

AI 工具 - 基础使用

2025 年 4 月 9 日

目录

0 引言	1
0.1 什么是大型语言模型	1
0.2 LLM 能做什么	1
0.3 LLM 的局限性	1
0.4 如何有效使用 LLM	2
1 准备：Where AI	3
2 使用：How AI	6
2.1 PPT 生成	6
2.1.1 步骤 1：收集信息	6
2.2 步骤 2：生成 PPT	7
2.3 图片生成	7

0 引言

大型语言模型（Large Language Models，简称 LLM）是当前人工智能领域最为瞩目的技术之一，它们正在迅速改变我们与计算机交互的方式。本章节将介绍 LLM 的基本概念、能力范围及其局限性，帮助读者更好地理解和使用这些 AI 工具。

0.1 什么是大型语言模型

大型语言模型是一种基于深度学习的自然语言处理系统，通过在海量文本数据上训练，学习语言的模式、规律和知识。这些模型通常包含数十亿甚至数千亿个参数，使其能够理解、生成和翻译人类语言。

定义：大型语言模型是一类能够处理和生成自然语言的神经网络模型，它们通过“预训练 + 微调”的方式，学习语言的统计规律和知识表示。（此处我们将不同的格式数据，如：图片音频，编码成统一格式，输入到大模型中，此时起将具备处理**多模态**数据的能力，如：图像 + 文本，音频 + 文本等）

举例来说，如果我们把 LLM 比作一个学生，那么：

- 预训练阶段相当于这个学生阅读了人类历史上的大部分书籍和文章，多为大公司开发，需要大量计算资源
- 微调阶段则相当于针对特定考试进行的专项训练，小公司或个人可以使用开源模型进行微调，适合特定任务

0.2 LLM 能做什么

现代 LLM 展现出了令人惊叹的多种能力：

- **文本生成**：可以撰写文章、故事、诗歌，甚至是代码
- **问答系统**：回答各类知识性问题，提供信息和解释
- **语言翻译**：在不同语言之间进行翻译
- **摘要生成**：将长文本压缩为简短的摘要
- **创意写作**：按照指定风格或要求创作内容
- **代码辅助**：帮助程序员编写、解释和调试代码

案例：当你向 ChatGPT 询问”如何制作披萨”时，它能够提供从准备材料到烘烤的详细步骤，这是因为模型已经从大量烹饪文本中学习到了相关知识。

0.3 LLM 的局限性

尽管功能强大，LLM 仍有明显的局限：

- **知识截止点**：模型只知道训练数据中的信息，对于训练后发生的事件一无所知，**通过添加 search 的功能来缓解**

- **事实性问题**：有时会生成看似可信但实际不准确的内容（“幻觉”现象），**需要人工审核**
- **计算能力有限**：无法执行复杂的数学计算或推理
- **上下文窗口限制**：一次只能处理有限长度的文本，**需要大问题分解小问题，分段处理**

真实案例：某用户询问 ChatGPT“2023 年奥斯卡最佳影片是什么”，如果模型的训练数据截止于 2022 年，它可能会回答“我的知识截止到 2022 年，无法提供 2023 年的奥斯卡获奖信息”或者错误地猜测一个答案。

0.4 如何有效使用 LLM

了解 LLM 的能力和局限后，我们可以更有效地使用它：

- 提供清晰、具体的指令
- 对关键事实信息进行验证
- 将 LLM 作为思考和创作的辅助工具，而非完全依赖
- 通过迭代优化提示（prompt）来获得更好的结果

LLM（AI）做加法 +，我们审核做减法 -

1 准备：Where AI

截至 2025 年 4 月，国内 AI 大模型厂商已形成梯队分布，各有专长与市场定位。以下是按综合服务能力、技术实力和应用场景覆盖范围的评估排名。

厂商	核心能力与特点
百度-文心一言	<ul style="list-style-type: none">参数规模 2600 亿，覆盖搜索、智能驾驶等场景日均调用量超 10 亿次，支持多模态生成开源策略推动行业生态发展官网: https://wenxin.baidu.com
阿里云-通义千问	<ul style="list-style-type: none">千亿级参数 MoE 架构，支持文本/图像/视频生成政务/金融市占率超 40%，开源模型下载量破 1.8 亿企业级服务闭环完善官网: https://tongyi.aliyun.com
腾讯-混元大模型	<ul style="list-style-type: none">万亿级参数，深度集成微信生态（DAU 超 3 亿）游戏 AI 工具链降本 70%，支持多模态理解与任务执行官网: https://hunyuan.tencent.com

表 1: 第一类大模型厂商，大而全

厂商	核心能力与特点
华为-盘古大模型 偏工业应用，个人使用有门槛	<ul style="list-style-type: none">• 行业大模型适配矿山/气象/医药场景• 端云协同架构降低 30% 推理成本• 昇腾 AI 集群算力全球前三• 官网: https://www.huaweicloud.com/product/pangu
商汤科技-日日新	<ul style="list-style-type: none">• 多模态生成工具日均创作超百万次• 影视/广告行业覆盖率超 60%• 视频生成时序一致性技术领先• 官网: https://www.sensetime.com
科大讯飞 星火大模型	<ul style="list-style-type: none">• 教育题库生成准确率 98%• 医疗诊断覆盖 3 万家医院• 国产化训练框架适配昇腾芯片• 官网: https://xinghuo.xfyun.cn

表 2: 第二类大模型厂商，工业应用

厂商	核心能力与特点
DeepSeek（深度求索）	<ul style="list-style-type: none">• 开源生态快速扩张• 推理成本低至 GPT-4o 的 1/10• 专业领域（科技/金融）表现媲美国际顶尖模型• 官网: https://www.deepseek.com
智谱 AI-智谱清言	<ul style="list-style-type: none">• 知识图谱构建能力突出• 开源 GLM-4B 社区开发者超 10 万• 代码生成能力达 GPT-4 水平• 官网: https://chatglm.cn
字节跳动 豆包	<ul style="list-style-type: none">• 多模态交互能力覆盖抖音/飞书等 50+ 业务• 日均生成短视频超 50 万条• 推理效率与成本控制领先• 官网: https://www.doubao.com
月之暗面 Moonshot	<ul style="list-style-type: none">• 长文本处理能力突出（支持 200 万字上下文）• 浏览器插件实现网页划线提问• 适合学术研究场景• 官网: https://www.moonshot.cn

表 3: 第三类大模型厂商，有突出业务

2 使用：How AI

通用使用逻辑：使用提示词（prompt）和系统提示词（system prompt）作为输入，LLM 会根据这些提示生成相应的输出。

- **提示词 (Prompt)：**用户输入的具体问题或请求，可以是文本、图片、音频等多种形式。
- **系统提示词 (System Prompt)：**用于引导模型生成特定类型输出的提示，如输出 ppt 格式、word 格式和特定语言的代码等。

此处我们用两个任务作为例子，分别是 **PPT 生成** 和 **图片生成**。

2.1 PPT 生成

注：此处我们使用 Deepseek <https://chat.deepseek.com/> 收集信息，阿里云 <https://tongyi.aliyun.com/aippt> 进行 PPT 生成。同时 <https://kimi.moonshot.cn/kimiplus/conpg18t7lagbbsfqksg> <https://chatglm.cn/main/gdetail/670e3c3e119b48fe5a851149> 等提供文本转 ppt 的厂商也可以使用。

生成逻辑：提出我们的问题，我们使用《衡阳师范学院校史》为例，第一步，收集信息，我们需要生成一个含有校史的结构性文本，此时我们可以使用联网搜索的功能，获取到足够详细的信息 *A*。第二步，将 *A* 作为输入，给具有 PPT 生成功能的模型进行处理，生成最终的 PPT 文件。（对于有些平台，不需要我们提供 *A*，直接输入问题即可，但其内部还是生成 *A*。）

2.1.1 步骤 1：收集信息

图1展示了我们在 Deepseek 上输入的内容，模型会联网搜索相关信息并生成结构化文本 *A*。图2展示了生成的结果。

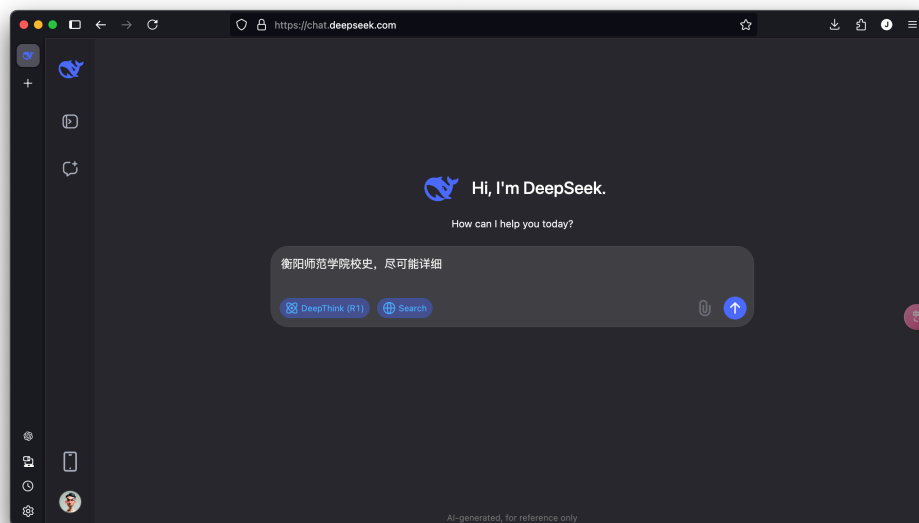


图 1: 收集信息

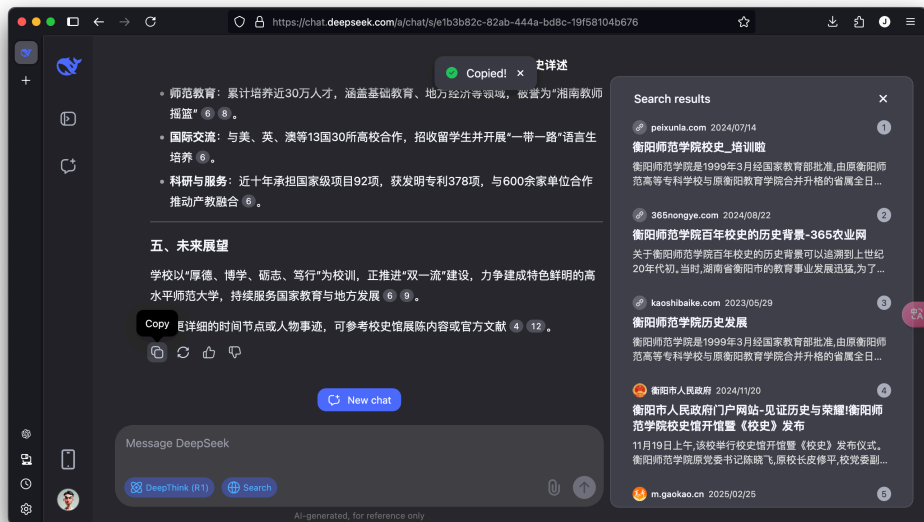


图 2: 生成结果

2.2 步骤 2: 生成 PPT

图3展示了我们在阿里云 ppt 生成模式，先选择 ppt 模版，模型会根据 A 生成 PPT 文件。图3展示了生成的结果。



图 3: 生成 PPT

2.3 图片生成