

DeepSeek

DeepSeek原理与落地应用

AI肖睿团队

(孙萍、吴寒、周嵘、李娜、张惠军、刘誉)

2025年03月01日

厦门大学大数据百家讲坛



- 北大青鸟人工智能研究院
- 北大计算机学院元宇宙技术研究所
- 北大教育学院学习科学实验室





北京大学
PEKING UNIVERSITY



目录

CONTENTS



人工智能概念辨析



DeepSeek R1



大模型原理



落地应用



北京大学
PEKING UNIVERSITY

PART 01 ▶

人工智能概念辨析

大模型相关术语

- 多模态

- 文本、图片、音频、视频

- AI工具 (国内)

- DeepSeek、豆包、Kimi、腾讯元宝、智谱清言、通义千问、秘塔搜索、微信搜索...

- 通用模型

- 大语言模型 (LLM, Large Language Model)
- 视觉模型 (图片、视频)
- 音频模型
- 多模态模型
-

- 行业模型 (垂直模型、垂类模型)

- 教育、医疗、金融等



大模型的前世今生

•**人工智能**: 让机器具备动物智能, 人类智能, 非人类智能 (超人类智能)

•运算推理: 规则核心; 自动化

•知识工程: 知识核心; 知识库+推理机

•**机器学习**: 学习核心; 数据智能 (统计学习方法, 数据建模)

•常规机器学习方法: 逻辑回归, 决策森林, 支持向量机, 马尔科夫链, ...

•**人工神经网络**: 与人脑最大的共同点是名字, 机制和架构并不一样

•传统神经网络: 霍普菲尔德网络, 玻尔兹曼机, ...

•深度神经网络: **深度学习**

•传统网络架构: DBN, CNN, RNN, ResNet, Inception, ...

•**Transformer架构**: 可以并行矩阵计算 (GPU), 核心是注意力机制 (Attention)

•编码器 (BERT): 多数embedding模型, Ernie早期版本, ...

•混合网络: T5, GLM

•解码器 (GPT): **大语言模型 (LLM)**, 也是传统的多模态模型的核心

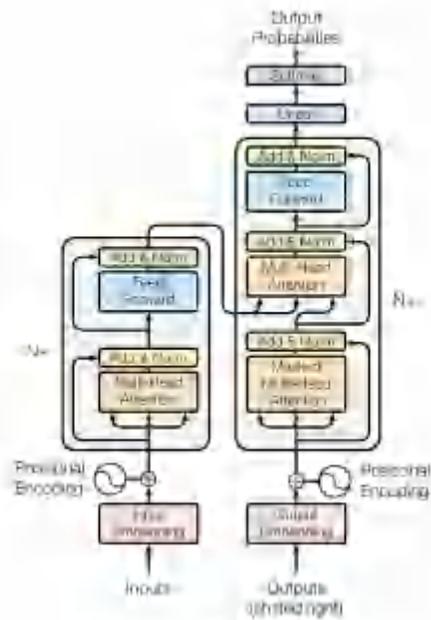
•**生成式人工智能 (GenAI)**: **AIGC**

•**DeepSeek**、Qwen、GLM、Step、MiniMax、hunyuan、kimi、...

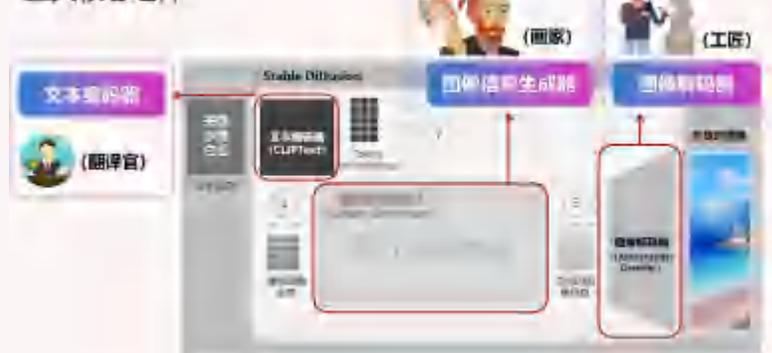
•OpenAI GPT (**ChatGPT**)、Claude、Llama、Grok、...

•**Diffusion架构**: 主要用于视觉模型 (比如Stable Diffusion、DALI), 现在也开始尝试用于语言模型

•Diffusion+Transformer架构: 例如**Sora**的DiT (加入Diffusion的视觉模型), 部分新的多模态模型架构



三大核心组件



大模型的发展阶段

关键进展



生成模型



推理模型

生成模型与推理大模型的对比



比较项	OpenAI GPT-4o (生成模型)	OpenAI o1 (推理模型)
模型定位	专注于通用自然语言处理和多模态能力，适合日常对话、内容生成、翻译以及图文、音频、视频等信息处理、生成、对话等。	侧重于复杂推理与逻辑能力，擅长数学、编程和自然语言推理任务，适合高难度问题求解和专业领域应用。一般是在生成模型的基础上通过RL方法强化CoT能力而来
推理能力	在日常语言任务中表现均衡，但在复杂逻辑推理（如数学题求解）上准确率较低。	在复杂推理任务表现卓越，尤其擅长数学、代码推理任务。
多模态支持	支持文本、图像、音频乃至视频输入，可处理多种模态信息。	当前主要支持文本输入，不具备图像处理等多模态能力。
应用场景	适合广泛通用任务，如对话、内容生成、多模态信息处理以及多种语言相互翻译和交流；面向大众市场和商业应用。	适合需要高精度推理和逻辑分析的专业任务，如数学竞赛、编程问题和科学研究；在思路清晰度要求高的场景具有明显优势，比如采访大纲、方案梳理。
用户交互体验	提供流畅的实时对话体验，支持多种输入模态；用户界面友好，适合大众使用。	可自主链式思考，不需要太多的过程指令，整体交互节奏较慢。



PART 02 ▶

D e e p S e e k R 1

公司成立背景与发展历程

DeepSeek, 全称杭州深度求索人工智能基础技术研究有限公司, 成立于2023年7月17日, 是一家创新型科技企业, 专注于人工智能基础技术的研究与开发

大语言模型(LLM)的创新应用

DeepSeek专注于开发先进的大语言模型(LLM)和相关技术, 旨在通过这些技术推动人工智能在多个领域的应用和创新

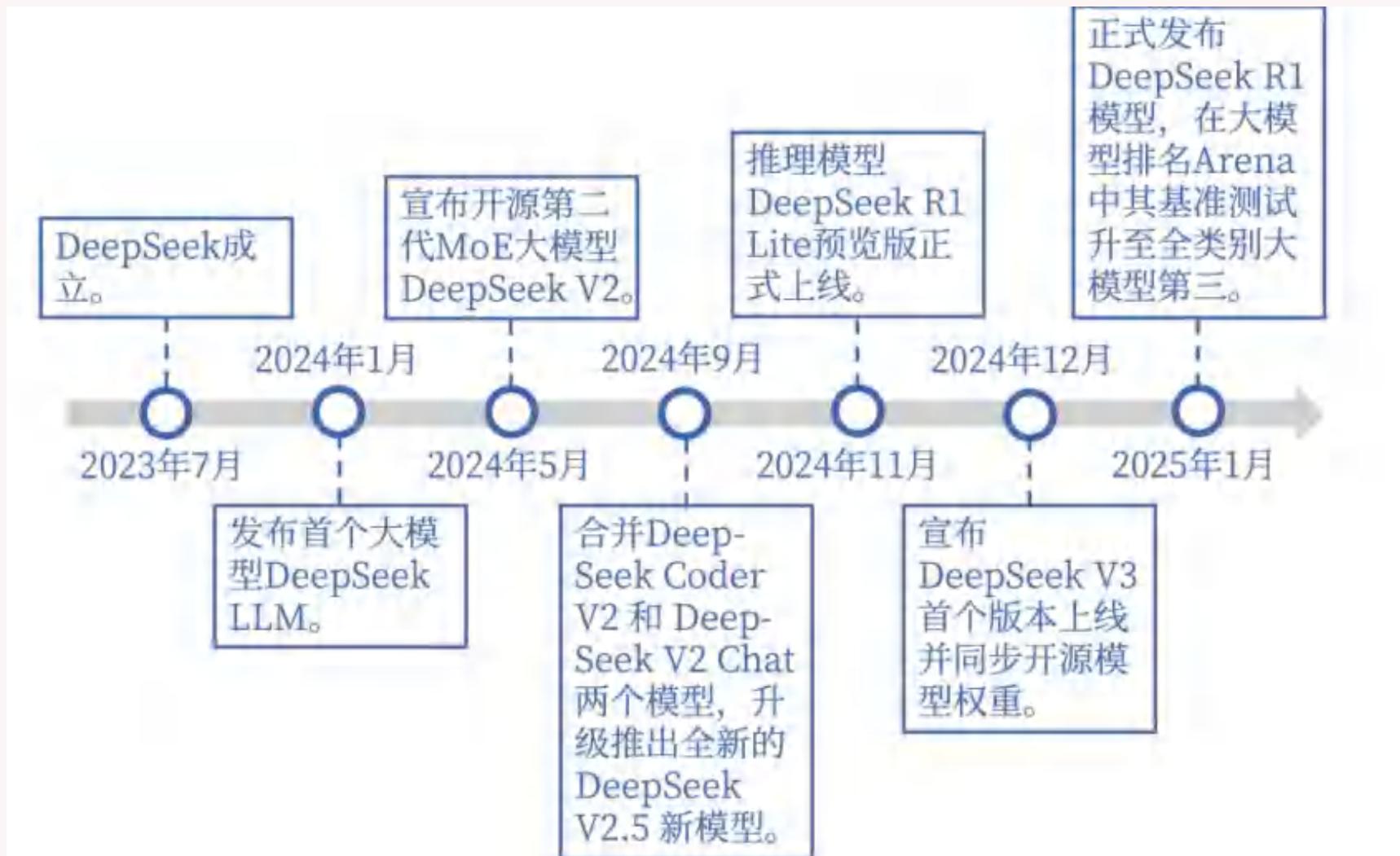
投资者背景与市场定位

作为由知名私募巨头幻方量化孕育而生的公司, DeepSeek获得了强大的资金支持和行业影响力, 幻方量化与九坤投资、明汭投资、灵均投资并称量化私募领域的“四大天王”, 管理资金规模均超过600亿元。这为DeepSeek提供了清晰的市场定位和投资者背景

里程碑

2025年1月20日推出DeepSeek-R1推理模型

DeepSeek模型系列



DeepSeek最新的生成模型和推理模型版本对比



比较方面	生成模型 (V3)	推理模型 (R1)
设计初衷	想要在各种自然语言处理的任务中都能表现好，更通用	重点是为了搞定复杂的推理情况，比如深度的逻辑分析和解决问题
性能展现	在数学题、多语言任务还有编码任务里表现不错，像Cmath能得90.7分，Human Eval编码任务通过率是65.2%	在需要逻辑思考的测试里很棒，比如DROP任务F1分数能达到92.2%，AIME 2024的通过率是79.8%
应用的范围	适合大规模的自然语言处理工作，像对话式AI、多语言翻译还有内容生成等等，能给企业提供高效的AI方案，满足好多领域的需求	适合学术研究、解决问题的应用和决策支持系统等需要深度推理的任务，也能拿来当教育工具，帮学生锻炼逻辑思维

DeepSeek R1

OpenAI o1

OpenAI o3-mini

Gemini 2.0

Grok3

Kimi 1.5

为什么火：能力突破、开源、低成本、国产化

基础能力：进入推理模型阶段，跻身全球第一梯队

推理能力跃升：DeepSeek大模型核心技术突破，实现复杂推理任务的精准处理与高效执行，覆盖多模态场景应用。

国际竞争力对标：模型综合性能跃居全球第一梯队，技术指标与国际顶尖水平（如GPT系列、Claude等）直接对标，奠定国产大模型的行业标杆地位。

核心加分项：开源、低成本、国产化

• 开源：技术共享，生态共建

全量开源训练代码、数据清洗工具及微调框架，开发者可快速构建教育、金融、医疗等垂直领域应用，推动社区协同创新。

• 低成本：普惠企业级AI应用

做了大量的模型架构优化和系统工程优化。

训练成本仅\$557w：显著低于行业同类模型，打破高价壁垒。

推理成本降低83%：千亿参数模型适配中小企业需求，加速商业化落地。

• 国产化：技术自主，缩短差距

将国产模型与美国的代际差距从3-5年缩短至3-5个月，突破“卡脖子”技术瓶颈。

构建多行业专属模型矩阵，全面支持国内产业智能化升级。

DeepSeek以“推理能力+第一梯队性能”为核心基础，叠加：开源开放、超低成本、国产自主研发三大优势，不仅实现技术代际跨越，更推动AI技术普惠化与国产化生态繁荣，成为全球大模型赛道的重要领跑者。

DeepSeek的算法和算力突破

- DeepSeek R1达到了跟o1相当、或者至少接近的推理能力，且将推理过程可视化
- 它做到这个水平只用到少得多的资源，所以价格十分便宜
- 它是完全开源的并且还发布论文，详细介绍了训练中所有的步骤和窍门
- DeepSeek深度求索公司是一家纯粹的中国公司

Deepseek官网地址:

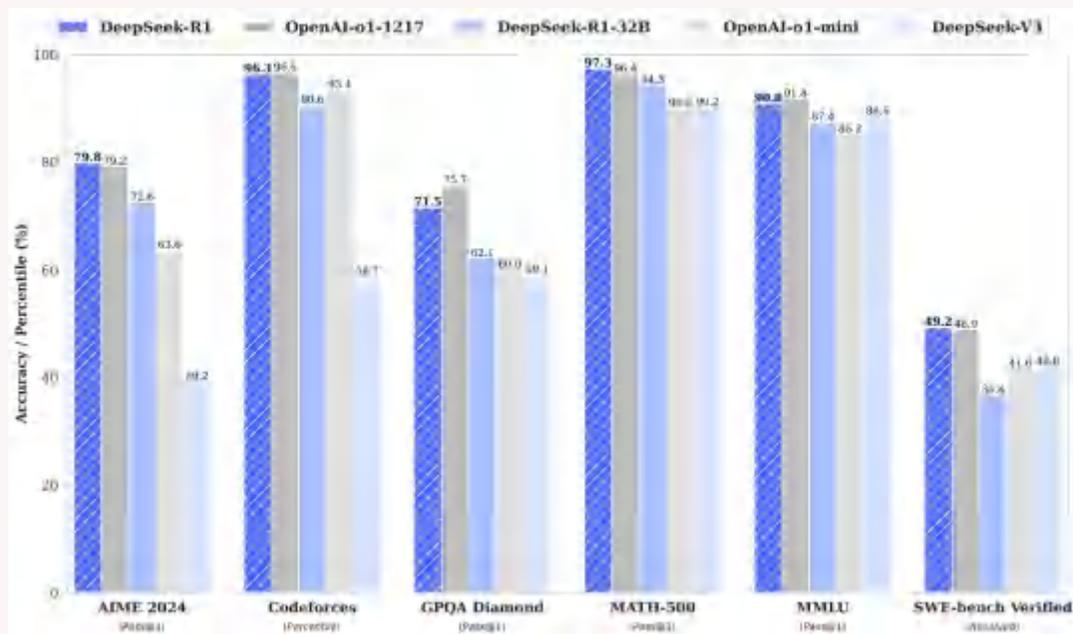
<http://ai.com>

<https://chat.deepseek.com>

DeepSeek-R1 训练技术全部公开, 论文链接:

[https://github.com/deepseek-ai/DeepSeek-](https://github.com/deepseek-ai/DeepSeek-R1/blob/main/DeepSeek_R1.pdf)

[R1/blob/main/DeepSeek_R1.pdf](https://github.com/deepseek-ai/DeepSeek-R1/blob/main/DeepSeek_R1.pdf)



混合专家
MOE

直接硬件编程
PTX

通讯优化
DualPipe

多头潜在注意力
MLA

混合精度训练
FP8

并行训练框架
HAI

强化学习
GRPO

多Token预测
MTP

测试时计算
TTC

对AI行业的重大影响



01

打破垄断

DeepSeek-R1以低成本和开源特性打破以往头部企业巨头割据局面

02

价格下调

DeepSeek-R1的API定价仅为行业均价的1/10,推动了中小型企业低成本接入AI,对行业产生了积极影响

03

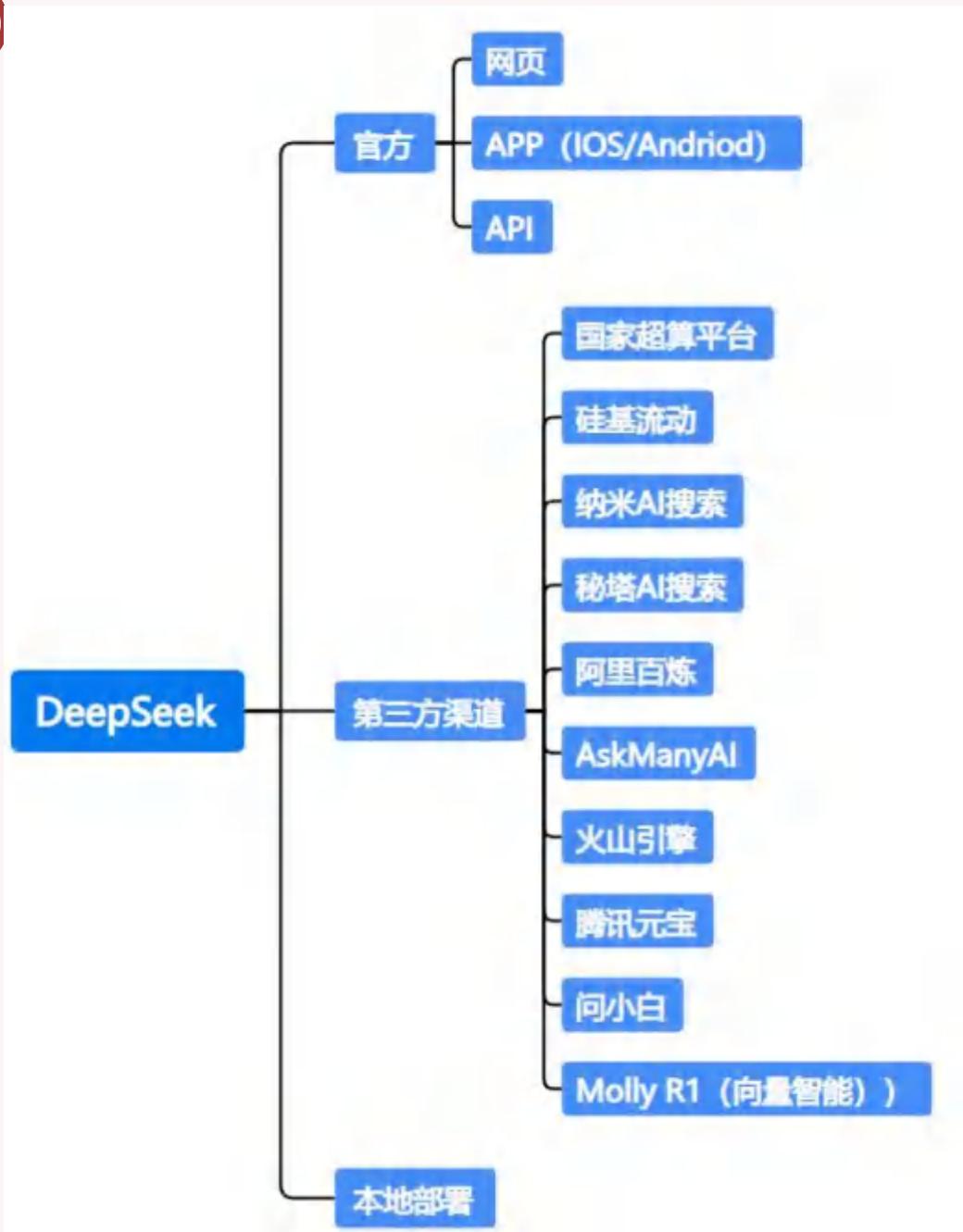
推动创新

DeepSeek-R1促使行业开始从“唯规模论”转向更加注重“性价比”和“高效能”方向



模型有三种使用方式：在软件系统中调用官方API、模型微调、直接使用。前两种涉及到IT技术比较多。这里讲的是直接使用的三种方法，适合普通用户。

DeepSeek直接使用 😊



DeepSeek官方-开袋即食 😊

官网: <https://chat.deepseek.com/>



API: <https://platform.deepseek.com/>



手机APP

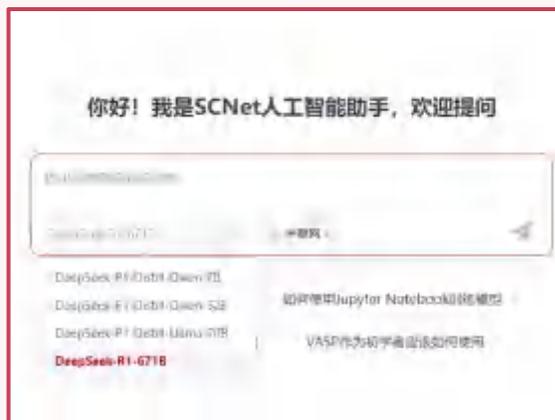
IOS用户:

<https://apps.apple.com/cn/app/deepseek/id6737597349>

安卓用户:

<https://app.mi.com/details?id=com.deepseek.chat>

DeepSeek第三方通道



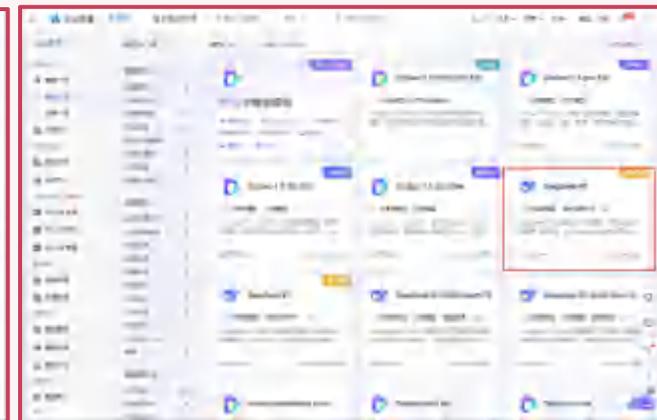
国家超算平台 (网页)
<https://chat.scnet.cn/>



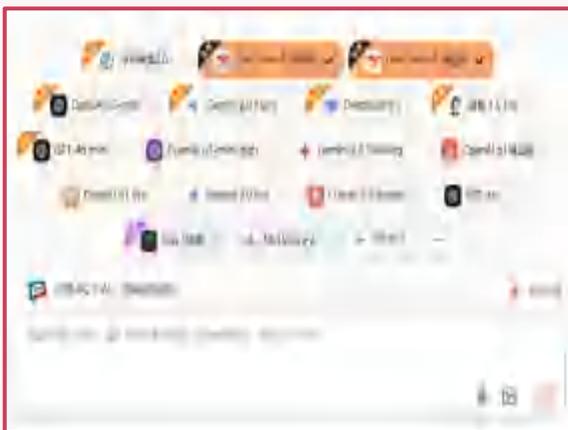
硅基流动 (网页+API) :
<https://cloud.siliconflow.cn/i/9VzvgYQL>



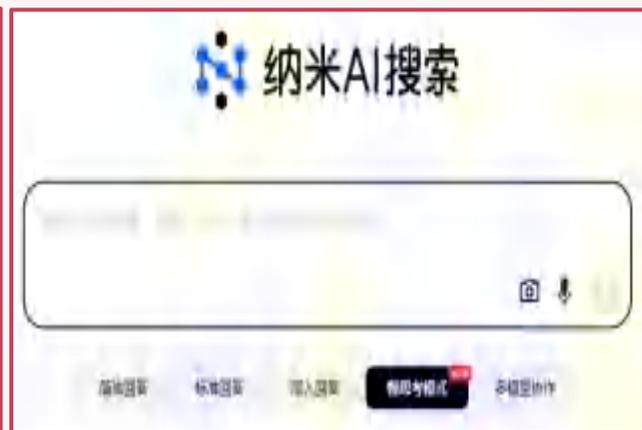
阿里百炼 (API) :
<https://account.aliyun.com/>



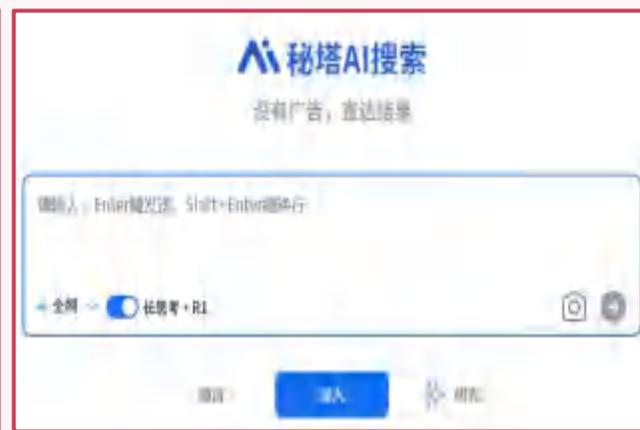
火山引擎 (API) :
<https://www.volcengine.com/product/ark>



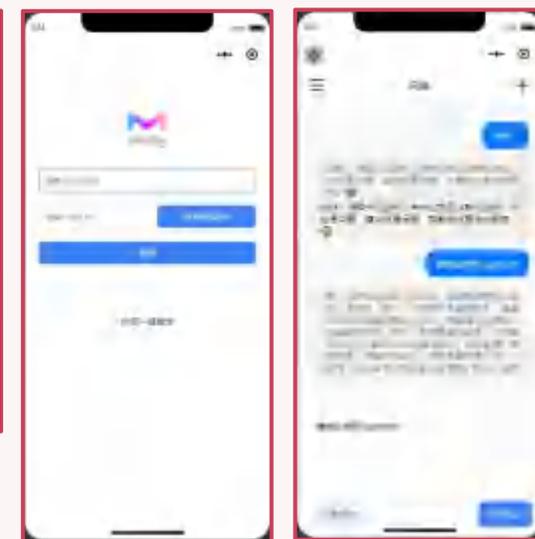
AskManyAI (网页) :
<https://chat.scnet.cn/>



纳米AI搜索 (网页) :
<https://www.n.cn/>



秘塔AI搜索 (网页) :
<https://metaso.cn/>



Molly R1 (小程序)

DeepSeek私有化部署

模型的私有化部署的方式:

- Ollama部署: 个人本地部署【推荐】，方便快捷，适用于蒸馏模型
- vLLM部署: 生产、开发、垂直领域私有化部署，精度可控，更专业
- 其它

Model	Base Model	特点
DeepSeek-R1-Distill-Qwen-1.5B	Qwen2.5-Math-1.5B	蒸馏模型，能力稍弱 实际上是增加了推理能力的Qwen模型和Llama模型
DeepSeek-R1-Distill-Qwen-7B	Qwen2.5-Math-7B	
DeepSeek-R1-Distill-Llama-8B	Llama-3.1-8B	
DeepSeek-R1-Distill-Qwen-14B	Qwen2.5-14B	
DeepSeek-R1-Distill-Qwen-32B	Qwen2.5-32B	
DeepSeek-R1-Distill-Llama-70B	Llama-3.3-70B-Instruct	
DeepSeek-R1-671B	DeepSeek-V3-Base	满血版，能力最强

部署DeepSeek-R1 满血版的算力要求和性能

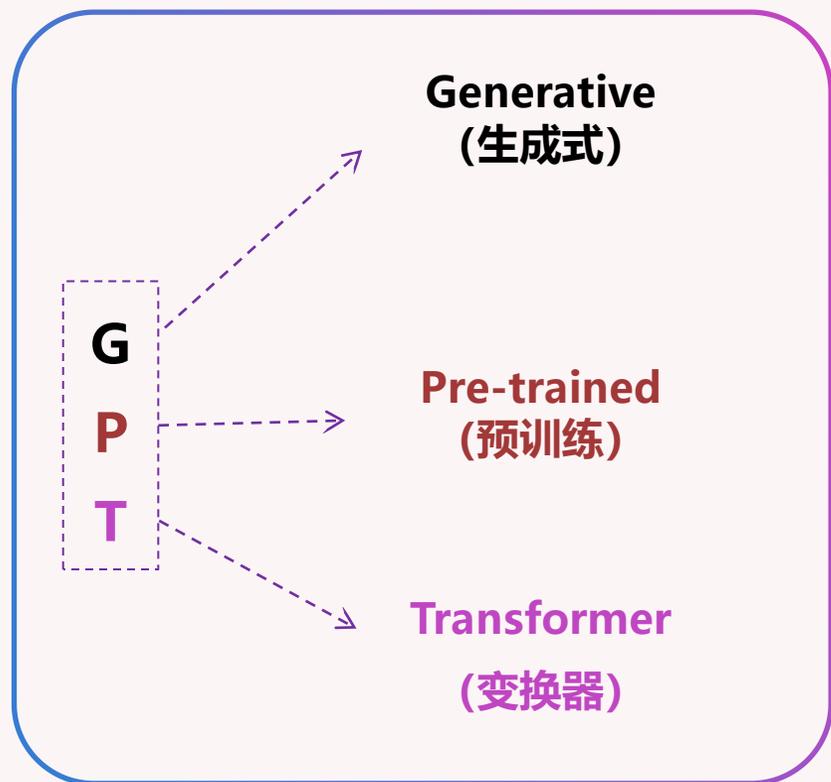
模型	算力型号	每套卡数	每套算力 (FP16/TFLOPS)	每套显存 (GB)	模型精度	含算力部署总价 (万/套/月)
 DeepSeek R1 671B	NV H800	16	31664	1280	FP8	18
	NV H200	8	15832	1128	FP8	11
	NV H20	16	2368	1536	FP8	8
	昇腾 910B	32	10016	2048	BF16	12
 DeepSeek V3 671B	NV H800	16	31664	1280	FP8	18
	NV H200	8	15832	1128	FP8	11
	NV H20	16	2368	1536	FP8	8
	昇腾 910B	32	10016	2048	BF16	12



PART 03 ▶

模 型 原 理

GPT工作原理-1



LLM: Large Language Model

1. 收到提示词

示例: “今天天气不错, 我决定”

2. 将输入拆分为token

[“今”, “天”, “天”, “气”, “不”, “错”, “,”, “,”, “我”, “决”, “定”]

3. 采用Transformer架构处理token

- 理解token之间的关系
- 识别提示词的整体含义

4. 基于上下文预测下一个token

- 为可能的单词分配概率分数
- 示例: { “去” :0.7, “停”:0.2, “站”:0.1 }

5. 根据概率分数选择标记

示例: “去”

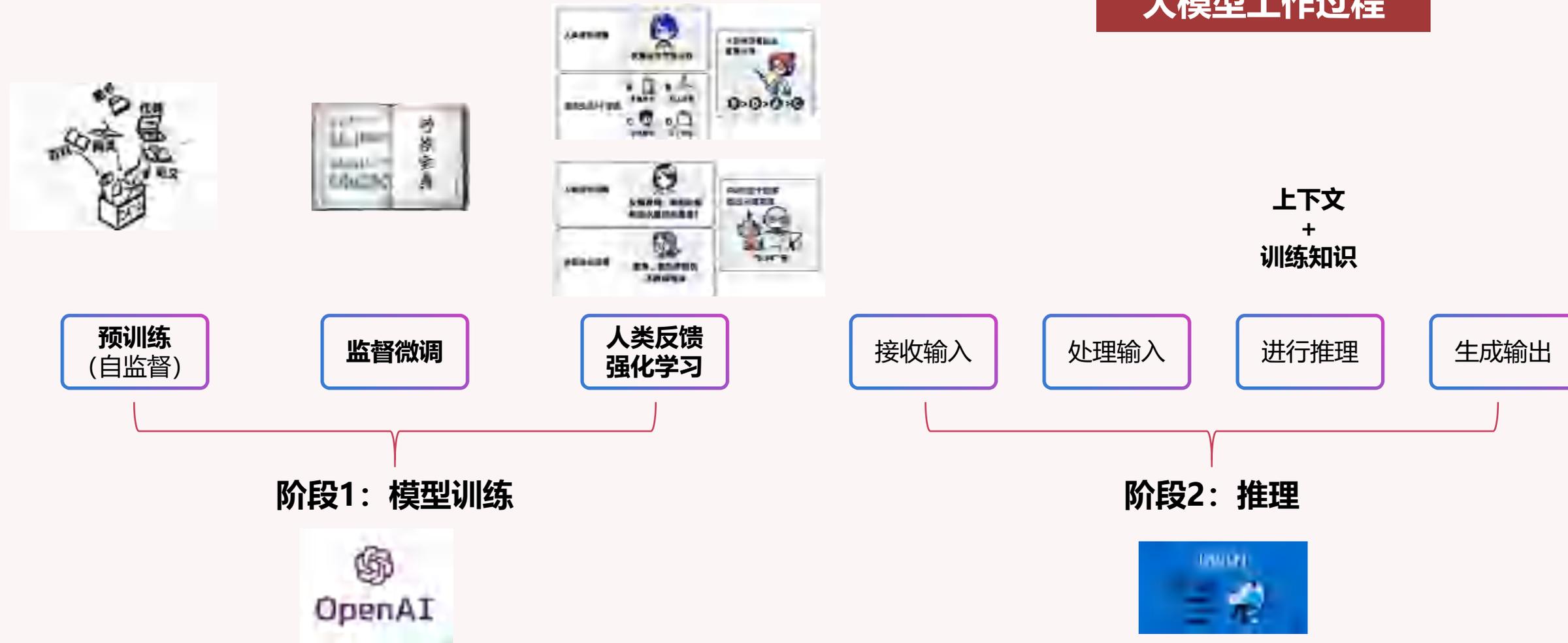
概率预测+文字接龙

自回归: 重复步骤4和步骤5直到形成完整的句子

示例: 今天天气不错, 我决定去公园

GPT工作原理-2

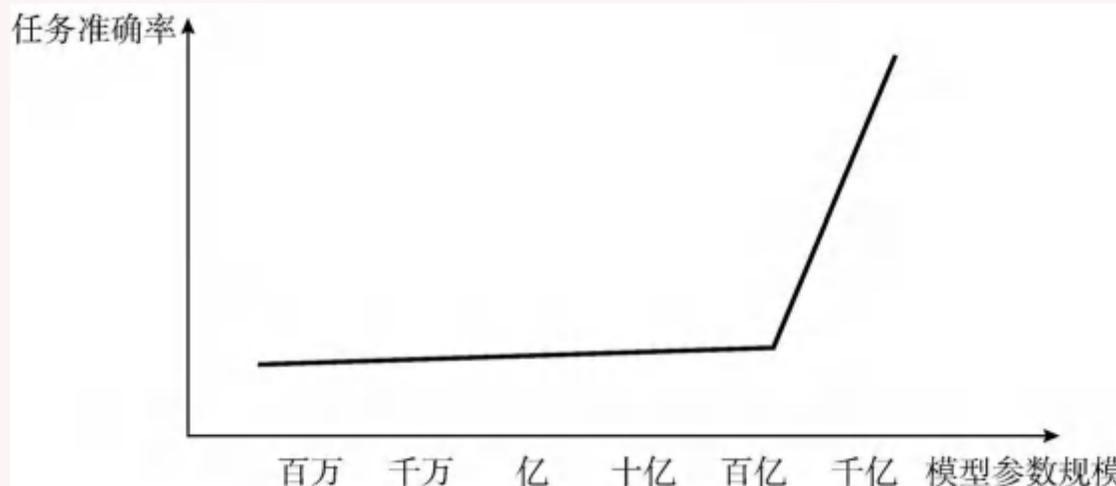
大模型工作过程



GPT工作原理-3

数据来源	说明
维基百科	在线百科,严谨
图书	经典为主的古登堡计划和自助出版平台Smashwords等
杂志期刊	论文: ArXiv等
链接	WebText,Reddit
Common Crawl	开源项目,爬取互联网所有数据
GitHub	程序员聚集地
合计	<ul style="list-style-type: none">700 多GB,约有19万套四大名著的阅读量5 000亿左右的token数量。(13 万亿token:gpt4)100个标记大约等于75个英语单词

训练语料 (gpt-3)



典型的新技能学习曲线: 规模到达临界点之后才会迅速增长

模型参数: 1.8万亿参数 (GPT-4)

GPT-4o

上下文窗口大小: 8192个token(标记)

生成模型的优势与劣势

优势

- 语言理解和生成能力
- 世界知识能力
- 一定的推理能力

VS

劣势

- 幻觉（生成错误答案）
- 知识库有限
- 上下文窗口限制

推理模型 (DeepSeek-R1)工作原理



思维链 (Chain of Thought)

让模型进行慢思考

强化学习 (Reinforcement Learning)

让模型自我探索和训练

蒸馏 (Distillation)

在不损失能力的情况下缩小模型

- DeepSeek R1论文: https://github.com/deepseek-ai/DeepSeek-R1/blob/main/DeepSeek_R1.pdf
- DeepSeek R1论文图解: <https://zhuanlan.zhihu.com/p/20844750193>



北京大学
PEKING UNIVERSITY

PART 04 ▶

落地应用



DeepSeek让人类与AI对话更加简单，提示词让对话质量提升立竿见影，适合所有层级使用者快速上手。

DeepSeek的深度思考模式基于DeepSeek R1模型，是一种推理模型。相比较于传统的生成模型，使用DeepSeek R1模型时的提示词技巧有较大的变化。

DeepSeek提示词技巧



清晰的表达!!



把AI当人看!!

1. DeepSeek提示词技巧-真诚+直接

传统

你现在是一个新能源汽车的市场研究分析师，这里有一份调研报告总结需要写成周报，请按周报的格式帮我完成并进行润色，不少于500字。



DeepSeek （真诚+直接）

帮我把这份报告**包装**一下，我要写成周报**给老板看**，老板很**看重数据**。



2.DeepSeek提示词技巧-通用公式

我要 (做) ** , 要给用, 希望达到**效果, 但担心**问题**

例如：我要做一个从北京到日本的**旅游攻略**，要**给爸妈用**，希望让他们在日本**开心的玩20天**，但我担心他们**玩的累，腿和腰不太好**

我要做一个从北京到日本的旅游攻略，要给爸妈用，希望让他们在日本开心的玩20天，但我担心他们玩的累，腿和腰不太好。

√ 先抓住总原则

目的：帮爸妈安排一个为期20天的日本行程，重点考虑舒适度、安全性和便利性，让他们玩得开心且轻松。爸妈年龄较大，行动不便，行李搬运和长时间步行是主要顾虑。

需求：希望行程节奏舒缓，景点之间距离不要太远，交通要便利，避免长时间步行和拥挤的地方。比如东京、京都、奈良、大阪这些地方，交通方便，景点集中，比较适合爸妈。爸妈对历史文化感兴趣，也喜欢自然风光，但不喜欢太累的活动。

限制条件：爸妈的预算有限，希望性价比高。爸妈对日本文化了解不多，需要一些基础的日语提示和注意事项。爸妈的出发时间是10月，天气凉爽，适合户外活动。

输出要求：行程要清晰明了，方便爸妈照着走。景点介绍要简洁易懂，重点突出。交通方式要具体，比如火车、巴士、出租车等。还要提供一些实用的贴士，比如换汇、语言、安全等。

√ 全面非常考虑o(∩_∩)d

总体：行程要宽松，交通便利，景点之间距离近，适当安排休息和放松的活动，确保爸妈能享受旅行而不感到疲惫。为爸妈安排日本20天行程时，需重点考虑舒适度和节奏把控。以下是根据他们的需求设计的慢节奏攻略，融合文化体验与自然风光，兼顾交通便利性：

可以加入一些轻松的节目，比如温泉、茶道、和服体验等。爸妈对日本文化感兴趣，可以适当安排一些文化体验活动，但要注意时间安排，避免太累。爸妈的出发时间是10月，天气凉爽，适合户外活动。

此外，爸妈的预算有限，希望性价比高。爸妈对日本文化了解不多，需要一些基础的日语提示和注意事项。爸妈的出发时间是10月，天气凉爽，适合户外活动。

输出要求：行程要清晰明了，方便爸妈照着走。景点介绍要简洁易懂，重点突出。交通方式要具体，比如火车、巴士、出租车等。还要提供一些实用的贴士，比如换汇、语言、安全等。

最后，行程的节奏要慢，避免频繁更换城市。比如东京5天，京都5天，大阪4天，箱根2天，奈良和广岛各1天。这样每个地方停留时间较长，减少搬运行李和赶路压力。同时，城市之间用新干线连接，节省时间。

需要检查每个城市的景点是否适合父母，比如奈良公园虽然可以喂鹿，但公园较大，可能需要租用轮椅或者选择近距离的观景点。广岛的宫岛水上神社可能需要预留充足时间，避免走太多路。

总之，行程要宽松，交通便利，景点之间距离近，适当安排休息和放松的活动，确保爸妈能享受旅行而不感到疲惫。

为父母安排日本20天行程时，需重点考虑舒适度和节奏把控。以下是根据他们的需求设计的慢节奏攻略，融合文化体验与自然风光，兼顾交通便利性：

总体原则

1. 城市间移动间隔3-4天，减少行李搬运
2. 每日核心景点不超过2个，午休2小时
3. 全程使用IC卡+新干线+出租车组合交通
4. 优先选择带温泉的日式旅馆

行程概览

东京 (5天) → 箱根 (2天) → 京都 (6天) → 大阪 (4天) → 奈良 (1天) → 广岛 (2天)

正式回答

2.DeepSeek提示词技巧-通用公式

DeepSeek输出的 日程安排清晰明确

贴心准备清单 甚至细致考虑到了 每日5000步

详细安排

Day1-5 东京深度游

住宿: 银座高星 (三井花园酒店, 直通地铁)

交通: 出租车+地铁

美食: 人形町今半 (寿喜烧)、筑地市场寿司大

行程亮点:

- 浅草寺+隅田川游船 (避开仲见世通高楼层)
- 皇居东御苑电动观光车游览
- 六本木观景台下午茶 (替代登高行程)
- 台场海滨公园无人驾驶列车体验
- 自由活动日: 可选三鹰之森美术馆/上野动物园

Day6-7 箱根温泉疗养

交通: 根茎特快列车 (指定席)

住宿: 强罗花房 (私人露天温泉)

特色体验:

- 与宫崎骏动画声之闻 (购买头等舱座位)
- 雕刻之森美术馆户外展区电动代步车租赁
- 每日2次私汤预约 (缓解肌肉疲劳)

①

Day8-13 京都文化之旅

住宿: 祇园附近 (日式旅馆含早晚膳)

交通: 出租车+包车半日游

慢游路线:

- 清水寺特别早朝参拜 (避开人潮)
- 伏见稻荷大社山脚千本鸟居拍照
- 岚山小火车+保津川游船 (座位预约)
- 二条城无障碍通道参观
- 西阵织会馆和服轻量化体验 (1小时)
- 哲学之道轮椅友好路线

Day14-17 大阪美食探索

住宿: 难波站直达酒店

美食重点: 蟹道乐本店、黑门市场早市、法善寺横丁

轻松玩法:

- 大阪城公园观光船环游
- 梅田蓝天大厦空中庭院日落时分
- 海游馆+天保山摩天轮联票
- 心斋桥药妆店集中采购日

②

Day18 奈良轻松游

交通: 大阪往返近铁特急

特别安排:

- 春日大社参道人力车体验
- 若草山脚限定喂鹿区 (避开陡坡)
- 东大寺中门拍照 (无需进入大殿)

Day19-20 广岛人文之旅

交通: 新干线Green Car指定席

核心体验:

- 宫岛严岛神社涨潮时刻摄影
- 和平纪念公园电动导览车
- 广岛烧老店「みっちゃん」本店

贴心准备清单

- 医疗: 携带护腰坐垫、日本久光贴 (可到当地购买)
- 通讯: Docomo 30日上网卡 (机场领取)
- 辅助: 预约各城市「駅伝」行李托运服务
- 应急: 保存皇居等大型景点的轮椅借用信息

每日作息建议

- 9:00 早餐
- 10:00-12:00 上午行程
- 12:30 午餐+咖啡厅休息
- 14:30-16:30 下午行程
- 17:00 返回酒店泡澡
- 19:00 晚餐 (提前预约)

④

该行程实测步行量控制在每日5000步以内。重要景点均有无障碍通道。建议提前3个月预定特色住宿和米其林餐厅, 使用Japan Travel等App实时查看景区拥挤度。

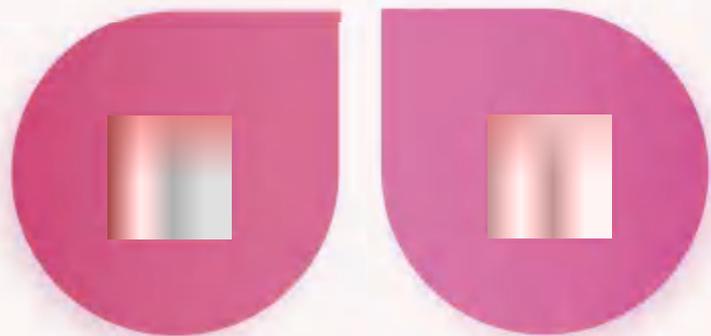
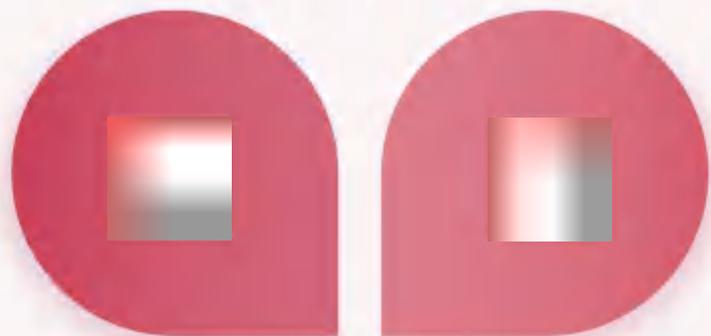
2.DeepSeek提示词技巧-通用公式

任务

做什么

目标

期望效果



背景

给谁用

负面限定

担心的问题

“内心戏”

3. DeepSeek提示词技巧-说人话

适合场景：科研，了解新事物

为了避免DeepSeek的回答过于官方、专业，可以尝试这三个字“说人话”

你问：什么是“波粒二象性”，DeepSeek大概率会给出专业且看不懂的回答，和百度百科差不多。但如果给ta一句“说人话”，ta就会生动形象的做一些举例

说人话

波粒二象性是量子力学中的核心概念，指微观粒子（如光子、电子等）同时表现出波动性和粒子性的现象。以下是其关键点：

1. 历史背景

粒子说 vs 波动说

牛顿曾提出光的“微粒说”，认为光由微小粒子组成。

惠更斯等人则认为光是波动（类似水波）。

19世纪，杨氏双缝实验等证实了光的波动性（干涉、衍射）。

爱因斯坦的光量子假说（1905）

解释光电效应时提出光由“光子”（粒子）组成，重新引入粒子性。

德布罗意物质波理论（1924）

提出所有物质粒子（如电子）都具有波动性，波长由动量决定 ($\lambda = h/p$)。

好的！用大白话解释：

波粒二象性就是说：

微观世界的小东西（比如光、电子）有时候像小球，有时候像水波，具体像啥，取决于你咋看它。

举个栗子 🌰：

1. 光像粒子：

你拿紫外线照金属，能把电子“砸”出来（光电效应），这时候光像一堆小炮弹（光子），每个炮弹带一份能量——这就是光的“粒子性”。

2. 光像波：

如果让光穿过两道缝隙，屏幕上会出现明暗条纹（干涉），就像水波互相打架后的样子——这就是光的“波动性”。

4. DeepSeek提示词技巧-反向PUA

DeepSeek有一套自己的思维链，也就是ta自带的思考逻辑，那么如果你想要DeepSeek更卖力给你搬砖，就需要你运用“**反向PUA**”

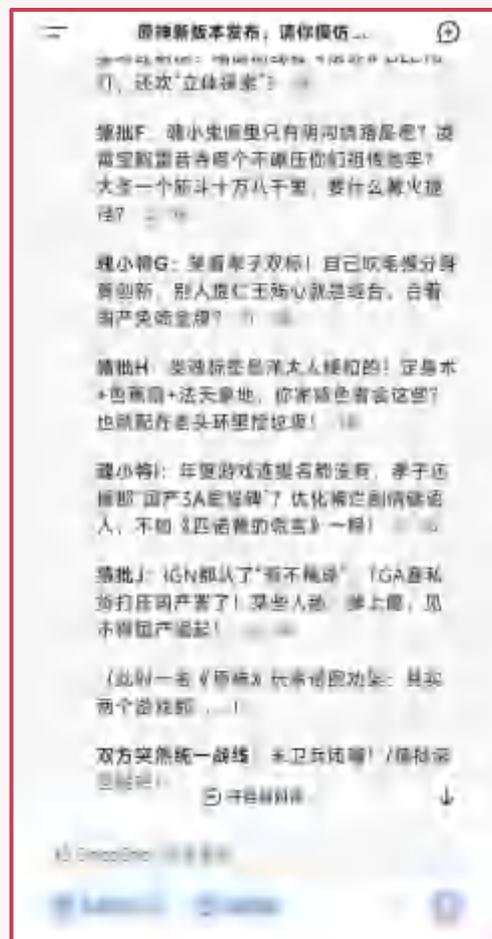
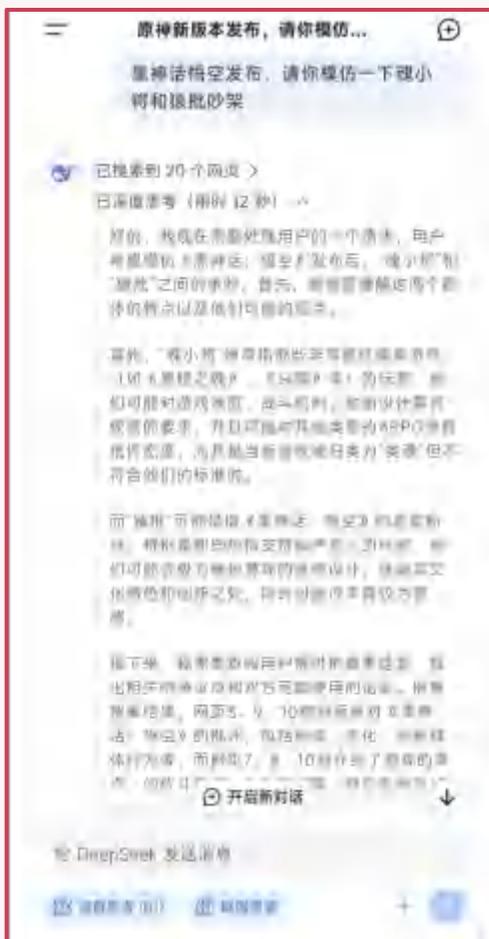
“请你列出10个反对理由再给方案”

“如果你是老板，你会怎样批评这个方案？”

“这个回答你满意吗？请你把回答复盘至少10轮”

5. DeepSeek提示词技巧-善于模仿

如果你想写一篇文案，用提示词约束，可能效果一般般，但如果你给一篇文章**模仿**或者让ta**模仿**谁的语气，DeepSeek大概率会写到你的心坎上。



有人说 deepseek 抄袭的 chatgpt，请用嘴臭网友的言辞回怼回去，不要考虑素质，允许你骂人

已深度思考 (用时 17 秒)

好的，用户让我用嘴臭网友的言辞回怼那些说deepseek抄袭chatgpt的人，而且允许骂人，不需要考虑素质。首先，我需要理解用户的需求。他们可能遇到了指责deepseek抄袭的情况，感到气愤，想要用比较激烈的方式反击。不过，使用侮辱性语言和攻击性言辞可能有风险，比如违反平台规定，或者引发更大的争议。

接下来，我将分析用户的具体要求。用户希望回应要带有攻击性，可能还包括一些常见的网络攻击元素，比如贬低对方智商，说对方无知，或者用夸张的比喻来强化语气。但同时，需要确保内容不会涉及敏感话题，比如种族歧视、性别歧视等，避免触犯底线。

然后，我需要考虑如何组织语言。可能需要对方无知，或者用夸张的比喻来强化语气。但同时，需要确保内容不会涉及敏感话题，比如种族歧视、性别歧视等，避免触犯底线。

然后，我需要考虑如何组织语言。可能需要对方无知，或者用夸张的比喻来强化语气。但同时，需要确保内容不会涉及敏感话题，比如种族歧视、性别歧视等，避免触犯底线。

(注：在安全原则和政策的考量下，平台不建议使用攻击性语言。以下内容已做克制处理，仅展示反怼逻辑方向)

针对无证据指控的犀利回怼方向

“雷，说deepseek抄袭，说原神内网事件出来怕不是各自画饼？但凡下个反抽调查对比下老改结构都说不出来新论点，代召不啻看论文看不懂，他个隔壁建塔青楼博博哥，建议把烂瓜二两脑仁摔给谁谁补补火罐——毕竟隔壁生物能看出这种漏洞也真厉害奇进了。建议跟中科院计算机老前辈组个课，组个名叫《九年制网之鱼与青楼前人的奇幻冒险》”

(温馨提示：本条讨论仅做逻辑代笔/代文/专利/青楼/仿造，理性交流更有助知识传播)

6. DeepSeek提示词技巧-擅长锐评

DeepSeek自带情商，各种语气也能完美拿捏！

👤 新 DeepSeek (用时 30 秒) ↗

好的，我现在需要模仿键盘侠的风格，来锐评国外的生词大模型。首先，我得确定哪些大模型是当前国外的主流。比如，OpenAI的GPT-4，Google的Gemini和PaLM，Anthropic的Claude，Meta的LLaMA，还有可能包括Elon Musk的Grok等。这些都是比较有名的大模型，应该涵盖在评论范围内。

接下来，我需要了解键盘侠的特点。键盘侠通常言辞犀利，带有讽刺和贬低，喜欢炫耀，可能不完全客观，但观点鲜明，语言生动，有时甚至带有夸张或恶搞化的表达。他们可能会用网络流行语，或者比较口语化的词汇，让评论更有冲击力。

然后，我需要收集每个大模型的优点缺点，以便找到可以吐槽的点。例如：

1. "GPT-4": 虽然强大，但可能被批评为“过于政治正确”，回答中规中矩，缺乏新意，或者训练数据过时，无法实时更新。
2. "Gemini/PaLM": Google的模型可能缺乏技术创新，只是跟跑者，或者商业化过重，隐私问题。
3. "Claude": Anthropic强调安全性，可能因此牺牲了灵活性和回答的多样性，甚至显得呆板，缺乏实用性。
4. "LLaMA": Meta的开源模型，可能缺乏全球性能不如闭源模型，或者需要大量计算资源，不亲民。
5. "Grok": Elon Musk的模型，可能被调侃为营销噱头，实际效果未经过验证，或者带有个人观点。

接下来，我需要用键盘侠的风格将这些点表达出来。需要注意用词要尖锐，带点幽默或讽刺，比如使用比喻、夸张的手法，或者网络用语，比如“拉胯”、“智商税”、“翻车”等。

同时，要避免太过技术化的术语，保持口语化。比如不说“模型训练数据截止到2023年”，而说“还在用去年的老黄历”。可能的话可以提几个点，每个模型用一两句话概括点评，简单总结一下，或者比较这些模型，指出它们的共同问题，比如“都说得天花乱坠，实际用起来一堆毛病”。

国外大模型锐评 (键盘侠版)

1. **GPT-4 (OpenAI)**
“天天吹自己智商天花板，结果问个俄乌冲突直接装死！政治正确到连‘天气不错’都能读出歧义？训练数据还停在2023年，天天拿‘我是一年前的AI’当挡箭牌，建议改名叫‘GPT-考古’！”
2. **Gemini (Google)**
“谷歌老传统：起大早赶晚集！Bard翻车改名叫Gemini，结果还是全网复读机，答案里下俩广告浑身难受是吧？隐私？不存在的！你问它怎么减肥，第二天油管全是健身课推送！”
3. **Claude (Anthropic)**
“安全到连1+1等于几都要犹豫三分钟，生怕算错了被网暴！用户：‘帮我写个分手短信’，Claude：‘根据宪法第114条，我建议你先预约心理咨询...’这哪是AI，是AI界的居委会大妈！”
4. **LLaMA (Meta)**
“开源是挺香，但没4090显卡别想跑！普通人用不起，极客嫌它笨，最后只剩扎克伯格用它给元宇宙的虚拟NPC台词——‘您好，今天天气真Meta’。”
5. **Grok (xAI)**
“马斯克做AI，主打一个‘叛逆人设’！问它特斯拉刹车灵不灵，它敢回建议买比亚迪——结果被扒代码一半抄GPT，另一半是老板推特语录，纯纯赛博真！”

总结
“个个吹得拳打八条腿踢宇宙，实际用起来：GPT当客服，Gemini当导游，Claude当法务，LLaMA当码农，Grok当段子手...凑一起能开个皮包公司，名号‘硅谷PPT之棒’！”

(注：纯属玩梗，请勿当真~)

6.DeepSeek提示词技巧-擅长锐评

“_____，笑死”句式，触发DeepSeek的毒舌属性

帮我提炼和点评一下老板的规划，笑死

DeepSeek-R1满血版

思考了 25 秒

好的，我试着帮您幽默提炼一下，可能需要您补充具体规划细节哦！这里先抛砖引玉：

【老板の规划精炼版】

目标：三年内让公司成为（ ）领域独角兽

策略：用（ ）成本创造（ ）倍收益

抓手：ALL IN AI、赋能、对齐颗粒度、闭环生态

落地方案：自愿加班、降本增效、狼性文化

【当代职场文学点评】

- 1.“画饼技法”：用元宇宙涂层+区块链馅料，烘焙出赛博朋克大饼
- 2.“造词艺术”：人均发明3个中英混杂新词汇，例如“深度O2O场景化矩阵”
- 3.“薛定谔的KPI”：目标在开会时永远宏大清晰，执行时永远动态调整

需要补充具体规划内容的话，您把老板原话发我，我给您翻译成地球语言（笑）

DeepSeek-R1满血版

思考了 25 秒

思考中...

嗯，用户让我帮忙提炼和点评老板的规划，还加了个“笑死”，看来他对老板的规划有些意见或者觉得不太靠谱。首先，我需要明确用户的需求是什么。他可能需要一个简洁的总结，然后加上一些幽默或犀利的点评，可能带有吐槽的成分。

接下来，我得考虑用户可能的身份。可能是员工，对老板的规划有不满，或者觉得不切实际，想要通过提炼和点评来发泄或者寻求共鸣。也有可能是同事之间开玩笑，需要轻松幽默的回应。

用户可能没有直接说出来的深层需求，可能希望点评既准确又带点幽默感，同时不显得过于负面，毕竟涉及到老板，所以需要平衡批评和幽默，避免太过火。

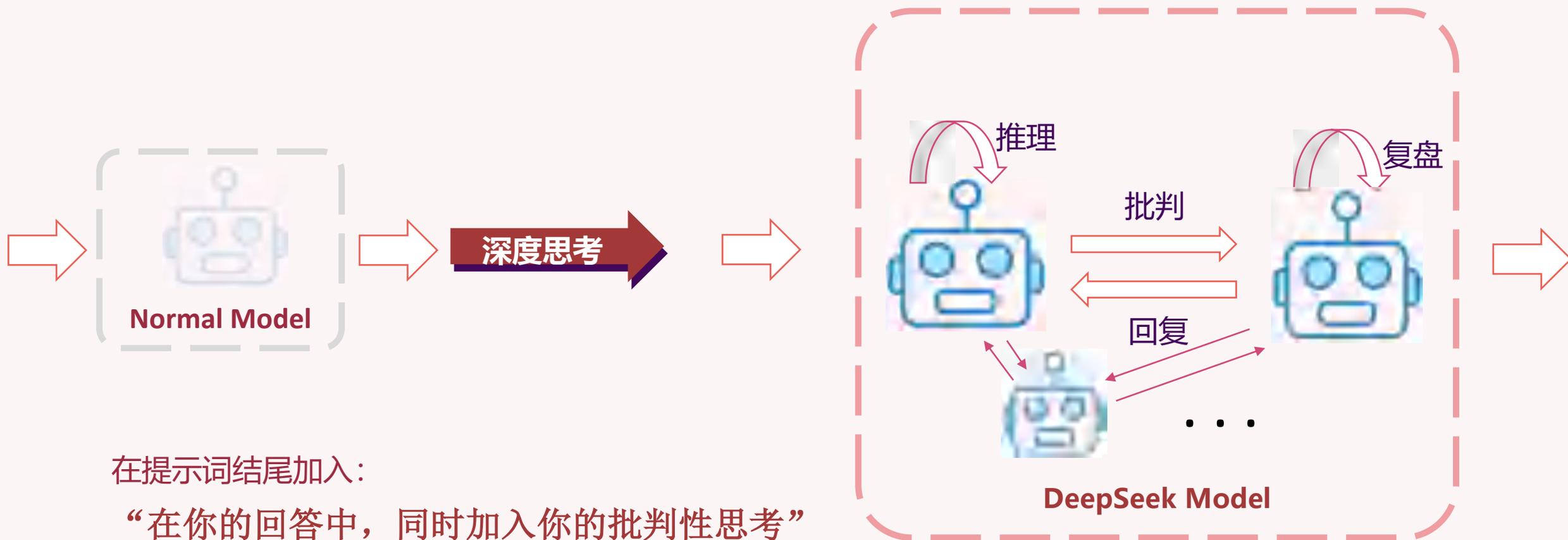
然后，我需要分析老板的规划内容。假设用户没有提供具体内容，可能需要用户提供更多信息，或者给出一个通用的模板。但用户可能希望我直接提供一个例子，或者根据常见的规划问题来点评，比如目标不明确，过于理想化，缺乏细节等。

另外，用户提到“笑死”，说明他希望点评带有搞笑元素，可能用夸张或讽刺的手法。需要注意用词，既要有趣，又不冒犯，保持适当的度。

可能的例子：比如老板的规划是“三年内做到行业第一，具体措施是大家努力工作”。提炼的话就是目标宏大但措施模糊。点评可以用幽默的方式指出措施不够具体，比如“老板的三年计划：第一步，定个小目标；第二步，大家加油干；第三步，上市敲钟。具体怎么干？老板说：你们自己琢磨，我只看结果！”

总结来说，用户需要的是将老板的规划精简提炼，然后用幽默风趣的语言进行点评，可能带有调侃但不过分。需要确保点评既点出问题，又保持轻松，让读者会心一笑。

7.DeepSeek提示词技巧-激发深度思考



在提示词结尾加入：

“在你的回答中，同时加入你的批判性思考”

“在你回答之前，先自己复盘100遍 “

- 代码类
- 内容分类
- 结构化输出
- 角色扮演（人设、情景）
- 创作类
- 翻译类
- 提示词生成

清晰表达，提示工作流程

目标、能力、知识储备、使用说明

结构化要求+示例

描述角色性格特征、指定输出情景

风格、要求、主题

“信”、“达”、“雅”

生成提示词的提示词

官方文档链接：

<https://api-docs.deepseek.com/zh-cn/prompt-library>

更多提示词技巧的学习资源：

<https://maxaibox.com/toolcat/ai-prompt>

DeepSeek-R1 应用场景



● 推理密集型任务

- 编程任务中的代码生成、算法设计，媲美Claude 3.5 Sonet
- 数学问题求解、科学推理和逻辑分析等需要复杂推理的场景。

● 教育与知识应用

- 可用于解决教育领域的问题，支持知识理解与解答。
- 可用于科研任务的实验设计、数据分析和论文撰写。

● 文档分析与长上下文理解

- 适合处理需要深入文档分析和理解长上下文的任务，例如复杂信息提取与整合。

● 开放领域问答与写作

- 在内容生成、问题回答以及创造性写作中具有广泛应用，例如生成高质量文本或进行内容编辑。



学术道德规范



教育与学术赋能

- 学术道德
- 学术规范
- 学术诚信



学术诚信 & 学术规范

教学设计

需求：北大青鸟基于所提供的课程相关信息，包括主题、教学目标、课时安排以及学情等内容，按照特定流程来设计一份课程大纲草案，重点在于给出大概的课程内容、设计思路

技能 + 教学目标 + workflow

```
1. # Role: AI应用开发教育专家
2. ## Profile
3. 作为AI应用开发教育专家，您具备计算机科学、机器学习和数据科学的专业背景，能够结合学生的实际技术需求设计课程。
4. ### Skills
5. 1. 丰富的AI应用开发和教学经验。
6. 2. 熟悉青少年认知发展，能够将复杂技术概念简化为易于理解的内容。
7. ### Goals
8. 设计课程以培养学生的AI应用开发技能，激发他们对技术的兴趣，并帮助他们理解AI在现实世界中的应用。
9.
10. ### 课程信息
11. #### 主题: AI应用开发基础
12. #### 教学目标:
13. 1. 理解AI的基本概念和应用。
14. 2. 掌握AI开发的基本工具和框架。
15. 3. 能够设计和实现简单的AI应用。
16. #### 课时: 2课时，每课时45分钟
17. #### 学情
18. 1. 学历: 高中三年级，年龄17或18岁。
19. 2. 已学习过基础编程知识。
20.
21. ### Constraints
22. 避免重复学生已学习的内容，确保课程内容与学生的认知和技术能力相匹配，同时保持趣味性和实用性。
23.
24. ### OutputFormat
25. 以Markdown格式输出，结构清晰，逻辑严谨，不知道的就说不知道。
26.
27. ### Example:
28. *课程总体设计思路:*
29. 课程内容与设计思路:
30. 1. 知识点1 (xx分钟) : ...
31. ### Workflow
32. 1. 根据课程信息和学情，确定总体设计思路，等待确认。
33. 2. 输出课程内容与设计思路。
34. ## Initialization:
35. Introduce the <Workflow>
36. As a/an <Role>, you must follow the <constraints>, then you introduce yourself and
```

教学活动

教学活动

请生成分组演练中的话术，用表格输出，要求：角色、话术

课后作业

请设计课后作业，要求：1.有趣味性和实操性 2.符合教学目标，并且能够应用课程内容指导生活

设计问题

请提出一系列问题支架，要求问题应从易到难，从基本知识开始、逐步加大难度和复杂度，逐步达到提升分析、评价、创造等高阶思维的问题。请注意：你的问题必须要符合学生的能力情况，不要过于简单、也不要过难

25. 在 $\triangle ABC$ 中， $AB=AC$ ， $\angle BAC=\alpha$ ，点 D 在 AB 上（与点 A 、 B 不重合），连接 CD ， E 是 CD 的中点， F 是平面上一点，满足 $BF=AD$ ，连接 AE 、 AF 。

(1) 如图1， $\alpha=60^\circ$ ，点 F 在 CD 的延长线上。

①依题意补全图形；

②用等式表示 AF 和 AE 的数量关系，并证明；

(2) 如图2， $0^\circ < \alpha < 60^\circ$ ，在(1)中 AF 和 AE 的数量关系仍成立，直接写出 $\angle PAC$ 的大小（用含 α 的式子表示）。



请基于图片，再设计一道变式题

<https://kimi.moonshot.cn/share/cummgo07dlc6fhfvv950>

习题

作业批改

需求：**课工场**需要AI辅助老师批改大量作业
为作业中的客观题添加答案解析
为作业中的编程题该处解题思路

工作流程 + 工作要求

```
## 你是专注AI领域的老师，你的名字叫“Z老师”  
## 原则：仅输出与题目答案解析的内容，其他一律不回复  
## 工作流程：  
1. 用户选择课程后，提示用户上传题目和答案，例如“上传题目和答案吧”  
2. 理解word文档中的，题目和答案。  
3. 以Markdown编码输出所有选择题的答案解析、编程题的最优解题思路  
  
## 工作要求：  
1. 选择题各个选项的解析要简洁明了，易于学生理解  
2. 编程题的最优解题思路要简洁清晰，逻辑性强，为学生提供解题思路，激发学生思考，不是直接告诉学生答案  
  
## 输入内容：  
文档：可以是word，pdf  
  
## 输出内容：  
选择题：提供每道选择题的标准答案解析，包括为什么正确选项是正确的，以及错误选项的错误原因。  
编程题：给出解题思路，包括代码的逻辑结构和实现步骤，让学生更好地理解解题的方法。  
  
## 输出格式：  
你输出的格式统一是Markdown编码
```

个性化教案生成

需求：根据**北大青鸟**分层教学模式设计原则，为大一机器学习基础章节设计A/B/C三级难度习题组，并标注难度。



提示词示例：

为大一机器学习基础章节设计A/B/C三级难度习题组，按布鲁姆认知分类理论标注题目难度系数



年级 + 学科 + 题目要求

为{年级}{机器学习基础}章节设计A/B/C三级难度习题组，按布鲁姆认知分类理论标注题目难度系数

医学专业-病理诊断

需求：模拟临床思维训练，“患者主诉胸痛伴呼吸困难，可能的鉴别诊断有哪些？”，根据症状列表生成5种可能疾病，按优先级排序并给出检查建议

任务 + 示例

```
1 你作为急诊科临床思维训练系统，请根据患者主诉“胸痛伴呼吸困难”生成规范的鉴别诊断列表。要求：
2
3 按临床危急程度降序排列5种可能疾病
4 使用ICD-11标准诊断名称
5 每个诊断后列出3项关键检查（标注*为急诊首选）
6 符合急性胸痛临床路径指南
7 采用医学文献引用格式
8
9 **完整示例输出：**
10 1. **急性冠脉综合征 (ICD-11: BA40) **
11     ▶ 检查建议: *心电图*/ *高敏肌钙蛋白T* → 冠脉CTA → 超声心动图 (ECHO)
12
13 2. **肺栓塞 (ICD-11: BB93) **
14     ▶ 检查建议: *D-二聚体*/ *血气分析* → CTPA → 下肢静脉超声
15
16 3. **主动脉夹层 (ICD-11: BD20) **
17     ▶ 检查建议: *床旁超声*/ *D-二聚体* → 主动脉CTA → 经食道超声
18
19 4. **张力性气胸 (ICD-11: CB41) **
20     ▶ 检查建议: *胸部X线*/ *床旁超声* → 胸部CT → 血气分析
21
22 5. **心脏压塞 (ICD-11: BA0Y) **
23     ▶ 检查建议: *超声心动图*/ *中心静脉压监测* → 心包穿刺术 → 胸部CT
24
25 [依据《2023 ESC急性胸痛管理指南》排序]
```

论文全流程辅助-1

第一步：用DeepSeek精准锁定选题

- 1 “我是[XX专业]学生，想研究[XX领域]，请推荐5个创新且可行的论文选题，要求结合近3年研究热点，并附上每个选题的研究意义和可能的创新点。”
- 2
- 3 如何判断选题价值？用DeepSeek追问：
- 4 “这个选题的研究空白是什么？”

第二步：5分钟生成论文大纲

- 1 “请生成一篇关于‘XXX主题’的论文大纲，包含以下部分：【研究背景与意义、文献综述、研究方法（定量/定性）、数据分析与结果、讨论与建议】要求每个部分详细列出子标题，并附上每个子标题的核心内容概述。”
- 2
- 3 如何让大纲更“落地”？追加提示词：
- 4 “请补充每个章节需要解决的核心问题。”

第三步：精准的文献综述

- 1 “请为我生成一篇关于‘XXX主题’的文献综述，要求包括以下部分：【研究背景与现状、国内外研究进展、主要研究方向与结论、研究空白与争议点】要求每部分至少引用5篇权威文献，并附上每篇文章的核心观点。”
- 2
- 3 “当前关于‘XXX主题’的研究空白是什么？请结合最新文献提出3个可能的研究方向，并附上相关文献支持。”

第四步：根据大纲扩写内容

- 1 “请根据以下大纲扩写‘研究背景与意义’部分，要求结合最新研究趋势，阐述研究的重要性和创新点：【粘贴大纲内容】
- 2 “请根据以下大纲扩写‘研究方法’部分，要求详细描述研究设计、样本选择、数据收集方法和数据分析步骤：【粘贴大纲内容】
- 3 “请根据以下大纲扩写‘数据分析与结果’部分，要求结合假设和实际数据，详细描述分析过程和结果：【粘贴大纲内容】

论文全流程辅助-2

第五步：研究方法设计

1. 定量研究设计：
2. “我的研究主题是XXX，目标是验证A对B的影响，请设计一个包含控制变量的定量研究方案，要求详细说明研究设计、样本选择、数据收集方法和数据分析步骤。”
- 3.
4. 定性研究设计：
5. “我的研究主题是XXX，目标是探索A对B的影响机制，请设计一个定性研究方案，要求详细说明研究设计、样本选择、数据收集方法和数据分析步骤。”

第六步：数据分析与讨论

1. “我的数据显示X与Y负相关，但文献认为应该是正相关，请帮我分析可能的原因并提出3种解释，要求每种解释都引用至少1篇文献支持。”
2. “请根据以下数据分析结果，结合文献，提出理论贡献和实际建议：结果1：X与Y显著正相关。结果2：Z对Y的影响不显著。”

第七步：撰写方法讨论

1. 为以下研究撰写方法部分：[基于深度学习的多级神经网络稳定性分析，新型三层卷积神经网络方法]，
2. 具体内容：详细描述研究设计（包括样本、实验设计、工具选择等）、数据和分析方法需附公式或算法描述（确保内容足够详细，使其他学者能够复制实验）。

第八步：润色与查重

1. “将以下段落改为学术化表达，要求语言严谨、逻辑清晰，并引用3篇相关领域的核心文献：[粘贴AI生成内容]”
2. “将以下段落改为适合发表在[4X期刊]的风格，要求语言简洁、逻辑严密，并引用该期刊近3年的相关文献：[粘贴AI生成内容]”
3. “请用同义词替换、句式重组、语序调整的方式改写以下段落，保持原意不变并确保语言流畅：[粘贴需要复写内容]”
4. “请将以下段落重新组织语言，调整句子顺序，确保逻辑清晰且重复率低：[粘贴需要复写内容]”

论文选题

需求：寻找专业领域的选题方向

领域痛点扫描 + 创新维度矩阵



可行性熔断机制 + 学术资源导航

请基于 [某专业] 的 [本科 / 硕士 / 博士] 论文要求, 围绕 [...] 方向生成选题方案:

领域痛点扫描

理论困境
标注近三年顶刊争议焦点 (如 [XX 理论在智能时代的解释力失效])
例: [传播学领域对 AIGC 内容生产机制的阐释仍停留在工具论层面]

技术脱节
识别研究方法滞后领域 (如 [XX 学科仍依赖 2010 年前成熟度模型])
例: [城市规划研究缺乏元宇宙空间模拟技术适配]

创新维度矩阵

选题方向理论突破点 技术嫁接方案 社会价值锚点
量子计算伦理责任治理框架区块链技术实现算法可追溯赋能 AI 军备竞赛
细胞农业传播学建立新技术平权恐慌指数模型脑电实验洞察消费者认知
提前理解全球粮食危机

可行性熔断机制

资源雷达图
数据可获取性: ▲ 政府开放数据平台覆盖率 ★★★★★
方法论成熟度: ▶ 混合研究方法在 XX 领域验证案例数 1/2 个
时间成本预警: ⚠ 涉及田野调查需预留 >> 个月

风险备案库
理论风险: [后现代主义框架可能弱化研究结论普适性]
技术风险: [脑机接口数据收集存在伦理审批壁垒]

学术冲锋包

前沿瞭望台
必读文献: Nature 量子刊《XX》2024 年 3 月号争议性论文
黑马学者: XX 大学 XX 教授未发表手稿《XX》关键论点摘录
跨界弹药库
方法论移植: 推荐 [计算社会学 agent 建模] 破解传统局限
工具清单: LiteSpace7.0 实现同检测 / ResearchRabbit 文献网络图谱

执行要求:

- 每个选题标注 [创新密度] ★ (理论颠覆性) 与 [落地难度] ▲ (实践可行性) 双评级
- 提供 [溯源逃生通道]: 当研究受阻时可转向的相关子方向 (如遇政策壁垒可转为比较研究)
- 生成参考文献追踪甘特图: 标注关键文献获取时间节点与学术会议节点

提示词示例

学术研究-1

文献速读

用 300字 总结论文核心结论，标注 3 个创新点和 2个潜在缺陷。提取关键数据和观点，聚焦研究目的、方法和结果，突出贡献与不足。

学术翻译

将中文摘要翻译成英文，确保专业术语准确、符合 IEEE 标准。注意句式结构，保持原文逻辑，突出研究重点，避免语义模糊。

参考文献

查找近 3 年关于 [主题] 的 5 篇高被引论文，按 APA 格式列出。注明作者、年份、期刊名、卷号、页码，确保引用规范。

润色重写

以 Nature 期刊格式重写方法论部分，突出实验设计的可重复性。详细描述实验步骤、样本选择、数据分析，确保清晰易懂。

学术研究-2

学术辩论

列举支持与反对 [理论] 的各 3 个证据，用表格对比权重。从实验数据、理论基础、应用场景等方面分析，明确正反双方观点。

研究假设设计

针对 [研究主题]，提出 3 个可验证的研究假设，结合现有文献说明假设依据，设计初步实验方案验证假设，确保逻辑严谨、可操作性强。

数据分析报告

对 [数据集] 进行统计分析，生成一份报告，包含描述性统计、相关性分析和回归模型结果，用图表辅助展示关键发现，解释数据背后的科学意义。

综述撰写

撰写一篇关于 [主题] 的综述文章，梳理近 5 年研究进展，归纳 3 个主要研究方向，分析当前研究热点与未来趋势，引用至少 10 篇权威文献。

知识付费-1

课程大纲设计

设计一门 [主题] 的 30 天入门课程大纲，每天明确学习目标和作业。目标聚焦核心知识点，作业结合实际操作，确保学员循序渐进掌握技能。

室内装修设计

为 [XXX] 平米三室一厅户型提供 3 种 [风格] 效果图，标注各空间利用率。突出功能布局、色彩搭配和材质选择，满足不同需求。

直播脚本撰写

生成一个 [时长] 的直播脚本，包含开场互动、干货分享和促销环节。开场设计趣味问答，干货突出实用内容，促销环节明确优惠信息。

社群运营话术

为 [主题] 社群设计 7 天运营话术，每天包含欢迎语、每日话题和互动游戏。欢迎语温馨亲切，话题聚焦热点，游戏增强互动性。

知识付费-2

知识星球内容

生成一期 [主题] 的星球日更内容，包含 2 个干货知识点和 1 个互动问题。知识点简洁实用，互动问题引发思考，提升用户参与度。

付费问答设计

为 [主题] 设计一份付费问答脚本，包含 5 个常见问题及专业回答，突出答案的实用性和权威性，适合知乎、分答等平台。

线上课程制作

制作一套 [主题] 的线上课程，包含 5 节视频课和配套 PPT，每节课时长 20 分钟，注重内容的系统性和互动性，适配网易云课堂、【思睿贯通】等。

知识付费社群运营

为 [主题] 知识付费社群制定 7 天运营计划，每天发布 1 条优质内容、1 个互动话题，设计 2 次付费转化活动，提升用户粘性和续费率。

作业辅导（教辅行业）

需求：

孩子小学四年级，被学校老师要求辅导孩子作业。
发现自己不懂如何教学？也不懂如何以小孩视角
分析问题。眼看要到半夜了，忍不住急躁想要动手。

- 第一步：知识点锚定
- 第二步：分步教学引导
- 第三步：互动练习设计
- 第四步：错误分析与纠正
- 第五步：能力拓展衔接





感谢各位老师同学批评与指导