



Search r/MachineLearning Search Reddit



295 | [N... News

Close

Posted by u/StellaAthena Researcher 3 days ago



# [N] Substantial plagiarism in BAAI's "a Road Map for Big Models"

News

BAAI recently released a [two hundred page position paper](#) about large transformer models which contains sections that are plagiarized from over a dozen other papers.

In a massive fit of irony, this was found by Nicholas Carlini, a research who (among other things) is famous for studying how language models copy outputs from their training data. [Read the blog post here](#)

60 Comments Share ...

98% Upvoted

Sort By: Best



mr\_dicaprio · 3 days ago

Damn, they really sirajed that paper

178 Reply Share Report Save Follow



Terkala · 2 days ago

They haven't tried to charge for their copied work. So it's a half siraj at best.

29 Reply Share ...



PlanetSprite · 2 days ago

I know, right? I was so confused when I read it.

7 Reply Share ...



TopGun\_84 · 2 days ago

What is 'sirajed' ?

3 Reply Share ...



r/MachineLearning Search Reddit



Citation booster

56 [up/down] [comment] Reply Share ...



godelski · 2 days ago

It is 200 pages...

13 [up/down] [comment] Reply Share ...

[Continue this thread →](#)



programmerChilli · 2 days ago

Researcher

BAAI's twitter just posted this:

<https://twitter.com/BAAIBeijing/status/1514062447165931529>

32 [up/down] [comment] Reply Share Report Save Follow



finitearth · 2 days ago

"We have noticed the discussion on "A Roadmap of Big Model" and are verifying the issues raised. BAAI encourages academic innovation and exchanges, and holds a zero tolerance policy toward academic misconduct. We will keep you updated."

25 [up/down] [comment] Reply Share ...



eyunzo · 2 days ago

This is not fair for people who put real effort. Same happens in other fields too and some put fake data which misleads people

13 [up/down] [comment] Reply Share Report Save Follow



nil- · 3 days ago · *edited 3 days ago*

Cheating in all forms is unfortunately rampant in China. I had a



r/MachineLearning Search Reddit



ArnoF7 · 2 days ago

You should bring it up. What bridge you're gonna burn by bringing it up lol? It's not like Tsinghua dictates your funding or anything.

Or are you a PhD student in China?

86 Reply Share ...

[Continue this thread →](#)



PlanetSprite · 2 days ago

That's a really unfortunate situation. Cheating and plagiarism are serious problems in any field, but especially in something as important as machine learning. It's good that you spoke to your advisor about it, and I hope that you'll continue to stand up for what's right, even if it means burning a few bridges along the way.

30 Reply Share ...



garlicpoop · 2 days ago

Just wanted to chime in and say that my ICLR paper was also plagiarized by Tsinghua last year.

17 Reply Share ...



The\_Mad\_Duck\_ · 2 days ago

You know what they say. Every great thing has a Chinese bootleg.

6 Reply Share ...

[Continue this thread →](#)

Comment deleted by user · 2 days ago

[Continue this thread →](#)



r/MachineLearning Search Reddit



↑ 1 ↓ Reply Share ...



purplebrown\_updown · 2 days ago

A co author did to a paper of mine. It really pissed me off and I hope that person gets fired. I even ran it through a similarity check and it found nothing at the time.

↑ 9 ↓ Reply Share Report Save Follow



SupportVectorMachine · 2 days ago

Researcher

█ A Roadmap for Big Model [sic]

Too bad they didn't also plagiarize the title.

↑ 8 ↓ Reply Share Report Save Follow



kilarampage · 2 days ago

Important distinction, this paper plagiarised from at least a dozen other papers, not plagiarized by.

↑ 24 ↓ Reply Share Report Save Follow



StellaAthena OP · 2 days ago

Researcher

Lol thanks

↑ 6 ↓ Reply Share ...



deep\_noob · 2 days ago

this is so sad!!!!!! They are dictating a roadmap by cheating! :/

↑ 6 ↓ Reply Share Report Save Follow



r/MachineLearning Search Reddit



we have noticed the discussion on "A Roadmap for Big Models" and are verifying the issues raised. BAAI encourages academic innovation and exchanges, and holds a zero-tolerance policy toward academic misconduct. We will keep you updated.

52 Reply Share Report Save Follow



lacifuri · 2 days ago

So BAAI specifically created a Reddit account just for this?

14 Reply Share ...



EzekielChen · 2 days ago

Its is said that some students who did the heavy lifting in writing this paper were only given one week to finish their job.

3 Reply Share Report Save Follow



coffee\_h · 2 days ago

That could do it.

2 Reply Share ...



fromnighttilldawn · 2 days ago



margot309 · 2 days ago



tentcamels · 2 days ago

only the weak copy others' work. this is disgusting behavior

1 Reply Share Report Save Follow



weiwchu · 1 day ago · edited 1 day ago



r/MachineLearning Search Reddit



**Technology reviews**, like this:[From Breaking Bad to Wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations](#)Subscribe me if you like my sharing, I will share more on how to analyze the authors' thoughts, and how to generate ideas of new papers.

To op: you're right, they should have used 'Big Models' in the title (definitely missed an s here). And I think they are trying to write a similar summary as Stanford's recent paper: 'On the Opportunities and Risks of Foundation Models' with many star authors, such that the Beijing Academy of Artificial Intelligence (BAAI) can give people the impression that: 'I am the Stanford in the East.'

1 Reply Share Report Save Follow