

EIJING DONGCHENG DISTRICT LIBRARY

东城区图书馆少儿 内部资料

网址：www.bjdclib.com

E—mail:dcet2@126.com



• 什么是ChatGPT

• ChatGPT的工作原理

• 对ChatGPT的一些评价

• ChatGPT的未来与趋势

|  |
| --- |
|  |

2023年 第三期

东 城 区 图 书 馆 少 儿 阅 览 室

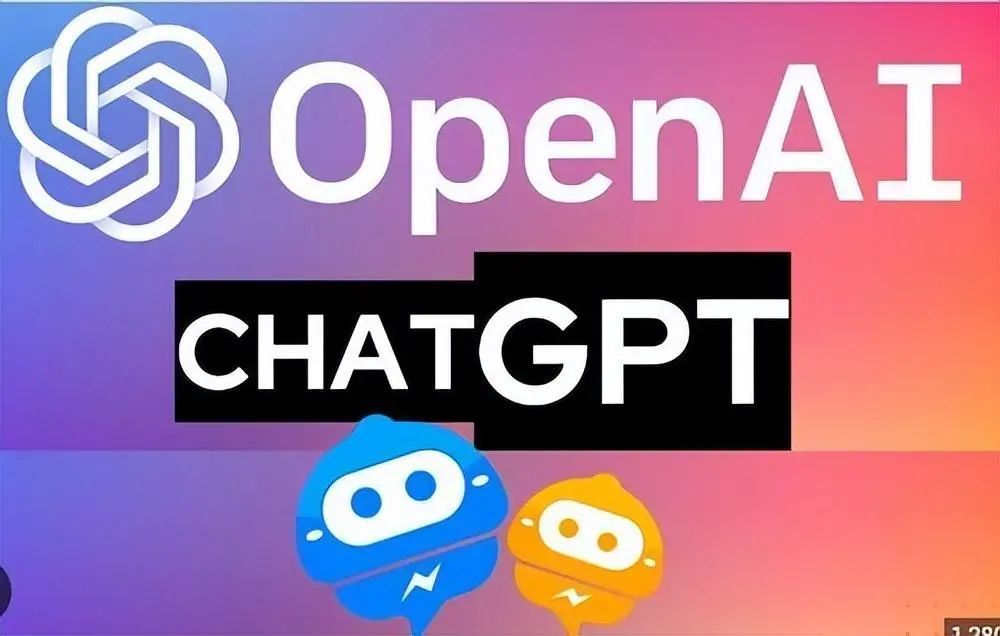


目 录

* 什么是ChatGPT··························2
* ChatGPT的工作原理······················4

• ChatGPT模型的主要训练流程···········6

* 对ChatGPT的一些评价 ··················8
* ChatGPT的未来与趋势···················11



1

什么是ChatGPT

ChatGPT是一种**自然语言处理模型**，ChatGPT全称**Chat** **G**enerative **P**re-trained **T**ransformer，由OpenAI开发。它使用了基于Transformer的神经网络架构，可以**理解和生成自然语言文本**。ChatGPT是当前最强大和最先进的**预训练语言模型**之一，可以生成具有**逻辑和语法正确性的连贯文本**。它在自然语言处理的各个领域，例如**对话生成、文本分类、摘要生成和机器翻译**等方面都取得了非常优秀的成绩。ChatGPT的成功表明，预训练语言模型已经成为自然语言处理领域的主流技术之一。

* ChatGPT在**语言理解、机器翻译、文本分类、信息抽取、文本生成**方面表现相当优秀。
* 目前NLP的主流解决技术方案是**人工智能**，人工智能的技术要素包括**数据、算法、算力、模型**。他们的关系为**通过数据、算法、算力求模型**，通俗地理解为**如同人类一样用数据找到规律**。人工智能区别于传统编程开发，传统编程开发是**用已知规律求数据。**
* 简单来说，ChatGPT是一种自动处理的AI智能系统，他可以根据人们说的话，来进行各种计算，从而得出对人类最有益的结论和建议。在某些实际领域上，例如照顾人类饮食起居，身体健康，工作生产都有着重大作用。

2

图片包含 背景图案

描述已自动生成它其实处处在我们身边

* 就像大家平常可能家里会有一些智能语音助手。
* 例如小爱同学，天猫精灵，小易。
* 大家和他们聊天的时候，大家总会发现，他们几乎能回答我们问的所有问题，而这个，就是他们的万能所在了。
* 例如你想听个音乐，你就可以让他帮你找首歌。有时候你睡觉前想听故事，你甚至也可以让他帮你找很多你喜欢的故事来陪伴你入睡。

3

二、ChatGPT的工作原理

* 日程表

  描述已自动生成ChatGPT是一个模型，是通过数据、算法、算力求得的一个模型，其中数据、算法、算力具体内容为：

1. 数据：ChatGPT使用了大量的自然语言文本数据进行预训练，包括维基百科、BookCorpus等。
2. 算法：ChatGPT使用了Transformer算法，这是一种基于自注意力机制的神经网络模型，能够有效地处理自然语言文本数据。
3. 算力：为了训练和使用ChatGPT模型，需要大量的计算资源，包括GPU和分布式计算框架等。具体来说，OpenAI在训练13亿参数的GPT-3模型时使用了数千个GPU和TPU。

4

* 其原理**主要**包括以下几个方面：

1. Transformer结构：ChatGPT使用了Transformer结构作为其基本架构，通过自注意力机制实现了对输入序列的编码和对输出序列的解码。
2. 预训练：ChatGPT使用了大规模语料库进行了预训练，从而学习到了大量的语言知识，包括词汇、语法和语义等。
3. 微调：ChatGPT在预训练的基础上，通过针对具体任务进行微调，从而实现了在特定任务上的优秀表现。
4. 无监督学习：ChatGPT通过无监督学习的方式进行训练，即在不需要人工标注数据的情况下，通过最大化语言模型的似然函数来训练模型，从而实现了对语言知识的自动学习。
5. 刚才那些话，大家可以将ChatGPT理解为一名学生，但是它太强大了，强大到可以不需要老师过多的指导。
6. 他可以自己泡在学校的图书馆一整天。但和我们不同的是，他泡在图书馆一天所读的书，可能是我们的几千倍，甚至几万倍，乃至几亿倍。
7. 他拥有一台庞大的计算系统，可以轻松计算出各种算术题，但是用的时间只需要人类的亿分之一，就相当于你刚刚抬笔，他就已经交卷出考场了。

5

1. 当然，他不是一日成才的，在他诞生之初，科学家们就利用输入数据库，语言库等而外导向链接等各种手段来进行预学习。
2. 就是他在大家还不认识他的时候，他就提前接受了我们学习的所有知识，每天要学习几十亿次，直到他完全吸收这些知识。
3. 他的处理系统和我们人不太一样，我们人类的大脑是由各种生物体质组成的，而他们则是由半导体和高精密电子零件组成的硅基大脑。
4. 他几乎处处都比人类强悍，但是别忘了，就算再厉害的人工智能，也需要学习，如果他连学习都不学了，最后的下场也是什么都不会，被人类打败。

那么ChatGPT模型是如何训练的呢

* ChatGPT模型的主要训练流程可以概括为以下几个步骤：

1. 数据准备：准备大规模的文本数据作为训练数据集
2. 模型设计：采用Transformer架构，构建多层的编码器-解码器结构，并采用自注意力机制实现对文本的建模
3. 模型初始化：使用随机初始化的参数，构建初始的模型
4. 模型训练：采用大规模的文本数据集对模型进行训练，以最小化损失函数为目标，让模型逐步学习输入文本的规律

6

1. 模型评估：对训练好的模型进行评估，通常采用困惑度（perplexity）等指标来衡量模型的性能

7

1. 模型微调：通过对模型参数进行微调，进一步提高模型的性能
2. 模型部署：将训练好的模型部署到应用场景中，实现自然语言生成、问答等功能

* 为什么ChatGPT在**语言理解、机器翻译、文本分类、信息抽取、文本生成**方面表现相当优秀？
* 其中重要的一个原因是**预训练**，相当于**人类的通识教育**
* 预训练的文本数据集包括**维基百科、书籍、期刊、Reddit链接、Common Crawl和其他数据集**， 主要语言为英文，中文只有5%，ChatGPT-3预训练数据量达45TB，参数量1750亿，对应成本也非常高，GPT-3 训练一次的费用是 460 万美元，总训练成本达 1200 万美元
* 注：参数量指的是模型中需要学习的可调整参数的数量，也就是神经网络中各层之间的连接权重和偏置项的数量之和。在深度学习中，参数量通常是衡量模型规模和容量的重要指标，**一般来说参数量越多，模型的表达能力也就越强。**

三、对ChatGPT的一些评价

* 现如今还有很多old-time NLPers停留在BERT时代，思考BERT训练的LM能做的一些任务，而他自己很少看到新加入的NLPers有这个问题。从我个人的理解，他这里说的[new-joiners](https://www.zhihu.com/search?q=new-joiners&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A2787763735%7D)，大概指的是在GPT3出现之后加入到NLP领域并进行探索的人，当然包括他自己。
* **而从这个角度，国内的情况更为夸张。**首先，GPT3的API需要收费，并且国内IP是没法直接访问的，我个人也和很多国内NLPers进行了交流，大家对于GPT3甚至后GPT3时代[语言模型](https://www.zhihu.com/search?q=%E8%AF%AD%E8%A8%80%E6%A8%A1%E5%9E%8B&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A2787763735%7D)的进展了解寥寥，大家对于GPT3、GPT3.5（davincci001/002）有什么样的能力了解都这么少，更遑论训练出这样的模型了。**这确确实实让我强烈的感受到了"卡脖子"的感觉。**
* 近期看到很多人不管是在朋友圈还是知乎回答，一遍又一遍的验证了这条推特的说法。我甚至还看到有人觉得训练出这样的模型“不算太难”，据我了解，在GLM130B出来之前，国内甚至没有一个能对标2020年6月OpenAI发布的GPT3的模型，更别说他们后续远远更强大的GPT3.5（davincci001/002）和近期新闻满天飞的GPT4了，让我自己真实的感觉到[信息茧房](https://www.zhihu.com/search?q=%E4%BF%A1%E6%81%AF%E8%8C%A7%E6%88%BF&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A2787763735%7D)的可怕。但ChatGPT的火爆让我觉得很幸运，至少在这个时刻，更多的人可以看到他们模型的能力，也就有更多的人可以看到差距，激发更多人去思考。

8

* **后面说一点我自己作为视觉领域的研究者（外行），对于GPT3时代的理解，以及近一两年的一些思考和思维变化。**
* 在理解了visual in-context learning之后，我才某种程度上对于我们落后openai的代际有了一定的认知（嘲笑自己在2022年11月才进一步读明白OpenAI在2020年6月发布的文章）。在这之后，又再次拜读了后GPT-3时代的一系列文章，[webgpt](https://www.zhihu.com/search?q=webgpt&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A2787763735%7D)、instructGPT、alignment research等，才发现在GPT-3之后，他们一直在探索的其实就是新的loss signal（驱动模型训练的信息），这可能也是因为他们先意识到[auto-regressive](https://www.zhihu.com/search?q=auto-regressive&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A2787763735%7D) LLM这种预测式的输出的局限性，比如对于一些之前文章没出现过的问题难以给出回答，例如WebGPT中长江还是尼罗河哪个更长的例子，哈利波特和指环王哪个字数更多等，这些问题都难以从LM这种任务中直接学到。而他们好像探索出了一条路径，基于GPT3对于语言的理解作为预训练模型，再使用强化学习去校正模型来拟合人对于很多知识的理解和语言的生成。
* 在我自己理解了这个LLM+RL的loop之后，我自己深切的感受到**好像真的找到了一条通往AGI的路径了**。就着这个问题和很多人交流过，让我印象最深的是和[黄铁军](https://www.zhihu.com/search?q=%E9%BB%84%E9%93%81%E5%86%9B&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A2787763735%7D)老师交流时他给出了一个例子：有研究表明，尽管盲人看不见，并且大多数是通过语言进行交流和反馈的，但他们和视力正常的人对颜色这种抽象视觉概念有着相似的理解，这个对色彩的理解是在盲人想象中的世界里的。LM+RL甚至也能帮助人类理解视觉，当然对于在LM里加入视觉信号，对OpenAI而言或许并不是一件难事，Flamingo已经给出了一个非常强的[solution](https://www.zhihu.com/search?q=solution&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A2787763735%7D)。

9

* ChatGPT的出现真的是很幸运，它打破了国内国外的信息壁垒、打破不同领域的壁垒，让更多人看到了这样的工作。
* 我真的希望各个领域的研究者，特别是相对比较[senior](https://www.zhihu.com/search?q=senior&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A2787763735%7D)的研究者们，能够去仔细阅读OpenAI的一系列文章，真的花些时间去理解现在世界最前沿的研究进展到了什么阶段，多看一些墙内墙外的examples，甚至自己去试一试，努力去理解后GPT3时代语言模型的能力，这真的比再在自己的领域中固步自封重要太多了。
* 期待国内也早日能有像OpenAI一样的公司出现，产出GPT3级别的工作。



10

四、ChatGPT的未来与趋势

* 图片包含 游戏机, 食物, 衬衫

  描述已自动生成从目前的体验来看，ChatGPT的边界就在「可信」上。你不能完全信任它，因此也要求你能验证它说出的内容。举个简单的例子，我让他帮我写邮件，这就是可以验证的——我只要读一遍，就知道它写的对不对。
* 举个反面的例子，就是向他咨询领域外的内容。比如说咨询法律：我能相信吗？我能验证吗？就算各个法律都是公开的，我也不能相信它。万一有什么[特殊条款](https://www.zhihu.com/search?q=%E7%89%B9%E6%AE%8A%E6%9D%A1%E6%AC%BE&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A2887348764%7D)我查不到，那后果可能会非常严重。
* 因此，在重要的问题上，目前的ChatGPT只能帮你做你能做的事情，只不过可以大大加快速度（后面我会提到，这个加速可不只是x2、x3这种级别）。比如修改语法、润色文章，我自己也能做，但它做的更快更好。ChatGPT写邮件更是一绝，落落大方，不卑不亢，只需要自己稍稍改动就可以发出。
* 虽然对我个人来说，ChatGPT干的最多的就是文秘的工作，但价值最大的，则是充当搜索引擎。
* 先说一个真实的笑话：1994年的一篇[生物论文](https://www.zhihu.com/search?q=%E7%94%9F%E7%89%A9%E8%AE%BA%E6%96%87&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A2887348764%7D)「重新发明了微积分」[1]。我们当然可以从学科壁垒、孤陋寡闻等角度去理解这件事情，但也可以从知识的角度去理解它：为何他没法检索到「微积分」已经存在了？
* 因为一段文字存在至少两个层级——文字的层级，以及知识的层级。在文字的层级上，它们要做的「determine total area under curves」和「微积分」、「calculus」差别非常大。而在知识层面上，它们是一件事。传统搜索引擎无法解决这一问题，稍有智能的搜索引擎可能可以解决，但绝对没ChatGPT做的那么好。就拿这个问题测试一下ChatGPT：

11

文本, 信件

描述已自动生成

* 不光让你去学积分，还给了详细解释。
* 第一件事是我一直在思考的一个问题：我们人类在面对不同[复杂度](https://www.zhihu.com/search?q=%E5%A4%8D%E6%9D%82%E5%BA%A6&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A2887348764%7D)的问题时，思考的时间长短不一。而传统的算法也是如此，会有随着问题规模增长的时间复杂度，例如O(n), O(n^2)等。凭什么[神经网络](https://www.zhihu.com/search?q=%E7%A5%9E%E7%BB%8F%E7%BD%91%E7%BB%9C&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A2887348764%7D)就能用同一个时间复杂度处理不同的输入呢？（例如[图像识别](https://www.zhihu.com/search?q=%E5%9B%BE%E5%83%8F%E8%AF%86%E5%88%AB&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A2887348764%7D)）有没有相关的工作呢？
* 必须得承认，我确实可能孤陋寡闻了，但我在[搜索引擎](https://www.zhihu.com/search?q=%E6%90%9C%E7%B4%A2%E5%BC%95%E6%93%8E&search_source=Entity&hybrid_search_source=Entity&hybrid_search_extra=%7B%22sourceType%22%3A%22answer%22%2C%22sourceId%22%3A2887348764%7D)上也确实搜不到满意的答案。而ChatGPT直接给出回答：

12

文本, 日程表

中度可信度描述已自动生成

文本, 信件

描述已自动生成

直接告诉：Dynamic Neural Network。ChatGPT当然可能会瞎扯，但我可以容易地用搜索引擎验证。如果我不知道这个关键词，花几个月重新造轮子也不无可能。从这个角度看，它极大地加快了我学习、研究的速度。

回到那个生物学笑话上，学科壁垒确实也是导致其出现的原因之一。但很多时候，词汇恰恰是导致壁垒的原因。消除这一壁垒，其价值会远远大于很多人的想象。

——以上内容摘自网络 百科 知乎 略有删减

13