

几个概念词解释

LLM的基础原理——以ChatGPT为例

基于模态转换的StartUp

从行业领域看LLM应用层机会

可进入行业+合作伙伴

# 几个概念词解释

**NLP：**自然语言处理，涉及计算机如何理解、处理和生成人类语言的技术

**LLM：**语言模型，NLP中的一个重要组成部分，用来预测自然语言文本中下一个词或字符的概率分布，可以看作是一种对语言规律的学习和抽象。如 **GPT**、**BERT**、**ELMo** 等都属于一种LLM。

**ChatGPT：**基于GPT模型技术的一种产品，以智能对话形式体现，本质上是一种生成模型

**AIGC：**人工智能生成，类比与PGC、UGC

**AGI：**通用人工智能，目标是创建一个全面智能的系统，可以解决广泛的问题并进行多种任务

**AI Agents：**人工智能体=LLM+记忆+规划技能+工具=“人工智能助理” ≠ 数字人/虚拟助手

# LLM的基础原理——以ChatGPT为例

OpenAI所引领的本次AI技术迭代是跨越性的，远不同于Web3传销概念

- 大语言模型不是聊天机器人，是一种生成模型。
- 大语言模型做到了让机器“理解”人类自然语言并产生“思维”能力。
- LLM的潜力不仅限于生成内容、故事、论文和应用等方面，还具有强大的通用问题解决能力，可应用于各个领域。
- LLM相当于是一道曙光/发动机引擎，为人工智能与计算机科学的后续发展推开一道门缝（很高确定性告诉世界这个方向的路可以走通），AI Agents将迎来大爆发。

# LLM的基础原理——以ChatGPT为例

## 工作原理（如何回答问题）

- 单词接龙：对于长文进行自回归生成
- 模型影响：针对于不同的模型会生产不同的内容
- 模型训练：给data学习材料（范例），通过调整模型，让模型能够根据范例学习材料的上文生成下文（否则模型会乱生成）
- 通过对模型的训练，让模型能够产生举一反三的“学习”（规律、概率）能力而非从数据库中调取某个内容，使内容泛化

# LLM的基础原理——以ChatGPT为例

## 制造过程（与搜索引擎具有本质上的差异）

- ChatGPT不具备搜索能力，在模型训练过程中学习材料并没有被保存在模型中，材料data的作用是调整模型来得到通用模型，以便之前没有学习过的内容也能被模型根据上文逐字生成，所以ChatGPT也叫生成模型
- GPT进行了的超大规模训练，通过三个阶段，产生涌现能力，成为了我们现在看到的基于GPT4的ChatGPT
- NLP：因为人类语言的复杂和逻辑关系，很难被机器理解，所以GPT经过了三个阶段的“学习”
- 第一阶段：给一堆材料，灌输给GPT。G=Generative生成，T=Transformer一种模型结构，P=Pre-training 资料学习/预训练（GPT从1到2到3的过程，学习量成十倍百倍级的加大，形成GPT3）
- 第二阶段对GPT3进行“模板规范”，第三阶段再进行“创意引导”
- 这三个阶段的专业名词是：无监督学习、监督学习、强化学习。

最终“**涌现**”能力产生，使得机器可以“理解”指令、“理解”例子、具备思维链

# LLM的基础原理——以ChatGPT为例

## LLM与AI Agents及应用

- LLM是“大脑”，给LLM配备外部工具可以显著扩大大模型的功能，使其能够处理更加复杂的任务，LLM驱动AI Agents。
- AI Agent就是一个代理人，它可以代替我们和ChatGPT之类的大语言模型进行交互。尽管当前的大模型可以帮助我们解决许多问题，但是对于复杂的任务来说，通常需要我们不断和LLM进行交互，不断地给LLM发送指令才能最终完成我们的目标。
- AI Agent技术目标就是帮助我们降低这部分繁琐的工作。例如，如果我们想要知道今天推特上最火的AI相关的讨论是什么？通常，这需要好几个步骤：理解问题，登陆推特网站，寻找今天与AI相关的推文，理解这些推文讨论的内容，按照讨论的数量找到最火的推文，最后对这些推文进行总结。只有经过这些反复的交互，我们才有可能完成我们的目的。
- 很多人基于大模型做应用，需要的是大模型的语言能力，而非知识。大模型自己是缺乏知识的。现在大的范式是，应用产生内容，然后我通过语言模型让内容和用户去做交互。比如微软的 Bing，其实是 Bing 搜索出结果，再让前端的语言模型讲给用户听。模型起到的是一个传声筒的作用。
- 语言 and 知识是分开的。ChatGPT 学习过很多语料，本身已经具备很多知识。但那些知识——第一它会有幻觉，会一本正经胡说八道；第二，它很难实时更新，因为训练需要时间。所以你可以把它理解成一个应用的「嘴替」，可能到最后你会发现用哪个大模型并不是那么重要。给用户提供服务的可能是整合模型，里面有 Google 的大模型、有 OpenAI 的模型、甚至有某种基于规则模型。
- 你提出问题后，应用会将这个问题给到不同的大模型，再把最好的回答放出来。或者整合几个回答后再告诉你，都是有可能的。因此，如果大模型只是一个引擎，数据并没有在大模型里，那在某种程度上大模型就是可切换的产品。

# 基于模态转换的StartUp

## 多模态：文本、图形、视频、音频、3D等

- 大型语言模型是能够分析和生成文本的计算机模型，它们接受了大量文本数据的训练，以提高它们在文本生成甚至编码等任务中的表现。
- 目前大多数LLM都是纯文本的，也就是说，它们只擅长基于文本的应用程序，而理解其他类型数据的能力有限。
- 纯文本LLM的例子包括GPT-3、BERT、RoBERTa等。
- 多模态LLM结合了其他数据类型，如图像、视频、音频和其他感官输入以及文本。将多模态集成到LLM中解决了当前纯文本模型的一些局限性，并为以前不可能实现的新应用程序提供了可能性。
- Open AI最新发布的GPT-4就是Multimodal LLM的一个例子，它可以接受图像和文本输入，并在许多基准测试中显示出人类水平的性能。

**AI四小龙时期的创业发展是基于AI的视觉能力、语言语义理解能力发生的，如安防、客服、ERP**  
**OpenAI时代下，大语言模型涌现出的能力将使智能化AI充分融入社会并重构改变社会**



# 基于模态转换的StartUp

## 多模态：文本、图形、视频、音频、3D等

- 文本：Jasper AI，2020年底立项，基于GPT-3，AI自动写文案，市场广告营销行业。第二轮融资15亿美元估值，但受ChatGPT冲击影响较大，普通C端大量迁移
- 图形：Midjourney，1年左右时间，基于discord社区，不依赖融资，10亿美元估值+，营收2亿美元，极少量员工；Stability AI，成立两年，2022年10月首轮融资1亿美元
- 视频：Runway AI、Synthesia等很多，首轮融资数千万及上亿美元（jasper、midjourney头部效应明显，差异化竞争，视频领域机会可能更大）
- 音频：语音克隆类产品，众多startup，融资数千万美元，国内有虚拟歌姬、CV声优方向等
- 3D：unity、unreal等，国内也有多个startup
- 聊天机器人：Inflection AI（Pi虚拟助手App），第二轮融了13亿美元；Character AI，虚拟名人聊天机器人，A轮融资2亿美元+

国内目前具有较高确定性的公司是MiniMax、光年之外（王慧文退出前），研发国产LLM阶段  
核心能力/竞争力/优势壁垒：**融资能力，资金储备**。相比移动互联网时代还要明显与直接。



# 从行业领域看LLM应用层机会

**心理**、教育、陪伴陪护（儿童/老人）、数字人卖货、文娱游戏社交

- 心理行业长期规模增长，需求强烈，具有一定市场潜力与规模
- 可从2B、2G切入，如大学及中学校，部队、国企等公职部门
- 市场分散，暂无头部效应，进入门槛不高，LLM对行业的改造能力强
- 如何产品化，拍脑袋举个case：针对于不同学校定制一款数字拟人代表学校，学生可与其交流互动，基于心理与情感需求做模型调整；该数字人嵌入学校某系统/如小程序、网站、微信公众号中





# 从行业领域看LLM应用层机会

心理、**教育**、陪伴陪护（儿童/老人）

数字人卖货、文娱游戏社交

- 右图是一个真实案例，而且还是计算机领域相对专业抽象的内容
- 仅是GPT-4，非基于教育场景下垂直模型
- 陆奇博士在奇绩创坛讲座时，有提到非常看好LLM在教育领域应用，其在微软几个很看好的高管朋友从微软辞职创业都在做教育
- 但国内做这个方向，除了监管问题外，国内的教育市场高度竞争，较难找到优势

I 我还是无法理解addObserver的用法以及为什么要这么用，请再详细清晰讲解，假如我是一个高中生，请不要用过于复杂的讲解

 当然，我会尽量用简单的语言来解释`addObserver`的用法。   

首先，让我们想象一下你在看一场足球比赛。你是观众（Observer），足球比赛是你正在观察的事件（Observable）。你想知道比赛的进展，比如谁进球了，比分是多少等等。但是，你不能一直盯着比赛看，你可能需要做其他的事情，比如吃东西或者和朋友聊天。所以，你需要一种方式来在比赛有重要进展时得到通知。



这就是`addObserver`的作用。当你到达球场时，你会告诉比赛的组织者你的座位号，这样当比赛有重要进展时，他们就可以通过广播系统通知你。在这个例子中，告诉组织者你的座位号就相当于调用了`addObserver`方法。

在编程中，我们经常需要处理这种"当某件事发生时，我想知道"的情况。Observer模式就是为了处理这种情况而设计的。在Observer模式中，我们有两种对象：Observable（被观察者）和Observer（观察者）。Observable是事件源，它的状态可能会改变。Observer是需要知道这些改变的对象。


当我们创建了一个Observer对象后，我们需要告诉Observable对象我们对它的状态改变感兴趣。这就是`addObserver`方法的作用。当我们调用`observable.addObserver(observer)`时，我们就告诉`observable`，当它的状态改变时，我们希望`observer`能得到通知。

所以，`addObserver`方法的作用就是注册观察者。它告诉Observable对象，有哪些Observer对象对它的状态改变感兴趣，希望在状态改变时得到通知。

I 我可以这么理解吗？在twitter里，我关注了马斯克，假设我是a, 马斯克是b. 那么马斯克的推文更新都会提醒我，那么现在是b.addObserver(a). 所以每个twitter account既是Observer又是Observable

 是的，你的理解是正确的。在Twitter这个例子中，每个用户（比如你和马斯克）都可以是Observable和Observer。   

- 当你关注马斯克时，你（a）是Observer，马斯克（b）是Observable。你对马斯克的推文更新感兴趣，所以你关注了他。这就像调用了`b.addObserver(a)`，告诉马斯克的账户，当有新的推文时，你希望得到通知。
- 同样，如果马斯克关注了你，那么马斯克就成为Observer，你就成为Observable。马斯克对你的推文更新感兴趣，所以他关注了你。这就像调用了`a.addObserver(b)`，告诉你的账户，当你有新的推文时，马斯克希望得到通知。

所以，是的，你可以把Twitter的用户  Regenerate response 又是Observer。他们可以发布推



# 从行业领域看LLM应用层机会

心理、教育、**陪伴陪护（儿童/老人）**

数字人卖货、文娱游戏社交

- 先做儿童市场（早教机器人），老年市场是第二曲线
- 智能硬件创业热潮时是一个热点，其主要利用机器视觉与自然语言处理能力做简单交互
- 硬件端成熟且市场高度分散（卖玩具），LLM能力融入其中可能会带来较大的优势（市场营销与产品效果双重优势）
- 考验产品集成能力与销售能力



# 从行业领域看LLM应用层机会

心理、教育、陪伴陪护（儿童/老人）、**数字人卖货/演艺**、文娱游戏社交

- 直播卖货是长期生意，虽然流量端（抖音、淘宝）高度控制行业利润
- 相比于真人（李佳琦、薇娅）卖货，数字人的优势是24小时+相比真人的绝对低成本，自动化直播带货
- 如果在应用场景端有绝对优势可以直接面向2C，做数字人卖货MCN；或者2B提供技术服务
- 数字人演艺：AI Vtuber MCN，目前已经做到在观众视角下高度成熟发达，效果拔群



# 从行业领域看LLM应用层机会

心理、教育、陪伴陪护（儿童/老人）、数字人卖货/演艺、**文娱游戏社交**

- 大杂烩阶段，可凭创意迅速制造热点形成某爆款
- AI+游戏对于大众短期高估，长期低估；游戏公司高度重视，降本增效初显成果
- 布局方式：AI创意创新中心
- 与业内AI技术公司合作做“技术成果转化落地”，将AI技术边界通过某产品点/游戏创意点以微信小程序形式快速落地推广（3天-1周）
- IAA广告变现可盈利，具有长生命周期的“点”可以产品化/游戏化完成商业闭环推出

可进入行业+合作伙伴