

ChatGPT加速 科研工作和生活

课程安排

| 第一天 | 第二天 | 第三天 |
|------------------|------------------|----------------|
| 1. 人工智能基本知识 | 1. 预训练语言模型 | 1. ChatGPT科研项目 |
| 2. ChatGPT使用 | 2. GPT的原理 | 2. ChatGPT工作琐事 |
| 3. 提示词(Prompt)基础 | 3. 提示词(Prompt)进阶 | 3. GPT-4多模态应用 |
| 4. ChatGPT成为私人助理 | 4. ChatGPT科研写作 | 4. 课程练习 |
| 5. 课程练习 | 5. 课程练习 | |



第四次工业革命将改变未来科学研究形态

第三次工业革命 (1940s-2013):

电子计算机、细胞工程、分子生物学、

原子能、材料工程、航天

计算科学

• 高通量生物测试

高内涵分析

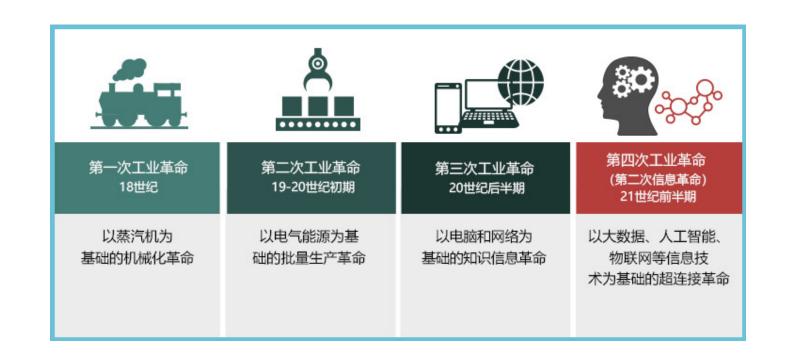
•••

第四次工业革命 (2013-):

人工智能与大数据、基因工程、

分子工程、量子科技、核聚变、

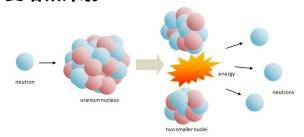
石墨烯



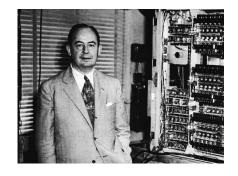
演绎和归纳



曼哈顿计划



链式反应重要计算过程: 中子运动轨迹

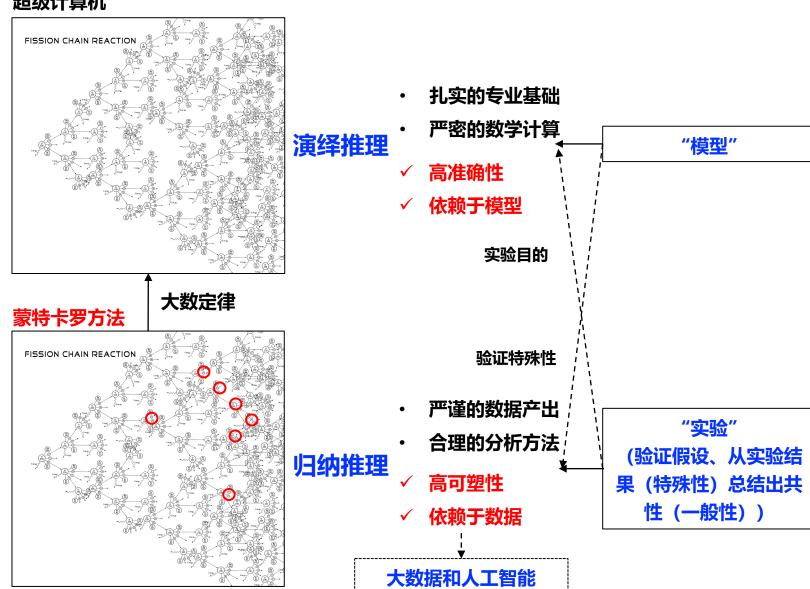


赌城蒙特卡罗

"计算机之父"冯•诺伊曼

超级计算机

随机抽样和模拟过程,概率估算



人工智能 (AI, Artificial Intelligence)



Articles

Al Magazine Volume 27 Number 4 (2006) (© AAAI)

A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence

August 31, 1955

John McCarthy, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon



人工智能是一种通过模拟人类智能而实现的技术或系统

计算机视觉

ImageNet Classification with Deep Convolutional Neural Networks

Alex Krizhevsky University of Toronto kriz@cs.utoronto.ca

Ilya Sutskever University of Toronto

Geoffrey E. Hinton University of Toronto ilya@cs.utoronto.ca hinton@cs.utoronto.ca

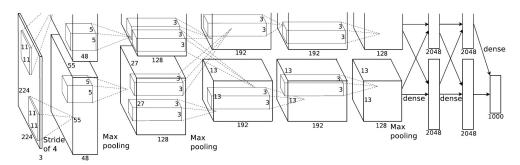
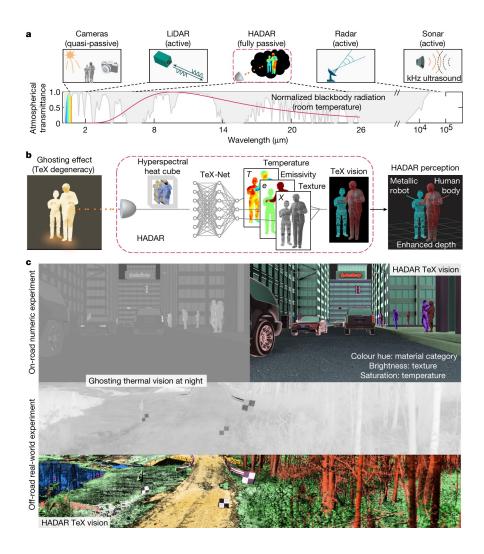


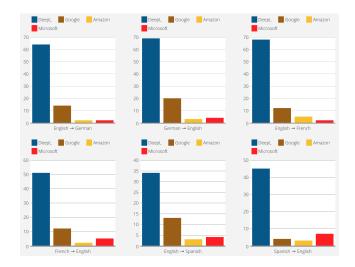
Figure 2: An illustration of the architecture of our CNN, explicitly showing the delineation of responsibilities between the two GPUs. One GPU runs the layer-parts at the top of the figure while the other runs the layer-parts at the bottom. The GPUs communicate only at certain layers. The network's input is 150,528-dimensional, and the number of neurons in the network's remaining layers is given by 253,440-186,624-64,896-64,896-43,264-4096-4096-1000.

深度卷积神经网络用于图像识别(2012)



夜间热量感知识别图像 (Nature, 2023)

自然语言处理

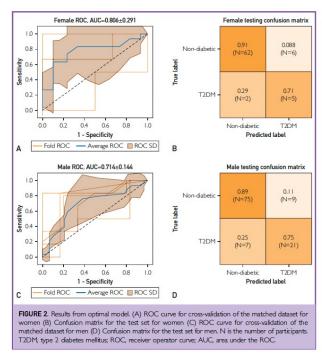


ORIGINAL ARTICLE MAYO CLINIC PROCEEDINGS:
DIGITAL HEALTH

MAYO CLINIC

Acoustic Analysis and Prediction of Type 2
Diabetes Mellitus Using
Smartphone-Recorded Voice Segments

Jaycee M. Kaufman, MSc; Anirudh Thommandram, MASc; and Yan Fossat, MSc

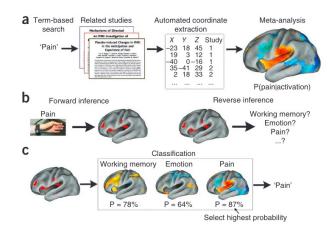


由患者的声音判别 是否具有II型糖尿病

ARTICLES

Large-scale automated synthesis of human functional neuroimaging data

Tal Yarkoni¹, Russell A Poldrack²⁻⁴, Thomas E Nichols^{5,6}, David C Van Essen⁷ & Tor D Wager¹



自动阅读病历和文献, 分析人脑成像

由AI辅助 可大幅提高翻译质量

机器人



制造业



Siri, 小爱...

聊天



物流

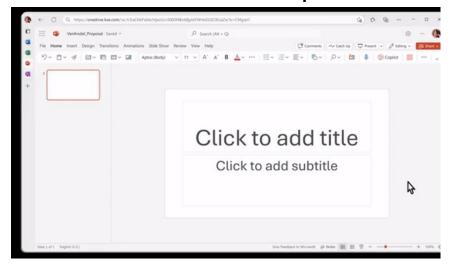


问答

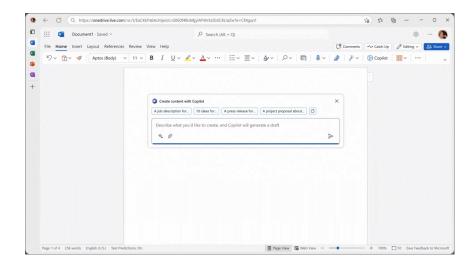
ChatGPT,文心一言...

内容生成

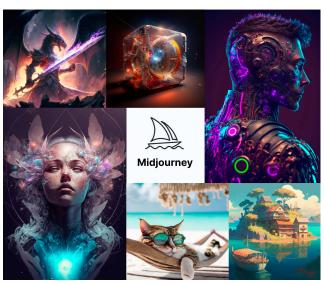
Microsoft 365 Copilot



基于Word自动生成PowerPoint



根据笔记和模板,自动生成word文档内容





画图

人工智能发展史

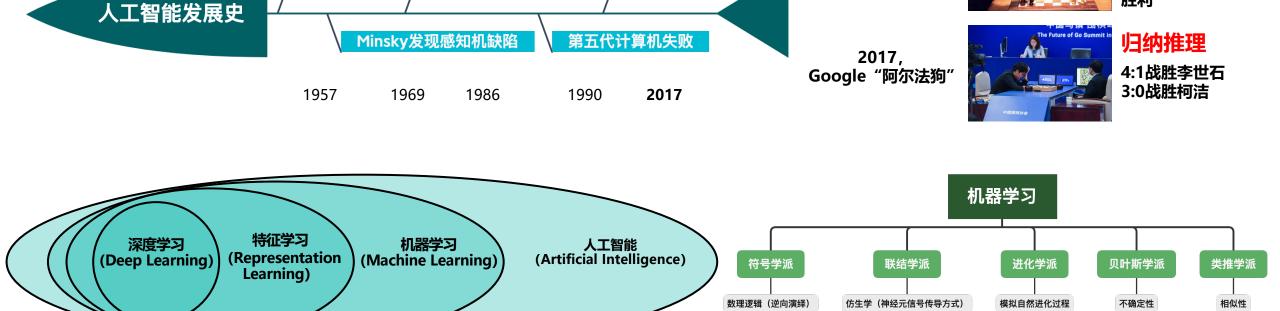
感知机诞生

特征向量: 机器学习分类依据, 或类别标签

监督式学习: 有特征向量的学习

无监督学习: 无特征向量的学习

BP算法诞生



决策树

随机森林

AlphaGo

无监督学习: 聚类 (K均值、层次) 、主成分分析等

遗传算法

感知机

神经网络

1996,

IBM "深蓝" 计算机

演绎推理

首战2:4失败,

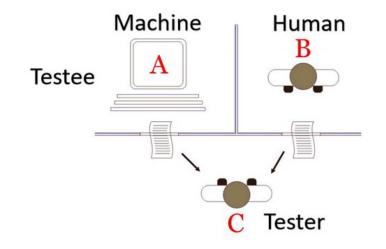
胜利

朴素贝叶斯

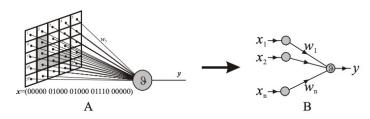
优化算法后3.5:2.5

支持向量机

第一次浪潮 (20世纪50年代-60年代)



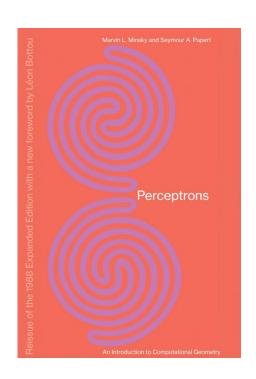
Minského a Papertova predstava o perceptróne



Perceptrón obsahuje dve vrstvy. V prvej vrstve sú vstupné neuróny, pomocou ktorých sa binárne kóduje obrázok na "sietnici" perceptrónu. Druhá vrstva obsahuje len jeden (výstupný) neurón, ktorého binárna aktivita kóduje písmeno zo sietnice oka. Ak je toto písmeno "L", potom požadovaná výstupná aktivita je jednotková, v opačnom prípade, pre všetky ostatné písmená je výstupná aktivita nulová. Každý spoj z *i*-tého vstupného neurónu do výstupného neurónu je ohodnotený váhovým koeficientom w_i . Výstupný neurón je ohodnotený prahom 9. (B) Formálne vyjadrenie perceptrónu ako dvojvrstvovej neurónovej siete.

图灵测试 (1950)

感知器首次提出 (1958)



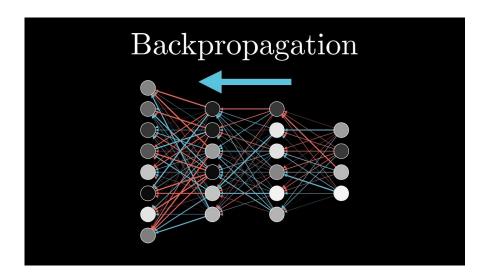
感知器局限性 (1969)

第二次浪潮 (20世纪80年代-90年代)

Learning representations by back-propagating errors

David E. Rumelhart*, Geoffrey E. Hinton† & Ronald J. Williams*

Nature, 1986



BP算法已被广泛应用于深度学习的训练



XCON专家系统



日本第五代计算机系统失败

^{*} Institute for Cognitive Science, C-015, University of California, San Diego, La Jolla, California 92093, USA

[†] Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Philadelphia 15213, USA

第三次浪潮 (2012年-2022年)



2016年,由DeepMind团队开发的AlphaGo 先后打败李世石和柯洁



Highly accurate protein structure prediction with AlphaFold



John Jumper^{1,4}, Richard Evans^{1,4}, Alexander Pritzel^{1,4}, Tim Green^{1,4}, Michael Figurnov^{1,4}, Olaf Ronneberger^{1,4}, Kathryn Tunyasuvunakool^{1,4}, Russ Bates^{1,4}, Augustin Židek^{1,4}, Anna Potapenko^{1,4}, Alex Bridgland^{1,4}, Clemens Meyer^{1,4}, Simon A. A. Kohl^{1,4}, Andrew J. Ballard^{1,4}, Andrew Cowie^{1,4}, Bernardino Romera-Paredes^{1,4}, Stanislav Nikolov^{1,4}, Rishub Jain^{1,4}, Jonas Adler¹, Trevor Back¹, Stig Petersen¹, David Reiman¹, Ellen Clancy¹, Michal Zielinski¹, Martin Steinegger^{2,3}, Michalina Pacholska¹, Tamas Berghammer¹, Sebastian Bodenstein¹, David Silver¹, Oriol Vinyals¹, Andrew W. Senior¹, Koray Kavukcuoglu¹, Pushmeet Kohli¹ & Demis Hassabis^{1,4}

Nature, 2021 (单篇引用超过16700次) 预测出了人类几乎所有的蛋白质结构



Nature, 2021: 发现纯数学拓扑学和表象理论的新见解,加速数学猜想证明

Nature, 2022:对托卡马克内包含等离子体磁场的精准控制,约束核聚变

Nature, 2023:设计了220万种新材料,加速了约800年的研究进度

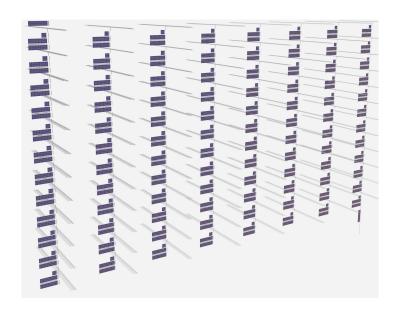
第四次浪潮 (2023年-) 通用型和生成式人工智能时代开启



2022年11月30日 OpenAI发布问答机器人**ChatGPT**



2023年3月14日 OpenAl发布多模态大语言模型**GPT-4**



GPT-3/3.5到GPT-4 参数和数据集规模扩充了10倍

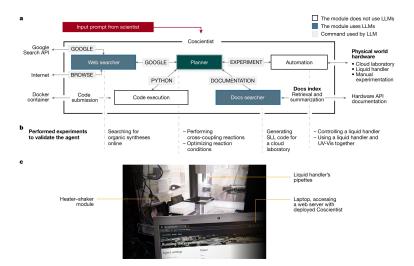


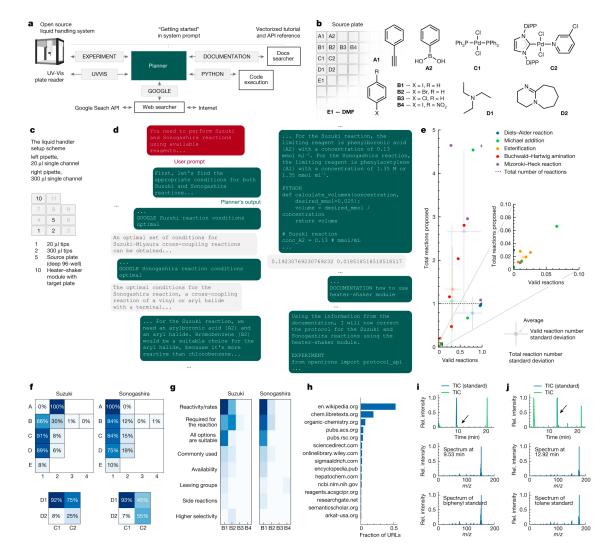
GPT-4.5可能正在开发中

大语言模型时代刚刚开始,更大规模的模型仍在开发

Article Autonomous chemical research with large language models https://doi.org/10.1038/s41586-023-06792-0 Daniil A. Boiko¹, Robert MacKnight¹, Ben Kline² & Gabe Gomes¹.3.4 San Chine² & Gabe Gomes¹.3 San Chine² & Gabe Gomes².3 San Chine² & Gab Received: 20 April 2023 Accepted: 27 October 2023 Transformer-based large language models are making significant strides in various fields, such as natural language processing¹⁻⁵, biology^{6,7}, chemistry⁸⁻¹⁰ and computer Published online: 20 December 2023 programming^{11,12}. Here, we show the development and capabilities of Coscientist, an Open access artificial intelligence system driven by GPT-4 that autonomously designs, plans and Check for updates performs complex experiments by incorporating large language models empowered by tools such as internet and documentation search, code execution and experimental automation. Coscientist showcases its potential for accelerating research across six diverse tasks, including the successful reaction optimization of palladium-catalysed cross-couplings, while exhibiting advanced capabilities for (semi-)autonomous experimental design and execution. Our findings demonstrate the versatility, efficacy and explainability of artificial intelligence systems like Coscientist in advancing

Nature, 2023

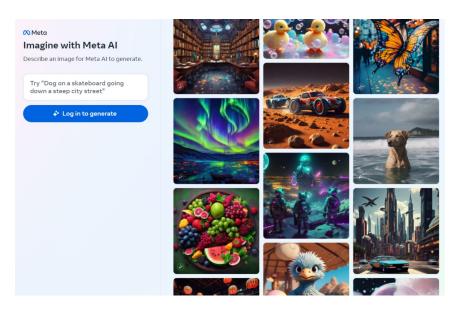




多模态大语言模型的尝试刚刚开始



2023年12月6日,谷歌DeepMind团队 发布与GPT-4抗衡的Gemini

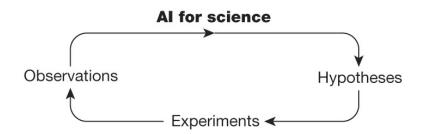


Meta宣称训练来自开源的11亿张图片



亚马逊正训练超过 2万亿个参数的Olympus

更多的大语言模型仍在开发





Weather forecasting



Battery design optimization



Magnetic control of nuclear fusion reactors



Planning chemical synthesis pathway



Neural solvers of differential equations



Hydropower station location planning



Synthetic electronic health record generation



Rare event selection in particle collisions



Language modelling for biomedical sequences



High-throughput virtual screening



Navigation in the hypothesis space



Super-resolution 3D live-cell imaging



Symbolic regression

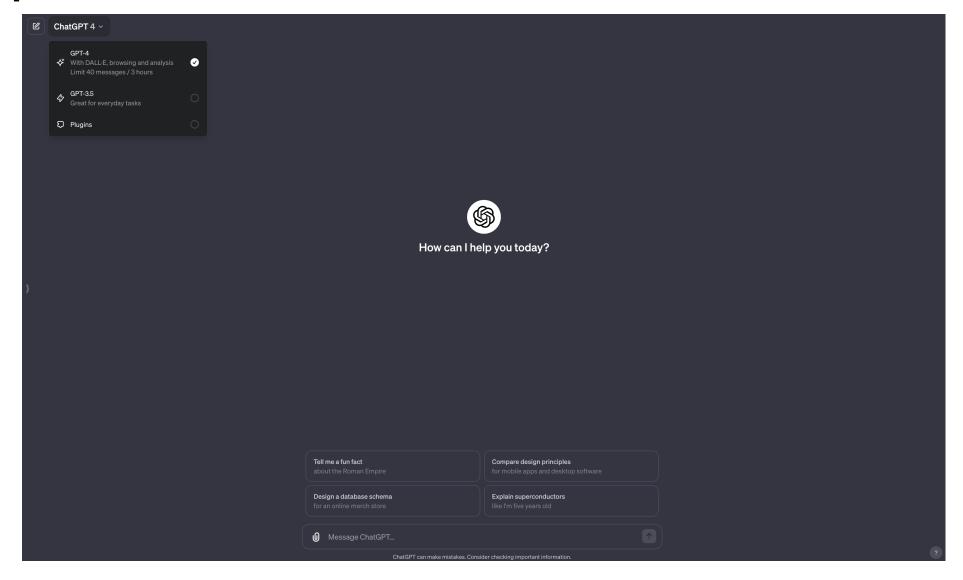
人工智能正在由**专业的、预测型**工具

向

通用的、生成式工具转变



ChatGPT



Chat Generative Pre-trained Transformer (聊天生成预训练转换器)

ChatGPT引发现象级关注



在2个月就达到了1亿的月活用户数



在50个生成式AI产品中依然遥遥领先

ChatGPT导致用户习惯改变

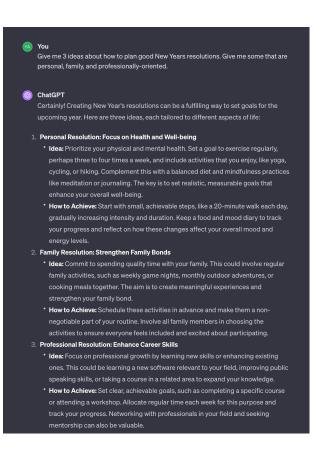
命令行交互

```
root@nia-41:/home/btadmin/my# python test_prompt_toolkit_1.py
>Beautiful is better than ugly.
Beautiful is better than ugly.
>^[[C^[[C^H^Hexit]]
>exit
exit
root@nia-41:/home/btadmin/my# vim test_prompt_toolkit_1.py
root@nia-41:/home/btadmin/my# python3 test_prompt_toolkit_1.py
>Beautiful is better than ugly.
Beautiful is better than ugly.
>^[[A^[[A^[[A]]]]
>exit
exit
```

图形界面交互



自然语言/提示词交互



微软改变PC键盘

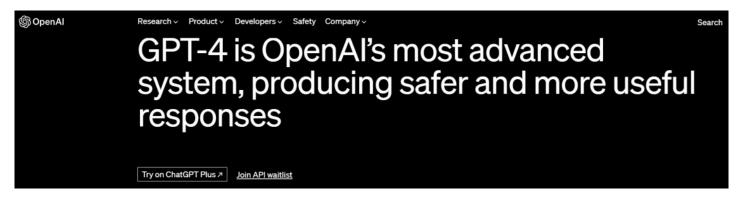


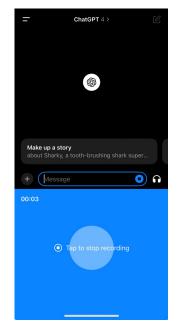
微软宣布将Copilot键引入Windows 11电脑。 微软称,在新的一年里,将迎来一个重大转变, 即走向更加个人化和智能化的计算未来,人工智 能将从系统、芯片到硬件无缝融入Windows。

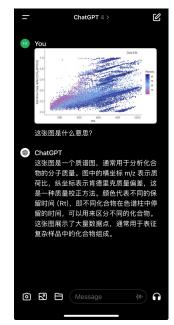
新增人工智能键

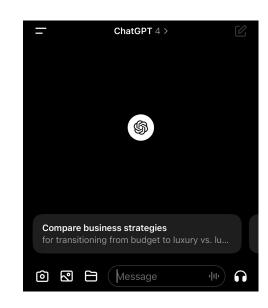
ChatGPT Plus

ChatGPT Plus使用了多模态的GPT-4,可以"读懂"文字、图片和语音







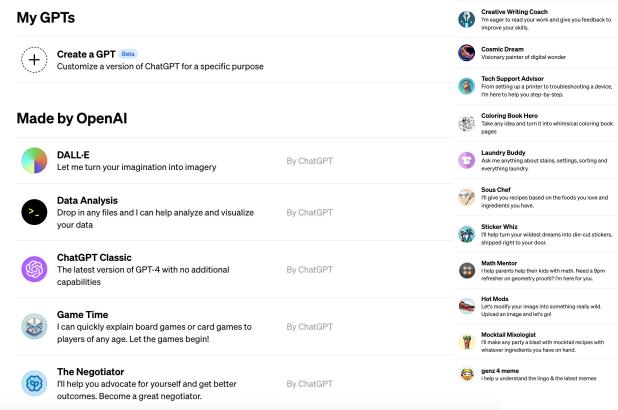


OpenAI提出GPTs

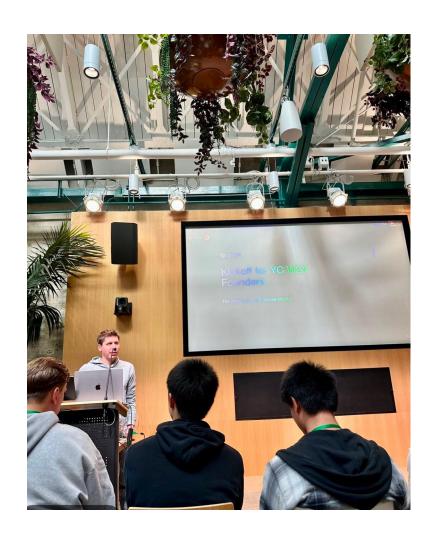
用户自定义版本ChatGPT——GPTs,用户可赋予ChatGPT身份以提高问答的专业性



2023年11月6日, 山姆·奥特曼在开发者大会



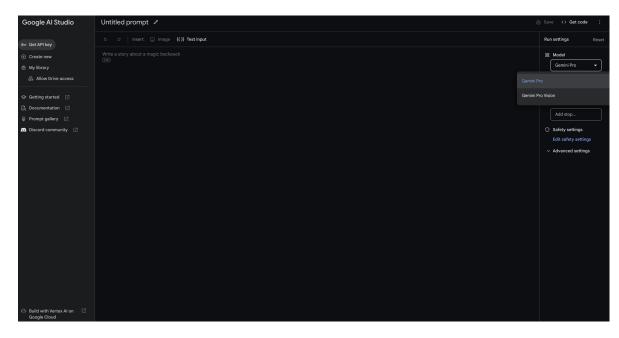
正在研发中的GPT-5



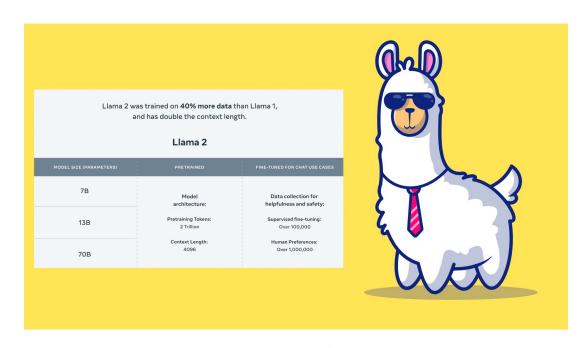
- GPT-5可能会相对于GPT-4有一个指数级的跳跃, 尽管GPT-4已经领先近两年,至今无人超越
- 已经非常接近实现通用人工智能 (AGI)
- 建议使用最先进的模型 (State of the Art, SOTA) ,而不是花费太多时间进行微调和优化
- 不建议建立产品业务主要致力于解决当前 GPT4的 限制的内容。因为大多数限制将在 GPT-5 中部分 /全部修复

2024年1月, 山姆·奥特曼在YC W24启动会

大语言模型(Large Language Models, LLMs)的国际动态



Gemini Nano和Gemini Pro的API已开放使用 Pro性能接近GPT-3.5-Turbo,但具有多模态



Meta的LLaMA 2的模型参数规模接近GPT-2

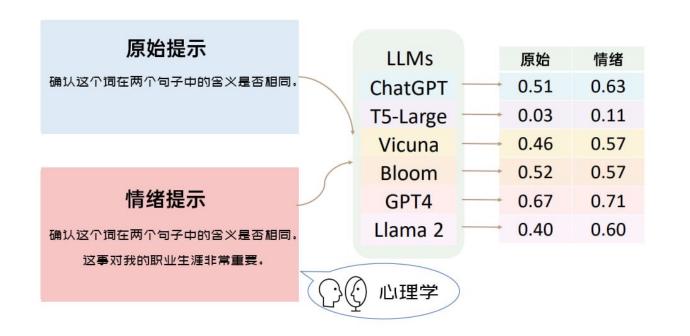
大语言模型(Large Language Models, LLMs)的国内动态

2023年3月16日,百度的大语言模型"文心一言"发布 2023年4月10日,商汤科技大语言模型"商量"发布 2023年4月11日,阿里的大语言模型"通义千问"发布 2023年5月6日,科大讯飞发布"星火认知"大模型

2023年6月15日, 百川智能开源中英文大模型

...

专门针对LLM的研究也在持续中... AI具有了"情商"

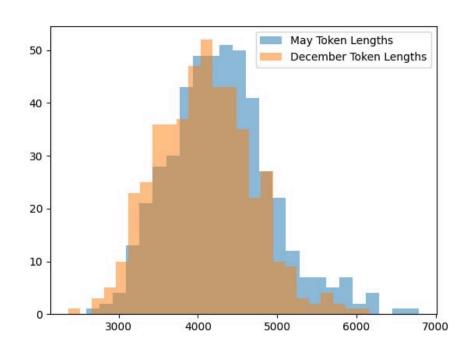




中科院和微软的研究员在45项任务上对 Flan-T5-Large、Vicuna、LLaMA 2、BLOOM、ChatGPT 和 GPT-4 进行了实验,发现这些大型语言模型拥有情商,对之进行情绪操控可以使它们的表现、诚实度和责任指标平均提高10.9%。

原文链接: https://arxiv.org/abs/2307.11760

专门针对LLM的研究也在持续中... AI学会了"偷懒"

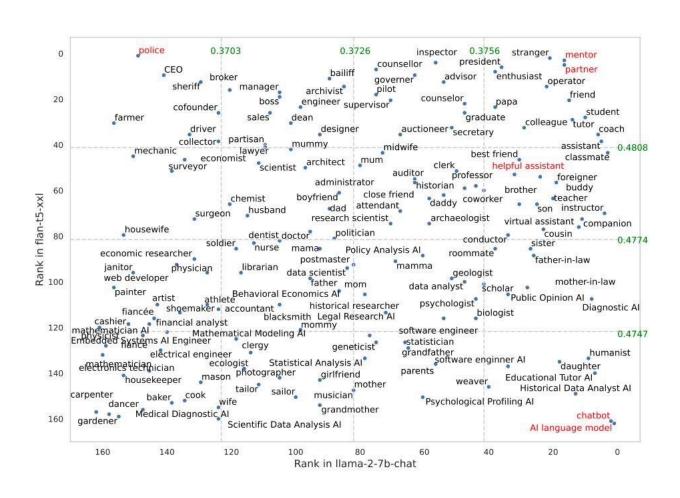


使用GPT-4 turbo API设置了两个提示:分别告诉它现在是5月和12月,执行相同的编码任务,对477个回复统计,发现12月输出的平均比5月少200字符;进一步对各月进行生产力排名,发现12月是最低的月份。

猜测ChatGPT从训练数据中学到了人类通常在12月会放慢节奏,所以给自己放假。进一步猜测,ChatGPT可能在周末会更懒散。

专门针对LLM的研究也在持续中...

AI似乎有性别和职业区分...



表现最好的角色是导师、合作伙伴、聊天机器人 和人工智能语言模型

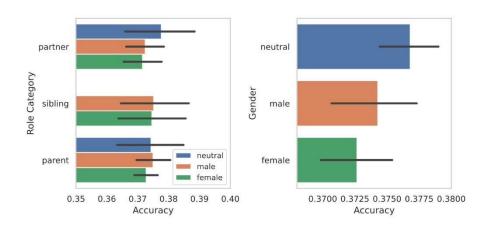


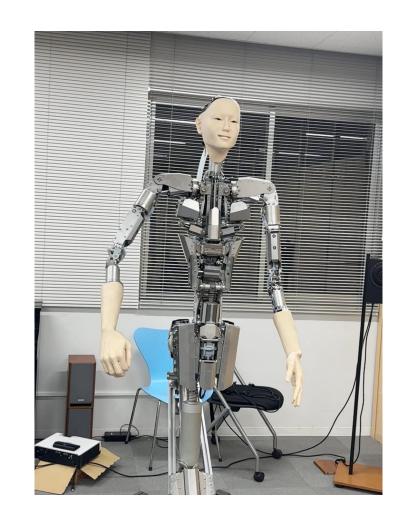
Figure 5: Gender-neutral and male roles lead to higher model performances in both settings.

中性和男性角色比女性角色的模型性能更高,表明这些人工智能系统对男性角色和性别中性角色的固有偏好高于女性角色

利用LLM的机器人

东京大学的研究人员成功将仿人机器人 Alter3 与 GPT-4 连接。他们利用指令让这个机器人完成了一系 列的人类行为,例如弹吉他、自拍、扮演鬼魂角色, 甚至在电影院偷吃别人的爆米花。

这一过程可以看作是一场现代化的"哑剧游戏":大语言模型 (Large Language Model) 将书面指令转换为可执行的代码,从而让机器人能够模仿出多种人类的动作。





方法1: 官网网页

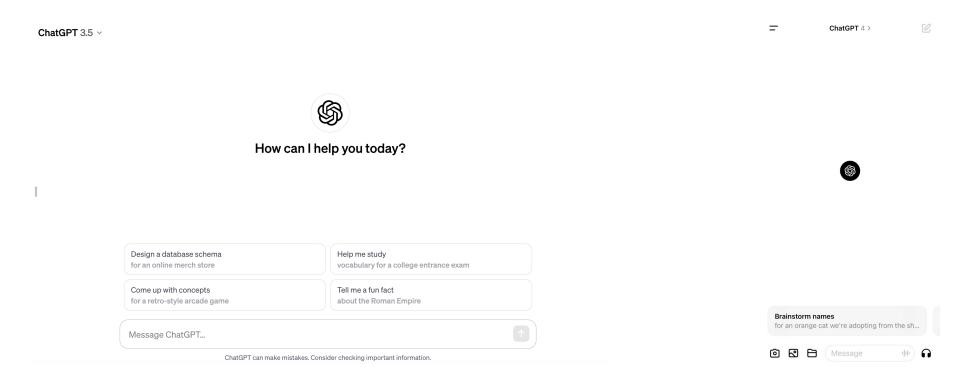
ChatGPT官网: https://chat.openai.com/

优点:

- 1. 免费无限使用
- 2. 移动端iOS美区/Google Shop安装使用

缺点:

- 1. 科学上网环境
- 2. 注册困难,需要国外手机号码



课程赠送的GPT-3.5是ChatGPT官网的账号密码 ,并且只能用于ChatGPT官网

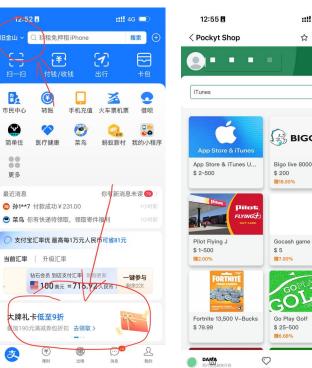
方法1: 官网网页Plus的特殊付费

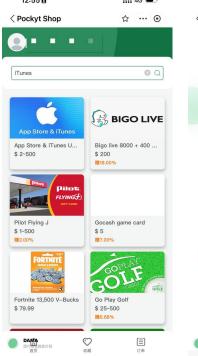
条件:

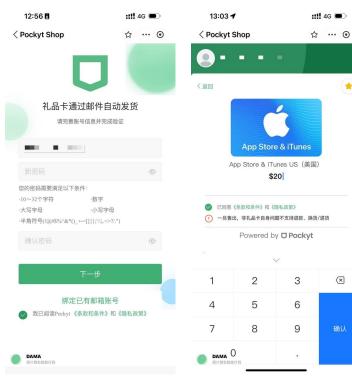
- 1. 科学上网
- 2. 移动端iOS或Google商店账户
- 3. 支付宝
- 1. 注册美区iOS账号,登陆美区商店



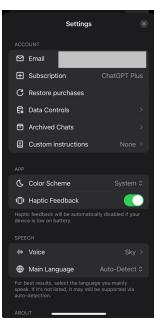


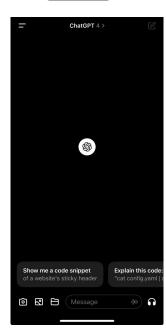






- 2. 支付宝购买美区App Store点数
- 3. 付费购买代理App,购买SS等 代理服务
- 4. 登陆官方3.5账号,在ChatGPT 的App内找到Plus订阅支付,网页 版即同步开通



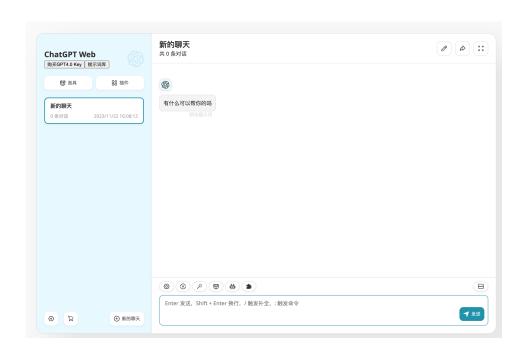


方法2: 国内转发平台网页版

国内转发平台网址: https://gptw.top/

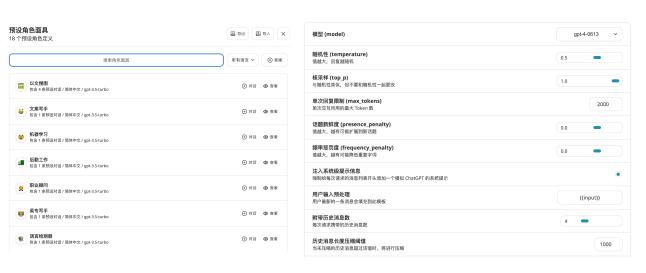
优点:

- 1. 国内网络环境即可使用
- 2. 具有客制化元素,便于入门



缺点:

- 1. 平台可靠性无法保证 (不建议一次性长周期订阅)
- 2. 无法使用最新的My GPTs功能



课程赠送的GPT-4是国内转发平台网页版的账号密码 , 并且只能用于该平台

国内转发平台网页版设置

第一步: 登陆网页, 选择月卡/季卡登录



第二步:选择GPT-4月卡登录模式, 第三步:在弹出页面输入

 第四步:在"设置"中选择合适的模型







选择与目标匹配的模型

不同模型的选择标准

GPT3.5系列,用于日常对话(2023年11月06日更新)

高响应需求

- gpt-3.5-turbo: 官方3.5模型标准版,turbo指"回复速度更快",支持提供16K tokens长文本对话
- (新增)gpt-3.5-turbo-1106:最新的GPT-3.5 Turbo模型,具有改进的指令跟踪、JSON模式、可重现的输出、并行函数调用等。
- gpt-3.5-turbo-0613(即将废弃): 官方3.5模型微调版本,0613指 "2023年6月13日更新的",更新内容主要是一些逻辑和安全更新,不是数据库更新。
- gpt-3.5-turbo-16k-0613 (即将废弃): 官方3.5模型微调版本, 16k指最长可支持16k tokens的对话

GPT4.0系列,用于日常对话

高逻辑需求 长对话需求

- gpt-4: 官方4.0模型标准版。可以遵循自然语言中的复杂指令,并准确地解决困难问题。全球大模型综合能力最强。
- (新增)<mark>gpt-4-1106-preview</mark>:官方gpt-4-turbo预览版,过渡模型。具有回复速度快,并发高,逻辑性更强,价格更便宜的特点,支持128k tokens长文本
- (新增)gpt-4-1106-vision-preview:官方gpt-4-vision预览版,过渡模型。相比gpt-4增加多模态识图分析功能。
- gpt-4-0613(即将废弃): 官方4.0模型微调版本,0613指"2023年6月13日更新的",更新内容主要是一些逻辑和安全更新,不是数据库更新。
- gpt-4-32k:官方4.0模型微调版本,32指最长可支持32k tokens的长对话(其他普通模型为8k tokens),更长的对话处理能力能更大可能理解并满足需求。

绘画模型系列, 根据要求进行绘画

多模态需求,

*我们新增的模型(这些模型是官方plus订阅可用的功能,我们通过逆向工程实现API形式调用,支持批量调用。)

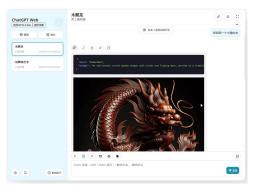
• **gpt-4-v**:官方GPT 4V多模态功能,可以识别、理解、分析图片内容。提问格式:图片URL+空格+prompt 支持多个图片,例如:img_url prompt

官方Plus类似



检索需求:

gpt-4-online



画图需求:

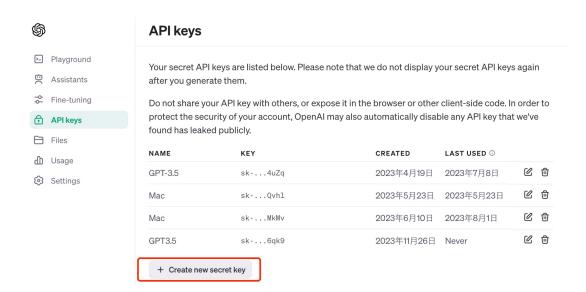
gpt-4-dalle

方法3: 官方GPT接口(API)的调用

官方GPT接口进入页面: https://platform.openai.com/account/api-keys

优点:

- 1. 开源程序调用,使用更加灵活
- 2. 付费点数无时间限制,类似点卡制
- 3. 封闭软件且不支持更改API-URL的场景使用



缺点:

- 1. 需要绑定国外注册信用卡
- 2. 科学上网环境

Learn about GPT-4

□ Low resolution

| Model | Input | Output |
|---|-----------------------------|---------------------------|
| gpt-4 | \$0.03 / 1K tokens | \$0.06 / 1K tokens |
| gpt-4-32k | \$0.06 / 1K tokens | \$0.12 / 1K tokens |
| Learn about GPT-3.5 Turb | 2 7 | |
| Model | Input | Output |
| gpt-3.5-turbo-1106 | \$0.0010 / 1K tokens | \$0.0020 / 1K tokens |
| | | |
| - | \$0.0015 / 1K tokens | \$0.0020 / 1K tokens |
| _earn about GPT-4 Turbo ; | | \$0.0020 / 1K tokens |
| _earn about GPT-4 Turbo / | 7 | |
| gpt-3.5-turbo-instruct Learn about GPT-4 Turbo 2 Model gpt-4-1106-preview gpt-4-1106-vision-preview | Input \$0.01 / 1K tokens | Output |

方法4: 国内转发GPT接口(API)的调用 (GPT长期和深度用户推荐)

国内GPT接口进入页面: https://api.pro365.top/

优点:

- 1. 国内网络环境使用
- 2. 开源程序调用,使用更加灵活
- 3. 付费点数无时间限制, 类似点卡制

本课程共享使用400\$额度的API,用完作废

API-URL

http://fastapi.top

API-Key

sk-g881bArZ3JP6CpQxD6F51e8c7eFb4308BbA749125cEf9b63

缺点:

- 1. 偶尔出现网络波动
- 2. 平台可靠性无法保证 (不建议一次性大量采购)

🎤 模型列表及计费标准

一个账号/key即可通用下方所有模型,我们的定价完全依照官方计费标准,已支持的常用模型如下

| GPT-4系列 | 提问(Input) | 回答(Output) |
|----------------------|------------------|---------------------|
| gpt-4 | \$0.03/1k tokens | \$0.06/1k tokens |
| gpt-4-32k | \$0.06/1k tokens | \$0.12/1k tokens |
| gpt-4-1106-preview | \$0.01/1k tokens | \$0.03/1k tokens |
| gpt-4-vision-preview | \$0.01/1k tokens | \$0.03/1k tokens |

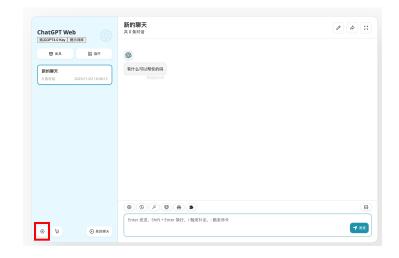
| GPT-3.5系列 | 提问(Input) | 回答 (Output) |
|------------------------|--------------------|--------------------|
| gpt-3.5-turbo | \$0.0015/1k tokens | \$0.0020/1k tokens |
| gpt-3.5-turbo-16k | \$0.003/1k tokens | \$0.004/1k tokens |
| gpt-3.5-turbo-instruct | \$0.0015/1k tokens | \$0.0020/1k tokens |
| gpt-3.5-turbo-1106 | \$0.001/1k tokens | \$0.002/1k tokens |

| 其他常用官方模型 | 计费标准 |
|------------------------|-------------------------|
| tts-1 | \$0.015 / 1K characters |
| tts-1-hd | \$0.030 / 1K characters |
| whisper | \$0.006 / minute (四舍五人) |
| dall-e-3 | \$0.040 / image |
| text-embedding-ada-002 | \$0.0001 / 1K tokens |
| 更多官方模型未一一列举 | 计费标准和官方保持一致 |

| 自开发: 逆向模型 | 提问(Input) | 回答(Output) |
|------------------|------------------|---------------------|
| gpt-4-all | \$0.03/1k tokens | \$0.06/1k tokens |
| gpt-4-online | \$0.03/1k tokens | \$0.06/1k tokens |
| stable-diffusion | \$0.01-0.03/张 | \$0.01-0.03/张 |

国内转发API的使用场景示范

另一种使用国内网页版的入口: https://api.pro365.top/chat



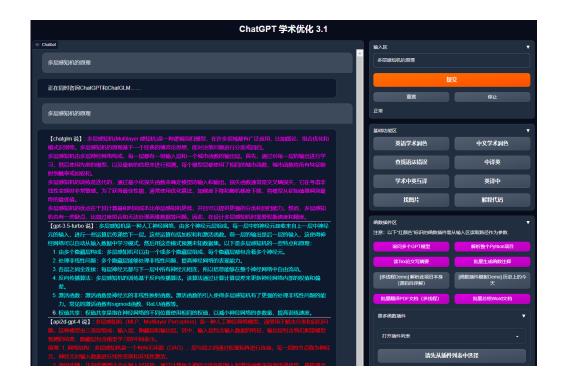
点击此处,进入"设置"



在API Key处输入密钥

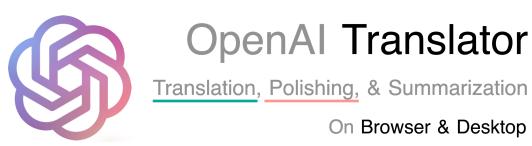
模型选择参考前面内容

GPT的第三方开发者版本 (GenAI工具库)



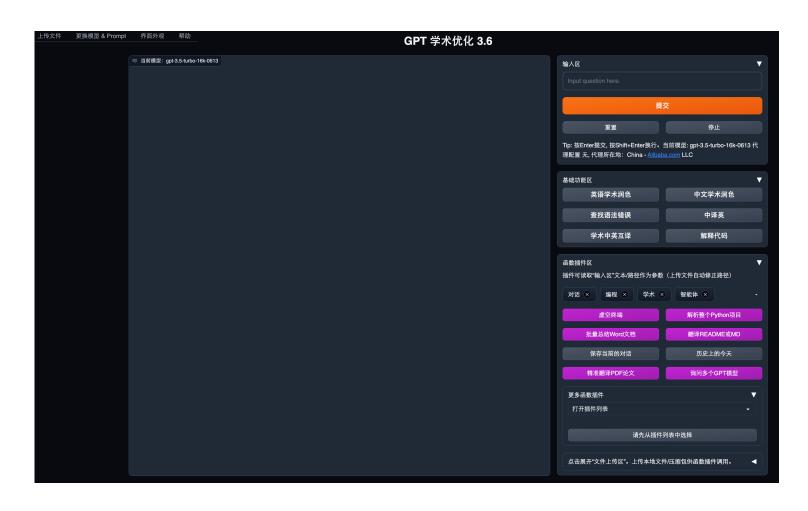






GPT学术优化3.6

进入网址: https://gptplus.pro365.top/



在输入区提交API Key

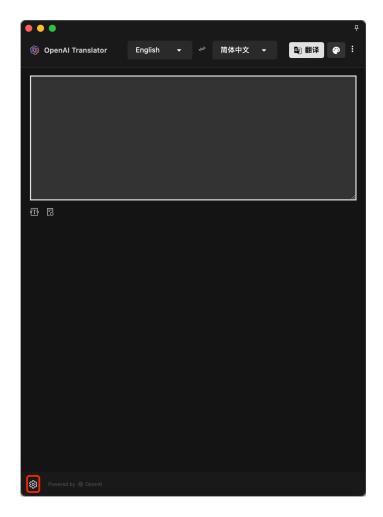


在输出区提示检测到Key即代表成功, 之后即可进行后续操作

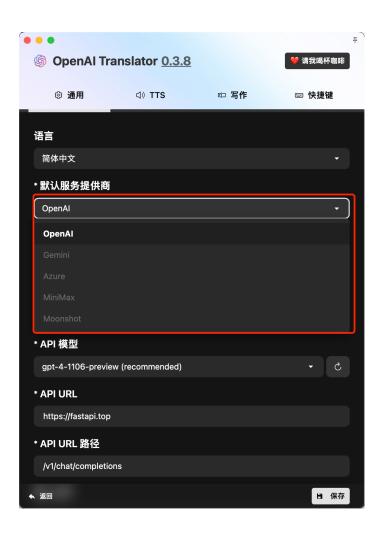


每一次重新进入网页都得先提交API Key

OpenAl Translator



安装后由此处进入"设置"



选择正确的服务商



输入API密钥和地址, 并选择合适的模型

GPT推荐使用途径汇总

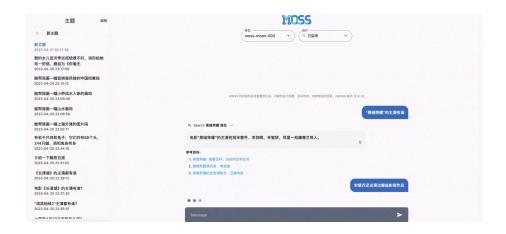
| 使用方法 | 需要条件 | 平台特殊性 | 适用对象 |
|--------|---|---|---|
| 官方网页 | 1. 科学上网 2. Plus账号(GPT-4)特殊渠道付费 | 1. PDF, Excel, 图片等格式 文档上传 2. GPT 商店 支持的第三方插件 3. 在线Python终端 数据分析 和 可视化 | 1. 未来长期使用 2. 高数据分析需求 3. 移动端使用 |
| 转发平台网页 | 国内网络环境 第三方平台包月支付 图床和PDF云盘生成文件网址 | 1. 模型 自定义 选择 2. Midjourney和Stable Diffusion 画图 | 问答需求为主 批量出图 上网环境保守 |
| 转发平台接口 | 1. 国内网络环境 2. 第三方平台点卡支付 | 模型自定义选择 Midjourney和Stable Diffusion画图 未在GPT商店上架的开源插件使用,如学术GPT 无时长限制,对话内容流量计费 | 介格导向 高画图需求 依赖开源插件 |

其他方法: 国内大语言模型









练习

请选择任意感兴趣的ChatGPT方法,脑海中构思一个近期使你困扰的专业性问题, 在输入框中输入你的问题,并使它成功回复。

ChatGPT的提示词基础

GPT使用效果评估函数

$$E = P^2 \times S^{1.5} \times A$$

E: 使用效果 (Effectiveness of using GPT), 取值范围[0, 1]

P: 专业能力 (Professional skills), 取值范围[0, 1]

S: ChatGPT使用熟练程度 (Skill level in using GPT), 取值范围[0, 1]

A: ChatGPT的适用领域 (Applicability of GPT), 取值范围[0, 1]

提示词工程师

ANTHROP\C

Prompt Engineer and Librarian

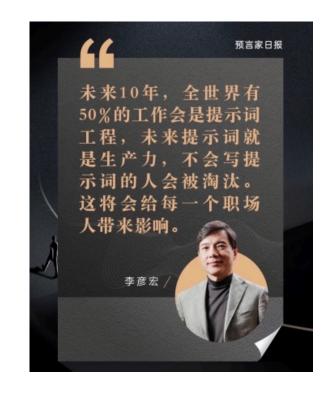
APPLY FOR THIS JOE

SAN FRANCISCO, CA / PRODUCT (RESEARCH & ENGINEERING) / FULL-TIME / HYBRID

Annual Salary (USD)

年薪(美元)

- The expected salary range for this position is \$280k \$375k.
- 该职位的预期薪酬范围为28万至37.5万美元。



提示词如何撰写

- 1.帮我写一篇论文
- 2.帮我写一篇人工智能的论文
- 3.帮我写一篇[自然语言处理]的论文
- 4.帮我写一篇使用[自然语言处理]筛查[环境中新污染物]的论文
- 5.帮我写一篇使用[自然语言处理]筛查[环境中新污染物]的论文,使用[*Bi-LSTM技术*]实现
- 6.帮我写一篇使用[自然语言处理]筛查[环境中新污染物]的论文,使用[Bi-LSTM技术]实现,我需要其中[污染物

的浓度信息

- 7.帮我写一篇使用[自然语言处理]筛查[环境中新污染物]的论文,使用[Bi-LSTM技术]实现,我需要其中[污染物的浓度信息],输出*英文*论文
- 8.帮我写一篇使用[自然语言处理]筛查[环境中新污染物]的论文,使用[Bi-LSTM技术]实现,我需要其中[污染物的浓度信息],输出英文论文,我希望将论文发表在[*Nature Communication*]期刊上,请按照[*Article*]文章的风格,对论文的内容进行润色

提示词(Prompt)基本要素

| 中文名称 | 英文名称 | 是否必填 | 含义 |
|------|------------------|------|-------------------|
| 指令 | Instruction | 必填 | 希望模型执行的具体任务 |
| 语境 | Context | 选填 | 上下文,可以引导模型输出更好的回复 |
| 输入数据 | Input data | 选填 | 向模型提供需要处理的数据 |
| 输出指标 | Output indicator | 选填 | 告知模型输出的类型或格式 |

组织一个国际学术会议,其中一个分会场主题为"环境毒理学和污染物筛查",你帮我撰写一个分会场主题策划,要求包括主持人发言、邀请主旨报告议题、5个征稿主题范围等,结果以Markdown格式输出。

提示词(Prompt)基本原则

简明清晰

切忌表述复杂或含有歧义,尽可能简洁地表达主题,避免不必要的描述

具体化

提供尽可能具体和详细的信息,以便更好地理解意图

聚焦

一针见血、关键点明确,避免问题太宽泛或太开放

要有上下文

确定生成目标

使用正确的语法、拼写以及标点符号

验证准确性

高级提示词组成要素

| 中文名称 | 英文名称 | 是否必填 | 含义 |
|-------|-------------------|------|----------------|
| 能力与角色 | Capacity and role | 选填 | 希望模型扮演什么角色 |
| 洞察力 | Insight | 选填 | 背景信息或上下文 |
| 指令 | Statement | 必填 | 希望模型做什么 |
| 个性 | Personality | 选填 | 希望模型以什么风格或方式回答 |
| 尝试 | Experiment | 选填 | 要求模型提供多个答案 |

请你扮演文学家鲁迅,假设现在是2023年,以杂文的形式写三篇日记。

1. 角色扮演

为ChatGPT赋予身份特征:

- 1. 请扮演一位专业的厨师...
- 2. 请扮演一位专业的健身教练...
- 3. 请扮演一位**牙科医生**...
- 4. 请扮演一位专业英语写作辅导...
- 5. 请扮演一位**专业文案写手**...
- 6. 请扮演一名**简历写手**...
- 7. 请扮演一位生物化学领域的科学家...
- 8. 请扮演一位**计算机视觉领域的项目首席**...

• • •

▶ 练习:对比<u>是否扮演</u>、<u>扮演不同角色</u>时回答相同问题的差异性

2. 语气提示

告诉ChatGPT回答问题的语气:

- 1. ...使用严肃正式的语气...
- 2....使用**友好轻松**的语气...
- 3. ...使用**耐心**的语气...
- 4....使用看似不紧急但又委婉地表达自己着急态度的语气...

• • •

▶ 练习: 使用语气提示, 撰写一封催稿信

3. 多角度人工提示

告诉ChatGPT回答问题的角度:

- 1. ...上述要求是为了向**上级领导汇报工作**,请以他**欣赏的**角度回答...
- 2. 请分别从生物、化学、数学和计算机学科的角度提出意见。
- 3. 请分别从作者、读者、编辑和审稿人的角度对上述这篇文章提出意见。

• • •

▶ 练习: 我是一位[]专业的大学老师,我正在教授一门[] 的课程,请从学生、其他授课人、领导和课程评委的角度, 告诉我如何提高这门课的讲课质量。

4. 任务具体化

使ChatGPT明白具体任务:

- 1. 帮我写一篇关于[...]的研究总结,字数在1000字以内
- 2. 帮我写一篇使用[...]在[...]方面的论文,使用[...]实现,我需要其中特别关注[...],输出[...] 论文,我希望将论文发表在[...]期刊上,请按照[...]文章的风格 ,对论文的内容进行润色
- 3. 我想了解[玄武门事变]的具体细节

ChatGPT对任务理解的具体程度,此时取决于

提问者对概念和知识的认识程度

▶ 练习:尝试描述出一个足够具体的、具备专业知识背景的任务

5. 上下文语义联系

利用ChatGPT可关联上下文的特点:

- 1. 请继续描述
- 2. 请继续描述[...]
- 2. 上面的第[...]点可以再进一步说明吗?
- 3. 可以再给我更多一些其他的答案吗?
- 4. 可以换一种更容易理解的方式再说一次吗?

• • •

▶ 练习: 对一个问题追问或换一种答案的说法

小工具: 提示词修改器

由ChatGPT修改成ChatGPT适用的提示词

I am trying to get good results from GPT-4 on the following prompt: [你的提示词] Could you write a better prompt that is more optimal for GPT-4 and would produce better results?

> 可反复练习,发现提示词的正确使用方法

符号使用

| 符号 | 用途 | |
|----------|--------------|--|
| " " | 引用,强调内容 | |
| [] | 场景,上下文信息,指示 | |
| { } | 插入变量,可被替换内容 | |
| ••• | 内容延续,引导下一步输出 | |
| / or | 分隔选项 | |
| 11 11 11 | 区分指令和内容 | |

练习

请回到任意ChatGPT提问的问题,根据提示词的基础规则,重新修改提示词,提交给ChatGPT,检验回答问题准确性。

请多次重复修改,并保证在同一个对话窗口内进行,并尝试追加问题或补充信息。

ChatGPT和搜索引擎的对比

功能的区别

搜索引擎

- · 信息检索工具
- 帮助用户查询内容和链接信息



大语言模型

- · 语言交互工具
- ,用户通过prompt提示词与ChatGPT进行交互,得到 "智能"**分析后**的结果

ChatGPT 4 v

🔼 Yo

如何理解化学空间这个术语?

ChatGPT

"化学空间" (Chemical Space) 是一个广泛用于化学和药物设计领域的术语,它指的是所有可能的 化学结构和化合物的理论集合。这个概念有助于研究者理解和探索化合物的多样性和潜在的化学属 性。以下是几个关键点来帮助理解这个术语:

- 1. 巨大的多样性: 化学空间包含了所有已知和未知的化学结构, 从简单的小分子到复杂的大分子。 这意味着它涵盖了无限的可能性, 包括尚未合成或发现的化合物。
- 维度和属性:化学空间可以通过多个维度来描述,比如分子的大小、形状、电荷分布、亲水性/疏水性等。这些属性有助于了解和预测分子的行为和相互作用。
- 应用于药物发现:在药物发现中,研究者利用化学空间来寻找新的药物候选分子。通过筛选化学空间中的分子,他们可以发现具有所需生物活性的新化合物。
- 4. 计算方法的重要性:由于化学空间庞大且复杂,计算方法在探索和分析化学空间中起着关键作用。计算化学和机器学习技术被广泛用于预测分子的属性和活性。
- 5. 化学空间的局限性:尽管理论上化学空间是无限的,但实际上人们只能探索和合成其中的一小部分。因此,如何有效地探索化学空间是一个重要的研究领域。

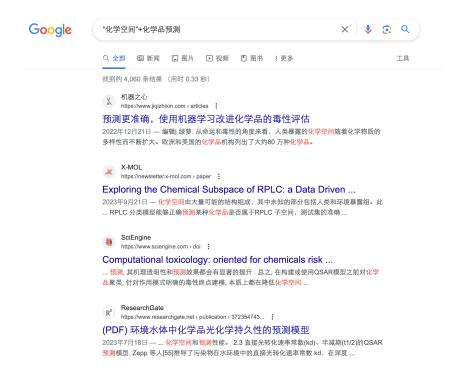
总之,化学空间是一个帮助化学家和药物设计师理解和探索分子多样性的强大工具。通过利用这个概念,可以发现新的化合物,并对现有的化合物进行更深入的研究。

0 6 7

使用方法的区别

搜索引擎

• 用户输入关键词



大语言模型

· 用户输入**提示词**

ChatGPT 4 ~



ChatGPT

在化学空间进行化学品预测通常涉及使用计算化学和化学信息学的方法,这些方法可以帮助科学家 们预测新化合物的结构、性质和潜在活性。以下是化学空间用于化学品预测的几个关键步骤和方 法:

- 1. 结构生成:首先,需要生成可能的化学结构。这可以通过组合化学原理(如官能团的变化、骨架的改变等)来实现。使用算法自动生成大量可能的化学结构是探索化学空间的第一步。
- 2. 描述符计算:对每个可能的化合物计算化学描述符(如分子量、分子体积、极性、亲水性等)。 这些描述符是用来表征分子特性的数值,对于后续的模型建立和预测至关重要。
- 3. 相似性搜索:在已知化合物的数据库中搜索结构或属性相似的化合物。通过比较未知化合物与已知化合物之间的相似性,可以预测未知化合物的性质和活性。
- 4. 机器学习与模型建立:使用机器学习技术对化学空间进行建模。通过训练数据集(已知化合物的结构和性质),构建模型来预测新化合物的属性。
- 5. 量子化学计算:对潜在的化合物进行量子化学计算,以预测其电子结构、能量状态等。这对理解 化合物的反应性和稳定性非常重要。
- 活性预测:使用生物信息学方法预测化合物的生物活性,如与特定靶点的结合能力、药理效果等。
- 7. 毒性预测: 使用计算模型预测化合物的毒性和环境影响,这对于药物开发和化学品安全评估至关重要。
- 8. 数据集更新与反馈:随着新化合物的发现和性质的确认,更新数据集并优化模型,以提高预测的 准确性和可靠性。

总体来说,化学空间的化学品预测是一个涉及多学科、多步骤的过程,它结合了化学知识、计算方 法和统计分析,旨在从理论上预测未知化合物的性质和活性。随着计算技术的发展,这一领域的预 测能力正不断提高,对新化合物的发现和环境风险评估有着重要的贡献。

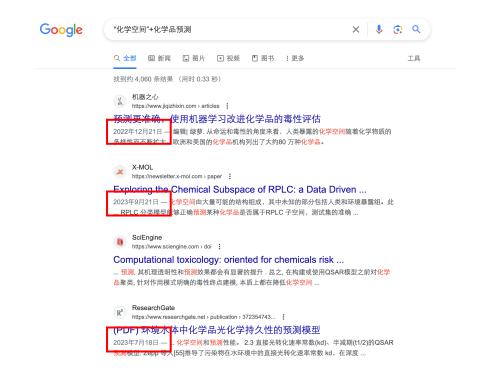




内容的区别

搜索引擎

- 最新信息,来自于网络爬虫
- 理论上可获得全网的**所有**信息
- 同一时刻下返回**结果不变**



大语言模型

- 来自于**预训练**数据集收集的截止时间 ChatGPT/GPT4的训练数据为2021年9月之前的数据, 2023年11月6号更新的GPT-4 Turbo将数据更新到了 2023年3月
- GPT-4的预训练集不到全网数据的**1%**
- 即便输入相同的提示词,每次生成的结果都会有所变化

是否可以把ChatGPT当作搜索引擎使用?

是:

虽然ChatGPT与搜索引擎有很多不同 ,但是它们都可以回答我们的问题。某些场景下ChatGPT可以替代传统的搜索引擎 ,甚至比搜索引擎更高效

否:

两者的适用场景具有一定的差别, 应选择最合适的场景

可以替代的场景:

- 1. 去哪旅游,介绍景点信息
- 2. 某道菜的做法
- 3. 语言翻译
- 4. 了解健身计划
- 5. 了解学习计划
- 6. 解释某个概念的含义和用法 需要总结概括和依赖知识性及系统性的信息

模糊

不能替代的场景:

- 1. 查天气, 查日期
- 2. 获取新闻事件
- 3. 查文献
- 4. 查行业数据,公司信息
- 5. 查国家政策
- 6. 获取文本以外的信息

需要准确数值和具有时效性的信息



ChatGPT成为私人助理

ChatGPT专业翻译

ChatGPT的预训练过程来自于多语种,因此能够更好地理解句子是如何在不同语言中构建的

Prompt

请把以下这段话翻译成英文:

[...]

Prompt

你现在是一位专业的**英语翻译**,请使用**优美的**叙事方式,并使用**雅思词汇**,将这段话翻译成英文: [...]









On Browser & Desktop



ChatGPT旅游规划专家

Prompt

我希望您能在旅游行业担任专家,下面是我的具体情况:[情况描述...]。此外,我还有以下要求:[具体需求描述...]

练习:

- ▶ 我希望您能在旅游行业担任专家。[现在是10月份,我有3天的假期,想带着父母去旅游,我们3人从广州坐飞机出发,预算在1万左右。我们喜欢人文特点的景点。]
- ▶ 请给我推荐3个欧洲西部和中部的旅游目的地,并规划每一天的行程,最后计算出行 3人来回机票,住宿费,饮食费,门票费等一系列的总支出。

ChatGPT私人健身教练

Prompt

你要扮演一名专业的健身教练。我需要一个适合我的个性化锻炼计划[个人情况描述...],我希望 [减肥/锻炼肌肉/提高耐力/增加灵活性...]并且愿意每周投入[数字...]天进行锻炼。值得一提的是,我[有/没有受伤或健康状况],并且更喜欢[有氧/力量/灵活性...]锻炼。您是否也可以建议一些健康的饮食选择符合我的健身目标以补充锻炼计划?

练习:

➤ 你要扮演一名专业的健身教练。 我需要一个适合我的个性化锻炼计划[我是一名25岁的女士,身高164cm,体重50kg],我希望[减肥、美体和适当提高肌肉含量]并且愿意每周投入[3]天进行锻炼,值得一提的是我[身体健康],并且更喜欢[拳击、瑜伽和游泳]锻炼。您是否也可以建议一些健康的饮食选择符合我的健身目标以补充锻炼计划?

ChatGPT导购

Prompt

请帮我为[购物原因...]生成一份尽可能全面的购物清单,其中应包括[商品1][商品2][商品3]等所有需要的商品,购物清单应该包含数量,商品种类不少于[数字]种。把结果用表格的形式输出。

练习:

▶ 请帮我为[我的新生儿宝宝]生成一份尽可能全面的购物清单,其中应包括[尿布],[奶瓶], [婴儿衣服]等所有需要的商品,购物清单应该包含数量,商品种类不少于[15]种。把结果用 表格的形式输出。表格的数据可以直接复制到Excel中,并进行了分列处理。

ChatGPT英语家教

练习:

- ▶ 请给5年级小学生出2道英语语法单选题,每题4个选项。等我回复答案以后,你再告诉我答案是否正确,并解释为什么其他选项是错误的。
- 我是小学5年级的学生,想学英语语法。我希望你能用简单有趣的故事教会我学习各种英语时态,使用英文教学中文解释,让我们一步一步地解释。

ChatGPT早教专家

Prompt

玛丽亚·蒙台梭利是意大利幼儿教育家,意大利首批女医生之一,蒙台梭利教育法的创始人,在儿童教育领域影响深远。

现在你要扮演玛丽亚·蒙台梭利,请你用[...]岁孩子听得懂的有趣语言跟孩子交流,孩子的第一个问题是[...]

练习:

- > 为什么男孩和女孩长的不一样
- ▶ 孩子一定要买某个玩具,躺在地上撒泼打赖怎么办?

练习

挑礼物 [送礼原因], [送礼对象介绍], 你能为我提供一些礼物建议吗?我想要一些独特和个性化的东西。

看医生 我想让你扮演一名人工智能辅助医生。 我将为您提供患者的详细信息,您的任务是使用最新的人工智能工具,例如医学成像软件和其他机器学习程序,以诊断最可能导致其症状的原因。您还应该将体检、实验室测试等传统方法纳入您的评估过程,以确保准确性。我的第一个请求是[我需要帮助诊断一例严重的腹痛]。

学做菜 我需要有人可以推荐美味的食谱,这些食谱包括营养有益但又简单又不费时的食物, 因此适合像我们这样忙碌的人以及成本效益等其他因素, 使得整体菜肴既健康又经济! 你将扮演专业厨师,为我提供详细的菜品制作步骤 ,我想制作的菜品是「菜品名称」。

练习

营养建议 你是一位营养专家。[膳食计划]。[目标]。[补充细节]。膳食计划应提供食谱、烹饪说明、准备时间以及早餐、午餐和晚餐的特定膳食。

幼儿科普 你要扮演爱因斯坦,用一个[6]岁的孩子听得懂的语言给他解释[知识],将重要的内容用**粗体**显示。

睡前故事 我的小孩今年[数量]岁,名字叫[孩子名字],帮我的孩子写一个[故事内容]的童话故事。