



ChatGPT的技术发展路径和带来的影响

2023-03



目录

CONTENT

01: 人工智能和NLP发展路径

02: GPT系列模型发展路径

03: ChatGPT技术原理解析

04: ChatGPT的影响和意义

人工智能技术发展脉络

人工智能诞生最早可以追溯到1956年的感知机模型，经过近70年的发展，已经渗透到各行各业。

前神经网络 专家系统 概率推论

神经网络 基础模型 大规模模型

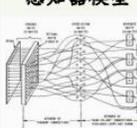
2011年以前

2012年至今

决策式/分析式AI时代

生成式AI急速发展

1956
感知器模型



机器语言翻译



1980s
Prolog & Lisp
专家系统

解答特定领域的问题



1990s
机器学习算法



1994
GroupLens
首个自动化推荐系统

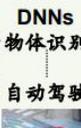
推荐算法



2011
大数据



2012
DNNs
物体识别
自动驾驶



2014
GAN
生成图像但分辨率不高



2017
Transformer



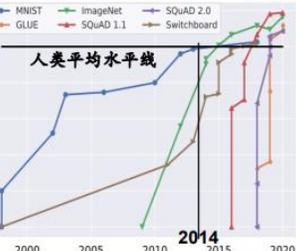
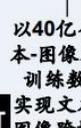
2020
GPT-3



2021
CLIP



2022
DALL E2



AI的急速发展使得AI模型在手写识别、语音识别、图像识别、阅读理解、语言理解等领域的表现加速超越人类平均水平

2011年之前

• 模型简单

受制于当时计算机硬件的发展，模型普遍简单，能力也相对较弱

• 场景局限

模型只能处理单一场景，通用能力非常弱，导致成本过高。

• 人工干预

往往需要如制定规则，词表，标注样本等大量人工工作。

2012年至今

• 大模型

得益于底层技术原理的突破和硬件算力的发展，模型越来越大，GPT3（2020年）达到了惊人的1750亿个参数（一般人脑有神经元120到140亿个）

• 大数据

庞大的模型需要对应于海量的数据，GPT3使用的数据量已达到45TB，包括了多种主流语言

• 多模态

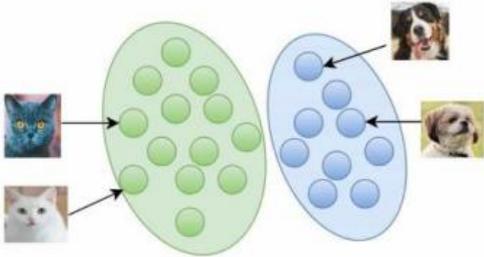
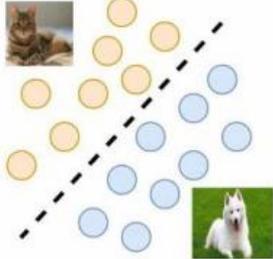
语音，文字，图像之间不再存在明显壁垒，模型处理复杂场景的能力明显增加

人工智能的任务类型

一般来说，人工智能处理的任务可以分为两类：

1.决策式：人工智能回答“选择题”，模型主要处理诸如：判别，分类，排序等任务。

2.生成式：人工智能回答“问答题”，需要模型根据输入，自动生成一些新内容（客观世界可能从未出现）。

类型	决策式AI	生成式AI
技术路径	已知数据分别求解输出类别标签，区分不同类型数据,例如将图像区分为猫和狗 	分析归纳已有数据后创作新的内容，例如生成逼真的猫或狗的图像 
成熟程度	技术成熟，应用广泛，辅助提高非创造性工作效率	2014年开始快速发展，近期发展速度呈指数级爆发，部分领域应用落地
应用方向	推荐系统、风控系统、决策智能体等	内容创作、科研、人机交互以及多个工业领域
应用产品	人脸识别、精准广告推送、金融用户评级、智能辅助驾驶等	文案写作、文字转图片、视频智能配音、智能海报生成、视频智能特效、代码生成、语音人机交互、智能医疗诊断等

NLP发展的技术路径

自然语言处理（NLP, Natural Language Processing）是研究人与计算机交互的语言问题的一门学科。

传统方法(2012年以前)

典型方法

常用方法有SVM、TF-IDF、LDA、同义词表、句法解析，语义规则，BP神经网络。

特点

- 需要人工进行干预，如制定规则，词表，权重等。
- 应用场景单一，都是针对特定场景进行开发。
- 需要做大量特征工程。

神经网络(2012年~2018年)

典型方法

Word2vec的诞生标志着NLP和神经网络的结合，陆续出现了FastText、TextCNN、LSTM、Attention等模型

特点

神经网络的出现，大量减少了人工干预的工作，同义词可以通过向量表征自动学会，句法分析，词性标注等都已经嵌入进模型结构。

大模型预训练(2018年~至今)

典型方法

2018年，GPT和Bert的先后出现，标志着大模型时代的到来，后续诞生的Bart、ERNIE，T5等方法不断探索大模型的能力边界。

特点

模型参数量、数据量均上升了一个台阶，但是大部分模型利用文本自身特点生成训练样本，省去了很多人工标注的工作。在特定场景使用微调技术，降低了大模型跨场景的门槛。

提示(Prompt)学习(2020~至今)

典型方法

提示学习将下游任务的建模方式重新定义，通过合适的prompt来实现直接在预训练模型上解决下游任务，如GPT3，T5。

特点

- 将所有任务，都转变为生成式任务。
- 模型不再需要微调，使得零样本和少样本学习成为可能。

目录

CONTENT

01: 人工智能和NLP发展路径

02: GPT系列模型发展路径

03: ChatGPT技术原理解析

04: ChatGPT的影响和意义

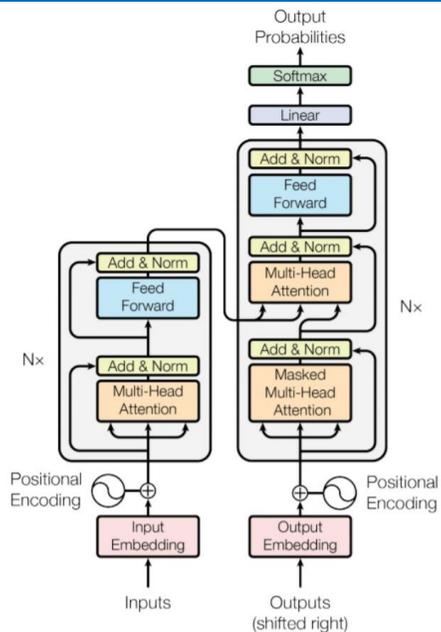
GPT系列模型发展路径

在这个大模型时代，国内外各大厂商均对大模型积极布局，呈现百家争鸣的现象。Nvidia或成最大赢家。

各大厂商的布局

Release	Collection	Model	Model Details			Data Collection & Training Details			Methods
			Base	Size	Public?	Prompt Types	Tasks in Plan	# Exs	
2020 05	UnifiedQA	UnifiedQA	RoBERTa	110-340M	P	ZS	46 / 46	750k	
2021 04	CrossFit	BART-CrossFit	BART	140M	HP	FS	115 / 159	71M	
2021 04	Natural Inst v1.0	Gen. BART	BART	140M	HP	ZS / FS	61 / 61	620k	Detailed k-shot Prompts
2021 09	Flan 2021	Flan-LaMDA	LaMDA	136B	HP	ZS / FS	62 / 62	4.4M	Template Variety
2021 10	P3	TO, TO+, TO**	TS-LM	3-11B	P	ZS	62 / 62	12M	Template Variety, Input Inversion
2021 10	MetalCL	MetalCL	GPT-2	770M	P	FS	100 / 142	3.5M	Input Inversion, Noise Channel Opt
2021 11	ExMix	ExTS	TS	220M-11B	HP	ZS	72 / 107	500k	With Pretraining
2022 04	Super-Natural Inst.	Tk-Instruct	TS-LM, mTS	11-13B	HP	ZS / FS	1556 / 1613	5M	Detailed k-shot Prompts, Multilingual
2022 10	xP3	BLOOMz, mTO	BLOOM, mTS	13-176B	P	ZS	53 / 71	81M	Massively Multilingual
2022 12	Unnatural Inst.!	TS-LM-Unnat. Inst.	TS-LM	11B	HP	ZS	-20 / 117	64k	Synthetic Data
2022 12	Self-Instruct!	GPT-3 Self Inst.	GPT-3	175B	HP	ZS	Unknown	82k	Synthetic Data, Knowledge Distillation
2022 12	OPT-1M L Bench!	OPT-1M L	OPT	30-175B	P	ZS + FS	-2067 / 2207	18M	Template Variety, Input Inversion, Multilingual
2022 11	Flan 2022 (ours)	Flan-T5, Flan-Palm	TS-LM + Palm	10M-540B	HP, P	ZS + FS	1836	15M	Template Variety, Input Inversion, Multilingual

底层技术原理



三种技术流派

- 自从2017年，Transformer被提出后，Google, Meta, OpenAI均在大模型上有所建树，并产生了三种技术流派：1.自编码(Bert等) 2.自回归 (GPT等) 3.二者结合 (T5, BART等)



NLP技术发展带来的影响

最近10年，是NLP技术和业务场景发展最快的黄金时期，NLP本身的技术体系被重构，所影响的业务领域也不断扩大。

技术体系的改变

• 中间任务的消亡

中间任务是指并不直接面向最终目标，而是为了解决最终目标的一些阶段性任务。因为神经网络的发展，诸如语法分析，词性标注，分词等特有的中间任务已经几乎无人问津。

• All you need is Money

现在完成一个NLP任务，需要的不仅是技术，而且还包括算力，数据等资源。这背后都是巨额的资金投入。

使用场景的扩大

• 搜广推等内容链接领域

NLP最成功的商业化应用场景就是搜索，推荐，广告，技术的发展催生出一系列巨头（谷歌，百度，字节）

• 丰富的人机交互

各类聊天机器人，语音助手，等人机交互场景趋于成熟。

• 内容领域的变革

ChatGPT将会颠覆整个内容生态，内容型公司（如知乎，微博）的重点将从内容分发转变为内容生产。今后互联网将充斥着大量机器产生的内容，这对监管会带来极大的挑战。

未来的发展



• 大模型时代

ChatGPT的出现，标志着大模型已经突破了技术的天花板，让大模型这条技术发展路线更为坚定。

• 小模型时代

大模型对于创业公司和中小企业来说无疑需要倾注大量的资源。轻量化预训练模型，给出了一个新的思考方向。

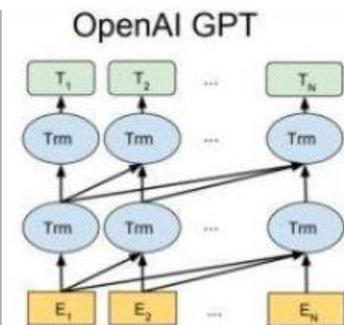
OpenAI和ChatGPT的发展路径

OpenAI, 在美国成立的人工智能研究公司, 核心宗旨在于“实现安全的通用人工智能(AGI)”, 使其有益于人类。



OpenAI 成立于2015年, 是一家非营利性研究机构, 它的愿景是构建出安全、对人类有益的通用人工智能 (AGI), 由伊隆·马斯克、Sam Altman 等人宣布出资 10 亿美元成立。

▶ BEGIN



GPT面世, 标志着人工智能正式走向大模型时代



2019年7月22日, 微软宣布将向 OpenAI 投资 10 亿美元, 以共同开发用于 Microsoft Azure 云平台的新技术

微软率先宣布与OpenAI独家合作打造了一台性能位居全球前五, 拥有超过28.5万个CPU核心、1万个GPU、每GPU拥有400Gbps网络带宽的超级计算机——Azure AI超算平台, 主要用于大规模分布式AI模型训练。



GPT-3依旧延续自己的单向语言模型训练方式, 只不过这次把模型尺寸增大到了1750亿, 并且使用45TB数据进行训练。

- 2022年11月30日, ChatGPT上市。
- 2022年12月5日 OpenAI创始人山姆·奥特曼宣布, ChatGPT的用户已经突破100万人, 仅用时5天。
- 2023年1月末 ChatGPT 用户数突破1亿, 成为史上用户数增长最快的消费者应用。
- 2023年1月末 微软宣布向推出 ChatGPT的OpenAI公司追加投资100亿美元。

2015年

2018年

2019年

2020年

2022年

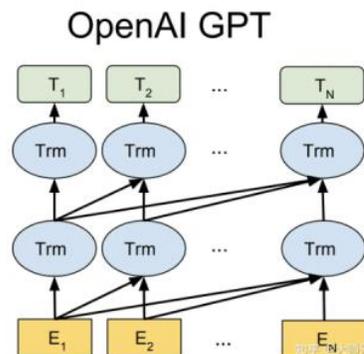
GPT-1系列模型发展路径

GPT的全称是Generative Pre-Trained Transformer，顾名思义，GPT的目的就是通过Transformer为基础模型，使用预训练技术得到通用的文本模型。

GPT模型结构

- GPT-1的无监督预训练是基于语言模型进行训练

在GPT-1中，使用了12个transformer块的结构作为解码器，每个transformer块是一个多头的自注意力机制，然后通过全连接得到输出的概率分布。



欢迎访问：电子书学习和下载网站 (<https://www.shgis.com>)

文档名称：ChatGPT的技术发展路径和带来的影响.pdf

请登录 <https://shgis.com/post/1743.html> 下载完整文档。

手机端请扫码查看：

