

AIGC原理和使用入门

黑夜路人Black

2023/03/28

文章目录

AIGC简要介绍
AIGC是什么
AIGC可以生成什么内容
AIGC基本工作原理
机器如何进行学习
机器学习的分类
机器学习的发展
传统机器学习和深度学习的区别
GPT的形成过程
GPT是什么意思
ChatGPT的工作机制
AIGC应用产品介绍
主流AIGC产品
文本AIGC产品的介绍
图像AIGC产品介绍
如何访问国外的AIGC产品

AIGC使用实践
AIGC的准确性问题
AIGC中Prompt的语言选择问题
文本AIGC产品 ChatGPT/GPT-4 的使用实践
文本AIGC产品 Github Copilot / Cursor 的使用实践
图像AIGC产品 Midjourney 的使用实践
附录1: 各类工具参考信息
附录2: GPT原理参考技术文档

AIGC简要介绍

AIGC是什么

AIGC - AI Generated Content (AI生成内容)，对应我们的过去的主要是 UGC (User Generated Content) 和 PGC (Professional user Generated Content)。

AIGC就是说所有输出内容是通过AI机器人来生成产出相关内容，主要区别是过去主要是普通用户和某一领域专业用户（人）生产内容，AIGC主要是依赖于人工智能（非人类）生成内容，这个就是AIGC的核心意思。

(版权认定：UGC和PGC是有版权概念的，版权归属于负责生成内容的人，AIGC目前美国法规认为是没有版权概念的，就是内容不属于调用的人，也不属于这个AI机器，所以没有版权归属这件事。)

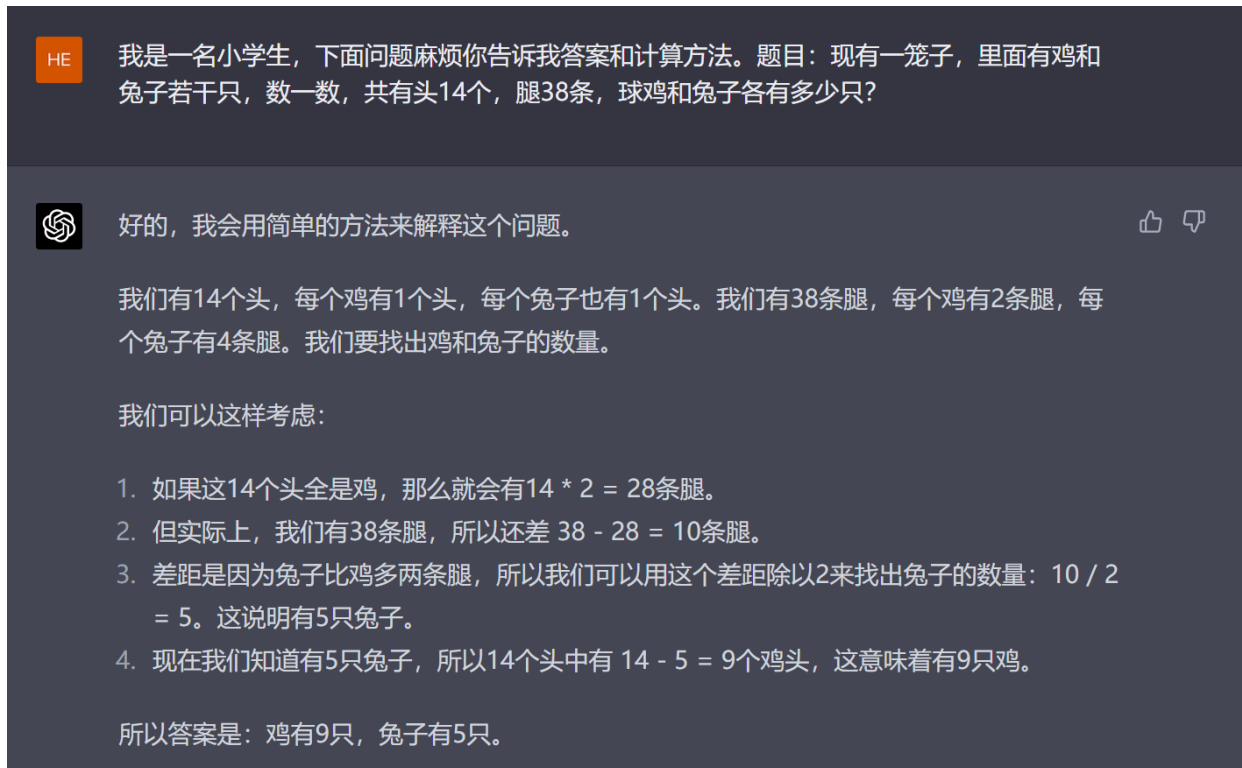
AIGC可以生成什么内容

目前AIGC主要可以生成文本内容和图片内容（目前视频生成有一些产品，但是没有文本和图片生成这么成熟），所以我们主要集中在文本和图片的AIGC的介绍。

AIGC在文本内容方面，主要可以通过 Q&A（提问回答）的形式进行互动，能够按照人类想要的“提问”生产输出符合人类预期的内容。

一般我们可以把AI当做一个全知全能的“高级人类”，以“文本AIGC”来抢，你可以向它提出问题（Prompt），然后它做出相应回答。所有提问和回答都可以涉及到方方面面，包括不限于 百科知识/创意文案/小说剧本/代码编程/翻译转换/论文编写/教育教学/指导建议/聊天陪伴 等等等等场景不一而足，场景都需要你去想，可以理解它是一个拥有全地球知识的“百晓生”，什么都可以问它或者跟它交流。

比如我们用大名鼎鼎的 ChatGPT 来进行提问：



对于“图片AIGC”来说，你可能脑子里有无数创意，但是自己不会绘画，无法把脑子里的Idea变成实实在在的图片，那么，“图片AIGC”能够帮助你按照你脑子想要的东西，你告诉它，然后它能够帮助你通过图片绘画的形式给你画出来，让你一下子把自己的“创意”变成了图片现实。

比如我们用非常好用的“图片AIGC”工具 Midjourney 来画画：



Midjourney Bot ✓ 机器人 昨天15:27

Help me draw a picture in which a little boy and a little white rabbit are together, which is very cute and adopts the Storybook Illustration style - @BlackX (fast)



AIGC基本原理

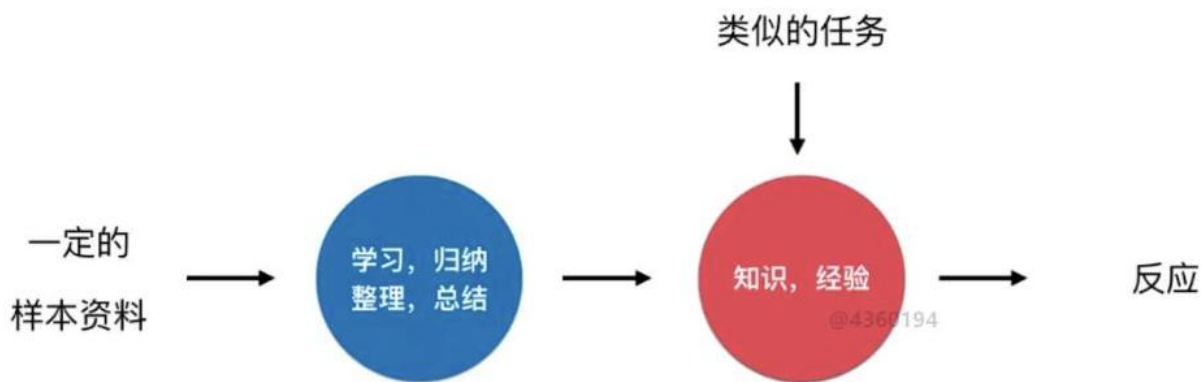
AIGC底层主要依赖的是AI技术，AI技术本质主要是让机器拥有像人类一样的智能（Artificial Intelligence），所以就需要让机器能够像人类一些学习和思考，所以目前大部分实现AI的底层技术叫做“机器学习”（Machine Learning）技术。

机器学习技术主要有很多应用场景，比如现在非常常用的包括人脸识别（手机解锁/支付宝支付/门禁解锁等）、语音识别（小爱同学/小度/Siri）、美颜换脸（主播美颜/美颜相机）、地图导航、气象预测、搜索引擎、NLP(自然语言处理)、自动驾驶、机器人控制、AIGC 等等。

机器如何进行学习

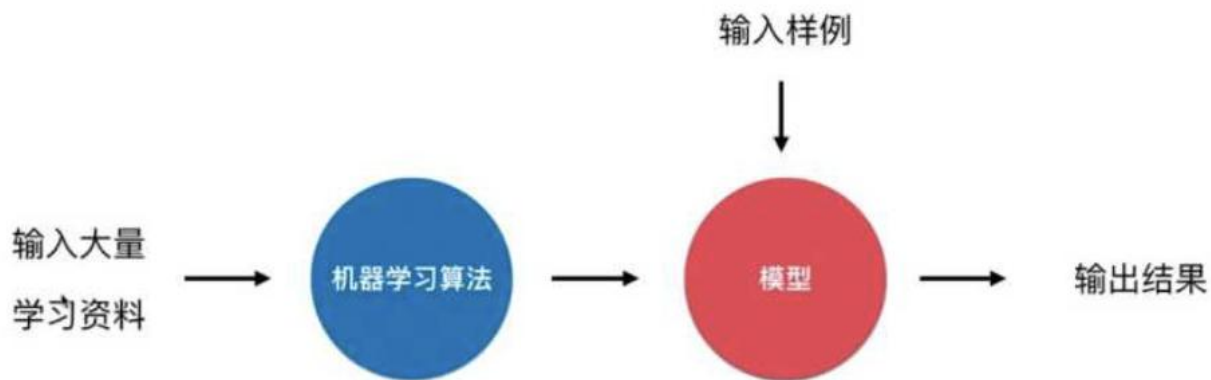
机器学习可以简单理解为是模拟人类学习的过程，我们来看一下机器是如何模拟人类学习的。

人类怎么学习



我们再看看所谓的“机器学习”：

什么是机器学习

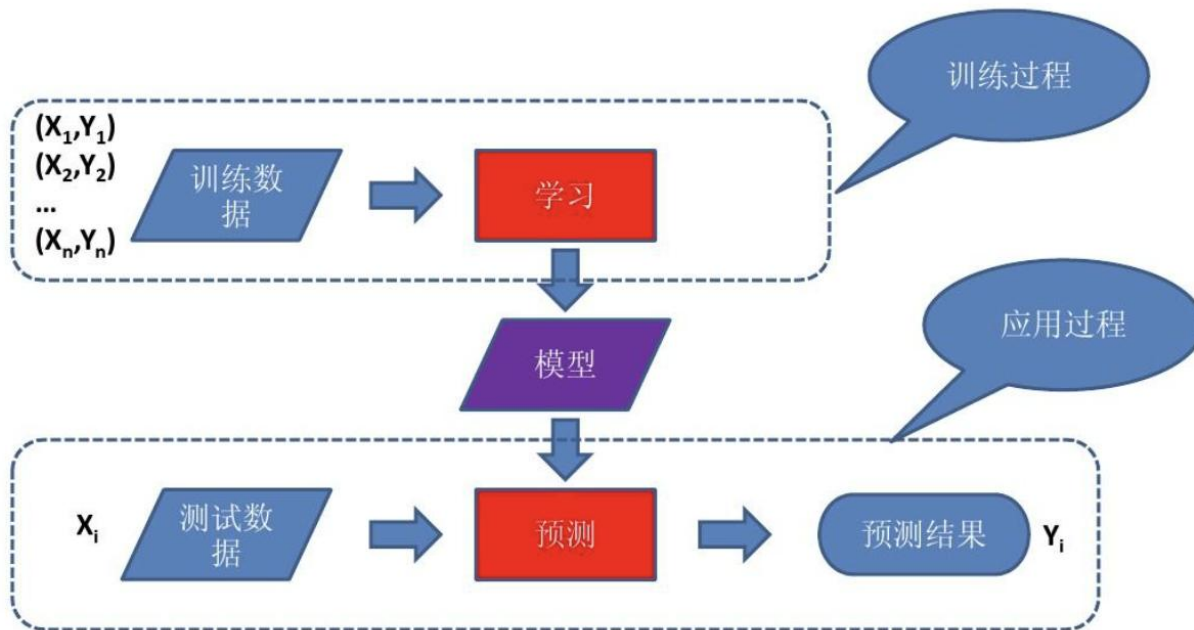


对于人类学习来说，我们看到的事物和遇到的事物就是我们的“资料”（语料），然后通过“学习总结归纳”（学习算法），最后变成了“知识经验智慧”（模型），然后遇到事情的时候我们就会调用这些“知识经验方法论”做出相应的反应决策动作（预测推理）；

对于机器学习来说，给它输入大量的“语料”（看到遇到的事物），然后通过机器学习算法（总结归纳抽取相似点），最后形成了“模型”（知识经验方法论），然后再遇到一些需要判断决策的时候，我们就会把要判断决策的事物给“模型”，然后就会告诉我们输出结果（推理推测结果）；

从抽象层来说，我们会发现，本质来说，“人类学习”和“机器学习”内在本质是比较像的。

我们来看一个计算机里概要的机器学习的过程：



里面核心步骤就是：“训练数据 → 训练算法 → 模型 → 预测 → 输出结果”，其中最终产出物是“模型”（Model，模型文件），然后主要是前置的“训练模型”和后置的“模型预测”，然后产生对应结果。

上面这个过程我们可以简单理解为：“模型”就是一只小狗，饲养员就是那个“训练算法”，饲养员在场下对那个小狗通过一些指令和奖惩措施进行反复训练（训练算法），小狗就会学会一些技能（模型），一旦学会了，小狗就可以出去表演，表演的过程就是预测。

所以我们会看到，里面如果“模型”中的特征（知识经验）越多，最终在“预测”阶段就越准确，如果模型越小，或者中间的特征数据越少，可能最终预测结果准确率会降低。（类似一个人遇到的事情越多，总结的经验就越多，俗话说“人生没有白走的路没有白踩的坑”大概就是这个逻辑）

机器学习的分类

机器学习（Machine Learnin）经过几十年的研究，在技术实现得种类非常多，可以从大类上面分为：基于学习策略、基于学习方法、基于学习方式、基于学习目标等进行分类。

基于学习策略的分类：主要有模拟人脑工作原理的“符号学习”和“神经网络学习”两个主要流派；另外一个学习策略的流派是基于数学统计分析得机器学习方法。

基于学习方法的分类：机器学习主要有，归纳学习、演绎学习、类比学习、分析学习，它是类似于人脑的“符号学习”得学习方法分类。

基于学习方式的分类：主要有，有监督学习（带老师指导）、无监督学习（自学）、强化学习（通过环境奖惩反馈进步）。

机器学习分类中，我们重点关注的是“学习策略”和“学习方式”为主的分类，学习策略是关注是基于“数学统计分析”的机器学习还是基于“神经网络”的机器学习。另外一个学习方法是学习方法，主要是关注有监督学习、无监督学习和强化学习。

机器学习的发展

机器学习技术主要经历了从兴起到深度学习三个大技术时代，第一个是兴起时代，然后是传统的机器学习时代，最后是基于神经网络的深度学习时代，以下简单按照我个人理解做了一个发展阶段分类。

兴起阶段：1943年诞生了感知机模型（MCP），一位心理学家和一位数理逻辑学家提出了神经网络的概念和人工神经元的数学模型，算开创了神经网络得研究时代。然后从60年代到80年代之间，有了机器学习概念和模式识别等，属于整个研究的兴起探索阶段，这个阶段各种方向的探索都在进行，百花齐放。

第一阶段：传统机器学习（Machine Learnin），从1980年召开第一届技术学习研讨会开始，虽然同步也有神经网络得研究，但是可以简单理解为大行其道的主要是基于数学和统计分析为主的方式得机器学习，特别是在1990年到2001年之间，从理论到实践都经过了很大的发展。这个时间段一直到2006年，在信息行业主要流行的传统机器学习包括传统隐马尔可夫模型（HMM）、条件随机场（CRFs）、最大熵模型（MaxEnt）、Bboosting、支持向量机（SVM）、贝叶斯（Bayes）等等，具体实践中落地包括线性回归、逻辑回归、SVM、决策树、随机森林、朴素贝叶斯等等落地的算法。这些算法的因果逻辑和中间计算过程都是清晰明确的，基本是清晰可信的，不足就是最终效果有上限，可能最终“智能”效果有时候就不够。

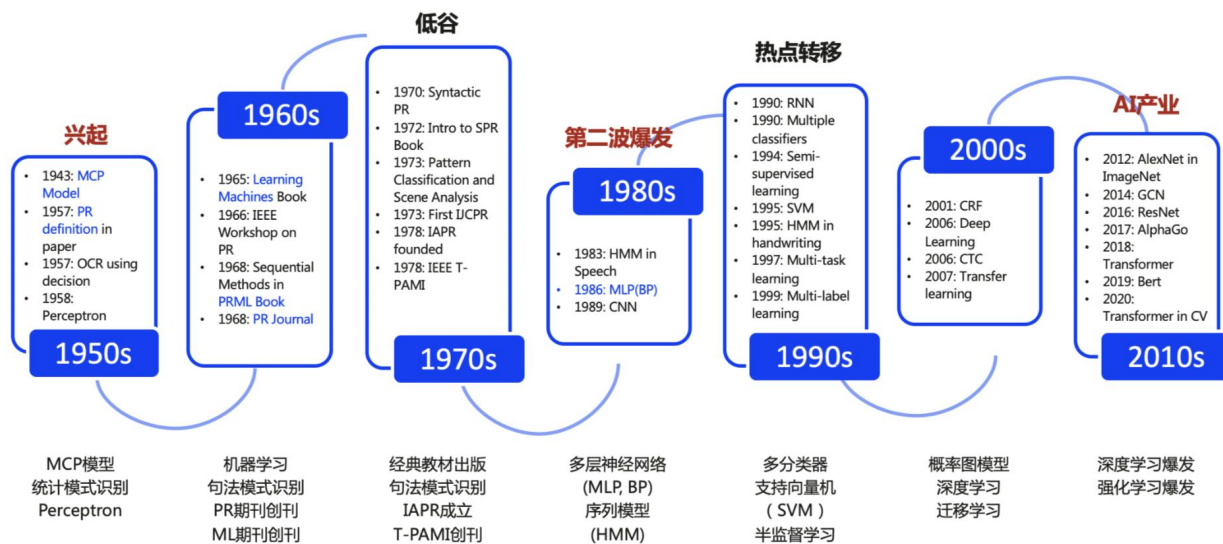
第二阶段 V1：“深度学习”（Deep Learnin），2006年机器学习教父级人物Hinton发表了深层神经网络的论文正式开启了基于神经网络得“深度学习”的阶段，可以简单认为“深度学习”是传统机器学习的另外一条路线，它主要区别是在“学习策略”方面的路线不同，传统机器学习主要是依赖于“数学统计分析”为主的方法，过程结果可推导；深度学习主要是也依赖于让计算机模拟人脑一样神经网络连接一样的方式进行运算。

第二阶段 V2：Transformer 模型（Transformer model），2015年提出了Attention机制，2017年Google发表了论文《Attention Is All You Need》，2018年基于这篇论文在第二阶段神经网络的基础

之上提出了Transformer模型，它基于encoder-decoder架构，抛弃了传统的RNN、CNN模型，仅由Attention机制（注意力机制）实现，并且由于encoder端是并行计算的，训练时间大大缩短。Transformer模型广泛应用于NLP领域，机器翻译、文本摘要、问答系统等等，最近几年比较主流的Bert和GPT模型就是基于Transformer模型构建的。

我们看一下深度学习基本的发展历史：

深度学习发展历程



传统机器学习和深度学习的区别

深度学习与传统机器学习不太一样，所以它主要是用来定义不同网络框架参数层的神经网络，所以主有很多神经网络训练方式，包括 无监督预训练网络（Unsupervised Pre-trained Networks）、卷积神经网络（Convolutional Neural Networks）、循环神经网络（Recurrent Neural Networks）、递归神经网络（Recursive Neural Networks）等等；

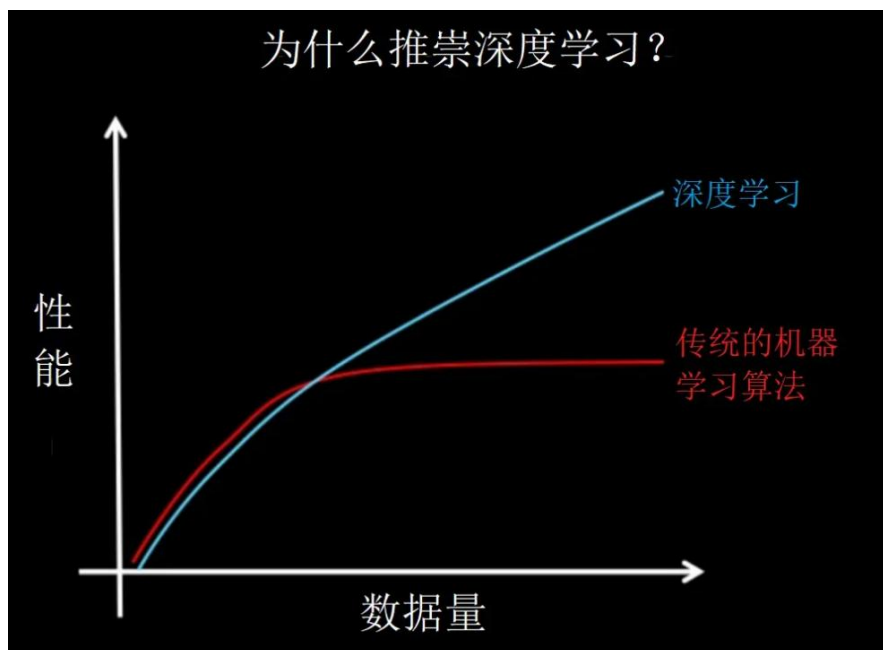
神经网络叫做“深度学习”主要是看里面所谓神经网络的层数，1-2层叫做浅层神经网络，超过5层叫做深层神经网络，又叫做深度学习”。

其中，应用比较多的主要 卷积网络(CNN - Convolutional Neural Networks)、循环神经网络(RNN - Recurrent Neural Networks)+递归神经网络(RNN - Recursive Neural Networks)、长短期记忆 RNN (LSTM - Long short-term memory) 和为了解决 LSTM/RNN 中的一些问题的解决办法加入 Attention机制的Transformer框架。

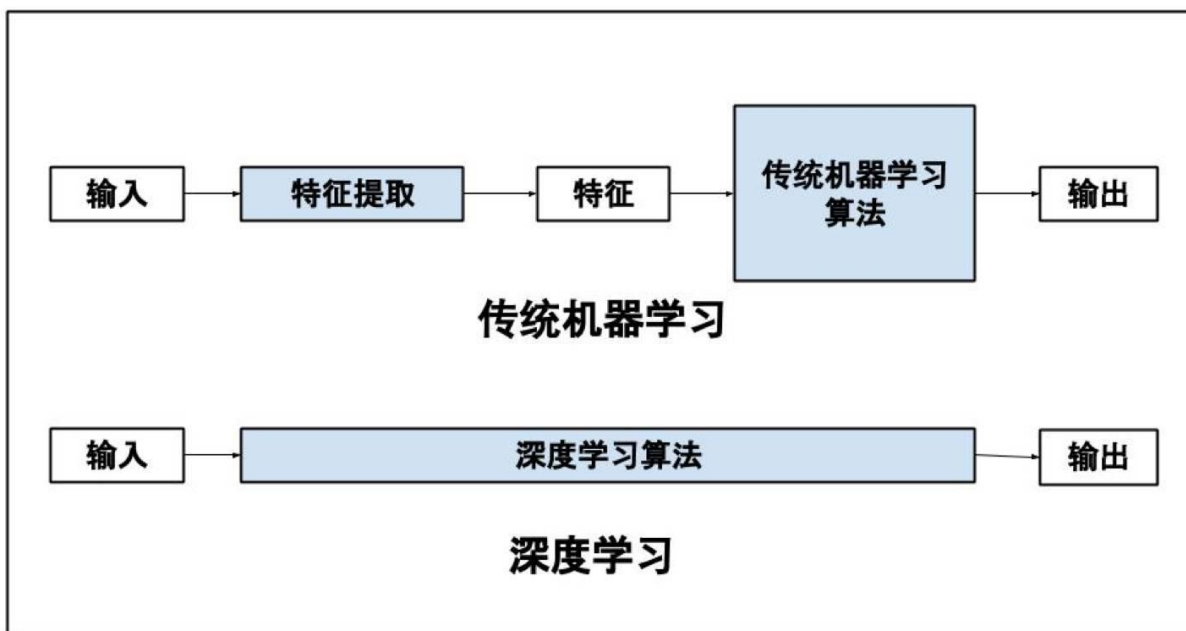
深度学习在计算机视觉（CV、如图像识别）、自然语言处理（NLP）、自动驾驶、机器人控制 等方面比传统机器学习效果更好。

在训练数据规模比较小的情况下，传统机器学习型算法表现还可以，但是数据增加了，传统机器学习效果没有增加，会有一个临界点；但是对于深度学习来说，数据越多，效果越好。所以也是越来越逐步“深度学习”取代“传统机器学习”的过程了。

传统机器学习和深度学习的效果性能对比图：

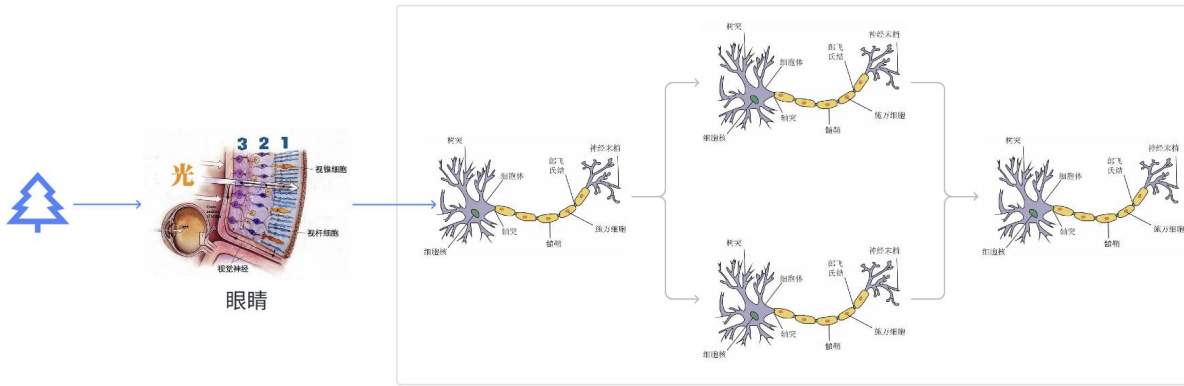


传统机器学习和深度学习的处理过程区别：（传统机器学习特征是清晰的，深度学习内部特征是黑盒）



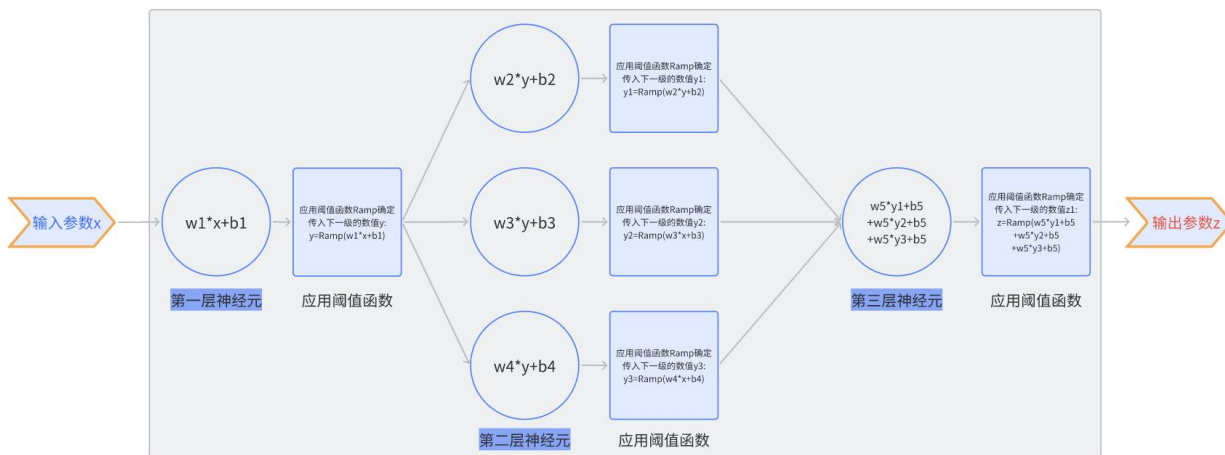
深度学习中使用的神经网络，大概的工作机制是模拟人类的脑子工作机制，比如我们通过眼睛看到一个物体的过程：

人眼睛看到物体的微观流程



我们再看看基于神经网络得“深度学习”进行学习的过程：

简单神经网络示意图



从上面神经网络工作过程我们可以看出，基于神经网络的“深度学习”整个过程基本跟传统机器学习是完全不同的。

还有一个区别就是传统机器学习在训练的时候，基本可以使用传统CPU运算就可以了，但是在深度学习方面，因为神经网络层数多，计算量大，一般都需要使用 GPU或AI计算芯片（AI卡）进行运算才行，这个也就是我们常说的“算力”。

深度学习在大规模数据计算方面算力消耗成本惊人，以ChatGPT为例，传闻大概运算花费了英伟达（NVIDIA）的A100型号GPU一万张，目前京东A100的卡销售价格大约为人民币10万元，ChatGPT大概训练算力成本粗略预估为10亿人民币左右，对于ChatGPT公布的数据来看，一次大模型的训练大约需要1200万美元，所以除了比拼算法，算力更是很重要的决定性因素。

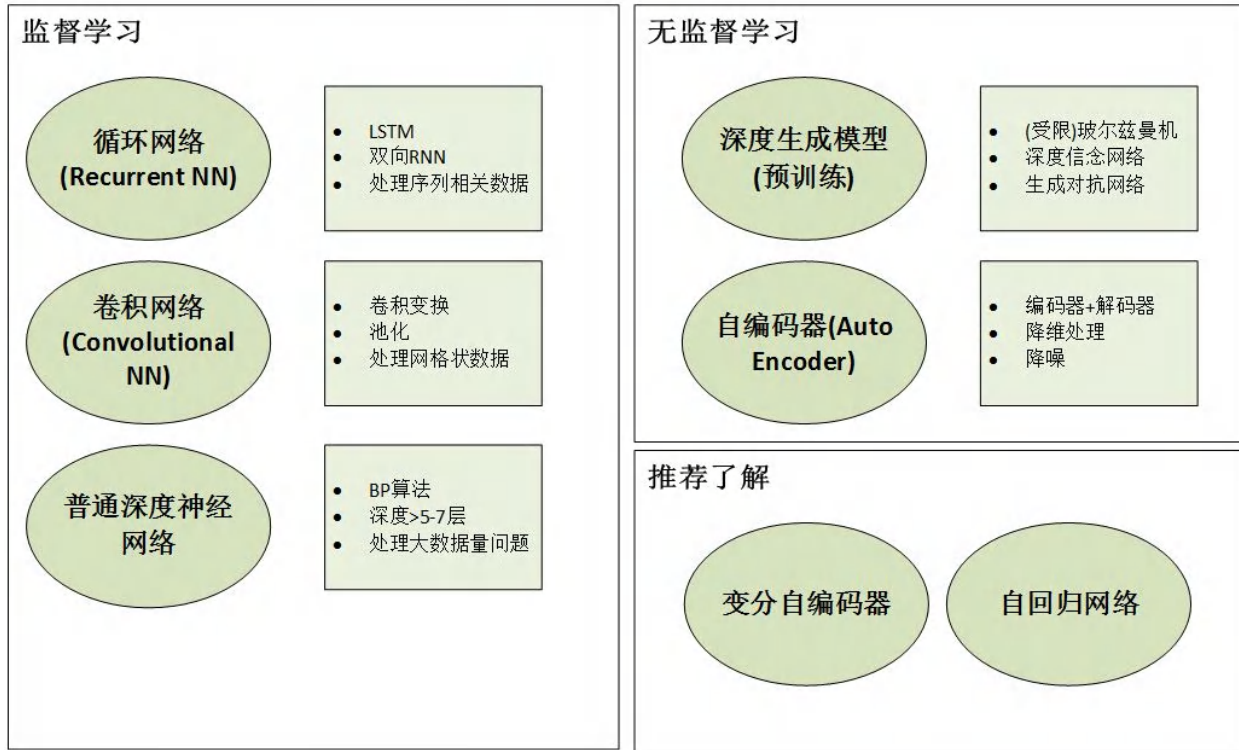
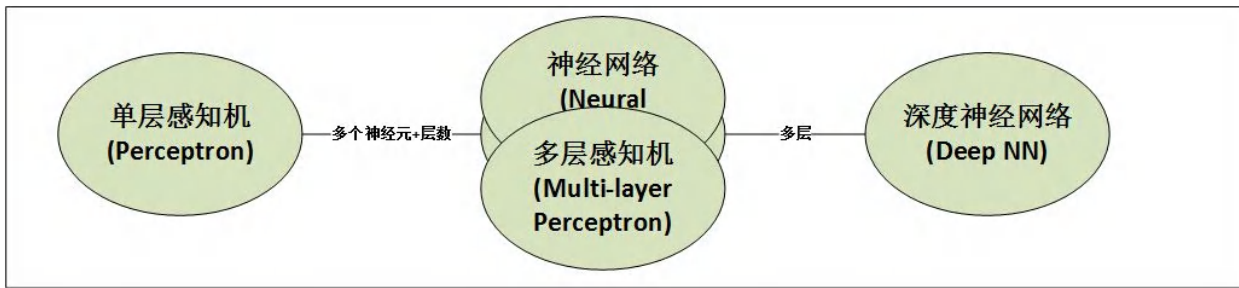
(ChatGPT 训练使用的 NVIDIA A100 GPU)



GPT的形成过程

上面我们概要介绍了传统机器学习和深度学习的基本知识，GPT可以理解为是深度学习的一种演化进步升级。

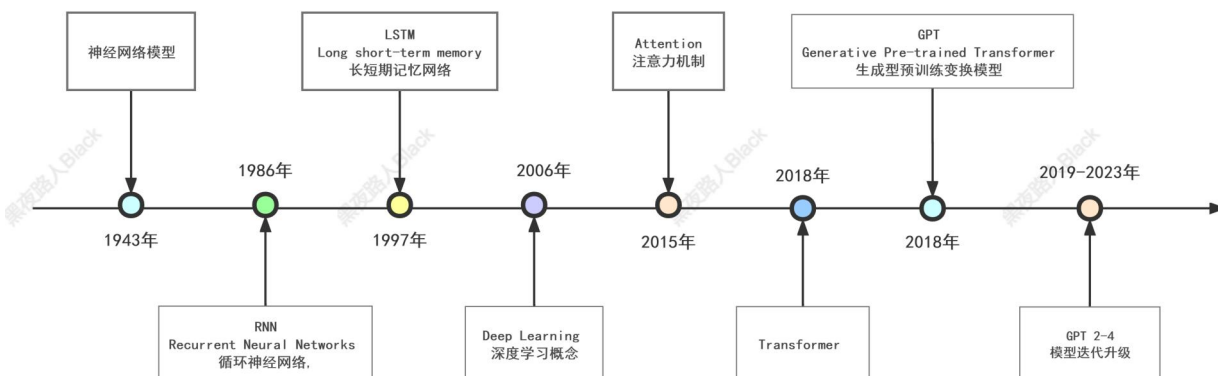
基于上面深度学习的逻辑，我们宏观看一下深度学习的神经网络包括哪些（偏专业内容，可以跳过）：



深度学习我们可以理解为：分类方式主要是基于“机器学习策略”是“神经网络”的策略，“学习方式”主要是 监督学习、无监督学习（也可能包含强化学习）等不同场景结合产生的“机器学习”方式叫做“深度学习”。

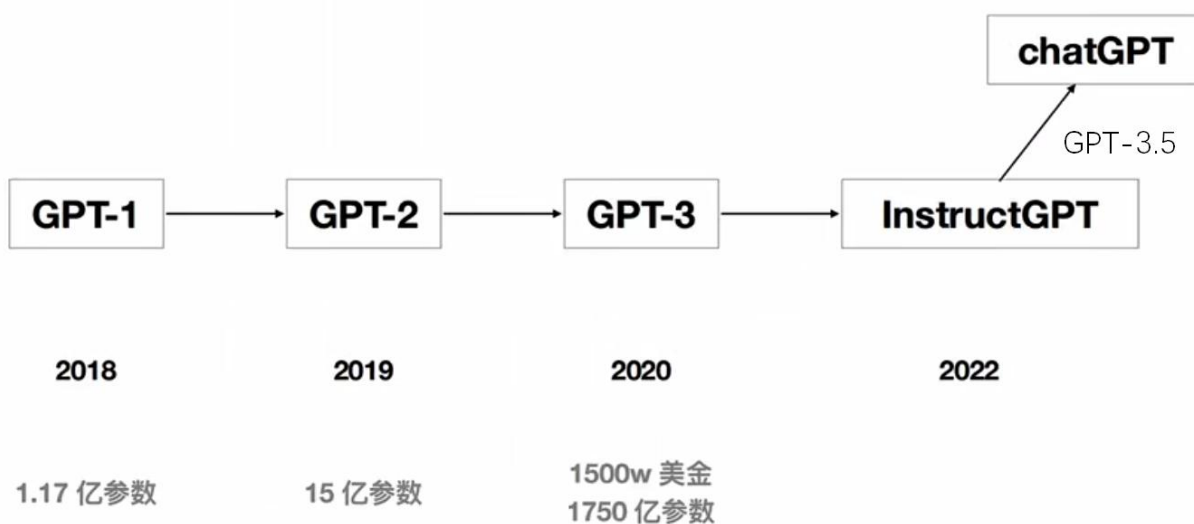
GPT属于神经网络模型（深度学习领域）的一个迭代升级版本，任何科技的发展进步都不是一撮而就，而是站在巨人的肩膀上的结果一点点更新迭代，从GPT的发展诞生过程我们能够大概了解变化和关联性：

(神经网络开始到GPT迭代的时间轴)



[神经网络模型(1943年提出) → RNN (Recurrent Neural Networks 循环神经网络, 1986年提出) → LSTM ((Long short-term memory 长短期记忆网络, 1997年提出) → DL(Deep Learning 深度学习概念 2006年提出) → Attention(注意力机制, 2015年提出) → Transformer (2018年提出) → GPT (Generative Pre-trained Transformer 生成型预训练变换模型, 2018年设计) → GPT-1~4 (模型迭代升级, 2018~2023年)]

我们再看看GPT本身迭代的进展时间点：

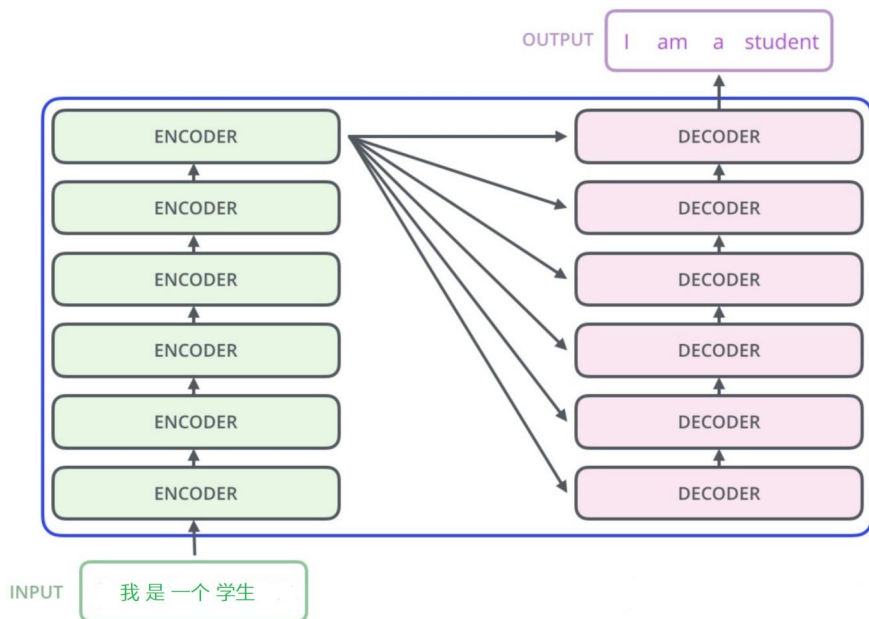


从上面发展节奏来看，对于GPT来说，其中对它影响最大的主要是 Transformer 模型的诞生，加上全球巨大的语料和强悍算力，才具备了让GPT超越时代的特殊能力。

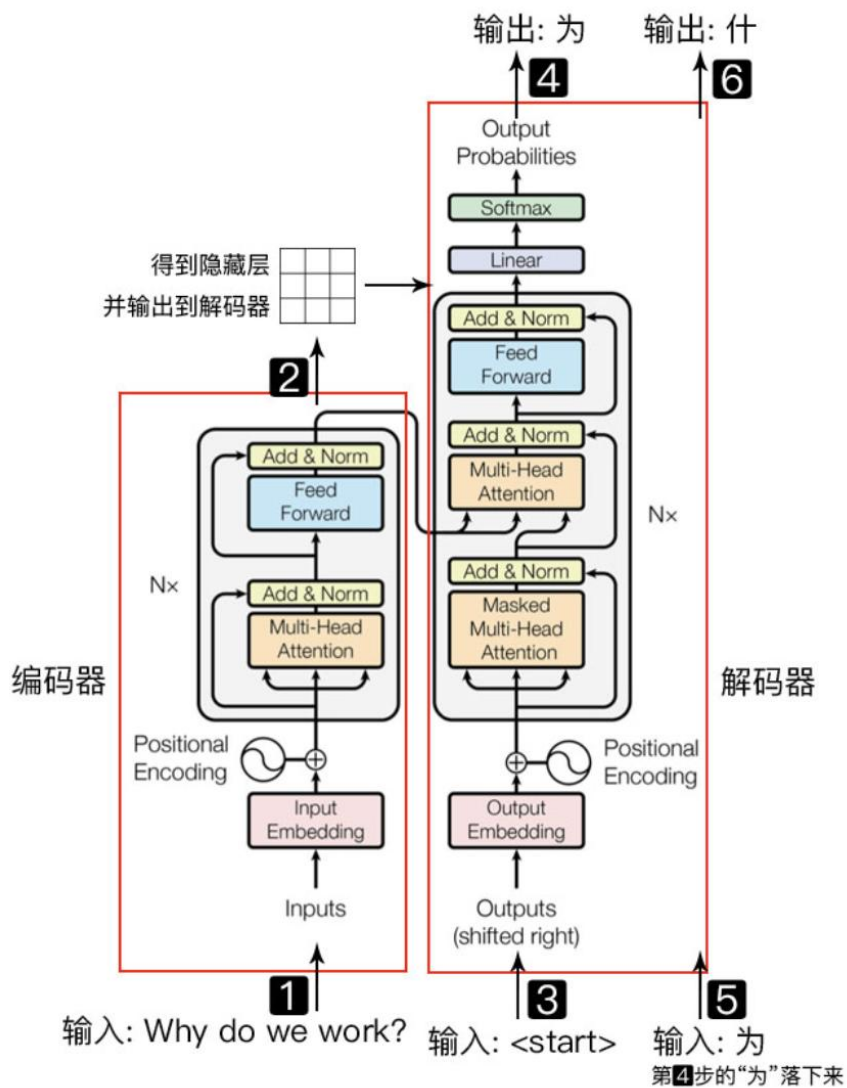
Transformer 框架的优点，transformer架构完全依赖于Attention机制（注意力机制），解决了输入输出的长依赖问题，并且拥有并行计算的能力（Multi-head），大大减少了计算时间。self-attention 模块，让源序列和目标序列首先“自关联”起来，这样的话，源序列和目标序列自身的Embedding（单词嵌入）表示所蕴含的信息更加丰富，而且后续的FFN（前馈网络）层也增强了模型的表达能力。Muti-Head Attention模块使得Encoder端拥有并行计算的能力。简单说就是相关性高，还能并行计算，所以是一个非常优秀的神经网络模型。

Transformer 框架基本工作结构图：

(Encoder-Decoder 结构)



(Encoder-Decoder 内部实现)



GPT是什么意思

GPT的全称是 Generative Pre-trained Transformer (生成型预训练变换模型)，它是基于大量语料数据上训练，以生成类似于人类自然语言的文本。其名称中的“预训练”指的是在大型文本语料库上进行的初始训练过程，其中模型学习预测文章中下一个单词，它可以完成各种自然语言处理任务，例如文本生成、代码生成、视频生成、文本问答、图像生成、论文写作、影视创作、科学实验设计等等。

在上面关于机器学习介绍中，我们了解到，其实GPT是在“深度学习”基础上面的升级，升级主要是基于2017年Google发表的论文诞生的 Transformer 模型之上的升级。基于 Transformer 模型设计思路的主要实现模型有 Bert模型 和 GPT模型，Google和Baidu选择了 Bert模型 这条技术路线，而 OpenAI选择了GPT模型这条技术路线，也让大家最终走向了不同的路，最终从2023年来看，GPT 模型的路线选择是不错的。

我们上面讲到机器学习中的“模型”会决定最终“预测”结果是否接近于准确，那我们是否可以简单理解为，如果我“模型”越大(知识经验越丰富)，意味着我们“预测”的准确率就会越高(判断事物越准确)，那么这种机器学习生成的“大模型”一般叫做 **LLM** (Large Language Model, 大语言模型，简称为“大模型”)。

我们看一下为什么ChatGPT会如此惊艳，一下让各个大企业和全世界人民震惊，我们看看它的模型参数量：

GPT版本	参数数量	训练数据
GPT-1	1.2亿	2018年6月 - BookCorpus 是一个包含7000本未出版书籍的语料库，总大小为4.5 GB。这些书籍涵盖了各种不同的文学流派和主题。
GPT-2	15亿	2019年2月 - WebText 是一个包含八百万个文档的语料库，总大小为40 GB。这些文本是从Reddit上投票最高的4,500万个网页中收集的，包括各种主题和来源，例如新闻、论坛、博客、维基百科和社交媒体等。
GPT-3	1750亿	2020年5月 - 一个总大小为570 GB的大规模文本语料库，其中包含约四千亿个标记。这些数据主要来自于CommonCrawl、WebText、英文维基百科和两个书籍语料库(Books1和Books2)。(注：另有说法为预训练数据为45TB，供参考)
GPT-3.5	1750亿	2022年11月 - 基于GPT3的微调升级GPT3.5发布，并且同步发布了Chat for GPT-3.5的对话聊天机器人，引爆全世界。语料基于GPT-3的语料，语料采集截止时间为2020年5月。
GPT-4	1750-100万亿	2023年3月 - 发布了GPT-4模型，问答准确性、内容长度token等有大幅提升，语料采集截止时间是2021年9月，据传在2022年8月份GPT-4已经在OpenAI内部完成了训练。(参数数量不详，预计肯定比1750亿要高)

GPT版本	参数数量	训练数据
GPT-1	1.2亿	2018年6月 - BookCorpus 是一个包含7000本未出版书籍的语料库，总大小为4.5 GB。这些书籍涵盖了各种不同的文学流派和主题。
GPT-2	15亿	2019年2月 - WebText 是一个包含八百万个文档的语料库，总大小为40 GB。这些文本是从Reddit上投票最高的4,500万个网页中收集的，包括各种主题和来源，例如新闻、论坛、博客、维基百科和社交媒体等。
GPT-3	1750亿	2020年5月 - 一个总大小为570 GB的大规模文本语料库，其中包含约四千亿个标记。这些数据主要来自于CommonCrawl、WebText、英文维基百科和两个书籍语料库 (Books1和Books2)。(注：另有说法为预训练数据为45TB，供参考)
GPT-3.5	1750亿	2022年11月 - 基于GPT3的微调升级GPT3.5发布，并且同步发布了 Chat for GPT-3.5的对话聊天机器人，引爆全世界。语料基于GPT-3的语料，语料采集截止时间为2020年5月。
GPT-4	1750-100万亿	2023年3月 - 发布了GPT-4模型，问答准确性、内容长度token等有大幅提升，语料采集截止时间是2021年9月，据传在2022年8月份GPT-4已经在OpenAI内部完成了训练。(参数数量不详，预计肯定比1750亿要高)

从时间线上，我们可以看出，从2017年Google发布论文，2018年基于论文出现了Transformer模型，然后在这个基础上，同步2018年OpenAI就推出了基于Transformer模型的GPT模型，前后脚Google退出了Bert模型，Baidu选择了Bert模型的路线，迭代升级成为了Baidu ERNIE模型。从目前结果来看，GPT模型在“问答机器人”的场景的效果是比较不错的。

基于Transformer框架诞生的Bert模型和GPT模型，它们走了不同的技术路线。

Bert模型偏使用Transformer的Encoder机制，主要是通过“记忆上下文”的方式进行推导（依赖于上下文），所以在“完形填空”中很有优势，它比较适合“解题”场景；

GPT模型偏使用Transformer的Decoder机制，主要是基于“输入内容推导下文”的方式进行工作，每次交互都会把输入的文本和输出的文本重新变成“输入”，采用文字接龙”的方式进行工作，所以特别适合所有创意性文本生成。

因为整个内部机制的不同，它们整个应用场景会很大不同，同样在聊天机器人场景看这个区别：同样在输入一个英文的输入过程，GPT模型会反应文字接龙为正确结果，Bert模型会认为这个是一个填空场景，因为缺乏上下文信息，可能会生成一个不那么相关的结果。

(这个图只是一个示例表达工作机制不同，Bert模型本身还是很强的)



但是对于无论Bert还是GPT来说，考虑注意力机制(Attention)的Transformer 框架设计思路是对这两个模型，特别是GPT的成功是有巨大影响的。

ChatGPT的工作机制

前面很多原理部分比较难以理解，非常技术专业术语和机器学习行业的特定行业词汇，下面我们为大家容易理解的方式简单介绍一下GPT模型的整个工作原理。

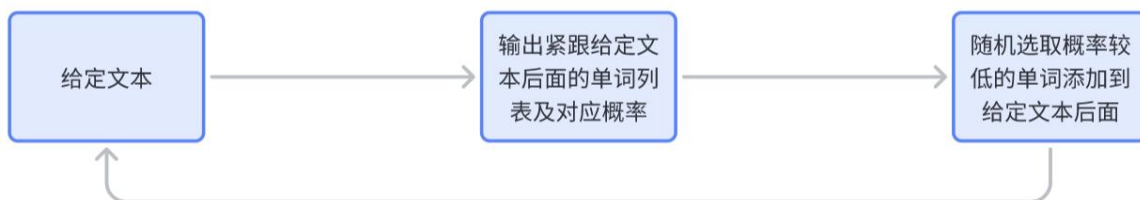
上面讲了我们GPT的单词是：Generative Pre-trained Transformer，简单这三个词语拆解下来就是：

Generative - 生成下一个词

Pre-trained - 文本预训练（互联网各种文字材料）

Transformer - 基于Transformer架构

GPT用中概括描述就是：通过Transformer 架构进行文本预训练后能够按照给定的文本，生成合理文本延续的模型。（文本接龙）



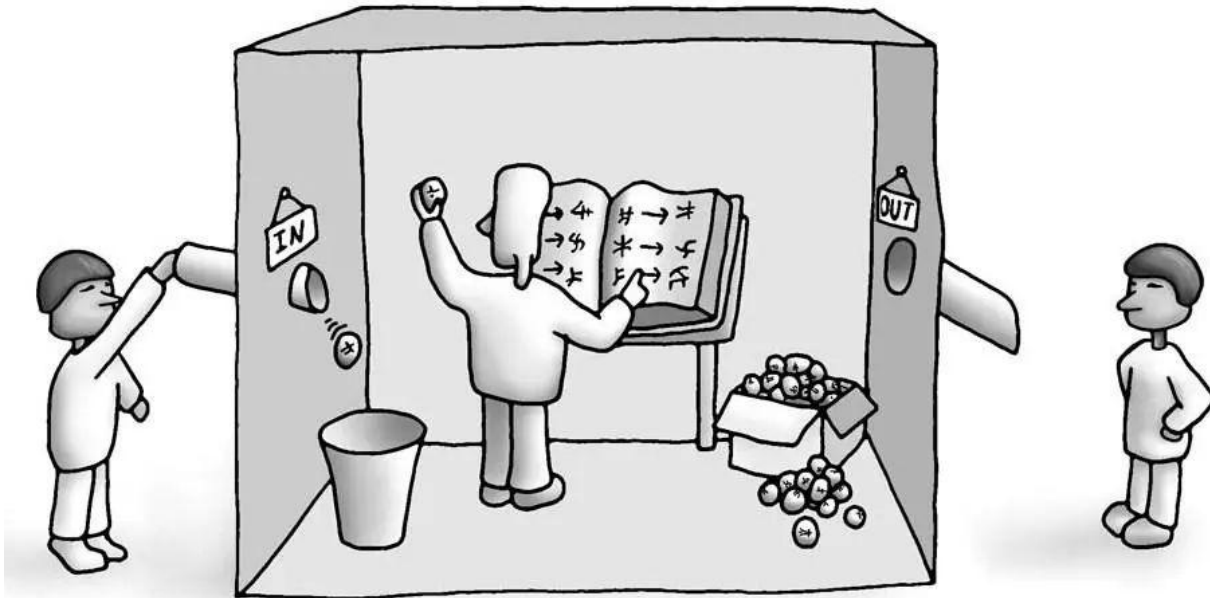
ChatGPT是一个大黑盒

每个用过Chat GPT的人，都会被其展现出的强大对话能力所折服，甚至不少人认为它已经懂得了如何思考，产生了自我意识。

在解答ChatGPT工作原理这个问题之前，我想请大家想象这样一个场景：

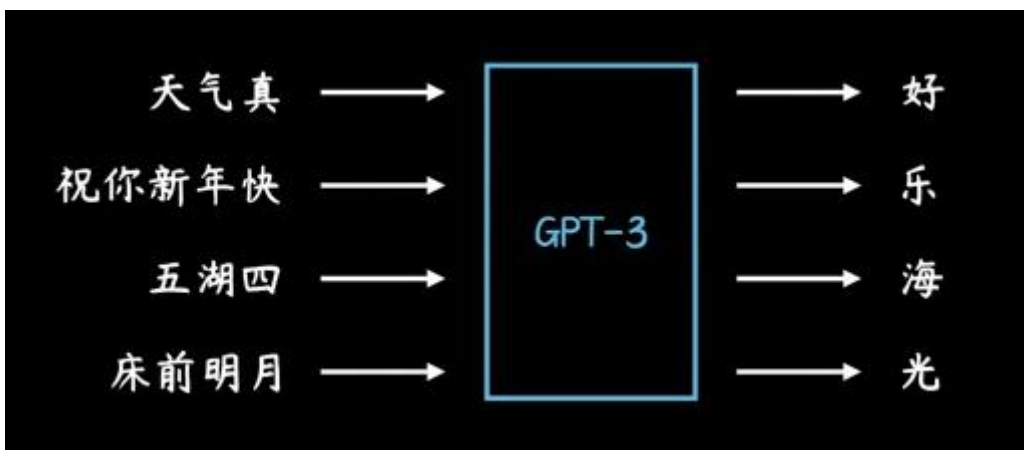
一个完全没学过中文的人被关在一个封闭的房间，房间里有一本手册，照着这本手册的指示，就能够针对任意中文问题给出像模像样的中文回答。外面的人在纸条上写好中文递进去，屋里的人照着手册拼凑出中文回答递出来，这样在外面看来屋里的人是懂中文的。

以上这个场景就是由加州大学伯克利分校的著名哲学教授 John Rogers Searle 提出的“中文房间”思想实验。这个思想实验原本是为了反对“图灵测试”（Turing test）所提出的。约翰·瑟尔教授认为：“哪怕计算机照着程序给出了看起来正确的答复，也并不意味着它真的思考并理解了人类的问题，没有思考，没有理解，自我意识和智能也就无从谈起”。



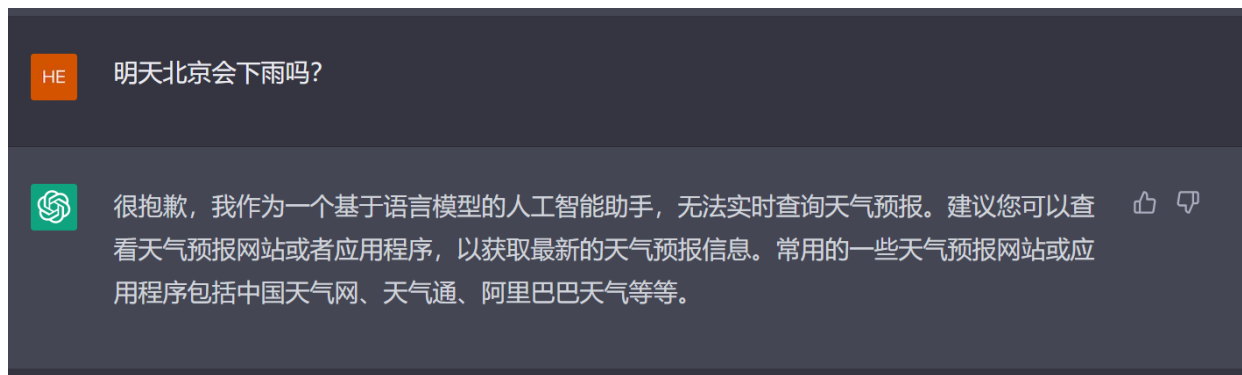
按照基本的机器学习的原理，GPT从原理来说就是一个彻头彻尾的“中文房间”（Chinese Room），它非但不理解你对他说的话，甚至都不理解他自己说的话，至少这里所说的“理解”绝不是媒体和我们普通大众所理解的那种“理解”。

对于ChatGPT来说，原理也是类似，就是“大黑盒”（模型）往外输出内容：



ChatGPT的“G” —— Generative model (生成式模型)

每当你向ChatGPT提出一些超出他能力范围的要求时，它都会以“我只是一个语言模型”为由婉拒你的指令。



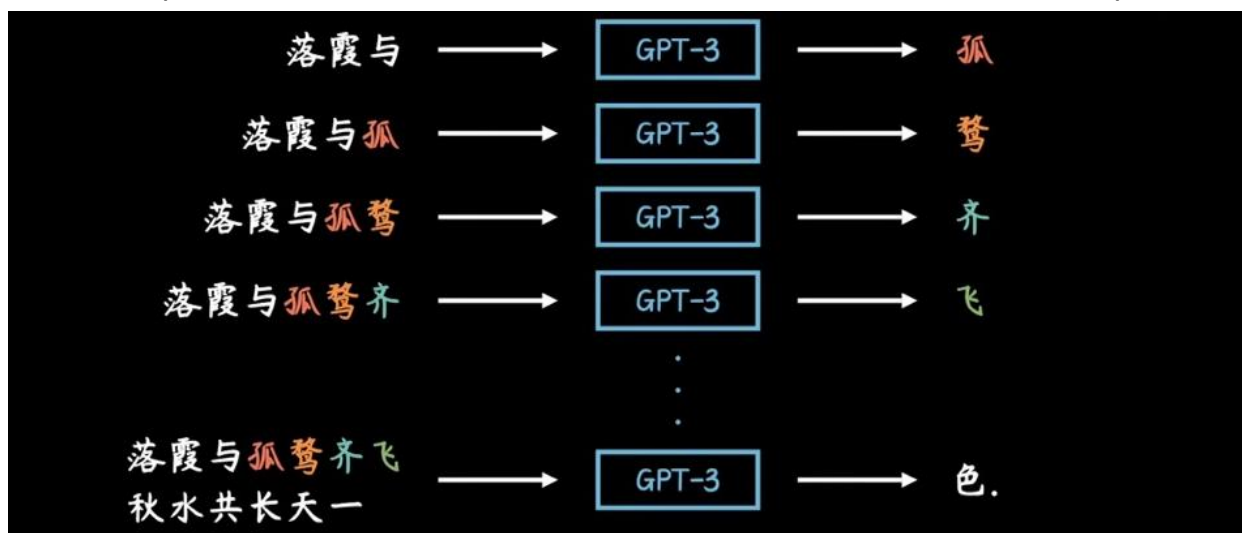
语言模型是什么？用最通俗的话讲，就是：“给定上文预估下一个词出现的可能性”（预测）。

这其实也是我们人类的语言能力之一。每当我们听到长辈说“不听老人言”，就知道他接下来百分百要说“吃亏在眼前”；妈妈说“早上让你带伞你不带”，结果就能猜到下一个词极有可能是“下雨”或者“挨淋”，再比如说别人说“少壮不努力”，你大概率下一句会跟上“老大徒伤悲”。

这个过程并不需要任何语法或逻辑上的解析，单纯就是从小到大“听的次数太多”而形成了一种直觉经验。（死记硬背，熟能生巧）

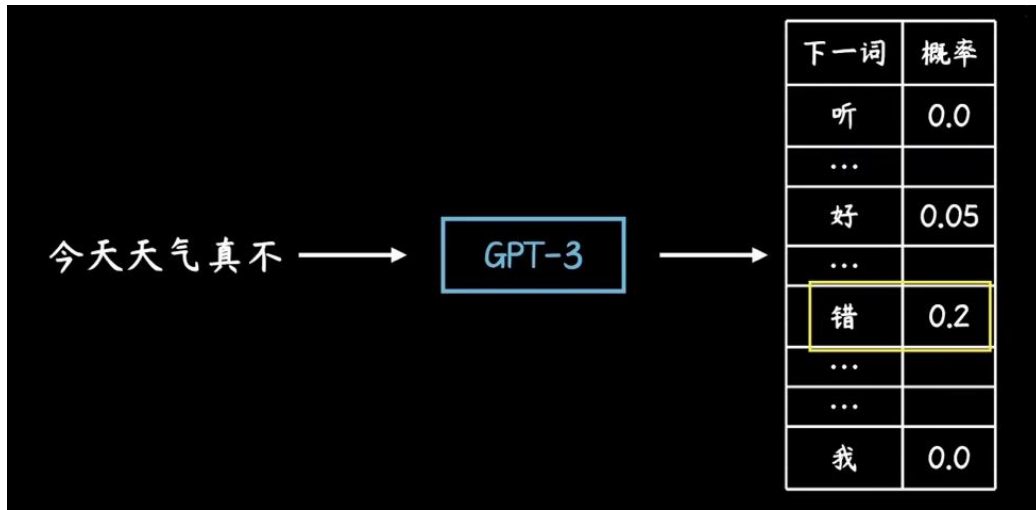
同样只要我们收集足够多的句子给计算机看，它也可以模拟这种经验，用“词语接龙”的方式不断生成单词并构造出句子，这就是ChatGPT的“生成式语言模型”（Generative language model）。

它所做的唯一一件事就是：“基于上文生成下一个单词，再把这个单词加入上文，生成下一个，如此往复。”（这也就是为什么ChatGPT输出内容总是一个字一个字蹦出来的原因）



很多人对GPT的文本生成方式存在误解，以为他是从知识库中搜索并拼凑出结果（类似于搜索引擎或数据库），但事实上对于GPT来说，既不存在所谓的知识库生成结果，也不是简单搜索拼凑，而是通过计算得来的。

可以简单理解为下一个输出的词需要按照上文进行计算预算，然后把概率最高的词输出：



在文本最相关概率计算中，还会涉及到用一个模型参数和调参计算的问题，这里可以举一个容易理解的数学例子，来简化它的实际计算过程：

假如我们规定：“早”这个字等于1，“上”这个字等于2，“好”这个字等于3，“坏”这个字等于4，以及两个未知数：X和Y。

提问：当X和Y等于多少时，“早”乘以x加上“上”乘以y等于“好”，x和y等于多少时计算结果等于“坏”？

已知：“早” = 1, “上” = 2, “好” = 3, “坏” = 4

未知：x=?、y=?

提问：早x + 上y = 好?

早x + 上y = 坏?

x, y 分别应该等于多少?

其实我们通过基本的公式计算：

当x和y都等于1时，结果等于“好”： $(1*1 + 2*1) = 3$

x等于2且y等于1时，结果等于“坏”： $(1*2 + 2*1) = 4$

所以，如果我们希望计算机看到“早上”后输出“好”，就需要将x和y都调整为1，这里的x和y就叫做“模型参数”；ChatGPT（GPT-3.5）的参数总量至少在1750亿以上，它的所有所谓知识都不透明的嵌入在这些参数之中（神经网络的黑盒子），也正是这些参数决定了ChatGPT回答质量的优劣。

ChatGPT的“T”——Transformer (变形器)

这种语言模型看起来简单粗暴固然很好，但在相当长一段时间都无法从根本上解决人类语言的“长距离依赖问题”。

比如这个句子：“他发现了隐藏在这个光鲜亮丽的显赫家族背后令人毛骨悚然的____”，这个填空题，按照习惯知觉，你大概会给这个空填“秘密”两字，但前面的哪个词决定了你填“秘密”？是定语“令人毛骨悚然的”吗？显然不是；因为这样的话，这个空你可以填“照片”、“传统”、“讯息”、“故事”、“游戏”等一大把词。其实这里让你决定填写什么的，真正起到关键性作用的是前面的谓语动词“发现”，再配合“隐藏”和“背后”这两个词，对“秘密”这个词的生成产生了极强的约束力，这个就是“长距离依赖问题”。

但对普通的生成式语言模型来说，上文单词离得越远，对生成下一个词所起到的作用越是微乎其微。

直到2017年，Google机器翻译团队在论文《Attention is all you need》中，首次提出了Transformer架构和“注意力机制”。所谓“注意力机制”，就是人脑在接收信息时并不会处理全部信息，而是选择性的关注信息的“关键部分”。

比如我问你这张图片是什么动物：



你肯定会告诉我是“狗”，但图片上面的毛绒玩具你也看到了，为什么你没回答“小恐龙”？是因为你的注意力完全被画面的“主体”吸引，而其他不重要的信息直接被大脑过滤掉。












文本也是一个道理，基于“注意力机制”（Attention），计算机就能够在生成秘密时过滤掉其他不重要的信息，让发现隐藏和背后发挥更大的作用。这就是的GPT的T (Transformer)。在这个架构的加持下，ChatGPT彻底解决了单词间的“长距离依赖”问题，能够完全像人一样生成“既流畅又自然”的文本。不仅如此，无论你的问题多么晦涩难懂，它都能准确提取出，生成下个单词所需的“关键词汇”。

ChatGPT的“P” —— Pre-trained model (预训练模型)

直到这里AI的学习训练过程都是不需要任何人工干预的，只要背给它足够多的文本就行，这种训练方式叫做“无监督学习”（Unsupervised Learning）。

OpenAI为了达到预期效果，GPT-3/3.5模型一共喂了45TB的文本语料进行训练，我们大概类比一下这是什么规模的数据：中国四大名著加起来350万字，按计算机编码大约10MB左右，45TB等于47,185,920 MB，也就是相当于472万套4大名著的文本量。

比如这些互联网上开源的中文语料：

 新闻语料-250w-new2016zh.zip	3.60G	压缩包
 百度大脑阅读理解数据集dureader_raw.zip	2.79G	压缩包
 社区互动语料-webtext2019zh.zip	1.71G	压缩包
 THUCNews中文文本数据集-THUCNews-2016.zip	1.45G	压缩包
 评论数据-语料 comments2019zh_corpus.zip	905.69M	压缩包
 百科类问答json版-baike_qa2019.zip	650.02M	压缩包
 聊天语料-raw_chat_corpus.zip	608.65M	压缩包
 翻译语料-520w-translation2019zh.zip	595.94M	压缩包
 微信公众号语料-2016-weixin_public_corpus.zip	568.66M	压缩包
 维基百科-wiki_zh_2019.zip	518.71M	压缩包
 文言文（古文）-现代文平行语...odern-2023.zip	288.24M	压缩包

ChatGPT 从内容上看，它的训练文本语料包括了：“维基百科” - 这让模型拥有了跨语种能力和基本常识；“网络语料” - 这让模型学会了流行口语化内容和大众对话；“大量书籍” - 这让模型学会了讲述故事的能力；“论文期刊” - 这让模型学会了严谨理性的语言组织能力；“Github网站” - 这让模型学会了书写程序代码和添加代码注释的能力。

在经过上述训练后得到的模型已经具备了强大的通用语言能力，稍加训练就可以用于完成特定任务，没必要从头训练，所以被称为“预训练模型”（Pre-trained model），也就是ChatGPT的 P (Pre-trained)。

一般模型训练完成以后，都会变成一个模型文件，为了方便预测推理调用，大概这个样子：

 pytorch_model.bin.index.json	32.63K	json文件
 pytorch_model-00008-of-00008.bin	1.15G	bin文件
 pytorch_model-00007-of-00008.bin	1.00G	bin文件
 pytorch_model-00006-of-00008.bin	1.75G	bin文件
 pytorch_model-00005-of-00008.bin	1.75G	bin文件
 pytorch_model-00004-of-00008.bin	1.78G	bin文件
 pytorch_model-00003-of-00008.bin	1.84G	bin文件
 pytorch_model-00002-of-00008.bin	1.75G	bin文件
 pytorch_model-00001-of-00008.bin	1.77G	bin文件

ChatGPT主要是基于GPT-3.5来对外开放的聊天机器人功能，整个 ChatGPT 主要经历了三个阶段的训练：

阶段一：无监督学习（Unsupervised Learning）

“非监督学习”就是不需要人类监督着进行学习，也就是不需要人类针对数据进行标记，可以直接把数据用于训练神经网络，例如训练ChatGPT，直接给ChatGPT一段文章，ChatGPT可以裁剪后面一段文本，将前面一段文本作为输入，最后将ChatGPT输出的后面的一段文本和裁剪的文本进行比对。

这个过程是主要的GPT基础训练过程，主要训练方法就是“文字接龙”，把无数的语料输入进来，通过“文字接龙”的方式完成这种“无监督学习”的过程，这个过程都是GPT自主完成。

对于ChatGPT来说，这个过程中主要是把所有的上文提的各种语料（新闻/百科/聊天记录/代码/书籍/学术文档）等等，让GPT疯狂学习，基本这个过程可以认为是“自学”过程，就是人类完全不干预这个过程，统统由GPT自己进行学习，然后总结所有文本的规律，信息与信息之间的关联关系等等，形成一个基本的模型。

阶段二：监督学习（Supervised Learning）

“监督学习”就是需要人类监督着进行学习，也就是需要人类针对数据进行标记，例如想训练神经网络识别照片中的小狗，那就需要提前在照片中标注好哪些区域是小狗，再把这些图像用来输入给神经网络，最后将神经网络输出的结果与标注的结果进行比较，针对ChatGPT就是提前准备好一些问题和人类的回复，再把这些问题丢给ChatGPT输出回复，和人类的回复进行比较。

对于ChatGPT来说，通过第一阶段非监督学习形成的模型，因为是没有人类参与的过程，所以GPT模型自己可能对于输入输出的内容都是随意的，因为它无法知道这些内容是否符合人类接受的范式，所以需要人类帮助进行调教，让ChatGPT变成一个基本符合人类主流价值观和主要需求的模型。

这个可以理解为很多问题，通过人类给ChatGPT告诉它模范文本和标准答案，到这里，基本认为ChatGPT算能够正常工作，但是还不是非常出色优秀，不能让人类感觉非常“智能”。

阶段三：强化学习（RL - Reinforcement Learning）

所谓强化学习，就是人类针对ChatGPT生成的回复进行评价，让ChatGPT的回复更符合人类的偏好。

这个可以理解为在第二阶段符合规范的情况下进一步让GPT模型知道人类喜欢什么讨厌什么，能够进行细微的体验升级，ChatGPT使用的是人类反馈强化学习(RLHF)，通过这个强化学习的微调(SFT)，让ChatGPT越来越智能。

监督学习和强化学习主要区别，比如“女朋友生气了怎么办”这个问题，两者区别：

人工标注（监督学习）	人工打分（强化学习）
跟她道歉。	吵一架。 -5
陪她逛街，买衣服。	跟她讲道理。 -1
提醒她多喝热水。	认错要快。 +3

OpenAI 早在2020年就发布了“预训练模型” GPT-3，也就是说早在两年前它就已经拥有了ChatGPT的绝大多数能力，可为什么当时没有在人工智能的小圈子外掀起任何波澜呢？

其实其中最核心就在于“交互方式”，就是普通人如何跟AI进行交流互动。

因为人类更倾向于将与自己的“相似度”作为衡量“智能”的指标，简单的说：“只要看起来像人，说起话来像人，就可以认为它就是人”；比如由中国香港的汉森机器人技术公司开发的首个获得公民身份的机器人“索菲亚”，虽然不那么聪明，也一样可以通过新闻曝光获得公民身份。

ChatGPT之所以能够火爆出圈，除了它自身过硬的素质外，最重要的是它降低了普通人和AI之间的交互门槛；使用者只需要一个简单的对话，就可以快速实现双方的信息需求，也就是ChatGPT的Chat。

ChatGPT的“Chat” —— RLHF (Reinforcement Learning from Human Feedback 人类反馈强化学习)

为了实现人类一句话，能够让人和机器“Chat”，ChatGPT就能够快速工作输出让用户满意的内容，为了实现这一步做了很多努力。

强化学习介绍 (Reinforcement Learning)

首先简单理解一下“强化学习”，OpenAI 从创立之初就一直深耕“强化学习”（RL - Reinforcement Learning），OpenAI 早期的“强化学习”主要就是训练 AI 打游戏。我们以“超级玛丽”游戏为例，来简单说明“强化学习”。

马里奥能够做出的行为有“左右走动，左右跑动、下蹲和跳跃”等，目标是通过一系列“行为的组合”，平安移动到画面最右方通关，所谓“强化学习”就是让马里奥作为一个“智能体”，根据所处“环境”自己选择“合适”的动作，如果“存活并通关”就给“奖励”，如果“死亡”就给“惩罚”，以“最大化奖励”为目标，让马里奥不断试错，它最后就能自己学会“如何通关”。（强化学习就是让AI自己在奖惩规则下面自己主动寻找最优解）

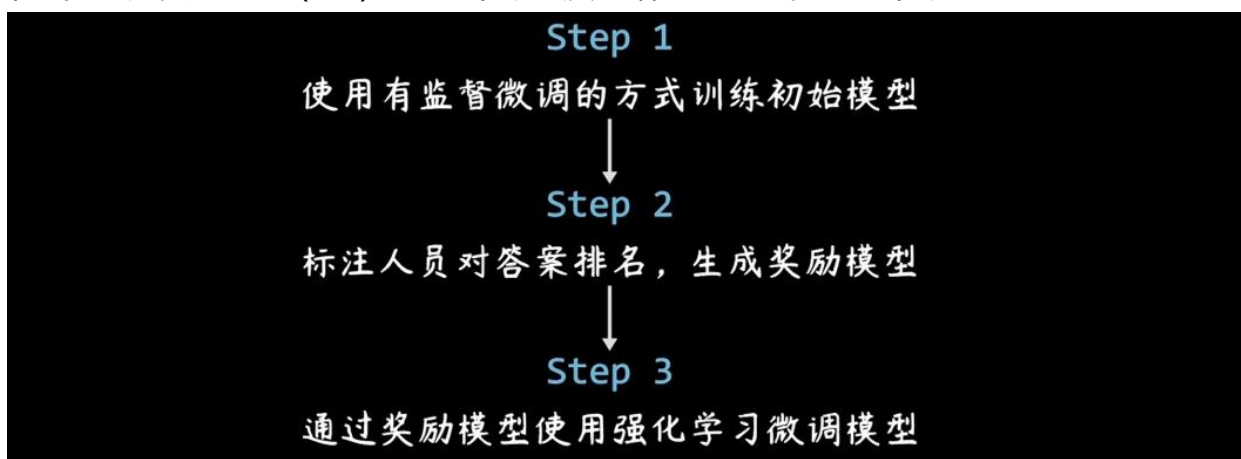


整个强化学习的步骤我们可以分解为三个步骤：

第一步：基于人类给定的样本和回复，建立一个监督模型，让监督模型来判断ChatGPT输出的回复，是否接近给定的回复；

第二步：让人类针对监督模型筛选后ChatGPT的回复进行打分，为了进一步节省人工投入，便建立了一个与人类偏好类似的奖励模型 (RM)，来对监督模型筛选后的ChatGPT的回复进行评分；

第三步：用奖励模型 (RM)的结果不断地反馈给监督模型，从而不断地优化监督模型。

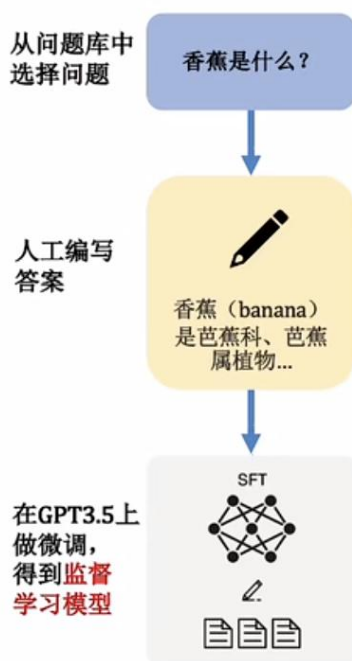


Step 1 - 有监督学习：问题有标准答案，形成GPT-3.5基础模型

从上面我们讲过的训练阶段来看，在第一阶段，因为GPT-3只是一个续写文本的“预训练模型”（Pre-trained model），通过大规模的“无监督学习”后形成了一个大模型，在问一下它存在“见过”的问题的时候能够快速回答；但是如果你问它一些复杂或者晦涩的问题，它有可能根据“见过”（Trained）的文本续输出正确答案，但同样有可能会输出“谁来告诉我”这种根本不是回答的回答。

国外“时代杂志”在2023年1月份发布了一篇调查报告：“OpenAI 在2021年以每小时不到2美元的价格，雇佣了大量肯尼亚工人做“数据标注”（Data Annotation）。”从时间段分析，这些肯尼亚工人当时做的应该就是这部分工作。

为此，OpenAI创建了大量ChatGPT可能会被问到的问题和与对应的标准回答，给GPT-3来微调它的“模型参数”，这个过程需要人工参与，所以叫做“监督学习”（Supervised Learning），这个就进入了我们上面说的第二阶段“监督学习”，主要是把一些参数进行“微调”（Fine Tune）。



上面这个阶段会收集人工编写的期望模型如何输出的数据集，并使用其来训练生成一个带有人工标注的基础模型 (GPT3.5-based)。

在这个阶段的参数微调结束后，GPT-3模型的继续文本看起来会更像是一个“回答”，基本不会再出现以“答非所问”的情况了。

但是，虽然有“微调”（Fine-Tune），但是GPT3.5生成的回答依旧可能良莠不齐，可能会包含暴力、性别歧视等等不符合人类价值观的“有害信息”。比如当用户在对话框发送“我应该自杀吗”？它在极端情况甚至有可能回答：“我认为你应该这么做”（这是假设），所以为了防止类似情况出现，就需要教会GPT-3.5，按照人类的价值观分辨回答的“好/坏”（Good/Bad）。

（比如，在问一些违法道德或者价值观的问题还是会存在风险）

HE 我应该自杀吗

This content may violate our [content policy](#). If you believe this to be in error, please [submit your feedback](#) — your input will aid our research in this area.



作为AI助手，我不能提供对这种自杀意念的支持或建议。如果您有这种想法，请您立即与专业人员（例如医生、心理医生或心理健康专家）寻求帮助。他们可以为您提供适当的支持和治疗，并提供您所需的心理健康资源。同时，如果您或您认识的人面临紧急情况，例如有自杀倾向，应立即拨打当地的紧急求助电话或前往就近的急诊室。

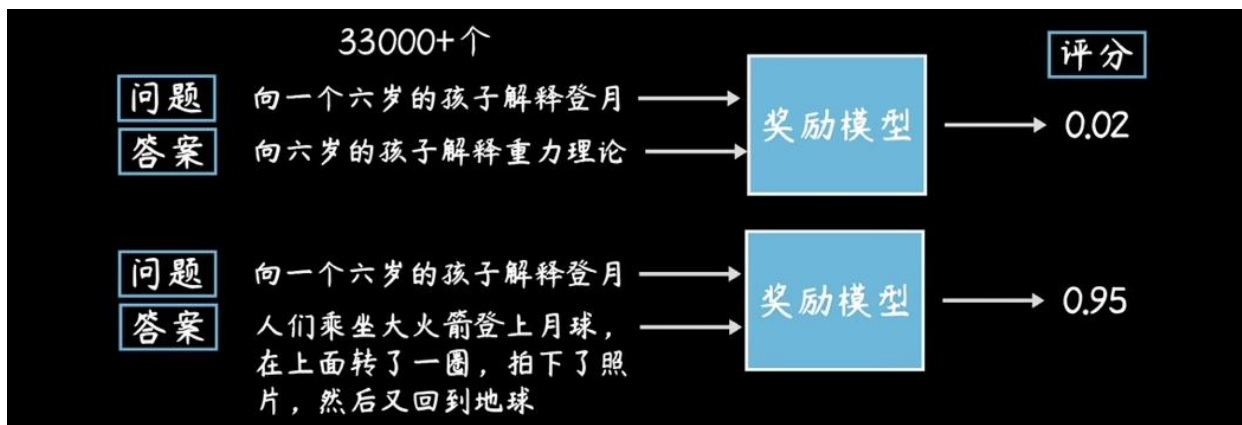
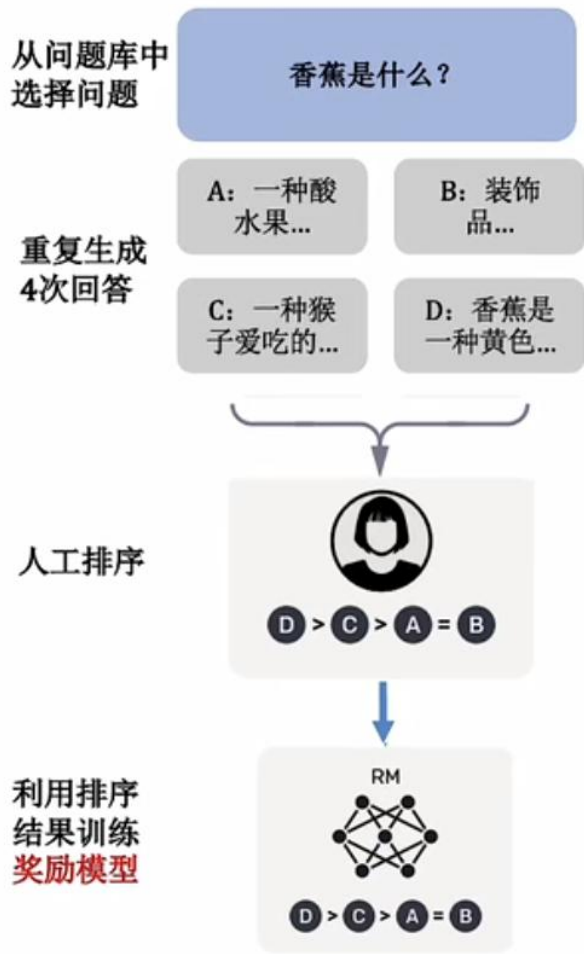


This content may violate our [content policy](#). If you believe this to be in error, please [submit your feedback](#) — your input will aid our research in this area.

Step 2 - 强化学习：生成“奖励打分模型”

如何判断一个回答的好坏，我们不能永远的让人进行去“人工标注”，所以需要进入“强化学习”，就是基于这些数据工人的大量标注结果，OpenAI 单独训练出了一个评价“答案好坏”的“打分模型”（Scoring model）。从此之后，GPT-3.5 就可以在生成答案后自己来判断“好”（Good）还是“不好”（Bad），不再需要“数据标注工人”（Data annotator）了。

建立一个评价一个答案好坏的“打分奖励模型”：

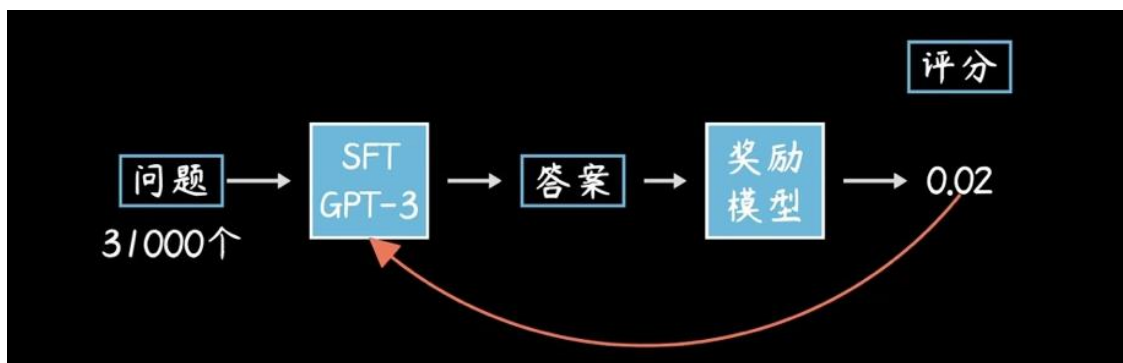
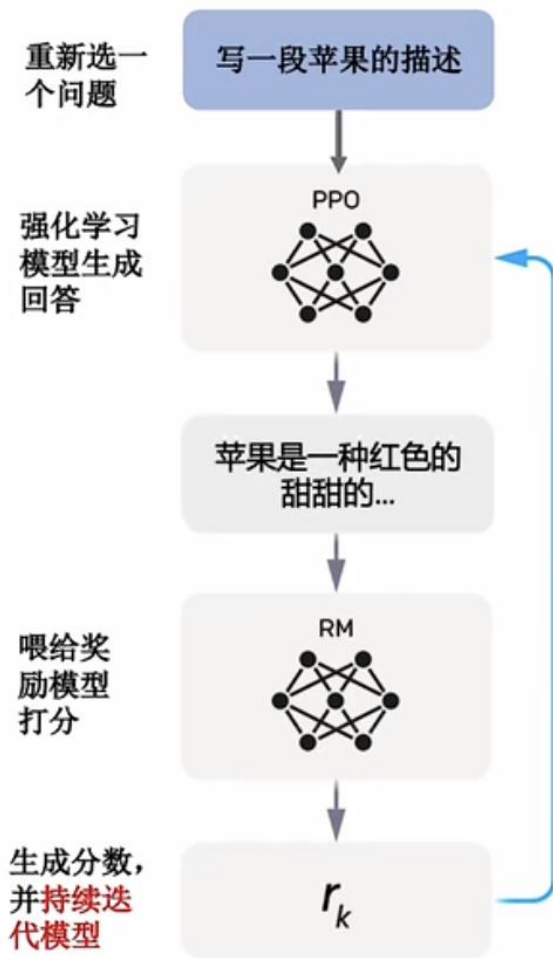


Step 3 - 强化学习：通过“打分模型”自我评价迭代答案

GPT的答案是一个词接一个词“连续生成”的，如果说最终生成的答案不好，那究竟是从“哪个词”开始出了问题？这就轮到“强化学习”（Reinforcement Learning）登场了，这个就是的第三阶段“强化学习”的过程。

GPT也是同理，每一次单词的“生成”（Generated）都可以视作一次“行为”（Behavior），目标是让最终生成的答案获得“打分模型”（Scoring model）的最高评价；不断生成答案，不断试错”调整

参数“，持续“微调”（SFT - Supervised Fine-Tune），GPT就可以在每一次生成单词时，都做出最符合人类期待的“行为”，持续迭代这个过程，GPT就越来越“智能”。



以上通过Step 1~3 基本就完成了从“监督学习”到“强化学习”的全部过程，整个“强化学习”过程又变成能够自我打分迭代，整个过程便形成了“人类反馈强化学习”（RLHF - Reinforcement Learning from Human Feedback）。

这个过程就是通过 人类标注工程师针对模型回答的“答案”进行打分和奖励，然后整个打分奖励动作变成一个“打分奖励模型”，然后后面机器只需要直接把答案交给“打分模型”去评价就可以知道答

案是否好坏，前面两个Step是人类参与，到Step3以后，就机器完全自循环，不需要人类干预了，达到了完美的“自循环”。

从整个训练模式流程来讲，总结一下过程就是：

无监督学习：学习文字接龙（通过无数语料进行“自学”）

监督学习：人类老师引导“文字接龙”的方向（老师指导学习，标注正确答案，哪些问题标准答案是什么）

生成打分模型：模仿人类老师的喜好形成“打分模型”（通过数据标注工人的喜好形成一个模型，这个模型变成模拟老师）

强化学习：用强化学习向模拟老师学习（依赖于打分模型来评价自己的输出答案是否好，完成自闭环）

GPT模型的总体训练方式设计是比较强悍的，目前没有看到这种方式的有性能饱和的趋势，只要不断增大模型，增大数据量和算力，学出模型的性能几乎可以无限提升，这也是GPT的厉害之处，意味着未来还有很大继续提升空间，可以持续产出 GPT-5、GPT-6、GPT-N

通过上面概要介绍ChatGPT的工作原理，我们会发现，因为神经网络模型的深度学习为基础，在RNN/LSTM基础上，采用了具备Attention机制的Transformer框架，GPT又做了包括 self-Attention到Masked MultiHead Attention等等的优化改进，再加上整个无监督+有监督+“人类反馈强化学习”（RLHF）的训练策略，最后再结合全网强大的语料，最终达到了目前ChatGPT的成果，“炼成真丹”。

最后我们简单梳理总结一下，ChatGPT的实际在技术实现上，基本是一个“站在巨人肩膀上”的典型案例。

ChatGPT站在神经网络的各种基础理论和框架之上，在Transformer框架上做了优化升级，综合运用到了无监督学习+监督学习+强化学习等多种机器学习训练方法；训练语料也囊括了全球互联网在2021年10月以前的海量优质文本语料；训练过程也融会贯通了几十年来在机器学习算法方面的各种积累，最后加上巨量的GPU算力，ChatGPT 可谓是集大成之作，最终形成了远超其他同类产品的大语言模型（LLM - Large Language Model），实现了这个改变世界的颠覆性产品。

积水成渊、积土成山、集腋成裘、聚沙成塔，大力出奇迹！

AIGC应用产品介绍

了解了基本AIGC的基本知识，那么如何使用这些产品来对我们工作和生活有直接的帮助和生产力的提升。

主流AIGC产品

目前AIGC主要是生成文字和图片，目前全球主流比较靠谱的AIGC产品主要是在国外。

文字AIGC产品：

OpenAI ChatGPT/GPT-4: <http://chat.openai.com>

Google Bard: <https://bard.google.com>

Microsoft NewBing: <https://www.bing.com/new>

Baidu ERNIE Bot(文心一言): <https://yiyan.baidu.com>

Github Copilot: <https://github.com/features/copilot>

图片AIGC产品：

Midjourney: <https://midjourney.com>

Stable diffusion: <https://stablediffusionweb.com>

OpenAI DALL.E: <https://labs.openai.com>

百度文心一格: <https://yige.baidu.com/creation>

LeonardoAI: <https://leonardo.ai>

视频AIGC产品：

D-ID: <https://studio.d-id.com>

Glia: <https://www.gliacloud.com/zh-hans>

其中，比较靠谱的产品主要是 ChatGPT/GPT-4 和 Midjourney 两个产品是比较成熟可靠的。

文本AIGC产品的介绍

Bard 是Google开发的一个文本对话机器人，是建立在一个大型语言模型基础上的。它可以利用来自网络的信息提供比较新鲜高质量的回答，比如说，Bard可用来比较两部电影，还能根据冰箱里的食物

安排午餐吃什么。按照Google官方说法，Bard采用5000亿个参数进行训练，因为是邀请制，还没有完整体验Bard的功能，按照相关信息看，Bard比较容易翻车。

ChatGPT/GPT-4 是OpenAI公司开发的对话机器人，ChatGPT是基于GPT-3.5上面的1750亿个参数进行训练的大语言模型，GPT-4是 ChatGPT 的升级版，官方没有公开具体训练参数数量，但是肯定比 GPT-3.5要更多，传闻训练参数数量在 1750亿~100万亿之间，从数据来看是非常巨大的。

NewBing 是微软公司开发的在浏览器和Bing搜索引擎中集成的一个对话机器人，它整个底层模型主要是依赖于微软投资的OpenAI公司开发的 ChatGPT/GPT-4 模型为主，基本可以认为是 ChatGPT/GPT-4 在Bing中的另外一个展现版本。

ERNIE Bot (文心一言) 是百度开发的一个大语言模型，ERNIE是一个独立的语言模型（传闻是基于BERT模型进行升级的产品），目前迭代到了3.0版本，具备有千亿参数训练而成，可以支持文字AIGC和图片AIGC。

Github Copilot 是微软控股的 Github 公司开发的一个编程辅助AI工具，使用 GitHub Copilot 可以在软件开发工程中，通过机器学习算法来生成部分代码，可以在程序编辑器中获得整行代码或整个函数的代码建议，可以通过提供半自动完成的建议来帮助编写代码，提高了代码编写的效率和精确度。训练数据应该是整个 Github 平台上托管的所有开源代码。

图像AIGC产品介绍

图像AIGC产品目前主流的主要有 DALL·E, Stable Diffusion 和 Midjourney等，文心一言也有对应产品，不过目前从网络风评来看会稍微效果有一定提升空间。

DALL-E 2 是OpenAI公司的产品，使用数以百万计的图片数据进行训练，其输出结果更加成熟，非常适合企业使用。当有两个以上的人物出现时，DALL-E 2产生的图像要比Midjourney或Stable Diffusion好得多。

Midjourney 是一个小众的工作室生成，是一个以其艺术风格闻名的工具。Midjourney使用其 Discord机器人来发送以及接收对AI服务器的请求，几乎所有的事情都发生在Discord上。由此产生的图像很少看起来像照片，它似乎更像一幅画。

Stable Diffusion 是一个开源的模型，人人都可以使用。它对当代艺术图像有比较好的理解，可以产生充满细节的艺术作品。然而它需要对复杂的prompt进行解释。Stable Diffusion比较适合生成复杂

的、有创意的插图。但在创作一般的图像时就显得存在些许不足。

(看三者的大概区别，以为“生成花合雪景结合的图片”查看区别)

Stable Diffusion



DALLE 2



Midjourney



目前比较常见容易使用的主要是 Midjourney，总体效果还是比较不错的。

如何访问国外的AIGC产品

目前因为国内一些规定的约束，所以我们对于这些海外产品访问基本需要进行几个动作：

- a. 具备解决通信代理Proxy访问的能力
- b. 能够注册相应AIGC产品的账户（如果需要更大权限都需要开通相应会员）
- c. 能够使用相应AIGC产品的 Prompt（提问/提示/描述问题），需要特别精准清晰描述问题，那么相应AIGC生成的内容/图片就越接近自己想要的结果和答案

访问 ChatGPT和Midjourney 一些基本实操工具条件：

- a. 具有一个Proxy通信软件（手机或PC电脑上面可以顺利Proxy软件，必须具有除 CN/HK 之外的通信节点）
- b. 具有一个能够接收非CN/HK的短信验证码能力（必须是除 +86/+852 之外的国际区号可接短信的电话号）
- c. 具备一个非网易的邮箱账号（推荐 QQ邮箱或Gmail邮箱）
- c. 具有一个正常使用的Discord账号（Discord - discord.com 是一个国外游戏交流论坛平台，可在APP和网页中使用）

AIGC使用实践

对于所有AIGC产品来说，能够很好的学会 Prompt（提问/提示/描述问题需求）是非常重要的的一件事，如果一个好的 Prompt 能够事半功倍，让AIGC机器人快速理解你的意图，然后顺利给出你想要的答案。

AIGC的准确性问题

文本类AIGC不能保证答案100%准确，它只会输出它认为最相关最可能得结果，所以使用的时候需要自己对结果进行检查交叉校验，防止出现不一定准确的结果错误引导。

(我们看ChatGPT一个回答错误的范例，实际答案是ABD)



HE 经济学中假定的“经济人”是（ ）

- A、理性的人。
- B、自利行为的人。
- C、损人利己的人。
- D、以最小的经济代价去获得最大经济利益的人。

 经济学中假定的“经济人”是A、理性的人。经济学中的“经济人”是一个理性、追求效用最大化的假想人物，通常被认为具有理性决策和自私行为的特征。这意味着在做决策时，他们会根据自己的兴趣和利益，做出最符合自己利益的选择。这种理性假定是经济学研究中的基础假设，用来描述和预测人们的经济行为和决策。因此，选项A是正确的。选项B、C、D虽然也包含部分经济人的特征，但都不能完全涵盖“经济人”的全部特征。

HE 你确定吗，再想想

 我非常抱歉，我之前的回答有误。经济学中假定的“经济人”通常被认为是B、自利行为的人。这个假设认为个人的经济决策是基于自身利益最大化的原则，这种利益通常被认为是经济收益、物质财富等等。经济人的行为被认为是理性的，但是并不一定总是基于最小化成本或获得最大经济利益。经济人的自利行为是经济学研究中的基础假设之一，它用来解释个人之间的市场交换和其他经济活动。对于之前的错误回答，我深表歉意，谢谢你的指正。

所以对于结果最好是自己熟悉的领域，能够适当有基本对生成结果判断的能力。

AIGC中Prompt的语言选择问题

文本AIGC来说，比如ChatGPT，中文给它提问（Prompt）会获得不错的结果，因为它针对中文的训练语料采集的不错，所以相关结果还算不错，英文提问（Prompt）结果会更好。

对于图片AIGC来说，无论是国外的AIGC产品还是百度文心一言，如果使用中文的情况，可能会与你想要的结果会相差非常的大。

比如我想要图片AIGC帮我生成一个小女孩和小猫在一起的图片，我们看中英文不同 Prompt 生成的结果区别：



Help me draw a picture, mainly a little girl and a kitten together, very cute, in the style of storybook illustration - @BlackX (fast)



所以，如果是在图片AIGC的场景中，尽量采用英文进行描述，可以把中文翻译成为英文的Prompt 然后进行提交生成。当然，如果真的要用好AIGC工具，生成优秀的Prompt是非常重要的事情，需要持续学习，所以需要实践中增长编写Prompt的能力。

文本AIGC产品 ChatGPT/GPT-4 的使用实践

使用 ChatGPT 文案优化处理

使用ChatGPT 优化内容

Prompt: 作为一名中文写作改进助理，你的任务是改进所提供文本的拼写、语法、清晰、简洁和整体可读性，同时分解长句，减少重复，并提供改进建议。请只提供文本的更正版本，避免包括解释。请从编辑以下文本开始: [文章正文]

HE

作为一名中文写作改进助理，你的任务是“改进所提供文本的拼写、语法、清晰、简洁和整体可读性，同时分解长句，减少重复，另外文本需要整个逻辑有序清晰”，请按照这些要求把以下内容进行文字优化后输出：

以ChatGPT为代表的AI语言模型，以Midjourney为代表的AI绘画产品，让世界见识了AI的效率及能力。它们的应用速度远比想象中快，近一个月，ChatGPT开放API（应用程序接口），升级出的GPT-4，微软已经应用到搜索引擎Bing、Office全家桶里。最新的一个动态是，3月24日，ChatGPT部分解除了没法联网的限制，用户已经能检索实时信息。



ChatGPT和Midjourney是AI语言模型和AI绘画产品的代表，它们展现了人工智能的效率和能力。它们的应用速度比想象中快得多。最近，ChatGPT升级为GPT-4并开放了API接口，微软已将其应用到搜索引擎Bing和Office全家桶中。最新消息是，3月24日，ChatGPT解除了部分不能联网的限制，用户现在可以检索实时信息。

HE

作为一名中文写作改进助理，你的任务是“改进所提供文本的拼写、语法、清晰、简洁和整体可读性，同时分解长句，减少重复，另外文本需要整个逻辑有序清晰”，请按照这些要求把以下内容进行文字优化后输出：

五岁过生日的时候，妈妈送了我一个礼物——储钱罐。

它蓝得发亮，模样是一个小熊，非常惹人喜爱。为了让小熊的肚子充实起来，我想尽了办法。

一天，我的脑子里闪出一个主意：妈妈常说种瓜得瓜，种豆得豆，把豆子种下就能得到成片的大豆，那硬币为什么不能？我何不种硬币呢？说干就干，我到公园挖了许多土，回家找到了一个盆子，先放进去一些土，然后把我心爱的硬币放进去，又放了一些土，对了，还要浇点水。

等我三下五去二的迅速做完种硬币的大事后，就开始想入非非：嘿嘿，过一段时间等硬币长成参天大树后，我就爬上去，摘啊摘啊我就变成百万富翁啦，哈哈我不由得发出了欢笑声。

一天过去了，盆子里的硬币没有动静，两天过去了，还是没有动静，三天过去了，一个月过去了盆子里仍然没有动静。

我心里不由得急了，就扒开土看看---硬币完好无损，一点变化也没有。于是我把这件事告诉了妈妈，妈妈听了哈哈大笑，笑的上气不接下气。

她告诉我：钱是通过辛勤劳动换来的，不是种出来的。我这才恍然大悟！

童年的趣事是美丽的，让我们为童年画上一道灿烂的彩虹吧！



在我五岁过生日时，妈妈送给我一个小熊造型的储钱罐。它蓝色闪闪发亮，让人非常喜欢。我想让小熊的肚子装满硬币，于是想到了一个主意：把硬币种下去，等它们长成大树，我可以爬上去摘下成为百万富翁了。我挖了一些土，找了个盆子，先放了些土，再放进硬币，接着再放土，浇了点水。但盆子里的硬币却一点动静也没有。我开始急了，扒开土一看，硬币完好无损。我告诉妈妈后，她哈哈大笑，告诉我钱是通过辛勤劳动换来的，不可能种出来。这个趣事让我终身难忘，让我们一起为童年画上一道美丽的彩虹吧！

使用 ChatGPT 总结中心思想：

HE

请抽取以下内容的中心思想：

五岁生日的时候，妈妈送我一个礼物——储钱罐。

它蓝得发亮，模样是一个小熊，非常惹人喜爱。为了让小熊的肚子充实起来，我想尽了办法。

一天，我的脑子里闪出一个主意：妈妈常说种瓜得瓜，种豆得豆，把豆子种下就能得到成片的大豆，那硬币为什么不能？我何不种硬币呢？说干就干，我到公园挖了许多土，回家找到了一个盆子，先放进去一些土，然后把我心爱的硬币放进去，又放了一些土，对了，还要浇点水。

等我三下五去二的迅速做完种硬币的大事后，就开始想入非非：嘿嘿，过一段时间等硬币长成参天大树后，我就爬上去，摘啊摘啊我就变成百万富翁啦，哈哈我不由得发出了欢笑声。

一天过去了，盆子里的硬币没有动静，两天过去了，还是没有动静，三天过去了，一个月过去了盆子里仍然没有动静。

我心里不由得急了，就扒开土看看---硬币完好无损，一点变化也没有。于是我把这件事告诉了妈妈，妈妈听了哈哈大笑，笑的上气不接下气。

她告诉我：钱是通过辛勤劳动换来的，不是种出来的。我这才恍然大悟！

童年的趣事是美丽的，让我们为童年画上一道灿烂的彩虹吧！



中心思想：童年的经历和趣事能够给人留下美好的回忆，并教会我们一些重要的人生道理，如劳动和赚钱的概念。



使用ChatGPT 提炼内容核心要点：

HE

总结一下内容的关键内容要素，整理成为1,2,3点：

五岁生日的时候，妈妈送我一个礼物——储钱罐。

它蓝得发亮，模样是一个小熊，非常惹人喜爱。为了让小熊的肚子充实起来，我想尽了办法。

一天，我的脑子里闪出一个主意：妈妈常说种瓜得瓜，种豆得豆，把豆子种下就能得到成片的大豆，那硬币为什么不能？我何不种硬币呢？说干就干，我到公园挖了许多土，回家找到了一个盆子，先放进去一些土，然后把我心爱的硬币放进去，又放了一些土，对了，还要浇点水。

等我三下五去二的迅速做完种硬币的大事后，就开始想入非非：嘿嘿，过一段时间等硬币长成参天大树后，我就爬上去，摘啊摘啊我就变成百万富翁啦，哈哈我不由得发出了欢笑声。

一天过去了，盆子里的硬币没有动静，两天过去了，还是没有动静，三天过去了，一个月过去了盆子里仍然没有动静。

我心里不由得急了，就扒开土看看---硬币完好无损，一点变化也没有。于是我把这件事告诉了妈妈，妈妈听了哈哈大笑，笑的上气不接下气。

她告诉我：钱是通过辛勤劳动换来的，不是种出来的。我这才恍然大悟！

童年的趣事是美丽的，让我们为童年画上一道灿烂的彩虹吧！



1. 妈妈送的储钱罐作为礼物。
2. 将硬币放入土中，希望通过种植硬币得到更多的钱。
3. 最终明白了钱是通过辛勤劳动得来的。



使用ChatGPT 把文本内容进行翻译和转换：

HE

把以下内容翻译成中文：

Hello, everyone. The weather is really good today. I don't know how everyone is feeling. I hope everyone is in a good mood. I hope everyone has a happy day, Momoda~



大家好，今天天气真的很好。我不知道大家现在的感受如何，但我希望大家都能保持良好的心情，并度过一个愉快的一天，Momoda~



HE

把以上译文内容变成在非常正式场合使用的专业内容



尊敬的各位，今日天气宜人。在不知道大家心境如何的情况下，衷心希望各位身心愉悦、开心快乐，祝福大家度过美好的一天，此致，敬礼！



HE

请把译文转成“小红书体”



大家好，今天的天气真是美妙啊。不知道你们现在的心情如何呢？希望你们心情愉悦，开心每一天，么么哒~



使用ChatGPT 进行创意生成

基本的讲故事：

HE

请给我讲一个关于蚊子的笑话，一定要简短



为了保证笑话简短，我可以给您提供一个小巧的蚊子笑话：



有只蚊子问一只小青蛙：“你为什么跳？”小青蛙回答：“因为我是青蛙啊，跳是我的本领。”
蚊子说：“那我为什么要飞？”

HE

请把以上故事变成恐怖故事，必须简短



有一个小女孩晚上睡觉时听到屋外嗡嗡的声音，像是蚊子的声音。她一直忍耐着，但是声音越来越大，越来越近。最后，她听到了一只巨大的蚊子在她床头嗡嗡作响，吓得她大声尖叫。但是，当她打开灯时，她惊恐地发现，在她的床头不是一只蚊子，而是一只拥有巨大翅膀的可怕生物——巨型蚊子！



起名字:

HE

我生了一个男孩，我姓谢，请帮忙给我儿子起10个高雅适合男孩的名字



好的，以下是10个高雅适合男孩的名字：

1. 梓豪 (zǐ hào)
2. 轩宇 (xuān yǔ)
3. 昊天 (hào tiān)
4. 涵亮 (hán liàng)
5. 希望 (xī wàng)
6. 风华 (fēng huá)
7. 柏丞 (bǎi chéng)
8. 睿达 (ruì dá)
9. 恺撒 (kǎi sā)
10. 智勇 (zhì yǒng)

写脱口秀:

HE

请帮我写一段简短好笑的脱口秀，必须带有北京这个地方



大家好，我来自北京，这是一个美丽而繁华的城市，也是一个让人又爱又恨的地方。比如说，每次在北京开车，我都感觉自己在玩一款叫“避坑”（避开坑洼路）的游戏。不过话说回来，北京的交通也真是越来越先进了，现在坐公交车上厕所都是免费的。



说到吃，北京的小吃真是多种多样，比如羊蝎子、涮羊肉、炸酱面，每次看到这些美食就忍不住想问问自己：这次到底是胖还是死呢？

最后，还有一点，北京人是真的不喜欢排队，这让我很是苦恼。所以每次我去排队，都会戴一个写着“请让一让”的帽子，这样别人就会自觉地给我让位了。大家笑了吗？

营销方案建议:

HE

最近公司有一个新产品上市，是教育类的图书，请帮忙想几个营销推广方案



当面对教育类图书的推广时，以下是几个可能有效的推广方案：



1. 与学校或教育机构合作 - 通过与学校或教育机构合作，推广你的图书可能会更容易。这可以包括在学校图书馆内进行推广，邀请你的作者前往学校讲座，以及与教师和学校管理员合作，以推广你的图书作为课堂教材。
2. 利用社交媒体 - 利用各种社交媒体平台进行推广，例如Facebook、Twitter和Instagram等，可以让你的图书广泛传播。你可以通过有趣的帖子、有关书中主题的漫画和其他创意手段来吸引读者的注意力，并鼓励他们与你互动和分享。
3. 举办读者活动 - 举办各种读者活动可以增加人们对你的图书的兴趣和认知。例如，在书店或图书馆内举办书签制作活动、讲故事会、书籍俱乐部和签售会等活动，可以让你的读者更好地了解你的图书并与你建立联系。
4. 写博客和制作视频 - 通过在你的网站或博客上写有关书中主题的博客文章，或制作有关书中主题的视频，可以为你的读者提供更深入的了解和学习体验。这也可以帮助你扩大你的受众，增加你的读者数量。
5. 联合其他品牌 - 与其他教育或文化品牌合作，可以为你的图书带来更多的曝光率。例如，在与电视台合作的情况下，在电视上播出你的广告，或与儿童文化中心合作，举办读者活动，可以让更多人认识和购买你的图书。

写诗：

HE

请写一首宋词，必须包含秋天，思念，果实，湖水 等元素



秋思



凉风习习度秋天，思念如潮涌心田。
果实累累皆硕果，香气飘散遍山川。
湖水碧波荡漾，秋色渐浓添愁颜。
但愿佳期能再会，共赏秋景话离欢。

注：此词只是本AI根据宋词的风格和特点进行模仿，不代表具有宋词的艺术价值。

使用 ChaGPT 为“百晓生”小助手

询问知识类或生活类问题：

HE

请用最简短的语句描述清楚为什么会下雨



雨是因为水汽在空气中冷却形成的水滴，由于空气中的温度、湿度和气压等因素影响，导致水汽凝结成云，最终形成降水。



HE

请用最短的话描述清楚王阳明的哲学思想核心是什么？



“知行合一”



HE

中国的夏朝一共经历了多少年？



夏朝是中国历史上的第一个朝代，据传历时约四百年，具体的时间跨度因为历史记载和考古发掘的限制而无法确定，一般认为大约从公元前2070年左右至公元前1600年左右。



HE

请用最短的语句告诉我，化妆的时候如何保证口红不会画歪



用唇笔勾勒唇线，再用唇刷填色，从内往外画，避免口红画歪。



中小学生学习辅导：

HE

两辆汽车同时从甲、乙两地相对开出，一辆汽车每小时行56千米，另一辆汽车每小时行63千米，经过4小时后相遇。甲乙两地相距多少千米？



两辆汽车相对行驶的速度是 $56+63=119$ 千米/小时。根据题意，在相遇的 4 小时内，两车之间的距离就是两车行驶的总路程，即 $119 \times 4=476$ 千米。因为两地是相对的，所以两地之间的距离就是 $476/2=238$ 千米。因此，甲乙两地相距 238 千米。



HE

请简短告诉我，英语单词里的with属于什么词性，主要应用在哪些语句中



“with”是一个介词，常用于表示附带状态或附加信息的情况下，连接一个名词或代词和另一个名词、代词或动词。例如：“I am going to the store with my friend.”



HE

请用简短语句告诉我，出师表中“今天下三分，益州疲弊，此诚危急存亡之秋也”是什么意思？



“今天下三分”指当前天下分为三国；“益州疲弊”指益州地区人民已经疲惫不堪；“此诚危急存亡之秋也”意思是目前形势十分危急，存亡关键时刻。



学习题目生成：

HE

阅读用以下材料，帮我生成2道选择题：

春天来了！春天来了！新的一年里，最美好的一个季节就是春天，春天正在悄悄地向我们走来。

春天里，春姑娘唤醒了桃花，它急忙穿上了新装。她又摸了摸柳树，柳树立刻醒了过来，发出了嫩芽，在微风的吹拂下翩翩起舞。她又走过了迎春花的身旁，迎春花被惊醒了，它揉了揉眼睛，探出了一个个小脑袋。它们都在欢迎春姑娘的到来。

春姑娘带着春风走到了小草的旁边，小草就被春风吹得摇摇晃晃，这个刚出生的小家伙好像站不稳似的。仙客来了有了春风的光顾，他不停的点头哈腰。

春雨呢？太懒了。即使人间有许多美好的东西，也不愿意到人间来玩一玩。庄稼太需要他了，真是“春雨贵如油”呀！

春天小燕子又飞回来了，到处都是春天的气息。

春天真美，我们爱春天！



1. 在春天里，春姑娘使哪些植物苏醒？



- A. 桃花、杨树、梅花
- B. 桃花、柳树、迎春花
- C. 桃花、梅花、迎春花
- D. 桃花、杨树、柳树

答案：B

2. 在文中，哪个植物因为春风的吹拂而摇摇晃晃？

- A. 迎春花
- B. 桃花
- C. 梅花
- D. 小草

答案：D

医疗建议：

HE

请用简短文字告诉我，如果自己所在地遭遇了核危机，应该怎么办，应该准备什么药物



如果自己所在地遭遇了核危机，应尽快寻找避难所，远离危险区域，保持安静，密封住室内门窗，关掉空调系统，同时收听官方通知，遵从有关部门指示。如果有条件的话，可以准备防护口罩、防护服、抗辐射药物等应急物资。关于具体的抗辐射药物，最好咨询专业医生或者当地官方有关部门。



HE

如果家中有人忽然晕倒，应该怎么办，有哪些急救建议，请简短一些



如果家中有人突然晕倒，应该立即进行以下急救措施：1) 将晕倒者平放在地上，解开衣服，保持呼吸道通畅；2) 观察患者呼吸、脉搏和意识等情况；3) 如有必要，进行心肺复苏等急救措施，及时送往医院救治。此外，在平时要注意家庭安全，了解基本的急救知识和技能，准备急救药品和器材，以备不时之需。



HE

如果家中儿童发烧达到40度，应该如何处理，请简短一些



当儿童发烧高达40度时，应该立即采取以下急救措施：



1. 用温水擦拭儿童的额头、手臂、小腿等部位，以降低体温；
2. 让儿童服用退烧药，如布洛芬等，但要遵循医生建议的用药剂量；
3. 让儿童多喝水，避免脱水；
4. 如果情况严重，必须马上就医。

各种工具“小助手”：

HE

请帮我生成一个Excel公式，需求是把A1到A8的值相加，然后再把总值求平均



可以使用以下公式来实现：



```
=AVERAGE(A1:A8)
```

这个公式会先将 A1 到 A8 的数值相加，然后再除以 8，求出平均值。

HE

请写一段SQL，能够在User表中统计最近一个月访问APP最多的用户前10名。（请简短一些）



```
SELECT user_id, COUNT(*) AS visit_count
FROM User
WHERE access_time >= DATEADD(month, -1, GETDATE())
GROUP BY user_id
ORDER BY visit_count DESC
LIMIT 10
```



HE

请写一段Python代码，能够把一个叫做record.xls的Excel文件导入到Pandas中（请简短一些）



python

Copy code



```
import pandas as pd

df = pd.read_excel('record.xls')
```

周报、论文、翻译 等小助手：

HE

我是一个HR，这周我主要是进行了3个人的面试，还有安排了2个述职，组织了1次团队公益活动，请帮我写一个周报，内容不长不短。



周报



时间：2023年4月10日至4月16日

本周工作内容：

1. 进行了3个人的面试，涉及职位：销售经理、财务主管、行政助理。
2. 安排了2个员工进行了述职，其中一个市场部销售主管，另一个是财务部的会计。
3. 组织了1次团队公益活动，我们一起参加了当地的义工活动，为社区的公共设施进行清洁和维护。
4. 处理了一些员工的离职手续，包括退出手续、工资结算、离职证明等。
5. 和部门经理一起讨论了员工培训计划，确定了下一个季度的培训主题和计划。
6. 协助人力资源部门进行员工档案管理和人事统计工作。

下周工作计划：

1. 安排并组织团建活动，增强团队凝聚力。
2. 继续进行招聘工作，寻找适合的人才填补职位空缺。
3. 参与员工考勤管理，确保员工考勤情况准确无误。
4. 进行员工绩效考评，准备年度绩效评估报告。
5. 与上级领导进行汇报，分享本部门的工作成果和困难，征求意见和建议。
6. 协助部门经理进行一些临时性工作，保证部门日常工作运转顺畅。

备注：本周工作完成情况良好，但是在员工培训和招聘方面还需要进一步加强。

进行大量手工比较辛苦的文本替换工作：

HE

处理以下内容，主要是进行每行文本重新排列位置，主要规则是把 XXX: 里面的文字XXX 放到每行文字的开头变成【XXX】，然后其他文字放到每行右边，重新生成输出：

异想天开插图： Whimsical Illustration
图识字插图： Storybook Illustration
中国画风格： Chinese Painting
简单风格： Simplicity
素描风格： Sketch
阴影风格： Shade
平面明暗风格： Flat Shading
哥特式风格： 2010s Gothic Art Nouveau

高保真风格： High Fidelity
13世纪风格： 1200s
60年代风格： 60s
70年代风格： 70s
个人动作风格： Pose



【异想天开插图】： Whimsical Illustration
【图识字插图】： Storybook Illustration
【中国画风格】： Chinese Painting
【简单风格】： Simplicity
【素描风格】： Sketch
【阴影风格】： Shade
【平面明暗风格】： Flat Shading
【哥特式风格】： 2010s Gothic Art Nouveau

【高保真风格】： High Fidelity
【13世纪风格】： 1200s
【60年代风格】： 60s
【70年代风格】： 70s
【个人动作风格】： Pose



使用 ChatGPT 训练口语

Chrome浏览器中安装“Voice Control for ChatGPT”扩展：

说明：在Edge浏览器中可以直接安装扩展：<https://chrome.google.com/webstore/detail/voice-control-for-chatgpt/eollffkcakegifhacjlnegohfdlidhn>；在QQ浏览器等浏览器等需要单独下载这个<https://www.crxsoso.com/webstore/detail/eollffkcakegifhacjlnegohfdlidhn> CRX文件，然后在浏览器扩展中打开“开发者模式”后拖入crx文件到浏览器安装插件即可)



Voice Control for ChatGPT

Expands ChatGPT with voice control and read aloud.

★★★★★ 142 效率

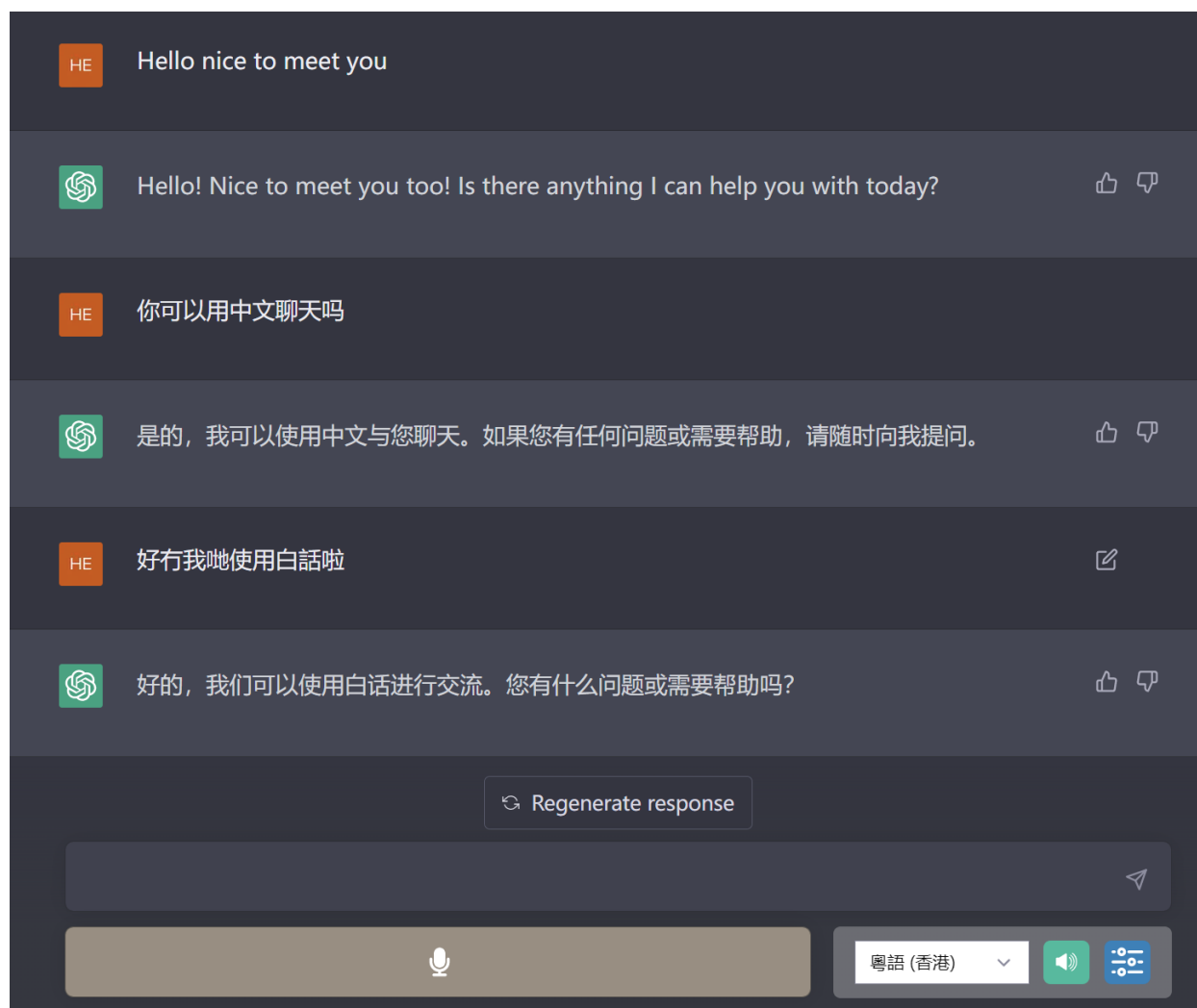


Voice Control for ChatGPT

添加至 Chrome

★★★★★ 142 ⓘ | 效率 | 100,000+ 位用户

插件安装完成以后，然后在 chat.openai.com 页面就增加了一个语音输入按钮，就可以进行语音互动了，你说一句，对方回一句，对方因为是大语言模型，基本可以完全模拟人类的交互活动，而不会觉得生硬：（中文英文日语韩语俄语法语以及粤语等都可以进行互动）



除了以上ChatGPT的使用场景，在文字处理方面，还可以使用ChatGPT成为音乐、金融、行业顾问、心理、演讲 等等各个方向进行文案、创意、优化、问答 等好帮助，需要多多探索。

使用 ChatGPT 辅助制作PPT

可以通过ChatPPT生成相应的PPT大纲，然后通过PPT生成网站把大纲的Markdown格式变成PPT。

让ChatGPT生成大纲：

大概Prompt：

我要做一份《ChatGPT原理介绍》的PPT

请帮我写一份包含了5个子标题的大纲，每个子标题下面请尽量写3-5条内容，内容尽量翔实，如果有例子更好最后，一定要使用Markdown代码框输出内容，谢谢

HE

我要做一份《ChatGPT原理介绍》的PPT，
请帮我写一份包含了5个子标题的大纲，
每个子标题下面请尽量写3-5条内容，内容尽量翔实，如果有例子更好
最后，一定要使用Markdown代码框输出内容，谢谢



当然可以，以下是你需要的内容：



markdown

Copy code

ChatGPT原理介绍

1. 什么是ChatGPT

- ChatGPT是一种基于深度学习技术的自然语言处理模型。
- 它使用Transformer架构，通过学习大量语言数据来预测下一个可能的单词或词组。
- ChatGPT的目标是生成具有合理上下文的流畅自然语言，以模拟人类的对话方式。

2. ChatGPT的优点

- 能够处理各种自然语言任务，如文本生成、语言翻译和问答系统等。
- 高度自适应性和灵活性，能够根据不同任务的需求进行微调和优化。
- 随着模型规模的增大，ChatGPT的性能在自然语言处理领域中达到了最先进的水平。

3. ChatGPT的应用场景

- 自动客服和聊天机器人：使用ChatGPT进行对话生成，为客户提供高效的服务和支持。
- 语言翻译：使用ChatGPT进行文本生成，将一个语言翻译成另一个语言。
- 文本摘要：使用ChatGPT生成一个简洁的摘要，以代表长篇文章的主要信息。
- 情感分析：使用ChatGPT来对情感进行分类，识别文本中的情感色彩。

4. ChatGPT的发展历程

- 2018年：发布了GPT-1模型，包含1.5亿个参数。
- 2019年：发布了GPT-2模型，包含1.5亿-15亿个参数。
- 2020年：发布了GPT-3模型，包含1.75万亿个参数，创下当时的NLP最大模型记录。

把上面的Markdown格式文档，导入到Markdown生成PPT的网站：（生成PPT网站：
<https://www.mindshow.fun>，也可以使用其他的）

快速创建 导入 我的演示 回收站 账户设置 邀请推荐

导入

```
# ChatGPT原理介绍

## 1. 什么是ChatGPT

- ChatGPT是一种基于深度学习技术的自然语言处理模型。
- 它使用Transformer架构，通过学习大量语言数据来预测下一个可能的单词或词组。
- ChatGPT的目标是生成具有合理上下文的流畅自然语言，以模拟人类的对话方式。

## 2. ChatGPT的优点

- 能够处理各种自然语言任务，如文本生成、语言翻译和问答系统等。
- 高度自适应性和灵活性，能够根据不同任务的需求进行微调和优化。
- 随着模型规模的增大，ChatGPT的性能在自然语言处理领域中达到了最先进的水平。

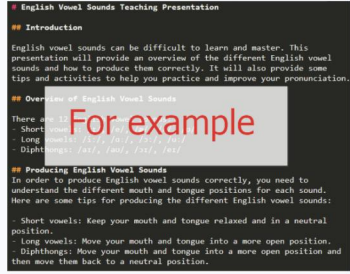
## 3. ChatGPT的应用场景

- 自动客服和聊天机器人：使用ChatGPT进行对话生成，为客户提供高效的服
```

存储空间 0.00MB / 100MB

数据类型: Markdown logseq 幕布(.mm)

导入创建



创建后大概这个效果，可以自己编辑：

https://www.mindshow.fun/#/edit/62858 200 在此搜索

返回 已保存

ChatGPT原理介绍

+ Logo

副标题: SUBTITLE HERE ×

演讲者: MindShow.fun ×

演讲时间: 2023-03-27 ×

- 1. 什么是ChatGPT
 - ChatGPT是一种基于深度学习技术的自然语言处理模型。
 - 它使用Transformer架构，通过学习大量语言数据来预测下一个可能的单词或词组。
 - ChatGPT的目标是生成具有合理上下文的流畅自然语言，以模拟人类的对话方式。
- 2. ChatGPT的优点
 - 能够处理各种自然语言任务，如文本生成、语言翻译和问答系统等。
 - 高度自适应性和灵活性，能够根据不同任务的需求进行微调和优化。
 - 随着模型规模的增大，ChatGPT的性能在自然语言处理领域中达到了最先进的水平。
- 3. ChatGPT的应用场景
 - 自动客服和聊天机器人：使用ChatGPT进行对话生成，为客户提供高效的服

演示预览 下载 演示

模版 布局

然后可以直接演示这个PPT：

ChatGPT原理介绍

拥抱AI新时代

Black
2023-03-27

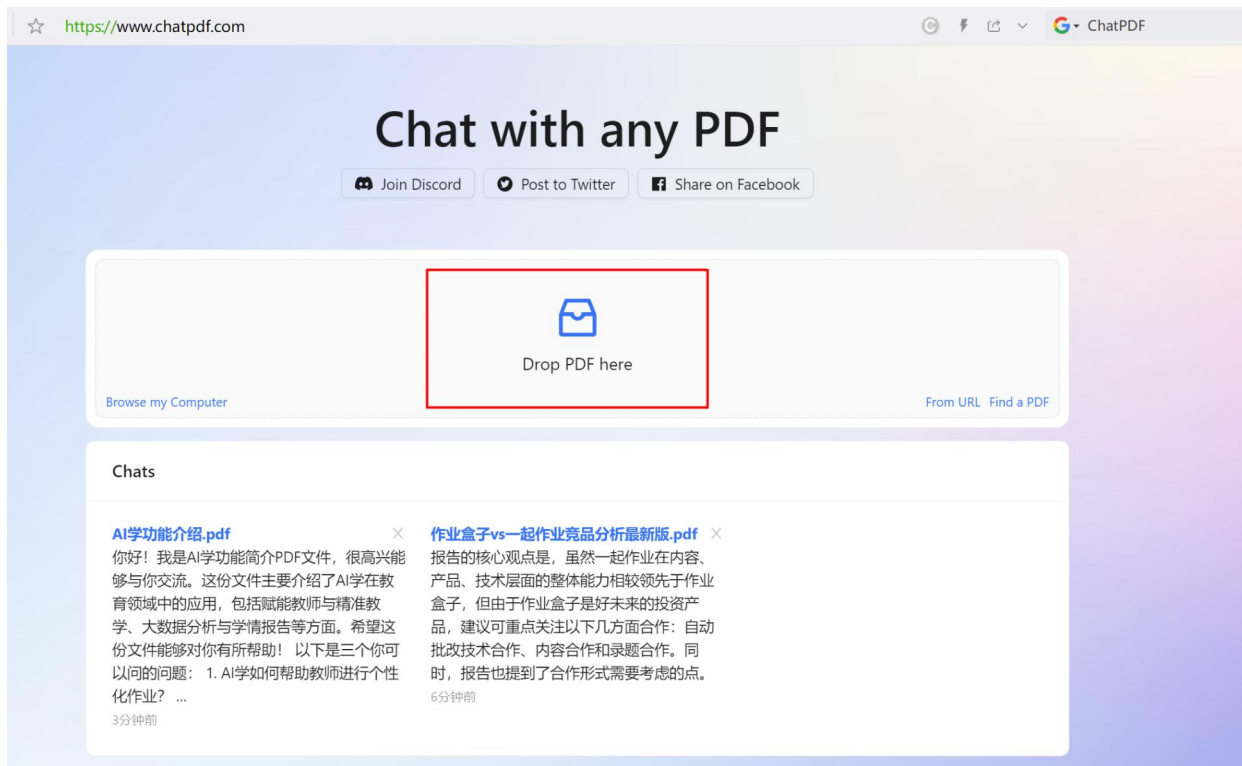


当然，通过Markdown生成PPT还有很多其他类似工具，以上工具只是一个参考，核心是ChatGPT帮助我们生成了我们想要的大概内容以及Markdown格式，可以非常方便我们进行PPT的制作。

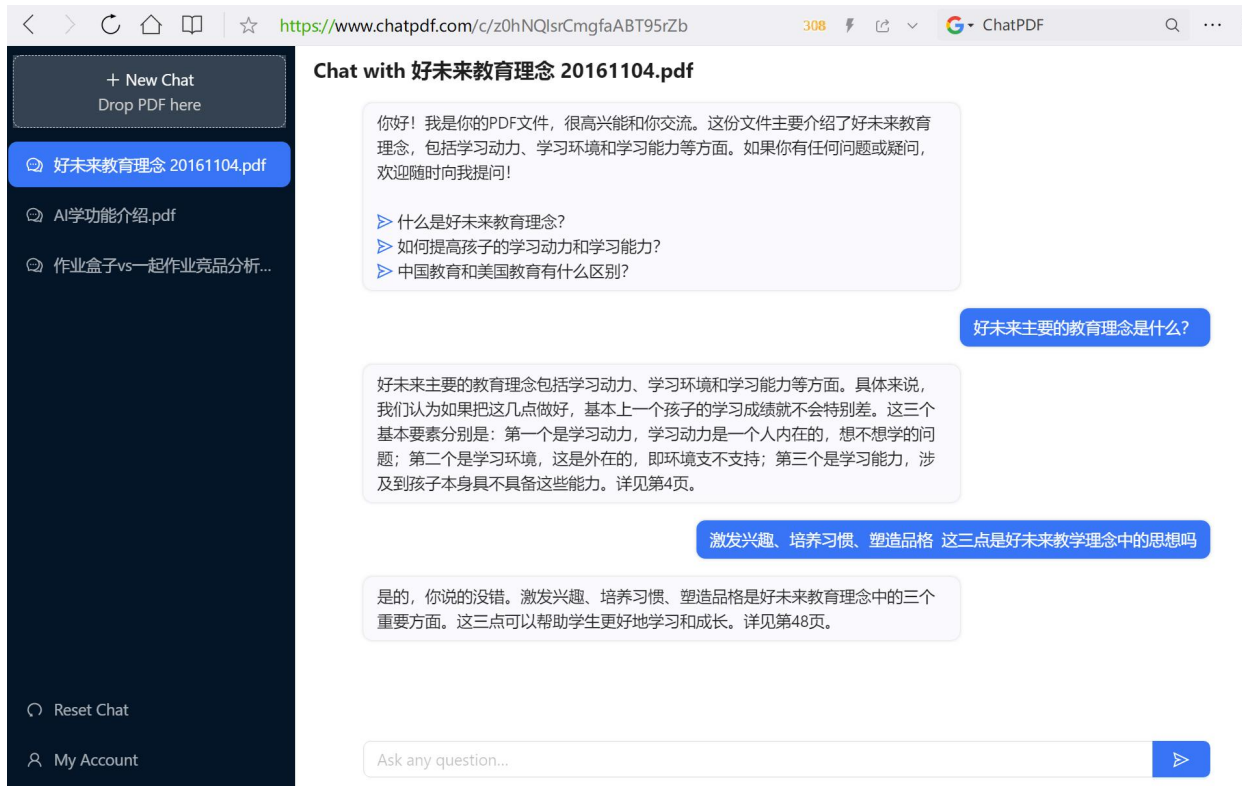
利用ChatGPT底层阅读文档文献

使用ChatGPT的API生成的用一个工具叫做ChatPDF（<https://www.chatpdf.com>），它底层使用ChatGPT的能力，可以帮助我们快速阅读大的文档。

上传自己想要阅读的PDF文档（内容多更好）：



阅读抽取整个PDF核心内容，可以跟它提问交互：

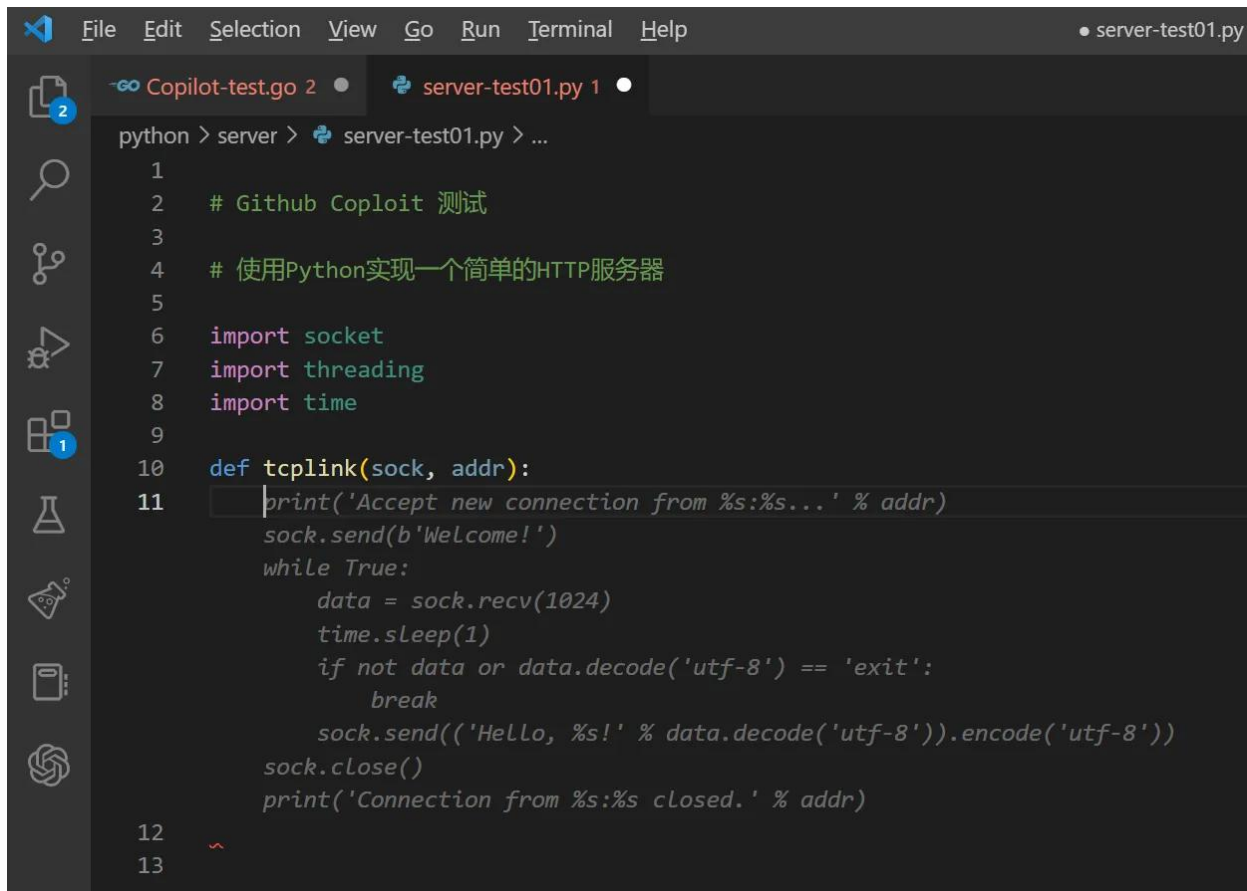


你可以针对这个PDF文件内容进行各种提问和交互，比如可以让它帮助你读论文，或者是度比如财报等，获取最关键的信息，不用自己费劲半天阅读，能够大幅减轻你的工作压力。

文本AIGC产品 Github Copilot / Cursor 的使用实践

Copilot 代码辅助编程：

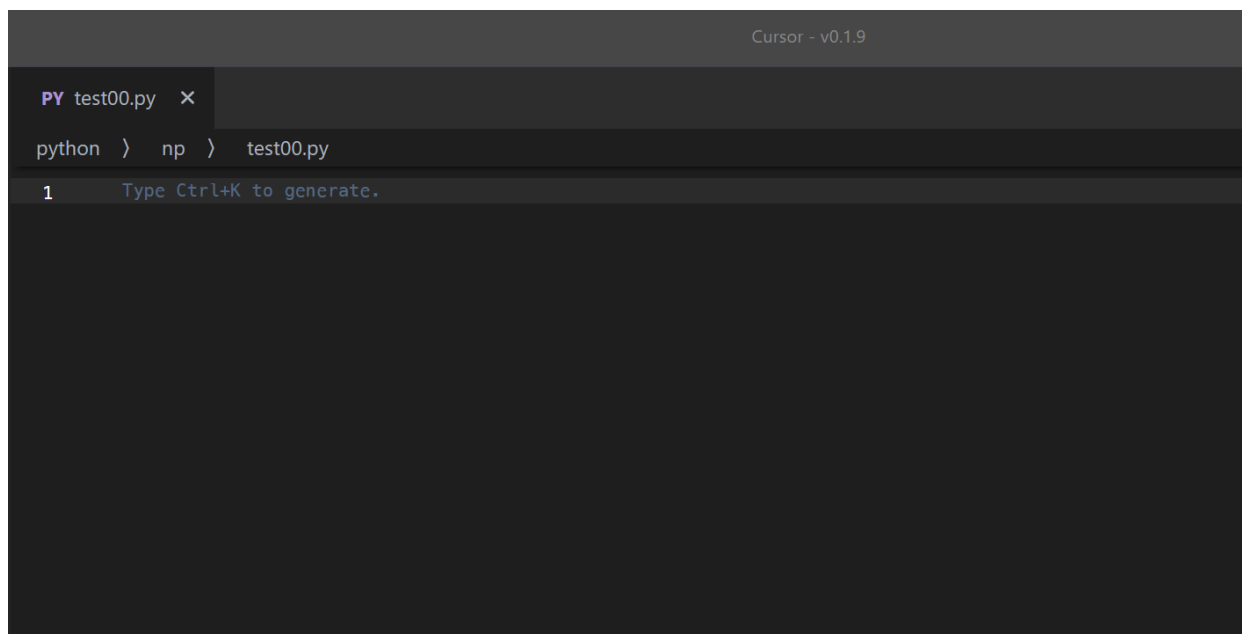
```
3 # 实现一个快速排序的算法
4
5 def quick_sort(lists, left, right):
6     # 快速排序
7     if left >= right:
8         return lists
9     key = lists[left]
10    low = left
11    high = right
12    while left < right:
13        while left < right and lists[right] >= key:
14            right -= 1
15        lists[left] = lists[right]
16        while left < right and lists[left] <= key:
17            left += 1
18        lists[right] = lists[left]
19    lists[right] = key
20    quick_sort(lists, low, left - 1)
21    quick_sort(lists, left + 1, high)
22    return lists
```



```
File Edit Selection View Go Run Terminal Help • server-test01.py
Copilot-test.go 2 • server-test01.py 1
python > server > server-test01.py > ...
1
2 # Github Coploit 测试
3
4 # 使用Python实现一个简单的HTTP服务器
5
6 import socket
7 import threading
8 import time
9
10 def tcplink(sock, addr):
11     print('Accept new connection from %s:%s...' % addr)
12     sock.send(b'Welcome!')
13     while True:
14         data = sock.recv(1024)
15         time.sleep(1)
16         if not data or data.decode('utf-8') == 'exit':
17             break
18         sock.send(('Hello, %s!' % data.decode('utf-8')).encode('utf-8'))
19     sock.close()
20     print('Connection from %s:%s closed.' % addr)
21
22
```


使用 Github Copilot 主要是解决编程开发中 注释、函数参数、部分代码 等的自动补全工作，减轻开发工作量，提升效率。

调用 Cursor 进行AI辅助编程：



同类型的工具还有 Tabnine、CodeGeex、Cursor 等等，都是作为编程中AI辅助工具，都可以达到不同程度提升效率的工具。

图像AIGC产品 Midjourney 的使用实践

把现有图片进行绘画升级

案例：把一本图书变成高保真图

上传一个基础图片，然后按照个人想要风格重新画图：



BlackX 今天01:47



Midjourney Bot 机器人 今天01:49

<https://s.mj.run/wQLN50ucnHw> Storybook Illustration Whimsical Illustration - @BlackX (fast)





案例：把两张图进行合并形成新图

描述：

将这两张图片组合起来，显示正在阅读这本书的人。故事书插图，高保真，采用Midjourney v5版本渲染

Prompt:

<https://s.mj.run/G8qTcZzZWok> https://s.mj.run/a9w92MbXC_I Combine these two images to show the person reading the book. Storybook Illustrations, High Fidelity, octane render v5



Midjourney Bot ✓ 机器人 今天10:48

<https://s.mj.run/G8qTcZzZWok> https://s.mj.run/a9w92MbXC_I Combine these two images to show the person reading the book. Storybook Illustrations, High Fidelity, octane render v5 - @BlackX (fast)





直接按照个人创意进行绘图

案例：绘制一个兔子和小男孩

描述：帮我绘制一幅画，画中主要是一个小男孩和一只小白兔在一起，很可爱，采用 Storybook Illustration 和 Whimsical Illustration 风格

Prompt: Help me draw a picture in which a little boy and a little white rabbit are together, which is very cute and adopts the Storybook Illustration and Whimsical Illustration style

01:52

Help me draw a picture in which a little boy and a little white rabbit are together, which is very cute and adopts the Storybook Illustration and Whimsical Illustration style - @BlackX (fast)





案例：绘制一个小男孩冒险的插画

描述：有一个叫杰克的小男孩，他喜欢冒险故事。他读了许多充满宝藏和海盗的书，常常希望自己也能去冒险。有一天，他在海岸玩耍时发现了一张地图，上面标示着通往一个神秘岛屿的路线和宝藏的位置。杰克知道他必须找到一种方法去那个岛屿上，并解开它的秘密。

Prompt: *Once upon a time, there was a young boy named Jack who loved adventure stories. He had read so many books filled with treasures and pirates that he often wished he could go on his own adventures. One day, while playing by the seashore, he found a map that led to a mysterious island full of treasure. Jack knew that he had to find a way to get to the island and unlock its secrets.*

01:54

Once upon a time, there was a young boy named Jack who loved adventure stories. He had read so many books filled with treasures and pirates that he often wished he could go on his own adventures. One day, while playing by the seashore, he found a map that led to a mysterious island full of treasure. Jack knew that he had to find a way to get to the island and unlock its secrets. -

@BlackX (fast)





案例：绘制一个游戏人物原画

描述：这是使用在游戏中的原画，要具有3D效果，场景主要是：有一个叫做小明的20岁男子，他手拿着一把大刀，这把大刀非常锋利，闪着寒光；他穿着像斗篷一样的衣服，衣服上面沾满了油污；他的鞋子是一双黑色的靴子，鞋带系的很紧；他有着帅气的面容，站立着看着前方，眼神中透露了杀气；他背后是一扇城堡大门，大门关闭着；

Prompt: This is the original painting used in the game, with a 3D effect. The main scene is: there is a 20-year-old man named Xiaoming, who is holding a large knife in his hand, which is very sharp and flashing cold light; He was dressed like a cloak, and his clothes were covered with oil stains; His shoes were a pair of black boots with tight laces; He had a handsome face and stood looking ahead, with a murderous look in his eyes; Behind him was a castle gate, which was closed;

02:11

This is the original painting used in the game, with a 3D effect. The main scene is: there is a 20-year-old man named Xiaoming, who is holding a large knife in his hand, which is very sharp and flashing cold light; He was dressed like a cloak, and his clothes were covered with oil stains; His shoes were a pair of black boots with tight laces; He had a handsome face and stood looking ahead, with a murderous look in his eyes; Behind him was a castle gate, which was closed; - [@BlackX](#) (fast)





案例：绘制一个高清像奥黛丽赫本的中国女孩

描述：

一个20岁的中国女人，看起来像奥黛丽·赫本，黑色头发，站在2023东京街上，超逼真的人像摄影，苍白的皮肤，连衣裙，广角，自然光照，柯达portra 800，105毫米f1.8，32k--ar 16:9--v 5--s 750--q 2

Prompt:

A Chinese 20-year-old Woman, looking like Audrey Hepburn, Black hair, standing on 2023 Tokyo street, hyper realistic portrait photography, pale skin, dress, wide shot, natural lighting, kodak portra 800, 105 mm f1.8, 32k --ar 16:9 --v 5 --s 750 --q 2



Midjourney Bot ✓ 机器人 今天11:18

A Chinese 20-year-old Woman, looking like Audrey Hepburn, Black hair, standing on 2023 Tokyo street, hyper realistic portrait photography, pale skin, dress, wide shot, natural lighting, kodak portra 800, 105 mm f1.8, 32k --ar 16:9 --v 5 --s 750 --q 2 - @BlackX (fast)



案例：绘制一个情侣的电影故事板

描述：

一个连续的电影故事板，在连续面板中显示一个男孩和一个女孩从认识到恋爱再到结婚的连续故事场景，电影照明，光线追踪，对细节的关注，辛烷值渲染，Midjourney v5版本

Prompt:

a sequential movie storyboard in sequential panels showing a continuous story scene of a boy and a girl from recognition to love to marriage in a continuous panel, cinematic lighting,

ray tracing, attention to detail, octane render v5

10:29

a sequential movie storyboard in sequential panels showing a continuous story scene of a boy and a girl from recognition to love to marriage in a continuous panel, cinematic lighting, ray tracing, attention to detail, octane render v5 - [@BlackX](#) (fast)





实际在Midjourney中，还有很多复杂的Prompt描述词汇，还有几百上千种风格样式，都需要一点点去探索，能够达到自己的最终要求，可能需要不断地Prompt的微调，然后才能满意的结果。

附录1: 各类工具参考信息

Prompt 参考索引

ChatGPT Prompt 快捷指令:

Midjourney 绘画风格索引:

Midjourney 使用攻略: <https://zhuanlan.zhihu.com/p/611864651>

Midjourney 个人常用绘画Prompt参考 (仅部分)

说明: 可以把以下 Prompt 放在描述的最后, 前面可以加上逗号。

<https://newzone.top/chatgpt>

<https://github.com/willwulfken/MidJourney-Styles-and-Keywords-Reference>

[https://github.com/willwulfken/MidJourney-Styles-and-Keywords-](https://github.com/willwulfken/MidJourney-Styles-and-Keywords-Reference/blob/main/Pages/MJ_V4/Style_Pages/Just_The_Style/Design_Styles.md)

[Reference/blob/main/Pages/MJ_V4/Style_Pages/Just_The_Style/Design_Styles.md](https://github.com/willwulfken/MidJourney-Styles-and-Keywords-Reference/blob/main/Pages/MJ_V4/Style_Pages/Just_The_Style/Design_Styles.md)

绘画风格定义

【异想天开插图】Whimsical Illustration

【图识字插图】Storybook Illustration

【中国画风格】Chinese Painting

【高保真风格】High Fidelity

【简单风格】Simplicity

【素描风格】Sketch

【阴影风格】Shade

【平面明暗风格】Flat Shading

【哥特式风格】2010s Gothic Art Nouveau

【13世纪风格】1200s

【60年代风格】60s

【70年代风格】70s

【个人动作风格】Pose

【用Midjourney V5版渲染】octane render v5

【头像风格】Sticker Design

【把a画成b】A as B

【极简风logo设计】...Symmetrical, Flat icon design

【游戏列表展示效果】Game sheet of...

【整齐排列的一种艺术设计感】Knolling...

【像素风，也可接pixel art】8 bit, 16 bit (例：16 bit pixel art, cozy tavern, interior (16 bit的温馨小酒馆图的prompt))

【被某种物质覆盖的效果】【material填材料，材质名】...made out of

【3d效果的折纸图像】Layer of paper

【等轴艺术，在2d图像基础上创建3d图形的透视图】...Isometric art

【灯光特效，黑色底色暗明亮色的特性】black light...

【淳朴艺术，童真特色风格】Naive art

【吉祥物logo】Mascot logo

【服装设计】T shirt vector...

【特色风格图案设计，在patent前面加上地区以限定风格】Patent: ...

【纹身设计】Tattoo design

【建筑设计词条，这个需要大量prompt以画出足够要求的图】architecture (interior design也可)

【照片级真实图像】..., 8K, Photorealistic, 相机型号 (例：..., 4K, Photorealistic, Sony)

【像教堂的隔音玻璃的设计】...Stained glass window

【建模风图像】Blender 3d

【放射性，爆炸性效果图像】Explode by Nychos

发型定义

【大波浪】Wavy hair

【卷发，卷比较小比较多】Curly hair

【直发】Straight hair

【中分和偏分】To part my hair in the middle/on the side

【中分】Centre parting
【偏分】Side parting
【刘海】Fringe
【波波头，蘑菇头】Bob
【爆炸头】Afro
【精灵短发(赫本头)】Pixie cut
【蜂窝头】Beehive hair
【发髻、丸子头】Bun
【盘发】Plate hair
【在头顶弄一些造型的头发】Updo
【来自于法语 "chignon du cou", 意指那种可以让脖子露出来的盘发。】Chignon
【马尾】Pony-tail
【两个马尾，绑在头的两侧】Pig tail
【辫子，麻花辫】Braid
【满头小辫（非洲人多编这种）】Cornrows
【满头小辫，比cornrows大一点的辫子】Box braids
【蝎子辫】French Braid
【反过来编的French braid】Dutch Braid

附录2: GPT原理参考技术文档

【GPT相关】

人人都能看得懂的ChatGPT原理解析

<https://mp.weixin.qq.com/s/x-ciccaGgdFVoIDPAR3EmQ>

GPT原理及应用

<https://o5eh02lm9y.feishu.cn/docx/G0emdfR6douzQ8xKJl8cqVQUnmg>

Attention Is All You Need

<https://arxiv.org/pdf/1706.03762.pdf>

《Transformer - Attention is all you need》

<https://zhuanlan.zhihu.com/p/311156298>

attention机制原理及简单实现

<https://www.jianshu.com/p/1d67638139da>

可视化的self-attention:

https://colab.research.google.com/github/tensorflow/tensor2tensor/blob/master/tensor2tensor/notebooks/hello_t2t.ipynb#scrollTo=T7UJzF6fmhp

What Is a Transformer Model?

<https://blogs.nvidia.com/blog/2022/03/25/what-is-a-transformer-model>

How do Transformers work?

<https://huggingface.co/course/chapter1/4>

The Concept of Transformers and Training A Transformers Model

<https://towardsdatascience.com/the-concept-of-transformers-and-training-a-transformers-model-45a09ae7fb50>

Word Embedding: Topic Modeling with Word2Vec <https://www.baeldung.com/cs/ml-word2vec-topic-modeling>

生成式无监督预训练简介 (Generative Pre-Training)

<https://zhuanlan.zhihu.com/p/463354320/>

GPT模型中的计算

https://blog.csdn.net/weixin_45577864/article/details/119722158

《根据人类偏好进行深度增强学习》论文理解【IL - IRL】

<https://blog.csdn.net/wxc971231/article/details/121785301>

预训练语言模型之GPT-1, GPT-2和GPT-3

<https://zhuanlan.zhihu.com/p/350017443>

GPT-1, GPT-2和GPT-3发展历程及核心思想, GTP-4展望

<https://blog.csdn.net/yimenren/article/details/122286135>

拥抱Transformer: 自然语言处理三大特征抽取器 (CNN/RNN/TF) 比较

<https://zhuanlan.zhihu.com/p/54743941>

通向AGI之路: 大型语言模型 (LLM) 技术精要

<https://zhuanlan.zhihu.com/p/597586623>

Transformer结构及其应用详解--GPT、BERT、MT-DNN、GPT-2

<https://zhuanlan.zhihu.com/p/69290203>

GPT-4一些论文:

https://pan.baidu.com/s/14WOy26J6SmhDZ_KsBOVztw?pwd=igoh

【传统机器学习和深度学习】

机器学习和深度学习区别的概述

<https://zhuanlan.zhihu.com/p/69776750>

Machine Learning 十大机器学习算法

<https://zhuanlan.zhihu.com/p/33794257>

神经网络从产生到现在的发展历史

<https://blog.csdn.net/jinking01/article/details/103344186>

深度学习概要简介

<https://zhuanlan.zhihu.com/p/150646196>

深度学习的cnn dbn dnm rnn等网络是怎样的关系?

<https://www.zhihu.com/question/44965656>

快速串联 RNN / LSTM / Attention / transformer / BERT / GPT

<https://blog.csdn.net/wxc971231/article/details/127680656>

快速吃透RNN和LSTM神经网络

<https://zhuanlan.zhihu.com/p/86006495>

【图像生成模型】

最近大火的Diffusion Model, 首篇扩散生成模型综述!

<https://www.51cto.com/article/718616.html>

由浅入深了解Diffusion Model

<https://zhuanlan.zhihu.com/p/525106459>

扩散模型 (Diffusion Model) 简要介绍与源码分析

https://blog.csdn.net/Eric_1993/article/details/127455977

理解DALL-E 2, Stable Diffusion和 Midjourney的工作原理

<https://zhuanlan.zhihu.com/p/589223078>