

alphaXiv

“Comments on arXiv”

What is alphaXiv?

[CS.AI] 31 Jul 2024

alphaXiv

Comment Anonymous Private Note 31 Search

Bookmark PDF Claim Authorship

Authors: Laurens van der Maaten
Published: 15 Aug 2024 | 3 versions | 8,235 views
transformers multi-modal-learning multi-task-learning vision-language-models

The Llama 3 Herd of Models
Llama Team, AI @ Meta¹
¹A detailed contributor list can be found in the appendix of this paper.

Modern artificial intelligence (AI) systems are powered by foundation models. This paper presents a new set of foundation models, called Llama 3. It is a herd of language models that natively support multilinguality, coding, reasoning, and tool usage. Our largest model is a dense Transformer with 405B parameters and a context window of up to 128K tokens. This paper presents an extensive empirical evaluation of Llama 3. We find that Llama 3 delivers comparable quality to leading language models such as GPT-4 on a plethora of tasks. We publicly release Llama 3, including pre-trained and post-trained versions of the 405B parameter language model and our Llama Guard 3 model for input and output safety. The paper also presents the results of experiments in which we integrate image, video, and speech capabilities into Llama 3 via a compositional approach. We observe this approach performs competitively with the state-of-the-art on image, video, and speech recognition tasks. The resulting models are not yet being broadly released as they are still under development.

Date: July 23, 2024
Website: <https://llama.meta.com/>

1 Response Share Highlight

Laurens van der Maaten • 6 months ago
Thanks for asking this question! We have seen a lot of confusion about this online, so the phrasing in the paper must somehow be confusing.

The 8B and 70B models are not distilled from the 405B model. However, we do use post-training data created using the 405B model in the post-training of the 8B and 70B models. So we did use the 405B model to improve the 8B and 70B models.

The blog post states that the community can use the 405B model to do their own model distillation experiments. This is enabled by a change in the license of the Llama 3.1 models, which now allows developers to use outputs of Llama models to train their own models.

Add a response

Bibliographic Tools Code, Data, Media Demos Related Papers About arXivLabs

Code, Data and Media Associated with this Article
☒ alphaXiv (What is alphaXiv?)
☐ CatalyzeX Code Finder for Papers (What is CatalyzeX?)
☐ DagsHub (What is DagsHub?)

Papers Comments People

Search for topics, research ideas, or papers...

Hot

3,843 Foundations of Large Language Models
transformers representation-learning text-generation
This comprehensive book from researchers at Northeastern University's NLP Lab and NuTrans Research provides a foundational overview of Large Language Models (LLMs), focusing on core concepts and technical approaches rather than cutting-edge methods.
Bookmark Like 157

222 UI-TARS: Pioneering Automated GUI Interaction with Native Agents
human-ai-interaction imitation-learning multi-task-learning
This paper introduces UI-TARS, a native GUI agent model that perceives screenshots to perform human-like interactions without relying on modular frameworks or external APIs.
Bookmark Like 9

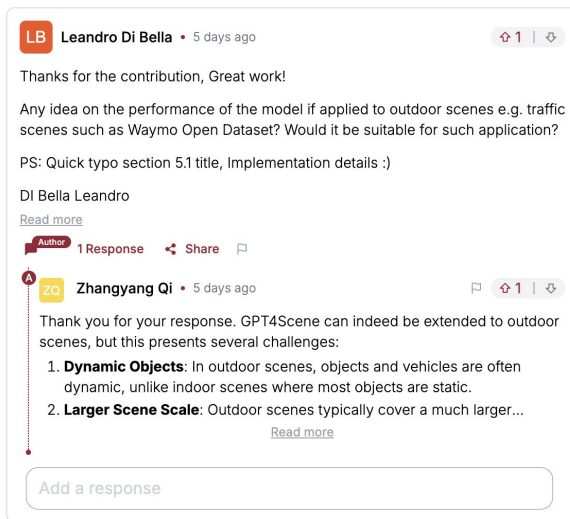
214 Test-time regression: a unifying framework for designing sequence models with associative memory
sequence-modeling attention-mechanisms representation-learning
This paper explains how sequence models with associative memory can be understood as performing regression at test time, unifying many recent architectures under a common framework.
Bookmark Like 13

102 DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning
deep-reinforcement-learning model-interpretation neural-coding

Inception

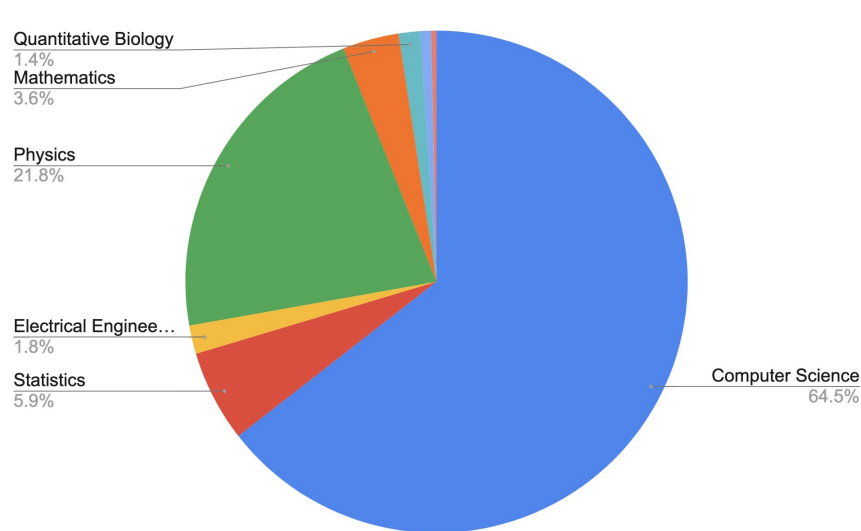
Undergraduates in research labs

- No idea what was going on in research papers
- pre-GPT
- “Stackoverflow for arXiv”




Inception

- Initially used as private discussion layer used at a couple classes and labs
- Around a year ago decided to open it up publicly
 - Since then ~5,000 comments 12,000 accounts
 - Interest across CS and Physics



Other arXiv Interfaces

 SciRate

Sign inSign up

Search

[arXiv.org](#)

Top arXiv papers

[sign in to customize](#)

Quantum computational advantage of noisy boson sampling with partially distinguishable photons

Byeongseon Go, Changhun Oh, Hyunseok Jeong

Jan 24 2025 quant-ph arXiv:2501.13433v1

Boson sampling stands out as a promising approach toward experimental demonstration of quantum computational advantage. However, the presence of physical noise in near-term experiments hinders the realization of the quantum computational advantage with boson sampling. Since physical noise in near-term boson sampling devices is inevitable, precise characterization of the boundary of noise rates where the classical intractability of boson sampling is maintained is crucial for quantum computational advantage using near-term devices. In this work, we identify the level of partial distinguishability noise that upholds the classical intractability of boson sampling. We find that boson sampling with on average $O(\log N)$ number of distinguishable photons out of N input photons maintains the equivalent complexity to the ideal boson sampling case. By providing strong complexity theoretical evidence of the classical intractability of noisy boson sampling, we expect that our findings will enable one to ultimately demonstrate quantum computational advantage with noisy boson sampling experiments in the near future.

Scite!5PDF

Date Published

Fri 24 Jan 2025 UTC

Prev day	Next day
1d	3d1w1m6m

Recent comments

Myths around quantum computation before full fault tolerance: What no-go ...
Zoe Holmes 8 days ago

Hey! Thanks for engaging!
Let me jump in on the fun.
In terms of Chae-Yeun's point 1 and Alex's point, I fully agree that "in a sense" is doing a lot of work but the aim of that paragraph was in a large part to point out the positive side of being able to often classically simulate/surrogate

Exponentially slow thermalization in 1D fragmented dynamics

Cheng Wang, Shankar Balasubramanian, Yiqiu Han, Ethan Lake, Xiao Chen, Zhi-Cheng Yang

Jan 24 2025 quant-ph cond-mat.stat-mech cond-mat.str-el math.GR arXiv:2501.13930v1

We investigate the thermalization dynamics of 1D systems with local constraints coupled to an infinite temper...

Scite!3PDF



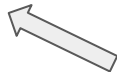
PUBPEER

The online journal club

Within CS...



**Open
Review
.net**



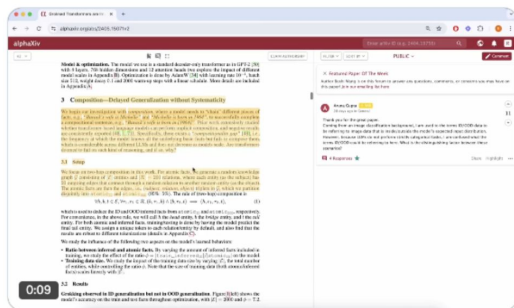
Open
Review
.net



hampton — e/acc
@hamptonism

Subscribe

Students at Stanford have built alphaXiv, an open discussion forum for arXiv papers. Think of it as X for research.



From **Stanford AI Lab**

4:31 PM · Sep 7, 2024 · 638.9K Views

94

1.5K

9.7K

5.5K



jack
@jack



alphaxiv.org
alphaXiv
Discuss, discover, and read arXiv papers.

3:12 AM · Sep 8, 2024 · 518.4K Views

121

361

2.1K

1K



Yann LeCun
@ylecun

Nice.

Stanford AI Lab · @StanfordAILab · Jul 31, 2024

arXiv -> alphaXiv

Students at Stanford have built alphaXiv, an open discussion forum for arXiv papers. @askalphaxiv

29

127

1.4K

233



Different subjects, different use cases



Jackson Burns • 1st

DOE CSGF Fellow in Chemical Engineering and Computation PhD at MIT
9mo •

My latest work in molecular property prediction achieves state-of-the-art results by taking a modern approach to an old process - so I am doing the same with publishing it!

Technological advances make the closed-door, journal-centered peer review approach nearly obsolete - its high-profile failures have also fueled skepticism among researchers and the public. For that reason my latest work is freely available online and I invite all readers to leave their review comments on alphaXiv, the new open-source peer review platform: <https://lnkd.in/gfGYGUHp>



Raymond Gasper • 9 months ago 5

So I like this as a benchmark of the general utility of using large sets of descriptors as input for statistical models, tho Pat's concerns about the quality of datasets and statistical rigor I think are super relevant here. In particular, random splitting of molecular datasets can often lead to an effective leakage where data points in the training and test set are extremely similar, resulting in misleading test set performance. I've highlighted that paragraph for this comment, but the rest...

[Read more](#)

Jackson Burns • 9 months ago 2

Thank you for the comments! Agreed RE dataset quality and statistical rigor.

I'm curious what splitting method(s) you would suggest for ensuring that data is not leaked from training to testing. We are familiar with some of the basics like scaffold or clustering on fingerprints, but are there others that you trust?...

[Read more](#)

[Click to View All](#)

Patrick Walters • 9 months ago 2

This paper has one of the best analyses I've seen. They include a heat map to show statistically significant differences. They also include a bolded table that highlights multiple entries if they are statistically equivalent.

<https://chemrxiv.org/engage/chemrxiv/article-details/65f6054e66c13817292312b3>

Raymond Gasper • 9 months ago 0

You're 100% right Patrick, more thorough analysis than just a table is definitely warranted. Thanks for linking the paper I hadn't seen that one before and yeah the tables/heatmaps and overall process are really impressively rigorous

[Add a response](#)



Jackson Burns • 1st

DOE CSGF Fellow in Chemical Engineering and Computation Ph...
4mo •

Thanks to [Raymond Gasper](#) and [Pat Walters](#) for their insightful comments (and their patience while I acted on them!) about my most recent work about fastprop for molecular property prediction. The attached figure tells the whole story - on this high quality ADME dataset, fastprop surpasses baseline methods and trades places with literature standard and SOTA approaches (see the code to reproducibly generate the figure here: <https://lnkd.in/gDfdzCvq>).

Different subjects, different use cases

ScienceDirect.com
https://www.sciencedirect.com › article › abs › pii

The unbearable slowness of being: Why do we live at 10 ...

by J Zheng · 2024 · Cited by 4 – Perspective. The unbearable slowness of being: Why do we live at 10 bits/s?

 The New York Times
<https://www.nytimes.com> › science › speed-of-thought

Human Thought Is Far Slower Than Your Internet Connection

Dec 28, 2024 – They titled their study, published this month in the journal *Neuron*, “The unbearable slowness of being.” It’s a bit of a counterweight to ...

Cell Press
https://www.cell.com › neuron › fulltext

The unbearable slowness of being: Why do we live at 10 ...

by J Zheng · Cited by 4 — The unbearable slowness of being: Why do we live at 10 bits/s? Jieyu Zheng.
Jieyu Zheng. Correspondence. Corresponding author. jzzheng@caltech.edu. Affiliations.


[illegible]

Acknowledgments

For discussions and comments on the early stage of the manuscript, we thank Frederick Eberhardt, Michelle Effros, Florian Engert, Mehrdad Jazayeri, Christof Koch, Ueli Rutishauser, and Anthony Zador. We also thank alphaXiv for providing a platform for open discussions on our preprint, where the comments received were intriguing and helpful for the revision. M.M. was supported by grants from the Simons Collaboration on the Global Brain (543015) and NIH (R01 NS111477).

Different subjects, different use cases

From Older Version


 **Rob** • 3 months ago 12 |

It's surprising that this paper should pass the check of **two Phys. Plasma reviewers!**


From the abstract, it's easy to realize that they did not mention the power to sustain hot ion mode or nonequilibrium distribution. Even a keyword search will reveal that they indeed ignored those terms. Now let's see whether these terms have been calculated previously, we found the corresponding author's book...Besides the hot ion mode discussed by the comment paper, in...

[Read more](#)

3 Responses Share


 **Yuanyuanzhang** • 2 months ago 7 |

Seriously, I begin to doubt the quality of the journal.

 **Yuanyuanzhang** • 2 months ago 6 |

Perhaps we can conceive other aproaches, neglecting the majority of input energy, and submit the roadmap to Phys. Plasma to see if it would be accepted. If they reject it, we could just refer to this paper.

(Edited)

 **Rob** • 2 months ago 5 |

That's probably what the fusion startups are doing exactly!

Different subjects, different use cases

AS

Aaron S • 29 days ago

3 |

"Futility" is far too strong of a term for the evidence presented. The evidence presented is mainly that a small number of H100s were used (at least 8); it's not clear if these are in China or rented. A100s and H20s could have come in before export controls (are H20s currently controlled officially? **Rumor** is that NVIDIA isn't selling them).

The arguments and evidence here point to conclusions like "the export controls happened too late..."

[Read more](#)

Author

3 Responses

Share

A

Ritwik Gupta

• 27 days ago

University of California, Berkeley

0 |

Aaron, your interpretation of the paper is incorrect. That is not what we are saying in the paper.

In this paper, we demonstrate that:

- Tencent is using NVIDIA **H20** GPUs effectively to train SOTA LLMs.
- They are building SOTA LLMs despite export controls not by using better or export controlled chips, but by using chips that are allowed for sale through the deployment of...

[Read more](#)

AS

Aaron S • 27 days ago

0 |

Thanks for clarifying. I continue to think this is a bad title.

The frontier of AI capabilities is constantly advancing, and whether Chinese firms can compete on this frontier in the future will depend on a combination of their software and hardware available, and how these relate to what's available in the US. On my read, all of the subsections on better software (MoE, mixed precision, quantization, vram, efficient networking) seem...

[Read more](#)

A

Ritwik Gupta

• 27 days ago

University of California, Berkeley

0 |


Aaron, the title of the paper is accurate. We show that software is a meaningful avenue by which non-controlled chips can be made useful to build frontier models. **Playing whack-a-chip with chips as they come up is a losing strategy.**





We are not saying export controls are bad. We are saying that this *specific* strategy of hardware export controls are built on flawed assumptions. The title explicitly refers to "hardware-centric"...

[Read more](#)

Different subjects, different use cases

07.21783v1 [cs.AI] 31 Jul 2024





>>

Comment

Anonymous

Private Note


31

Search

Bookmark


PDF

Claim Authorship

Authors:  Laurens van der Maaten

Published: 15 Aug 2024 | 3 versions | 8,235 views

transformers multi-modal-learning multi-task-learning vision-language-models



The Llama 3 Herd of Models

Llama Team, AI @ Meta¹
¹A detailed contributor list can be found in the appendix of this paper.


Modern artificial intelligence (AI) systems are powered by foundation models. This paper presents a new set of foundation models, called Llama 3. It is a herd of language models that natively support multilinguality, coding, reasoning, and tool usage. Our largest model is a dense Transformer with 405B parameters and a context window of up to 128K tokens. This paper presents an extensive empirical evaluation of Llama 3. We find that Llama 3 delivers comparable quality to leading language models such as GPT-4 on a plethora of tasks. We publicly release Llama 3, including pre-trained and post-trained versions of the 405B parameter language model and our Llama Guard 3 model for input and output safety. The paper also presents the results of experiments in which we integrate image, video, and speech capabilities into Llama 3 via a compositional approach. We observe this approach performs competitively with the state-of-the-art on image, video, and speech recognition tasks. The resulting models are not yet being broadly released as they are still under development.

Date: July 23, 2024
Website: <https://llama.meta.com/>

1 Introduction

Foundation models are general models of language, vision, speech, and/or other modalities that are designed to support a large variety of AI tasks. They form the basis of many modern AI systems.

The development of modern foundation models consists of two main stages: **(1)** a pre-training stage in which the model is trained at massive scale using straightforward tasks such as next-word prediction or captioning

 Cristian Gutiérrez Gómez • 6 months ago

19

About 'distillation'

The report references distillation as a method used in smaller models, citing the original Phi approach. However, reviewing the Phi papers (Textbooks are all you need) there is not any mention of traditional knowledge distillation... (Neither logit-level, token-level, or sequence-level).


From what I've seen, the Phi approach involved using GPT-3.5 and 4 to annotate and generate synth data. ...

Read more

1 Response

Share

Highlight

 Laurens van der Maaten • 6 months ago

10

Thanks for asking this question! We have seen a lot of confusion about this online, so the phrasing in the paper must somehow be confusing.

The 8B and 70B models are not distilled from the 405B model. However, we do use post-training data created using the 405B model in the post-training of the 8B and 70B models. So we did use the 405B model to improve the 8B and 70B models.

The blog post states that the community can use the 405B model to do their own model distillation experiments. This is enabled by a change in the license of the Llama 3.1 models, which now allows developers to use outputs of Llama models to train their own models.

Read less

Add a response

Road to Best Discussions

Visibility and author presence

- Authors share the link to their paper on alphaXiv
- A reader shares their comments on other platforms
- The paper is claimed on alphaXiv when a new reader visits

The image is a composite screenshot showing a Twitter thread and an alphaXiv paper page. The top part shows a tweet by Daniel Geng (@dangengdg) announcing a discussion about a paper on alphaXiv. Below it is a tweet by Konczer (@KonczerJ) commenting on the paper. The bottom part shows the alphaXiv paper page for 'Model Swarms' by Shangbin Feng, Yike Wang, and Zifeng Wang. The paper page includes a summary, authors, publication date, and a list of tags. A diagram of the 'Model Swarms' architecture is also visible.

Twitter Thread:

Daniel Geng @dangengdg
Hey all, I'll be answering questions about our "Motion Prompting" paper on alphaXiv (@askalphaxiv) (it's like arXiv, but adds a discussion section, and I think is quite well built!):
alphaxiv.org/abs/2412.02700...
9:37 AM · Jan 10, 2025 · 4,955 Views

Konczer @KonczerJ
I read Model Swarms by authors from @GoogleDeepMind & @UW and commented on @askalphaxiv: alphaxiv.org/abs/2410.11163
Fusing Neural Networks with Genetic algorithms is a promising direction, and this is a concrete implementation. I am curious what future work will bring.

alphaXiv Paper Page:

Model Swarms
Authors: Shangbin Feng, Yike Wang, Zifeng Wang
Published: 14 Oct 2024 | 1 version | 5,341 views

This paper introduces a collaborative search algorithm called MODELSWARMS that adapts Language Learning Models (LLMs) through swarm intelligence principles

Problem Method Results Takeaways Abstract

multi-agent-learning transfer-learning ensemble-methods collaborative-learning we

Diagram: The diagram illustrates the 'Model Swarms' architecture. It shows a central 'Model' box connected to a 'Swarm' box. The 'Swarm' box contains a network of nodes and edges, representing the collaborative search process. The diagram is divided into two main sections: 'Update' and 'Weight Update'. The 'Update' section shows a sequence of nodes and edges, with arrows indicating the flow of information. The 'Weight Update' section shows a sequence of nodes and edges, with arrows indicating the flow of information. The diagram is labeled with 'g' and 'g_u'.

Categorizing Comments

- *technical-clarification* – Questions seeking to understand specific technical details or procedures in the paper.
- *conceptual-discussion* – Exploration of theoretical ideas, implications, or abstract concepts from the paper.
- *error-report* – Pointing out specific mistakes, typos, or inconsistencies in the paper.
- *research-extension* – Suggestions or questions about potential future directions or expansions of the work.
- *methodology-question* – Questions about why certain research design choices were made.
- *literature-connection* – Comments relating the work to existing research or pointing out relevant papers.
- *praise* – comments highlighting specific strengths or contributions of the paper.
- *substantive-critique* – Detailed, evidence-based challenges to the paper's methods, results, or conclusions.
- *academic-dispute* – Contentious disagreements about attribution, priority, or academic integrity.

H

Helen • 19 days ago

2 | 0

When you use holographic principle on CFT2, does it naturally lead to AdS/CFT correspondence? Or instead this method apply to CFT2 states that have no corresponding AdS states?

Author

2 Responses

Share

A

H

hyang • 18 days ago

Sichuan University

2 | 0

AdS/CFT correspondence is a realization/example of holographic principle, but not completely covers holographic principle in our current understanding.

This paper was finished over 6 years ago. Though the approach is correct, we were discouraged by the referee reports to have it published. We publish it now because our recent papers solved the problem in a complete and ne...

[Read more](#)

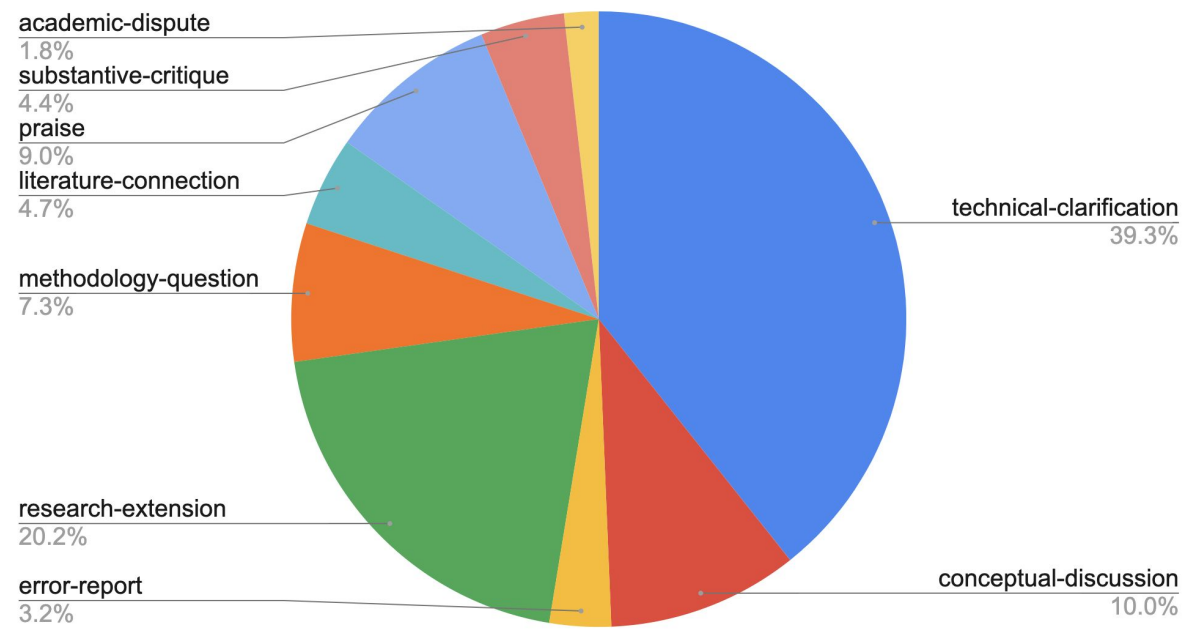
H

Helen • 18 days ago

1 | 0

Thanks! Will have a closer look.

Add a response



Moderation



na_cohomologist · 4mo ago

How are they going to moderate comments to filter out inappropriate things? I don't mean obviously inappropriate to random person on the street, but professionally inappropriate stuff that difficult to explicitly delineate, but stemming from bias, unconscious and otherwise?

↑ 42 ↓ Reply Award Share ...

+ 8 more replies



jieyou zhao · 3 months ago

Edit Delete 0 ↓

I honestly hope you could read the paper carefully and check the result on leaderborad.

Even a pupil could figure out the difference between L1 and L2 but you can't, which means our result is basically the same



jieyou zhao · 3 months ago

Edit Delete 5 ↓

Besides, about you are saying that we do not show your result just because your result is better than us.

I am happy to say that you could check LEGO(which is also an online method)'s result, It could achieve 80.75 by only using CasA detector. (We labeled it as Virconv incorrectly but it will be fix soon)

We DO NOT use your result is because your improvement on KITTI is really huge but you didn't show the same improvement on Waymo.

And considered you could not identify if CasTrack and VirconvTrack is online or offline (showed in your email), I highly suspect if your code is truly online.

As I said, show me the code, talk is cheap.

[Read less](#)



Mohamed Nagy Mostafa · 3 months ago

Edit Delete 0 ↓

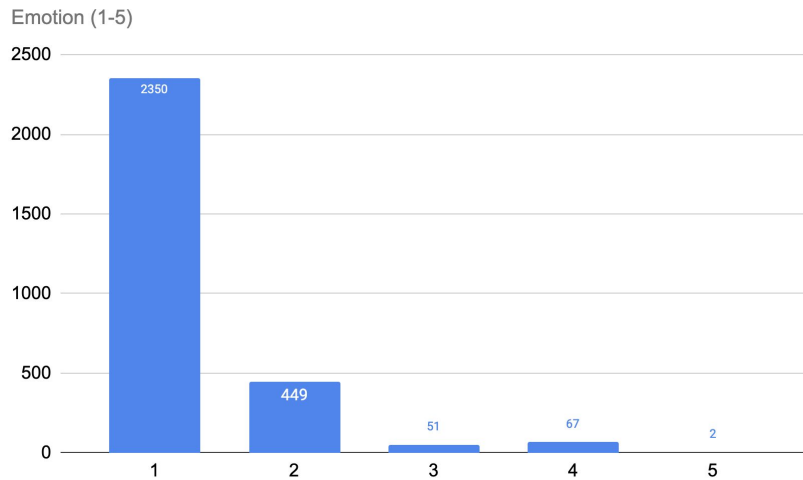
"You won't get away with this"

They didn't answer anything what I mentioned; instead, trying to change the topic as expected. I can answer you, but I know the purpose of your questions is to simply hide what I mentioned. So, let me copy/paste my comment above so everyone can see it. In summary, the authors manipulate the results of existing work and remove methods that outperform their work. **I will keep copying/pasting it as long as no response.**

"(First bullet point by @XiyangWang): After acceptance, the code of our method will be available on GitHub, regardless of the situation of this paper, **as per our input** to the journal we sent our work. However, let me response properly to this point:

I quote you: "The claimed online performance of non-open-source methods stems"; so you are saying your paper only considers open-source methods, and "somehow" you added it in V1 and removed in V2 because it is not open source! Good. **Let me list for you all the "non-open source" work you have in V1 in Table 1 and still there in V2.**

Anonymity



UA

User A5F • 1 hour ago

0 | 0

Question on contemporary application

Can SIMS be directly bolt on to Stable Diffusion fine-tuned models (or Flux and other open weight models)? What would be the expected challenges of adding a score function to a diverse set of models with slight differences in architecture?

[Add a response](#) [Share](#) [Flag](#)


Difficulty of browsing on arXiv

Computer Science

- [Computing Research Repository \(CoRR new, recent, search\)](#)
includes: (see [detailed description](#)): [Artificial Intelligence](#); [Computation and Language](#); [Computational Complexity](#); [Computational Engineering, Finance, and Science](#); [Computational Geometry](#); [Computer Science and Game Theory](#); [Computer Vision and Pattern Recognition](#); [Computers and Society](#); [Cryptography and Security](#); [Data Structures and Algorithms](#); [Databases](#); [Digital Libraries](#); [Discrete Mathematics](#); [Distributed, Parallel, and Cluster Computing](#); [Emerging Technologies](#); [Formal Languages and Automata Theory](#); [General Literature](#); [Graphics](#); [Hardware Architecture](#); [Human-Computer Interaction](#); [Information Retrieval](#); [Information Theory](#); [Logic in Computer Science](#); [Machine Learning](#); [Mathematical Software](#); [Multiagent Systems](#); [Multimedia](#); [Networking and Internet Architecture](#); [Neural and Evolutionary Computing](#); [Numerical Analysis](#); [Operating Systems](#); [Other Computer Science](#); [Performance](#); [Programming Languages](#); [Robotics](#); [Social and Information Networks](#); [Software Engineering](#); [Sound](#); [Symbolic Computation](#); [Systems and Control](#)

alphaXiv Browser Extension

- Offers a great sample of papers people are reading.
- Many people seem to want to know what papers their peers are reading or discussing.
- This data could be helpful!

 **alphaXiv - Open Research Discussion** [Remove from Chrome](#)

5.0 ★ (9 ratings)

Extension Tools 4,000 users

view license

alphaXiv
[View on alphaXiv](#)

👍 Like

🔖 Bookmark

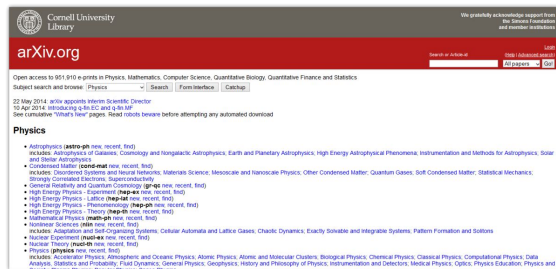
Overview

The official Chrome Extension for alphaXiv, the site for open research discussion directly on top of arXiv.

With this extension, you'll never miss out on any alphaXiv discussion! We'll gently let you know whenever an arXiv paper you're currently reading has comments on alphaXiv.

Join the (research) discussion!

Custom Topical Search



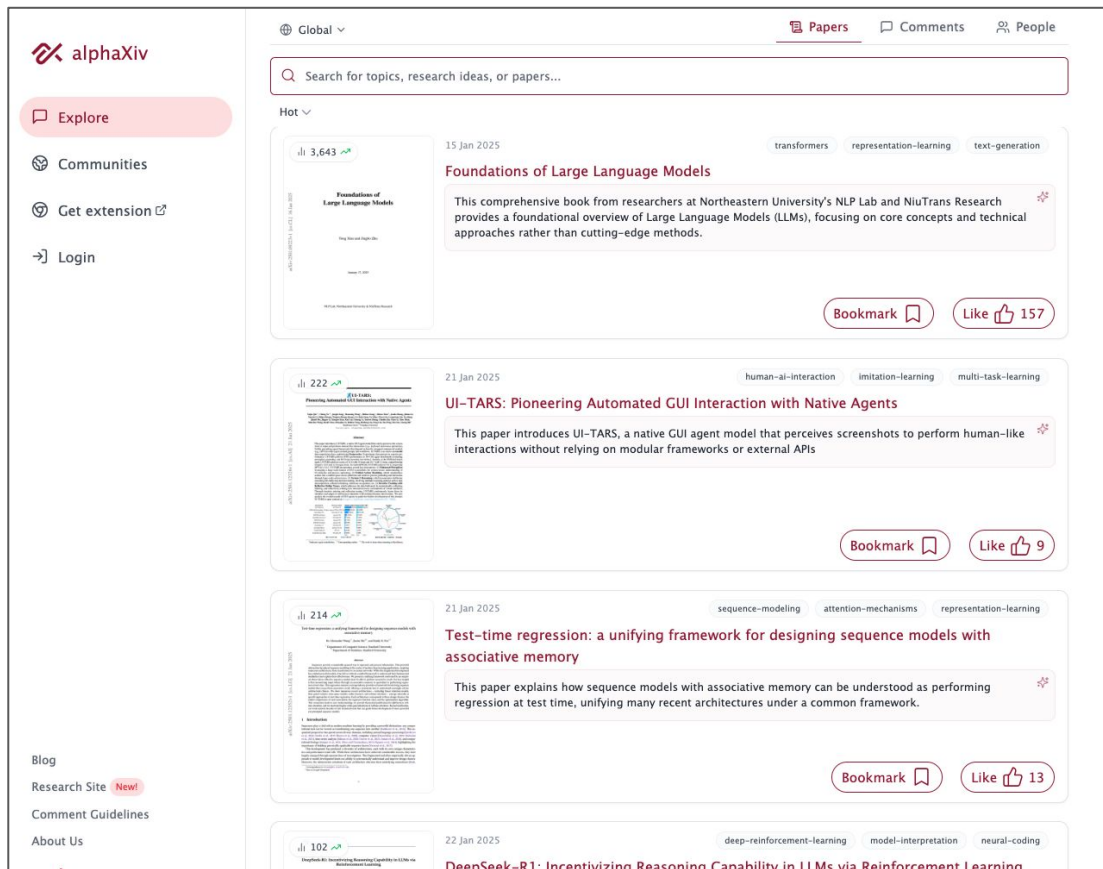
human-ai-interaction

multi-modal-learning

-
-
-

deep-reinforcement-learning

arXiv Sanity Reimagined



ResearchGate: An attempt at social media for research?

- Is alphaXiv social media?
 - **No.** not quite
 - Platform is built around papers, not people
- There are other examples of academic social networks
 - I.e. ResearchGate
 - For CS researchers it's just Twitter



Kevin Joey Chen  @kevخالchemy · Dec 10, 2024

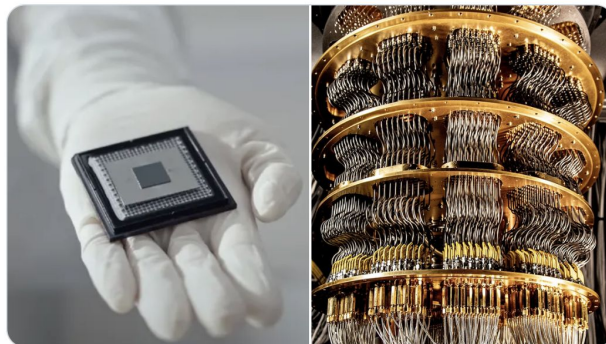
While everyone watches AI, **Google** broke time itself.



Their quantum chip just solved a problem that would outlast the universe
— in 5 minutes.

And this is just the start...

What you need to know about **Willow** — and how it'll transform humanity in 2030+: 



1.5K

7.1K

51K

9.7M




Future Directions

Communities


Explore Communities

Discover new communities in your field – [how it works](#)




Agents


Created by a moderator

 +270 more

Agent-based systems, multi-agent learning, model swarms


Credits offered by  **akash** together.ai

[Join](#)




Vision-Language Models

Created by StephenQS

 +208 more


Multimodal LLMs, VLMs, Video Foundation Models

[Join](#)




Interpretability, AI Safety & Explainability

Created by Varshini Subhash

 +132 more


- Understanding the behavior of ML models by interpreting and explaining them.
- Developing techniques for the safe usage and deployment of large models.

[Join](#)




High Energy Theory

Created by Helen

 +95 more


QFT, quantum gravity, string theory, particle physics etc

[Join](#)




Plasma Physics

Created by CHAO




[Join](#)




Cosmology

Created by Misael Arnaldo Espinal Valladares



[Join](#)

Communities




Agents

Created by Lino Le Van · 274 members

Request to Join

Agent-based systems, multi-agent learning, model swarms

\$20k credits pool for top contributors thanks to  together.ai


Papers

Comments

Members

30 Days

5,424



08 Jan 2025

deep-reinforcement-learning

self-supervised-learning

model-compression


rStar-Math: Small LLMs Can Master Math Reasoning with Self-Evolved Deep Thinking

This paper presents a self-evolving approach to enable small language models (SLMs) to achieve state-of-the-art math reasoning through Monte Carlo Tree Search-based deep thinking

Bookmark

Like 296

3,438




09 Jan 2025

information-extraction

knowledge-distillation

self-supervised-learning


Search-o1: Agentic Search-Enhanced Large Reasoning Models

 xiaoxi Li

This paper introduces a framework that enhances large reasoning models with agentic search and knowledge refinement capabilities

Bookmark

Like 215




Agents

Created by Lino Le Van · 274 members

Request to Join


Agent-based systems, multi-agent learning, model swarms

\$20k credits pool for top contributors thanks to  together.ai

Papers


Comments

Members




Enzo Doyen
Université de Strasbourg

202
Index




Nikil Ravi
Stanford University

172
Index




Chirag Shah
University of Washington

39
Index




Daniel Kim
Stanford University

26
Index




Lino Le Van
University of California, ...

21
Index




Adriel Saporta
New York University

20
Index




MiquelRamirez
University of Melbourne

18
Index




pkd
Ecole Normale Supérieur...

17
Index




Wanqi Xue
NTU Singapore

17
Index




Axel Brunnbauer
Technische Universität W...

16
Index




aaronshaw
Northwestern University

15
Index




Allen Kiriroath Chau
Stanford University

15
Index




Bidipta Sarkar
Stanford University

15
Index




Chau Thi Minh Pham
University of Maryland, ...

15
Index




cqz
Stanford University

15
Index




Emily Hsu
Stanford University

15
Index




Immanuel Peter
University of Chicago

15
Index




ke
University of California, ...

15
Index




kir0ul
University of Massachus...

15
Index



Lino Le Van
University of California, ...

15
Index



Rahul Chand
Stanford University

15
Index

To AI or not AI?

Global ▾

Papers


Comments

People

Search for topics, research ideas, or papers...

Hot ▾

3,643 ↗



15 Jan 2025

transformers representation-learning text-generation

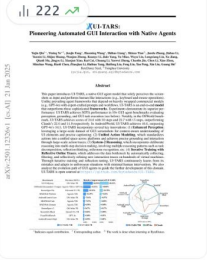
Foundations of Large Language Models

This comprehensive book from researchers at Northeastern University's NLP Lab and NiuTrans Research provides a foundational overview of Large Language Models (LLMs), focusing on core concepts and technical approaches rather than cutting-edge methods.

Bookmark

Like 157

222 ↗



21 Jan 2025

human-ai-interaction imitation-learning multi-task-learning

UI-TARS: Pioneering Automated GUI Interaction with Native Agents

This paper introduces UI-TARS, a native GUI agent model that perceives screenshots to perform human-like interactions without relying on modular frameworks or external APIs

Bookmark

Like 9

To AI or not AI?

SFT Memorizes, RL Generalizes: A Comparative Study of Foundation Model Post-training

Tianzhe Chu[✉] Yuexiang Zhai[✉] Jihan Yang^{*} Shengbang Tong^{*}
Saining Xie[✉] Dale Schuurmans^{*} Quoc V. Le^{*} Sergey Levine^{*} Yi Ma[✉]

Abstract

Supervised fine-tuning (SFT) and reinforcement learning (RL) are widely used post-training techniques for foundation models. However, their respective role in enhancing model generalization remains unclear. This paper studies the comparative effect of SFT and RL on generalization and memorization, focusing on text-based and visual environments. We introduce *GeneralPoints*, an arithmetic reasoning card game, and also consider *V-IRL*, a real-world navigation environment, to assess how models trained with SFT and RL generalize to unseen variants in both textual and visual domains. We show that RL, especially when trained with an outcome-based reward, generalizes in both the rule-based textual and visual environments. SFT, in contrast, tends to memorize the training data and struggles to generalize out-of-distribution in either scenario. Further analysis reveals that RL improves the model's underlying visual recognition capabilities, contributing to its enhanced generalization in visual domains. Despite RL's superior generalization, we show that SFT is still helpful for effective RL training: SFT stabilizes the model's output format, enabling subsequent

generalization (Bousquet & Elisseeff, 2000; Zhang et al., 2021) remain unclear, which makes it challenging to build reliable and robust AI systems. A key challenge in analyzing the generalization ability of foundation models (Bommasani et al., 2021; Brown et al., 2020) is separating data memorization¹ from the acquisition of transferable principles. We therefore investigate the key question of whether SFT or RL primarily memorize the training data (Allen-Zhu & Li, 2023a; Ye et al., 2024; Kang et al., 2024), or whether they learn generalizable principles that can adapt to novel task variants.

To address this question, we focus on two aspects of generalization: textual rule-based generalization and visual generalization. For textual rules, we study a model's ability to apply learned rules (given text instructions) to variants of those rules (Zhu et al., 2023; Yao et al., 2024; Ye et al., 2024). For vision-language models (VLMs), visual generalization measures performance consistency to variations in visual input, such as color and spatial layout, within a given task. For studying text-based and visual generalization, we investigate two different tasks that embody rule-based and visual variants. Our first task is *GeneralPoints*, an original card game task that is similar to the *Points24* task from RL4VLM (Zhai et al., 2024a), which is designed to evaluate a model's *arithmetic reasoning capabilities*. In *GeneralPoints*, the model receives four cards (pre-

Learn | O3 Mini

Switch Chat

Do these results align with @12948 DeepSeek-R1: Incentivizing Reasoning Capability in ... 's findings comparing their r1 and r1-zero experiments?

Yes, they do. The paper shows that while the pure RL approach (DeepSeek-R1-Zero) already achieves significant improvements over the base model, its performance is noticeably enhanced once additional high-quality (cold-start) data and further supervised fine-tuning steps are incorporated—this is what they refer to as DeepSeek-R1. For example, across several benchmarks (such as AIME 2024, MATH-500, and others), DeepSeek-R1 and its distilled versions clearly outperform the R1-Zero variant. This comparison supports their overall conclusion that although incentivizing reasoning via pure reinforcement learning is promising, the extra refinements (cold start data, multi-stage RL, and subsequent distillation) lead to substantially better reasoning performance and more user-friendly outputs.

I am confused about what the GlobalPoints environment in this paper is

The *GeneralPoints* environment is a custom-designed evaluation setting meant to test a model's arithmetic reasoning and its ability to generalize. In this environment, a model is given four cards (which can be provided as text or as an image) and must generate a valid arithmetic expression that uses each card's numerical value exactly once to reach a predetermined target (usually 24). Several aspects make *GeneralPoints* interesting for research

Bookmark PDF Claim Authorship

Authors: PeterTong Jihan Yang SimonZhai Tianzhe Chu

Published: 28 Jan 2025 | 1 version | 12,318 views

2501.17161v1 [cs.AI] 28 Jan 2025

Call to Action

We encourage you all to claim your papers and share the alphaXiv discussion links for your research. We're also happy to help promote your work :)

Feel free to use the Explore page/Communities to discover new and relevant papers. We'd love to hear your thoughts!

Questions

What are your biggest challenges in research today?

How do you discover relevant papers?

What can we add that would be useful?