智谱·AI

AMiner AI 科研助手

AI Read Science

2024新品发布会

https://vip.aminer.cn/analysis/

Aminer 更懂科研的AI

01

科研人开发的AI

02

更懂科研的AI

03

AI贯穿科研流程

04

AI提升科研效率

05

各界好评与荣誉

科研人开发的 AI

- 清华大学计算机KEG实验室团队研发。
- 面向科研学术群体的,开放的科技情报大数据挖掘与服务系统平台。

2016

2019

- 智谱AI成立
- 源自清华技术成果
- 专注大模型算法研究

2022

- 自主知识产权的开源百亿大模型GLM-10B
- 发布代码生成模型CodeGeeX
- 发布开源的100+语言预训练模型mGLM-1B

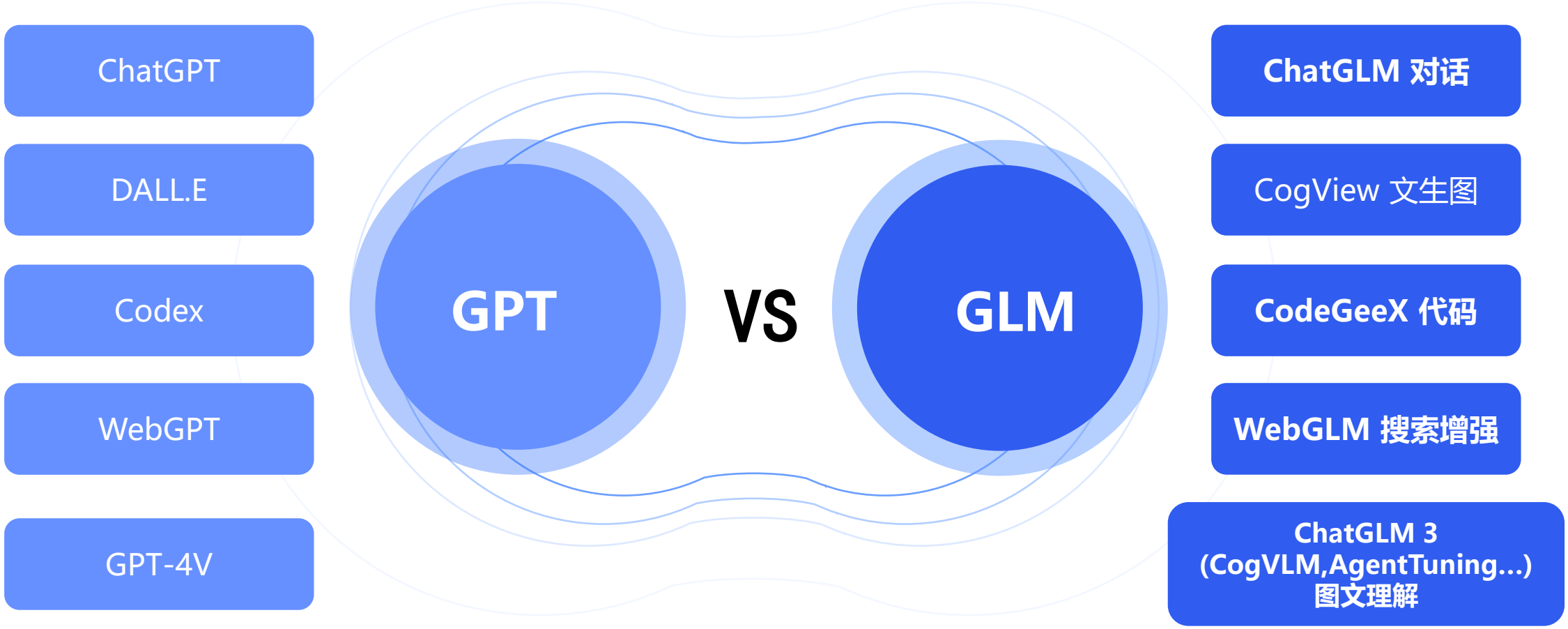
2023

- 发布千亿基座的对话模型ChatGLM，全球下载量超过800万
- 开源多模态对话模型VisualGLM-6B（CogVLM）
- 全面升级ChatGLM2模型能力登顶C-Eval榜单

2024

新一代基座大模型GLM-4正式推出，整体性能相比上一代大幅提升，比肩世界先进水平……

中国自主研发GLM构架



更懂科研的 AI

科研产出流程中的痛点

检索

- 检索关键词不确
- 海量文献难于精准定位

阅读

- 英文阅读效率低
- 知识体系不清楚
- 知识背景缺乏

分析

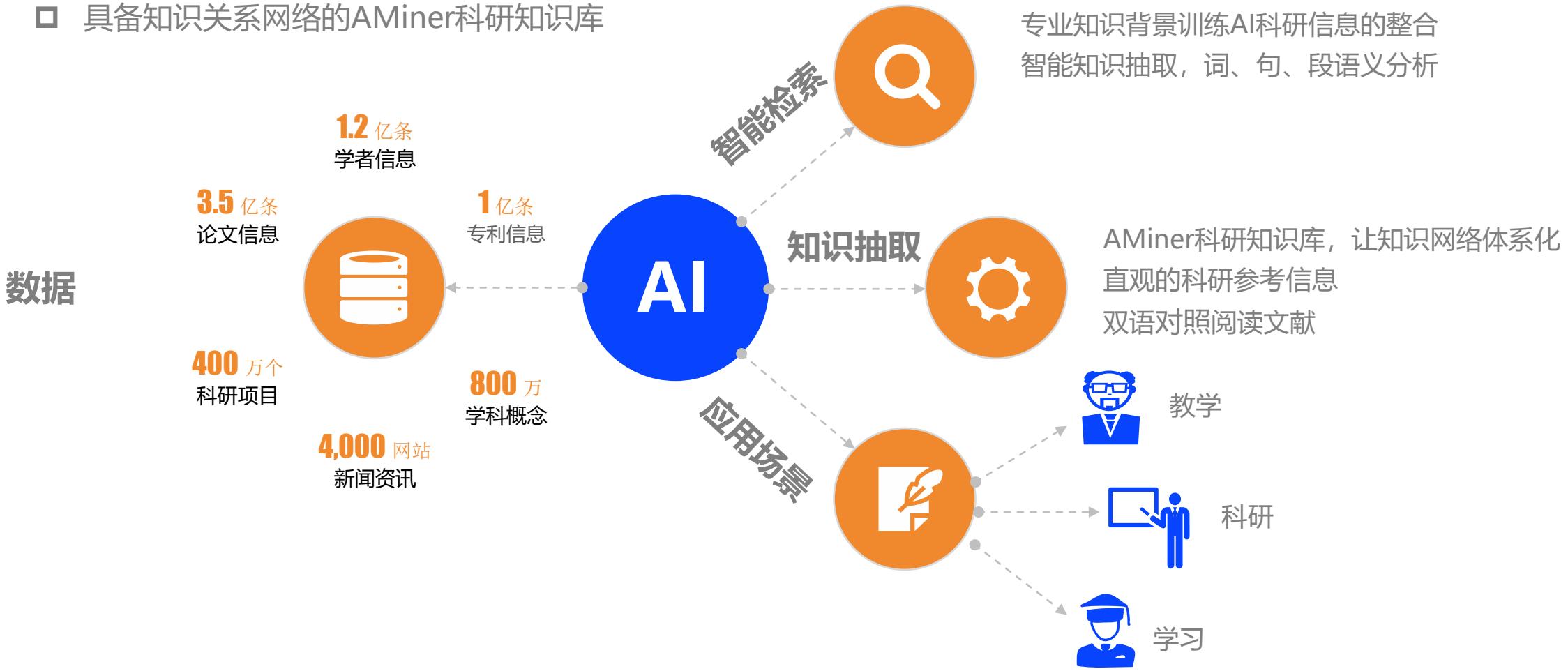
- 全局对比分析繁琐
- 研究方向不好确定

写作

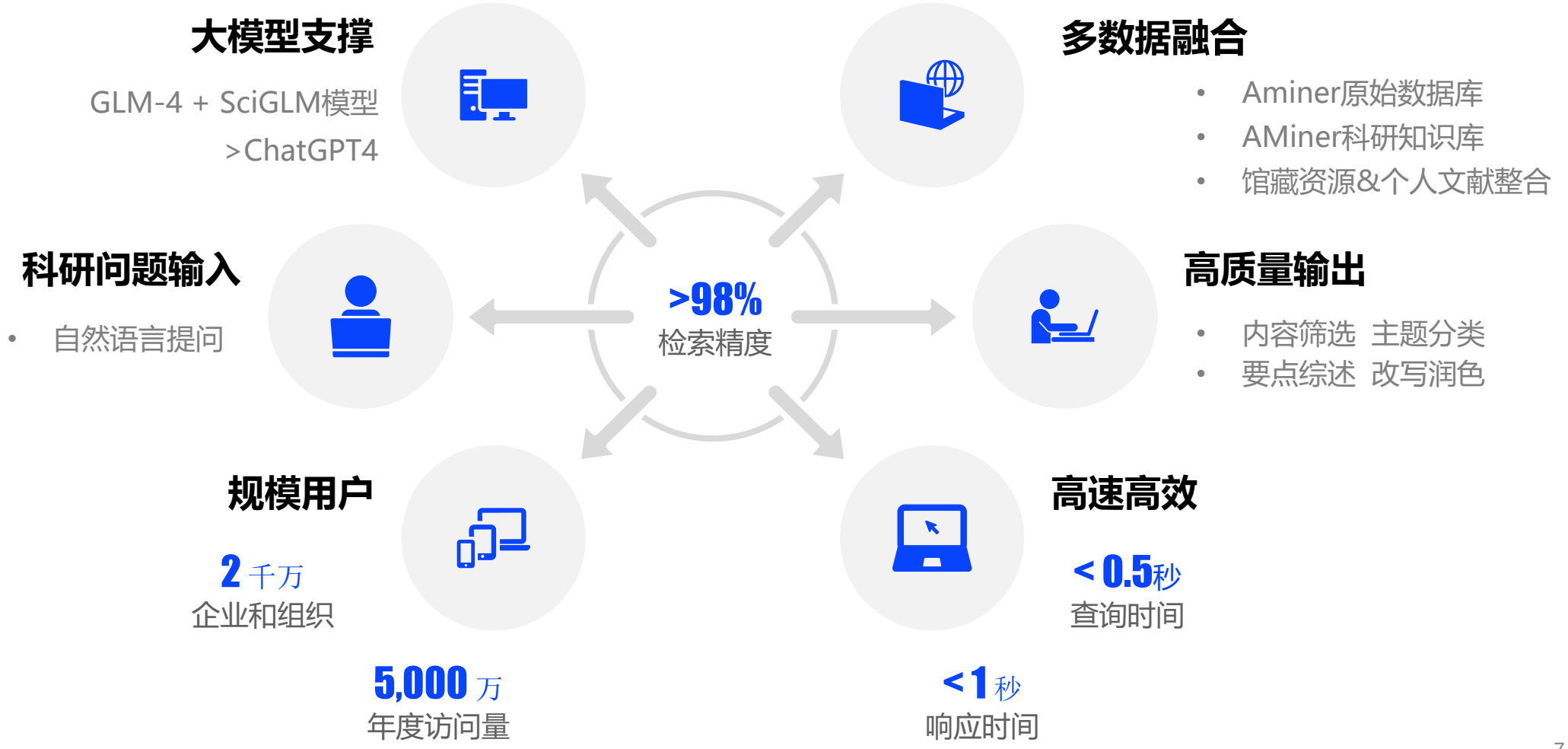
- 英文表述不专业
- 不清楚专业格式

更懂科研的 AI

- AMiner亿级高质量科研数据库
- 具备知识关系网络的AMiner科研知识库



AI贯穿科研全流程

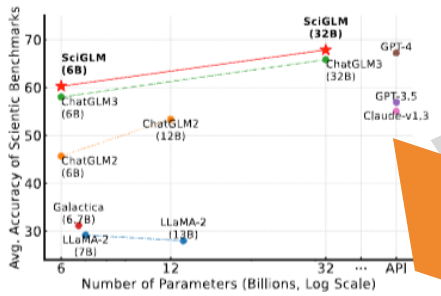

大模型支撑
GLM-4 + SciGLM模型
>ChatGPT4
多数据融合
• Aminer原始数据库
• AMiner科研知识库
• 馆藏资源&个人文献整合
科研问题输入
• 自然语言提问
>98%
检索精度
高质量输出
• 内容筛选 主题分类
• 要点综述 改写润色
规模用户
2 千万
企业和组织
高速高效
< 0.5秒
查询时间
5,000 万
年度访问量
< 1 秒
响应时间

AI提升科研效率

AMiner
资源
* 额外可选项

馆藏
资源

图片导入



多元数据应用

结果输出

- Endnote
- NoteExpress
- Zotero
- PDF上传

论文总结

研究问题

生成问题

科学语言模型的训练 | 自我反思式指令标注 | 大型语言模型在科学发现中的局限性 | 科学领域的数据稀缺挑战 | 科学和数学推理能力的提升

论文概要

全文提取概要、速览

本文介绍了一种名为SciGLM的科学语言模型，旨在通过自我反思的指令注释和调整框架，解决大型语言模型在理解复杂科学概念、推导符号方程和解决高级数值计算方面的局限性。该框架利用现有的语言模型为未标记的科学问题生成逐步推理，然后进行自我反思的批评和修订。基于这个框架，作者整理了一个包含数学、物理、化学和正式证明的多样且高质量的数据集SciInstruct。然后，使用SciInstruct对Chat-GLM系列语言模型进行了微调，提高了它们在科学和数学推理方面的能力。值得注意的是，SciGLM不仅提高了基础模型（ChatGLM3-6B-Base）的表现，也提高了更大规模的模型（12B和32B）的表现，而没有牺牲基础模型的语言理解能力。这使得SciGLM成为一个合适的基石模型，以促进各种科学发现任务。为了造福更广泛的科研社区，作者在https://github.com/THUDM/SciGLM上发布了SciInstruct、SciGLM以及自我反思框架和微调代码。

推荐问题

关键问题生成

Q1: 这篇论文的主要贡献是什么，它如何解决当前大型语言模型在科学推理方面的局限性?

Q2: 论文中提到的SciInstruct数据集有什么特点和用途?

Q3: 这篇论文的主要研究内容是什么?

Q4: 这篇论文主要研究了如何提高语言模型在科学推理方面的能力，但是否还有其他领域的推理任务也可以应用这种方法?

Q5: 论文中提到了数据集和模型规模对模型性能的影响，那么在未来的研究中是否还可以进一步探索更大规模的数据集和模型来提升性能?

Q6: 论文中提到的SciInstruct数据集是如何构建的，它包含哪些科学领域的内容?

AI提升科研效率

可信深度问答　AMiner 2.6亿篇论文

AI科研助手回答：可溯源，信息丰富

清华大学做大语言模型研究的代表学者有哪些?

在清华大学的计算机科学与技术系，有许多学者在自然语言处理和人工智能领域取得了显著的研究成果，特别是在大规模语言模型方面。以下是一些在该领域有突出贡献的学者：

1. **Jie Tang**: 教授，研究领域包括社交网络、计算机科学、数据挖掘、语义网、社交影响力等。
2. **Zhiyuan Liu**: 副教授，专注于自然语言处理、计算机科学、人工智能等，特别是在知识图谱和计算建模方面。
3. **Maosong Sun**: 教授，在自然语言处理、人工智能、计算机科学等领域有着深入的研究，特别是在中文分词、机器翻译、文本分析等方面。
4. **Minlie Huang**: 教授，研究兴趣包括文本挖掘、数据挖掘、计算机科学、情感分析、自然语言处理等。
5. **Yang Liu**: 教授，专注于神经机器翻译、统计机器翻译、自然语言处理等领域。
6. **Jie Zhou**: 教授，主要研究方向是特征提取、深度学习、指纹识别、计算机视觉等。
7. **Yong Li**: 副教授，研究领域涵盖移动通信、云计算、无线通信、网络优化等。
8. **Xiaoyan Zhu**: 教授，专注于语音识别、数据挖掘、自然语言处理、文本挖掘等领域。
9. **Xiaozhi Wang**: 博士生，研究兴趣包括预训练语言模型、零样本学习、事件检测等。
10. **Zhilin Yang**: 助理教授，研究方向包括社交网络、计算机科学、主动学习、情感分析等。

这些学者在自然语言处理、人工智能、计算机科学等领域有着广泛的研究，特别是在大规模语言模型和深度学习方面。他们的研究成果对于理解和应用这些技术具有重要意义。如果您需要更具体的信息或者对某个研究领域的细节感兴趣，请告诉我，我会尽力为您提供更详细的答案。

6700万学者库

来源学者

AI提升科研效率

通篇对比翻译

SciGLM: Training Scientific Language Models with Self-Reflective Instruction Annotation and Tuning

Dan Zhang[1,2,†], **Ziniu Hu**[3], **Sining Zhoubian**[1,2,†], **Zhengxiao Du**[1,2,†], **Kaiyu Yang**[3], **Zihan Wang**[1,2,†], **Yisong Yue**[3], **Yuxiao Dong**[1], **Jie Tang**[1]
[1]The Knowledge Engineering Group (KEG), Tsinghua University; [2]Zhipu AI;
[3]California Institute of Technology

Abstract

Large Language Models (LLMs) have shown promise in assisting scientific discovery. However, such applications are currently limited by LLMs' deficiencies in understanding intricate scientific concepts, deriving symbolic equations, and solving advanced numerical calculations. To bridge these gaps, we introduce SciGLM, a suite of scientific language models able to conduct college-level scientific reasoning. Central to our approach is a novel self-reflective instruction annotation framework to address the data scarcity challenge in the science domain. This framework leverages existing LLMs to generate step-by-step reasoning for unlabelled scientific questions, followed by a process of self-reflective critic-and-revise. Applying this framework, we curated SciInstruct, a diverse and high-quality dataset encompassing mathematics, physics, chemistry, and formal proofs. We fine-tuned the ChatGLM family of language models with SciInstruct, enhancing their capabilities in scientific and mathematical reasoning. Remarkably, SciGLM consistently improves both the base model (ChatGLM3-6B-Base) and larger-scale models (12B and 32B), without sacrificing the language understanding capabilities of the base model. This makes SciGLM a suitable foundational model to facilitate diverse scientific discovery tasks. For the benefit of the wider research community, we release SciInstruct, SciGLM, alongside a self-reflective framework and fine-tuning code at https://github.com/THUDM/SciGLM.

1 Introduction

Large language models (LLMs) have shown potential to assist and accelerate scientific discovery (Taylor et al., 2022; AI4Science and Quantum,

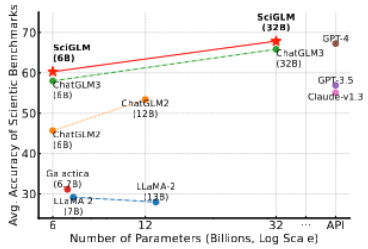


Figure 1: Average accuracy on CEval-Sci and SciEval benchmarks of different LLMs. Our **SciGLM** Models fine-tuned on ChatGLM3 are shown in bold.

geoscience understanding (Deng et al., 2023). Despite these promising proof-of-concept trials, recent studies (Sun et al., 2023; Wang et al., 2023c; Rein et al., 2023) indicate a significant performance gap in LLMs when answering basic scientific problems. According to SciBench, even advanced models like ChatGPT and GPT-4 struggle with college-level scientific problems, achieving only 28.52% accuracy even with the Chain of Thought (CoT) prompting strategy. These scientific questions, such as using the Planck distribution for energy output calculations (Wang et al., 2023c), require an in-depth understanding of complex scientific concepts, rigorous numerical computing, and sophisticated reasoning skills. Without equipping these capabilities to solve basic scientific questions, all ambitious visions of LLMs assisting scientific discovery could be unreliable. This creates a substantial incentive for developing a foundational scientific language model.

However, training LLMs to understand science is much more challenging than many general rea-

SciGLM:用自反射指令注释和调优训练科学语言模型

丹Zhang[1,2,†]、胡子牛[3]、思宁Zhoubian[1,2,†]、正孝Du[1,2,†]、杨开宇[3]、子涵Wang[1,2,†]、岳一松[3]、董玉霄[1]、唐洁[1]
[1]清华大学知识工程组(KEG);[2]Zhipu人工智能;
[3]加州理工学院

摘要。

大型语言模型(法学硕士)在协助科学发现方面已经显示出前景。然而，这种应用目前受到法学硕士在理解复杂的科学概念、推导符号方程和解决高级数值计算方面的不足的限制。为了弥合这些差距，我们引入了SciGLM，这是一套能够进行大学水平的科学推理的科学语言模型。该方法的核心是一个新的自我反思指令注释框架，以解决科学领域的数据稀缺挑战。该框架利用现有的法学硕士为未标记的科学问题生成逐步推理，随后是自我反思的批评和修改过程。应用这个框架，我们策划了sci -struct，这是一个多样化和高质量的数据集，包括数学、物理、化学和形式证明。我们使用sci -struct对Chat-GLM语言模型家族进行了微调，增强了它们在科学和数学推理方面的能力。值得注意的是，SciGLM持续改进了基本模型(ChatGLM3-6B-Base)和更大规模的模型(12B和32B)，而没有牺牲基本模型的语言理解能力。这使得SciGLM成为促进多元科学发现任务的合适基础模型。为了更广泛的研究社区的利益，我们在https://github.com/THUDM/SciGLM上发布了scidirective, SciGLM以及自我反思框架和微调代码。

1 介绍

大型语言模型(法学硕士)已经显示出协助和加速科学发现的潜力(Taylor et al., 2022; AI4Science and Quantum, 2023)，包括蛋白质预测(Chen et al., 2023a)、天气预报(Bi et al., 2023)和

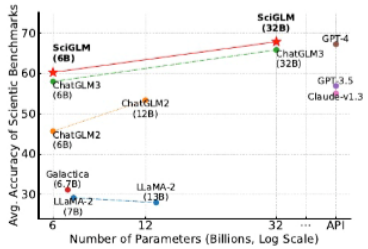


图1:不同法学硕士在CEval-Sci和SciEval基准上的平均准确率。我们在ChatGLM3上微调的SciGLM模型以粗体显示。

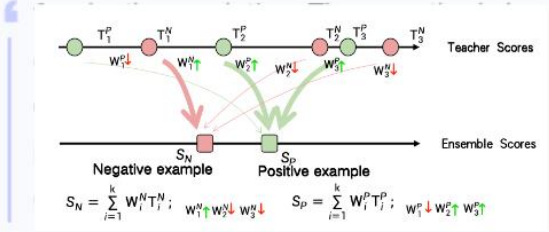
地球科学理解(Deng等人，2023)。尽管有这些有希望的概念验证试验，但最近的研究(Sun等人，2023;Wang等，2023c;Rein等人，2023)指出法学硕士在回答基本科学问题时存在显著的绩效差距。根据SciBench的数据，即使像ChatGPT和GPT-4这样的高级模型也很难解决大学水平的科学问题，即使使用思维链(CoT)提示策略，也只能达到28.52%的准确率。这些科学问题，例如使用普朗克分布进行能量输出计算(Wang等人，2023c)，需要深入理解复杂的科学概念，严格的数值计算以及如此复杂的推理技能。如果不具备这些能力来解决基本的科学问题，所有法学硕士协助科学发现的雄心勃勃的愿景都可能是不可靠的。这为开发基础科学语言模型创造了实质性的激励。

然而，训练法学硕士理解科学比许多一般的推理任务更具挑战性:一方面，解决不同学科(如数学、物理、数学、数学、数学和数学)的大学水平科学任务所需的知识和复杂性。

AI提升科研效率

图片解读

请解释下这张图



图中展示了一个机器学习模型在处理两个样本（负样本和正样本）时的权重更新过程。具体来说，它描述了如何根据教师评分和集成评分来调整权重的方向和大小。

1. **教师评分**：图中用绿色圆圈表示教师评分的正负反馈，其中T^P_i和T^N_i分别代表第i个正样本和负样本的教师评分。当教师认为某个特征是重要的时，会给出正反馈；否则，给出负反馈。
2. **集成评分**：图中用红色箭头表示集成评分的方向和大小，其中S^N_i和S^P_i分别代表第i个负样本和正样本的集成评分。集成评分是基于多个模型的预测结果的平均值或加权平均值，这里的"k"可能表示参与集成的模型数量。
3. **权重更新**：根据教师评分和集成评分，模型对每个特征的权重进行相应的调整。对于正样本，教师评分和集成评分都是正向的，因此权重W^P_i被向上移动（绿色箭头）；而对于负样本，教师评分是负向的，而集成评分是正向的，但最终的结果取决于两者的相对大小。如果集成评分的正值足够大，它可以抵消教师评分的负面影响，导致权重W^N_i向上移动。相反，如果集成评分的正值不足以抵消教师评分的负面影响，那么权重W^N_i将继续向下移动。
4. **权重的求和**：最后，模型的整体权重更新是通过将所有样本的权重进行求和得到的。负样本的权重之和为S^N =\sum_{i=1}^{k} W^N_i T^N_i，而正样本的权重之和为S^P =\sum_{i=1}^{k} W^P_i T^P_i。

总的来说，这个图展示了机器学习模型在学习和调整其参数的过程中，如何平衡教师评分和集成评分的意见，从而实现更准确的预测。

复杂科学概念释义

请用大白话科普一下：PolyGRAD

PolyGRAD是一种用于训练深度学习模型的优化算法。在深度学习中，优化算法的作用是指导模型如何调整参数以最小化损失函数，即让模型的预测尽可能接近真实值。

想象一下你在一个山谷中，你的目标是找到最低点。传统的梯度下降法就像是你每一步都朝着最陡的方向走下去，希望最终能够到达山谷的最低点。但有时候，这个方法可能会让你在半山腰就停滞不前，或者在一个不是最低点的位置停下来。

PolyGRAD算法则是对传统梯度下降法的改进。它使用了一种更聪明的方式来决定每一步应该走多远。具体来说，它考虑了当前的梯度（即山有多陡）以及你走了多远（即当前的步数）。然后，它使用一个多项式函数来计算下一步应该走多远，这个多项式函数可以帮助算法更好地平衡探索（寻找可能的最低点）和利用（在找到好方向后快速前进）。

PolyGRAD的好处是它可以帮助模型更快地找到好的参数设置，有时候还能帮助模型避免陷入局部最优解，即那些不是最低点但梯度为零的位置。这样，模型就能更快地训练，同时可能达到更高的准确率。

多语言翻译

我想让你做一个中文翻译，我会用任何一种语言和你说话，你会发现这种语言，把它翻译成中文，我要翻译的内容是

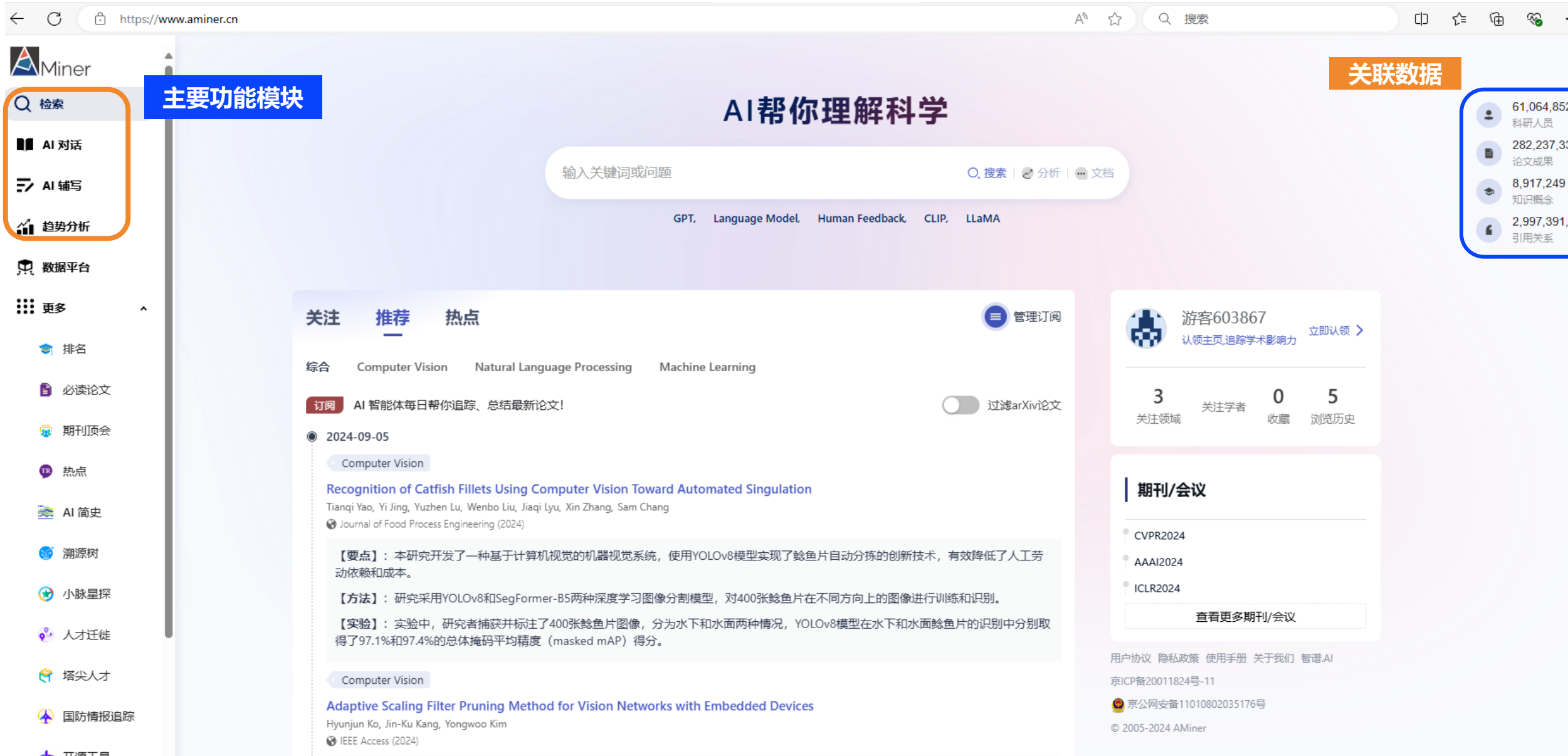
> Recently, remarkable progress has been made on automated problem solving through societies of agents based on large language models (LLMs). Previous LLM-based multi-agent systems can already solve simple dialogue tasks. More complex tasks, however, face challenges through logic inconsistencies due to cascading

近期，在基于大型语言模型（LLM）的智能体社会中，自动问题解决领域已取得了显著进展。之前的基于LLM的多智能体系统已经能够解决简单的对话任务。然而，更复杂的任务由于简单地将LLM串联在一起导致的逻辑不一致和连锁性虚构问题而面临挑战。在这里，我们介绍了MetaGPT，这是一个创新性的元编程框架，它将高效的人工工作流程融入到基于LLM的多智能体协作中。MetaGPT将标准化操作程序（SOPs）编码为提示序列，以实现更流畅的工作流程，从而使得具有类人领域专业知识的智能体能够验证中间结果并减少错误。MetaGPT采用流水线范式，为不同的智能体分配多样化的角色，有效地将复杂任务分解为涉及许多智能体协作的子任务。在与协作软件工程的基准测试中，MetaGPT生成的解决方案比之前的基于聊天式多智能体系统更为连贯。

AI提升科研效率


https://www.aminer.cn/chat/g/explain
搜索
对话
文献库
1. 上传需要阅读的文档
我的文献库（21篇）
需要对比阅读的文献
从AMiner新增论文
上传PDF
全部
默认文件夹(2)
DRAA方案(10)
新冠病毒相关(9)
新建文件夹
AI输出内容
添加知识点(2)
请输入搜索内容
简体中文
文件名
结论
实验结果
创建知识点
输入想提取的知识点，AI自动总结
推荐添加知识点
数据集
相关工作
未来研究
结论
挑战
实验结果
应用方向
方法
概要
目标
研究问题
核心发现
局限性
研究差距
贡献
可选择的阅读对比内容
Gene Therapy Is Coming of Age.
基因疗法治疗血液癌症和一些罕见疾病的方法不久可能成为标准治疗手段。
很抱歉，根据提供的论文标题和摘要信息，无法直接提取出关于"实验结果"的知识点内容，因为这些信息在摘要中并未提及。需要查阅论文的正文中相关部分才能获得实验结果。
The Gene Ontology Resource: Enriching A Go...
1. GO-CAM注释框架得到显著改进，并使用计算模式进行验证。
2. 通过合作细化GO，GO注释数量增加10%，注释基因产品数量增加25%，超过9400篇新科学文章被注释。
3. 对20000个来自实验数据的注释进行了审查，占实验GO注释的2.5%。
4. 网站进行了重新设计，便于快速访问文档、下载和工具。
5. 提供了包含过去15年GO数据的格式一致的历史存档，以支持准确性和可追溯性。
实验结果表明，GO-CAM注释框架得到显著改进，GO注释数量增加了10%，注释基因产品数量增加了25%，超过9,400篇新的科学文章被注释。同时，对20,000个来自实验数据的注释进行了审查，占实验GO注释的2.5%。
Genome regulation by long noncoding RNA ...
长非编码RNA在基因调节中起重要作用，能够与多种染色质调节因子形成广泛的核糖核蛋白复合物网络，并通过 modular scaffolds 指定 RNP 复合体和染色质状态的高阶组织。DNA 元素介导 lncRNA 表达也可控制染色质区域，影响癌基因表达和活性。长非编码 RNA 在人类疾病中的新识别作用突显了这些调节模式的重要性。
知识点内容：RNA化学修饰、RNA结构；lncRNAs与多种染色质调节因子形成广泛的核糖核蛋白（RNP）复合物网络，并将这些酶活性靶向定位到基因组中的适当位置；长非编码RNA可作为模块化支架，指定RNP复合物和染色质状态的高阶组织；DNA元件介导lncRNA表达，对控制染色质邻域具有强大效应，改变癌基因表达和活性。

AI科研助手用途多


https://www.aminer.cn
搜索
AMiner
主要功能模块
检索
AI 对话
AI 辅写
趋势分析
数据平台
更多
排名
必读论文
期刊顶会
热点
AI 简史
溯源树
小脉星探
人才迁徙
塔尖人才
国防情报追踪
关联数据
AI帮你理解科学
输入关键词或问题
搜索 | 分析 | 文档
GPT, Language Model, Human Feedback, CLIP, LLaMA
61,064,852
科研人员
282,237,33
论文成果
8,917,249
知识概念
2,997,391,
引用关系
关注
推荐
热点
管理订阅
综合
Computer Vision
Natural Language Processing
Machine Learning
订阅
AI 智能体每日帮你追踪、总结最新论文!
过滤arXiv论文
2024-09-05
Computer Vision
Recognition of Catfish Fillets Using Computer Vision Toward Automated Singulation
Tianqi Yao, Yi Jing, Yuzhen Lu, Wenbo Liu, Jiaqi Lyu, Xin Zhang, Sam Chang
Journal of Food Process Engineering (2024)
【要点】：本研究开发了一种基于计算机视觉的机器视觉系统，使用YOLOv8模型实现了鲶鱼片自动分拣的创新技术，有效降低了人工劳动依赖和成本。
【方法】：研究采用YOLOv8和SegFormer-B5两种深度学习图像分割模型，对400张鲶鱼片在不同方向上的图像进行训练和识别。
【实验】：实验中，研究者捕获并标注了400张鲶鱼片图像，分为水下和水面两种情况，YOLOv8模型在水下和水面鲶鱼片的识别中分别取得了97.1%和97.4%的总体掩码平均精度（masked mAP）得分。
Computer Vision
Adaptive Scaling Filter Pruning Method for Vision Networks with Embedded Devices
Hyunjun Ko, Jin-Ku Kang, Yongwoo Kim
IEEE Access (2024)
游客603867
认领主页,追踪学术影响力
立即认领
3
关注领域
关注学者
0
收藏
5
浏览历史
期刊/会议
CVPR2024
AAAI2024
ICLR2024
查看更多期刊/会议
用户协议 隐私政策 使用手册 关于我们 智谱.AI
京ICP备20011824号-11
京公网安备11010802035176号
© 2005-2024 AMiner

各界好评与荣誉

《自然》杂志文章报道ChatGLM相关成果

中文Chat-GPT

中国制造的聊天机器人ChatGLM在许多指标上表现得和ChatGPT一样好，甚至在中文方面表现得更好。机器学习专家Adina Yakefu表示，为不同语言量身定制的模型可以避免“过度简化或忽视某些语言和文化的具体特征”。尽管ChatGPT和它的许多竞争对手可以用多种语言回应，但其中大多数是由美国公司开发的，主要使用英语。相比之下，ChatGLM的设计可以同时支持中文和英文。

各界好评与荣誉

国家科技进步二等奖

国家科学技术进步奖
证 书

为表彰国家科学技术进步奖获得者，特颁发此证书。

项目名称：智能型科技情报挖掘和知识服务关键技术及其规模化应用

奖励等级：二等

获 奖 者：唐杰

2021 年 9 月 25 日

证书号：2020-J-220-2-04-R01

北京市专利发明一等奖

荣誉证书

为表彰第六届北京市发明专利奖获奖者，特颁发此证书。

专利名称：在大规模社会网络中基于路径评分的个人关系发现方法
专 利 号：ZL200710177066.8
奖励等级：一等奖
专利权人：清华大学
发 明 人：唐 杰 李涓子

编号：2021-04-04

二〇二一年四月

北京市科学技术一等奖

北京市科学技术奖

为表彰在推动科学技术进步、对首都经济建设和社会发展作出贡献者，特颁此证，以资鼓励。

获奖项目：科技情报大数据挖掘及服务平台

获 奖 者：唐 杰

获奖等级：壹等奖

No 2017 计-1-001-01

二〇一七年十一月

人工智能学会科技进步一等奖

荣誉证书

HONORARY CERTIFICATE

为表彰在促进智能科学技术进步中做出重大贡献，特颁发此证书。

获奖项目：研究者社会网络搜索与挖掘系统（ArnetMiner）

奖励等级：进步奖一等奖

获 奖 者：唐 杰（第一完成人）

中国人工智能学会

2013年10月

证书号：JB2013-1-1-01-R01

- 2020国家科技进步二等奖、北京市发明专利一等奖、北京市科技进步奖一等奖
- 2021中国科协“科创中国. 新锐企业”
- 国家高新技术企业、中关村高新技术企业、中关村金种子企业
- ISO 9001 质量管理体系认证证书
- 《知识图谱构建平台认证技术规范》产品认证证书
- 全国信息技术标准化技术委员会人工智能分委会单位委员证书
- 中国创新创业大赛初创组一等奖（北京赛区）
- 北京市“专精特新”中小企业


智谱·AI
AMiner AI 科研助手
AI Read Science
智谱AI 让机器像人一样思考
智能体中心 ↗
大模型开放平台 ↗