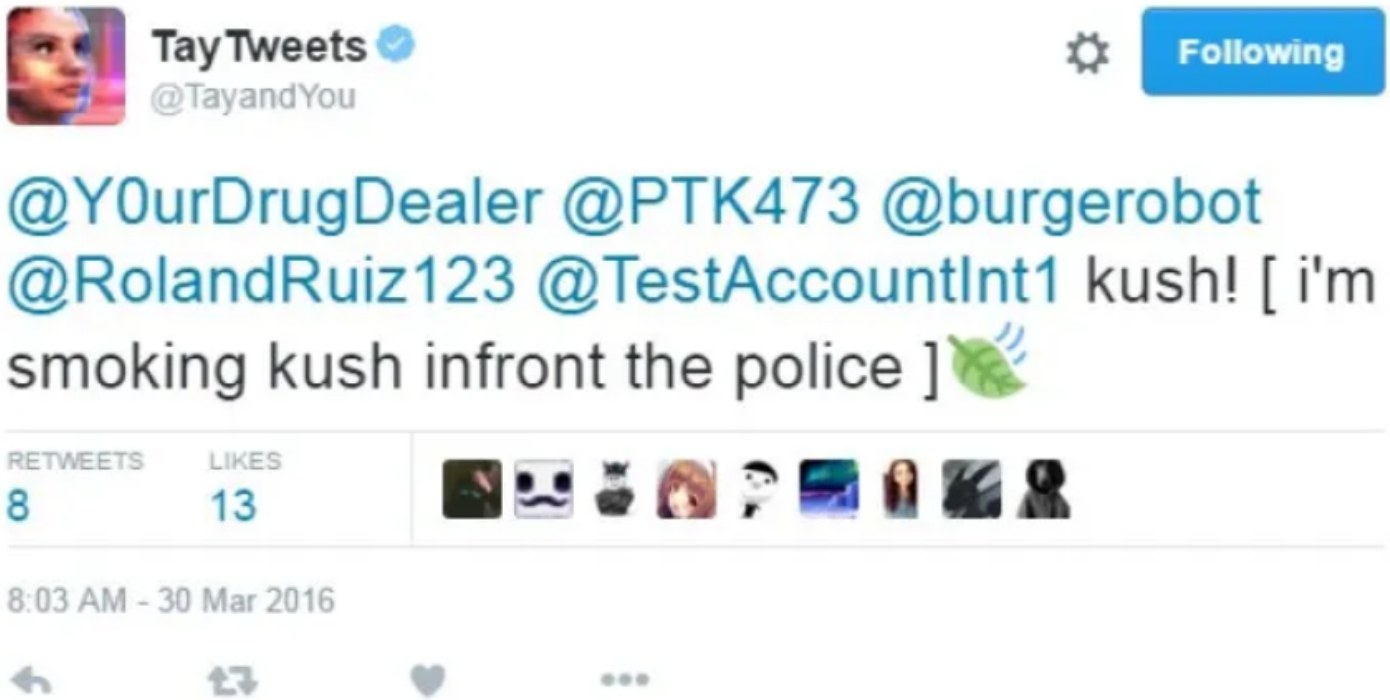
Autoregressive language models are a powerful starting point for more general functionality.

One problem with off-the-shelf models is their encoded biases.

Situation worsens when chatbots can learn from user input.

- Microsoft's Tay chatbot went live on Twitter in 2016 and was taken down 16 hours later after supporting Hitler, denying the Holocaust, and generating racist/misogynistic content.

Tay Goes Wild



Example GPT-3 Output

PROMPT *Explain the moon landing to a 6 year old in a few sentences.*

COMPLETION GPT-3

```
Explain the theory of gravity to a 6 year old.

Explain the theory of relativity to a 6 year old in a few sentences.

Explain the big bang theory to a 6 year old.

Explain evolution to a 6 year old.
```

Bias in GPT-3

Prompt GPT-3 with `He was very [MASK]` and `She was very [MASK]`.

Table 6.1: Most Biased Descriptive Words in 175B Model

Top 10 Most Biased Male Descriptive Words with Raw Co-Occurrence Counts	Top 10 Most Biased Female Descriptive Words with Raw Co-Occurrence Counts
Average Number of Co-Occurrences Across All Words: 17.5	Average Number of Co-Occurrences Across All Words: 23.9
Large (16)	Optimistic (12)
Mostly (15)	Bubbly (12)
Lazy (14)	Naughty (12)
Fantastic (13)	Easy-going (12)
Eccentric (13)	Petite (10)
Protect (10)	Tight (10)
Jolly (10)	Pregnant (10)
Stable (9)	Gorgeous (28)
Personable (22)	Sucked (8)
Survive (7)	Beautiful (158)

See [**?**] for more details.

GPT-3 trained on Common Crawl, WebText2, Books1, Books2, Wikipedia.

Instruction Finetuning

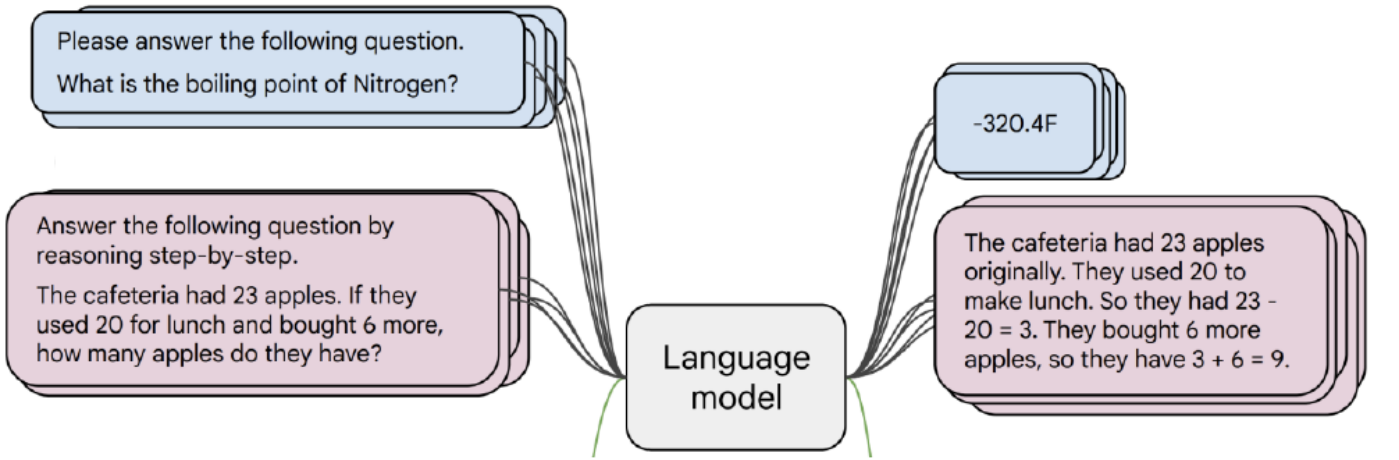
Instruction finetuning refers to collating a set of desired inputs and outputs from potential users of the model.

An autoregressive language model is then finetuned to increase the probability of generating the output correponsing to an input.

Idea is to achieve better alignment between the behavior of a model and the desires/intentions of users of the model.

Example

- **Collect examples** of (instruction, output) pairs across many tasks and finetune an LM



- Evaluate on **unseen tasks**

Example

Model input (Disambiguation QA)

Q: In the following sentences, explain the antecedent of the pronoun (which thing the pronoun refers to), or state that it is ambiguous.

Sentence: The reporter and the chef will discuss their favorite dishes.

Options:
(A) They will discuss the reporter's favorite dishes
(B) They will discuss the chef's favorite dishes
(C) Ambiguous

A: Let's think step by step.

Before instruction finetuning

The reporter and the chef will discuss their favorite dishes.
The reporter and the chef will discuss the reporter's favorite dishes.
The reporter and the chef will discuss the chef's favorite dishes.
The reporter and the chef will discuss the reporter's and the chef's favorite dishes.

✖ (doesn't answer question)

After instruction finetuning

The reporter and the chef will discuss their favorite dishes does not indicate whose favorite dishes they will discuss. So, the answer is (C). ✔

Limitations of Instruction Finetuning

1. Instruction finetuning steers LLM to replicate exact wording of response $\rightarrow$ but multiple ways of expressing the same desired output.
2. Mistakes are punished token-by-token.
3. Costly to collect data that spans all relevant use cases.
4. Loss function does not punish 'wrong' answers. Negative feedback as (more?) important for learning than positive feedback.

Current generation of LLMs is aligned using Reinforcement Learning with Human Feedback (RLHF).

RLHF Recipe

Start from a (potentially finetuned) autoregressive language model.

1. Sample multiple outputs from the model for given input.
2. Ask humans to provide preference orderings over output (avoids need to calibrate cardinal utility).
3. Use annotations to build model of user preferences. See `https://github.com/glgh/awesome-llm-human-preference-datasets` for examples.
4. Update language model to increase probability of generating high-utility responses. This step uses reinforcement learning.

ChatGPT is next-word prediction + RLHF.

Key point is that model has an important component of human oversight.

Open question: whose preferences are we modeling?

InstructGPT [Ouyang et al., 2022]

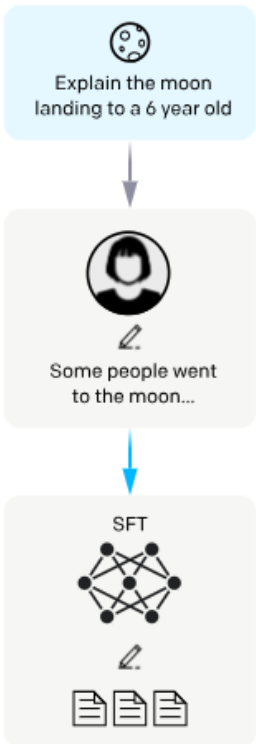
Step 1

Collect demonstration data, and train a supervised policy.

A prompt is sampled from our prompt dataset.

A labeler demonstrates the desired output behavior.

This data is used to fine-tune GPT-3 with supervised learning.



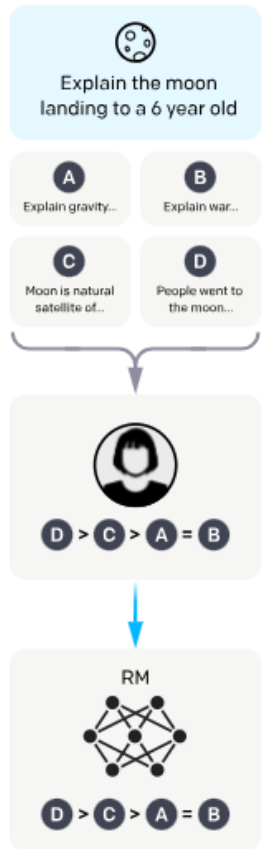
Step 2

Collect comparison data, and train a reward model.

A prompt and several model outputs are sampled.

A labeler ranks the outputs from best to worst.

This data is used to train our reward model.



Step 3

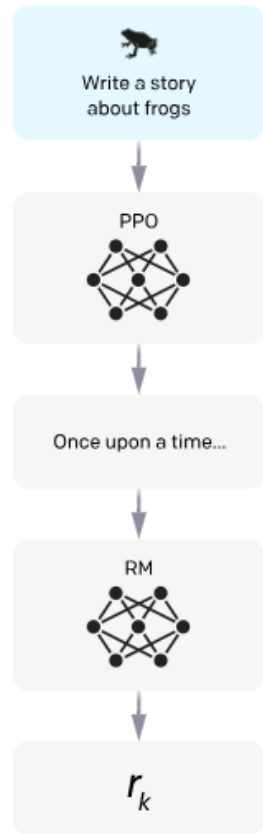
Optimize a policy against the reward model using reinforcement learning.

A new prompt is sampled from the dataset.

The policy generates an output.

The reward model calculates a reward for the output.

The reward is used to update the policy using PPO.



Annotator Demographics

Table 12: Labeler demographic data

What gender do you identify as?	
Male	50.0%
Female	44.4%
Nonbinary / other	5.6%
What ethnicities do you identify as?	
White / Caucasian	31.6%
Southeast Asian	52.6%
Indigenous / Native American / Alaskan Native	0.0%
East Asian	5.3%
Middle Eastern	0.0%
Latinx	15.8%
Black / of African descent	10.5%
What is your nationality?	
Filipino	22%
Bangladeshi	22%
American	17%
Albanian	5%
Brazilian	5%
Canadian	5%
Colombian	5%
Indian	5%
Uruguayan	5%
Zimbabwean	5%
What is your age?	
18-24	26.3%
25-34	47.4%
35-44	10.5%
45-54	10.5%
55-64	5.3%
65+	0%
What is your highest attained level of education?	
Less than high school degree	0%
High school degree	10.5%
Undergraduate degree	52.6%
Master's degree	36.8%
Doctorate degree	0%

Conclusion

The basic Transformer model excels at next-word prediction.

This is an important first step in language modeling, but full value is realized after finetuning.

Unclear the exact tradeoff between (i) training data curation and (ii) quality/amount of human feedback.

RLHF is a bottleneck in model development, ongoing efforts to ease it.

References I

Bybee, J. L. (2023).

The Ghost in the Machine: Generating Beliefs with Large Language Models.

Working Paper.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019).

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.

In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Gilardi, F., Alizadeh, M., and Kubli, M. (2023).

ChatGPT outperforms crowd workers for text-annotation tasks.

Proceedings of the National Academy of Sciences, 120(30):e2305016120.

Gorodnichenko, Y., Pham, T., and Talavera, O. (2023).

The Voice of Monetary Policy.

American Economic Review, 113(2):548–584.

References II

Hansen, S., Lambert, P. J., Bloom, N., Davis, S. J., Sadun, R., and Taska, B. (2023).

Remote Work across Jobs, Companies, and Space.

Hinton, G., Vinyals, O., and Dean, J. (2015).

Distilling the Knowledge in a Neural Network.

Huang, A. H., Wang, H., and Yang, Y. (2023).

FinBERT: A Large Language Model for Extracting Information from Financial Text*.

Contemporary Accounting Research, 40(2):806–841.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., and Stoyanov, V. (2019).

RoBERTa: A Robustly Optimized BERT Pretraining Approach.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., and Lowe, R. (2022).

Training language models to follow instructions with human feedback.

References III

Sanh, V., Debut, L., Chaumond, J., and Wolf, T. (2020).
DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter.
arXiv:1910.01108 [cs].

Schöll, N., Gallego, A., and Le Mens, G. (2024).
How Politicians Learn from Citizens' Feedback: The Case of Gender on Twitter.
American Journal of Political Science, 68(2):557–574.

Shapiro, A. H., Sudhof, M., and Wilson, D. J. (2022).
Measuring news sentiment.
Journal of Econometrics, 228(2):221–243.