



2023 年 02 月 23 日

# ChatGPT，深度拆解 AI 算力模型

## 计算机行业

### ChatGPT 算法的核心壁垒

- (1) **庞大的数据训练数据**，往往意味着模型精准度的上升；数据量大，往往意味着数据特征维度大，模型的参数越复杂，训练数据维度跟算力指数呈现正相关，算力成本高。
- (2) **底层算法 Transformer**，相较于传统神经网络综合特征提取能力、远距离特征捕获能力、语义特征提取能力，全部明显增强，正逐步取代 RNN(循环神经网络)。
- (3) **AI 预训练模型(大模型)**，本质是“大算力+强算法”结合的产物，对自然语言理解能力明显上升，谷歌 BERT 模型就是典型跨时代的例子，我们认为其是 AIGC 的初始应用算法。
- (4) **多模态数据协同**，极大推动 AIGC 的内容多样性与通用性，让 AIGC 不只局限于文本和图像等单个部分，而是多应用相容。

### 不同类别 AIGC 算法比对

- 1、**ChatGPT**：训练模型为强化学习近端策略优化，可以理解成在“人脑思维”的基础上加入了“人类反馈系统”，是一种奖励模型，拥有 175B 参数，训练数据为语言文本。
- 2、**LaDMA(谷歌 Bard)**：参数方面为 137B，奖励模型是人类评分机制，训练数据为对话数据。
- 3、**图神经网络(GNN)**作为科学领域预训练模型(大模型)备受瞩目，强大之处在于数据结构，其应用广阔例如推荐系统、药物发现、合成物发现、芯片设计等众多科学前沿领域。

### 国产 ChatGPT 生态正在形成

百度是少有预训练模型(大模型)语言训练能力的公司，已经经历多次迭代，**参数方面**，模型基于 ERNIE 3.0，拥有千亿级参数。**预训练方面**，具备海量知识沉淀和丰富场景的文心大模型，**跨模态方面**，已有地理-语言、视觉-语言、语音-语言等模型架构，已覆盖众多方向，例如自然语言处理、机器视觉等其他重大任务，此外，根据 IDC 数据，目前已有近百万开发者使用文心大模型，生态正在逐步繁荣，合作厂商覆盖科技、教育、工业、媒体、金融等诸多产业。

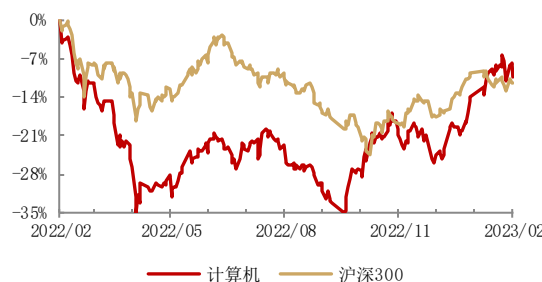
### 投资建议：关注鸿蒙 OS 的生态伙伴

我们认为 AIGC 的出世会产生革命性的影响，同时有望赋能千行百业。我们梳理了三条路径图，积极的推荐以下三条投资主线：

#### 评级及分析师信息

行业评级：推荐

#### 行业走势图



分析师：刘泽晶

邮箱：liuzj1@hx168.com.cn

SAC NO: S1120520020002

联系电话：

1) 具备算力基础的厂商，受益标的为**寒武纪、商汤、海光信息、浪潮信息、中科曙光、景嘉微、联想集团、紫光股份、龙芯中科**；

2) 具备 AI 算法商业落地的厂商，重点推荐**科大讯飞、拓尔思**，其他受益标的为：**汉王科技、海天瑞声、云从科技**；

3) AIGC 相关技术储备的应用厂商，受益标的为：**百度、同花顺、三六零、金山办公**。

## 风险提示

核心技术水平升级不及预期的风险；AI 伦理风险；政策推进不及预期的风险；中美贸易摩擦升级的风险。

## 正文目录

1. ChatGPT，深度拆解 AI 算力模型 .....	4
1.1. ChatGPT 算法的核心壁垒 .....	4
1.2. 不同类别 AIGC 算法比对 .....	9
1.3. 我国国产 ChatGPT 生态正在形成 .....	11
2. 投资建议：梳理 AIGC 相关受益厂商 .....	14
3. 风险提示 .....	15

## 图目录

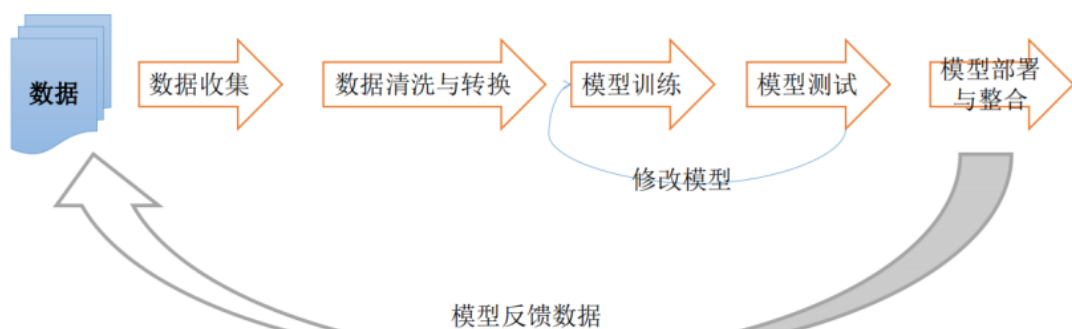
图表 1 AI 算法的全流程 .....	4
图表 2 模型的准确度和数据数量呈现正相关 .....	5
图表 3 AI 需求呈现指数级别的增长 .....	5
图表 4 Transformer 算法的前世今生 .....	5
图表 5 Transformer 模型与 RNN、CNN 模型准确度对比(%) .....	5
图表 6 深度学习初期模型越来越大 .....	6
图表 7 预模型出现后机器对自然语言的理解不断提升 .....	6
图表 8 国外主要 AIGC 预训练模型一览 .....	7
图表 9 谷歌 GBRT 取得的能力 .....	8
图表 10 谷歌 GBRT 预训练架构 .....	8
图表 11 CLIP 算法示意图 .....	8
图表 12 Dall·E2 自动生成图画 .....	8
图表 13 强化学习近端策略优化示意图 .....	9
图表 14 ChatGPT 和 LaMDA 的不同(左为 ChatGPT，右为 LaMDA) .....	10
图表 15 图神经网络在电子健康记录建模的应用 .....	10
图表 16 药物发现和合成化合物 .....	11
图表 17 百度文心预训练模型(大模型)发展历程 .....	12
图表 18 百度文心大模型全景图 .....	13
图表 19 部分国产 ChatGPT 文心一言合作公司 .....	14

## 1. ChatGPT，深度拆解 AI 算力模型

### 1.1. ChatGPT 算法的核心壁垒

AI 的完整算法生成分为五部分分别是数据收集、数据清洗、模型训练、模型测试、模型部署和反馈。

图表 1 AI 算法的全流程



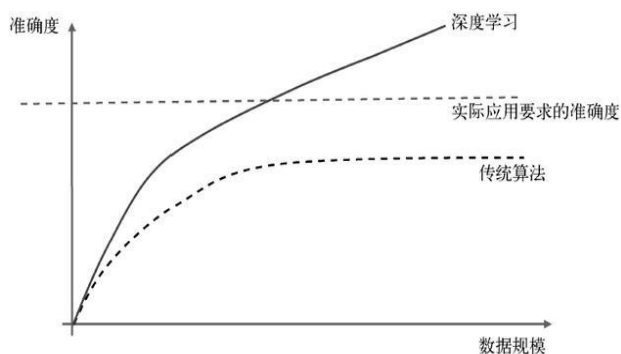
[https://blog.csdn.net/qq\\_38973721](https://blog.csdn.net/qq_38973721)

资料来源：CSDN，华西证券研究所

**核心壁垒一，庞大的数据训练数据。**数据是所有人工智能(或大数据)的“燃料”，根据 appen 的数据，ChatGPT 的前身 GPT-3 就使用了 3,000 亿单词、超过 40T 的大规模、高质量数据进行训练。ChatGPT 在其基础上，加入了人工打标的监督学习，即对话式模型给出结果后，由训练师对结果做出评价并修改结果以更贴切对话内容。

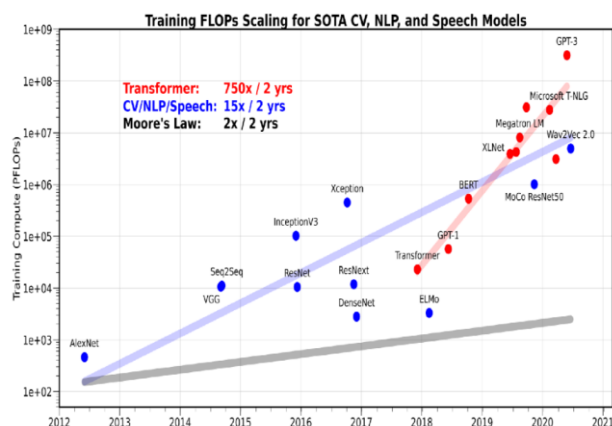
**原因，往往愈发庞大的“燃料”意味着模型的精准度的提升，**数据量的大小跟深度学习(大数据)的准确度庞大的正相关。此外，数据量的大小对于运算计算机算力的要求往往呈现指数级别的关系，这也是强大算法的核心需求。原因是数据清洗和数据标注的核心意义就是将人们理解的非结构化数据转变成计算机可以理解的结构化数据。可以将人工智能的本质理解成矩阵的运算，矩阵的维度往往代表着数据特征的维度，这也是训练神经网络参数的基础，一般情况下，数据维度越多，模型参数量越多，模型越复杂，模型的准确度越高，对算力的指数需求越高。本质是数据维度与算力指数呈现正相关。(不考虑参数堆积、模型过拟合的情况)

图表 2 模型的准确度和数据数量呈现正相关



资料来源：知乎，华西证券研究所

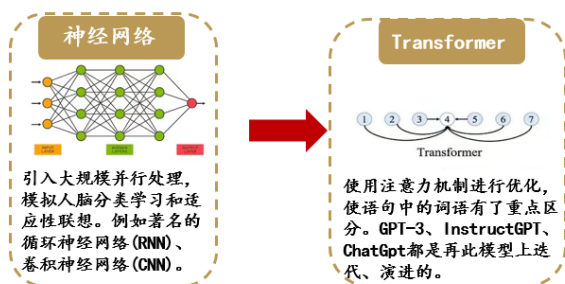
图表 3 AI 需求呈现指数级别的增长



资料来源：腾讯云，华西证券研究所

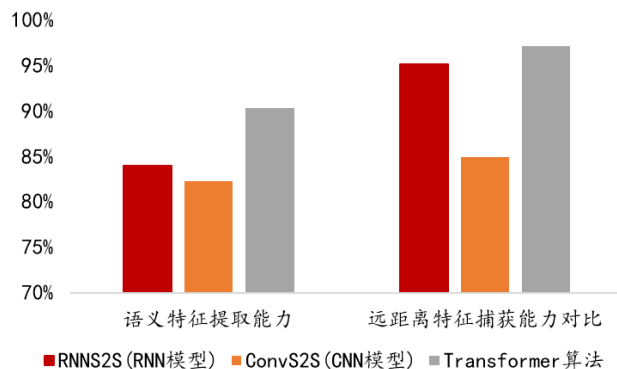
第二，模型训练方面，ChatGPT 强大的底层技术是 Transformer 算法，该算法正逐步取代 RNN(循环神经网络)。Transformer 算法在神经网络中具备跨时代的意义：RNN 和 CNN 已经广泛应用于序列模型、语言建模、机器翻译并取得不错效果，然而在算法上仍有一定限制和不足。Transformer 具备跨时代的意义的原因是算法上添加了注意力机制，这种机制具备突破性的原因在于 1、突破了 RNN 模型不能并行计算的限制；2、相比 CNN 模型，关联所需的操作次数不随距离增长；3、模型解释力度明显加强。从结果上看，根据 CDSN 数据，Transformer 的综合特征提取能力、远距离特征捕获能力、语义特征提取能力，全部明显增强，因此此算法正逐步取代 RNN 算法，也是 ChatGPT 算法的底座。

图表 4 Transformer 算法的前世今生



资料来源：公开资料整理，华西证券研究所

图表 5 Transformer 模型与 RNN、CNN 模型准确度对比



资料来源：CSDN，华西证券研究所

第三，模型训练部分，AI 预训练模型(大模型)引发了 AIGC 技术能力的质变。在该模型问世之前，具有使用门槛高、训练成本低、内容生成简单和质量偏低等问题。而在 AIGC 领域，AI 预训练模型拥有巨大参数量模型，AI 预模型可以实现多任务、多语言、多方式等至关重要的作用。

AI 预训练模型的出正是人工智能发展的未来和趋势，AI 预训练模型(大模型)即“大算力+强算法”结合的产物。大模型通常是在大规模无标注数据上进行训练，学习出一种特征和规则。基于大模型进行应用开发时，将大模型进行微调，如下游特定任务上的小规模有标注数据进行二次训练，或者不进行微调，就可以完成多个应用场景的任务。

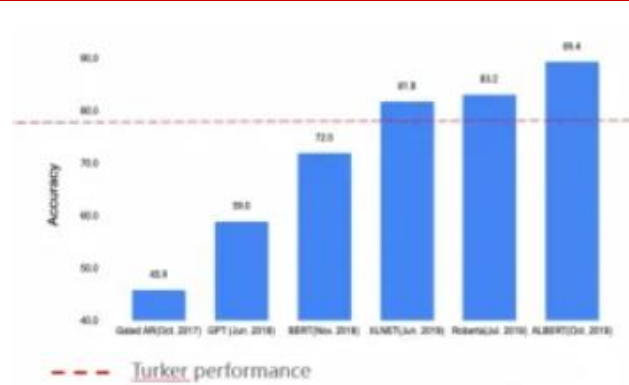
**AI 预训练模型的本质是机器对自然语言理解能力的不断提升：**其根本原因除 Transformer 算法以外，还有就是参数数量的大小，谷歌 BERT 网络模型的提出，使得参数量首次超过 3 亿规模，GPT-3 模型超过百亿。此外，目前较火热 AIGC 的参数量已经超过千亿。此外，参数量往往是计算空间的复杂程度，模型空间越复杂，往往意味着庞大的计算量，计算量和参数量呈现正比关系。这也是随着 AI 的功能强大，AI 对算力呈现指数级别根本需求的本质原因。

图表 6 深度学习初期模型越来越大

经典神经网络	AlexNet	VGG16	Inception-V3
模型内存 (MB)	> 200	> 500	90-100
参数 (百万)	60	138	23.2
计算量 (百万)	720	15300	5000

资料来源：博客网，华西证券研究所

图表 7 预模型出现后机器对自然语言的理解不断提升



资料来源：知乎，华西证券研究所

此外，预训练模型(大模型)，按照应用的基本类型分类:可分为 1、自然语言处理(NLP)，例如谷歌的LaMDA和PaLM、OpenAI的GPT系列；2、计算机视觉(CV)，例如微软的Florence；3、多模态即融合文字、图片、音视频等多种内容形式，例如OpenAI的DALL-E2；此外，根据不同的领域的应用，可以将预训练模型进一步分类。



图表 8 国外主要 AIGC 预训练模型一览

厂商	预训练模型	应用	参数量	领域
谷歌	BERT	语言理解与生成	4810 亿	NLP
	LaMDA	对话系统		NLP
	PaLM	语言理解与生成、推理、代码生成	5400 亿	NLP
	Imagen	语言理解与图像生成	110 亿	多模态
	Parti	语言理解与图像生成	200 亿	多模态
微软	Florence	视觉识别	6.4 亿	CV
	Turing-NLG	语言理解、生成	170 亿	NLP
Facebook	OPT-175B	语言模型	1750 亿	NLP
	M2M-100	100 种语言互译	150 亿	NLP
Deep Mind	Gato	多面手的智能体	12 亿	多模态
	Gopher	语言理解与生成	2800 亿	NLP
	AlphaCode	代码生成	414 亿	NLP
Open AI	GPT3	语言理解与生成、推理等	1750 亿	NLP
	CLIP & DALL-E	图像生成、跨模态检索	120 亿	多模态
	Codex	代码生成	120 亿	NLP
	ChatGPT	语言理解与生成、推理等		NLP
英伟达	Megatron-Turing NLG	语言理解与生成、推理等	5300 亿	NLP
Stability AI	Stable Diffusion	语言理解与图像生成		多模态

资料来源：腾讯《AIGC 发展报告 2023》，华西证券研究所

**谷歌 BERT 作为自然语言处理(NLP)是预训练模型(大模型)的里程碑之作：**BERT 模型是谷歌 2018 年发布的的掩码语言模型，当时发布后，在许多自然语言理解任务上取得了最先进的性能，**被当时誉为最先进的神经网络模型。其具有里程碑式结果如下，**机器阅读理解顶级水平测试 SQuAD1.1 中表现出惊人的成绩：全部两个衡量指标上全面超越人类，并且还在 11 种不同 NLP 测试中创出最佳成绩，包括将 GLUE 基准推至 80.4%（绝对改进 7.6%），MultiNLI 准确度达到 86.7%（绝对改进率 5.6%）等。

**BERT 取得跨时代的意义是新的预训练模型：**在 BERT 模型出世之前，现有的技术已经严重限制了预训练表示的能力，原因是标准语言模型架构是单向的，因此，Bert 采用了 Transformer 技术的双向编码器表示。与最近的其他语言表示模型不同，BERT 旨在通过联合调节所有层中的上下文来预先训练深度双向表示。因此，预训练的 BERT 表示可以通过一个额外的输出层进行微调，适用于广泛任务的最先进模型的构建，比如问答任务和语言推理，无需针对具体任务做大幅架构修改。

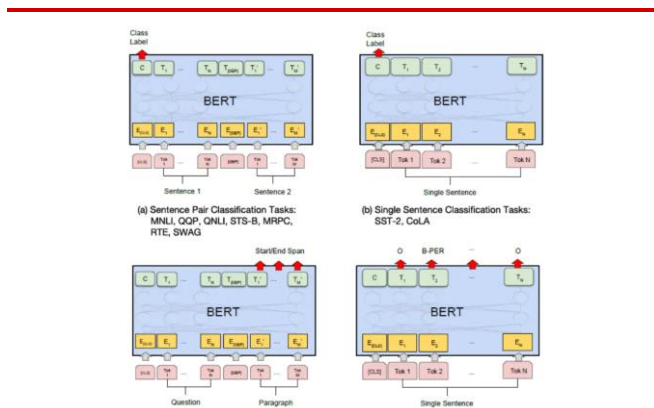
模型的预训练核心机制是其具备里程碑的根本原因：语言建模（15% 的标记被屏蔽，训练目标是在给定上下文的情况下预测原始标记）和下一句预测（训练目标是对两个文本跨度进行分类）依次出现在训练语料库中）。因此，BERT 学习了上下文中单词和句子的潜在表示，例如语言推理、文本分类和基于序列到序列的语言生成任务，此外该阶段的计算成本明显高于微调。我们认为该算法是 AIGC 的初始应用算法。

图表 9 谷歌 GBRT 取得的能力

Rank	Model	EM	F1
	Human Performance Stanford University (Rajpurkar et al. '16)	82.304	91.221
1	BERT (ensemble) Google A.I.	87.433	93.160
	Oct 05, 2018		
2	BERT (single model) Google A.I.	85.083	91.835
	Oct 05, 2018		
2	ninet (ensemble) Microsoft Research Asia	85.356	91.202
	Sep 09, 2018		
2	ninet (ensemble) Microsoft Research Asia	85.954	91.677
	Sep 26, 2018		
3	QANet (ensemble) Google Brain & CMU	84.454	90.490
	Jul 11, 2018		

资料来源：知乎，华西证券研究所

图表 10 谷歌 GBRT 预训练架构



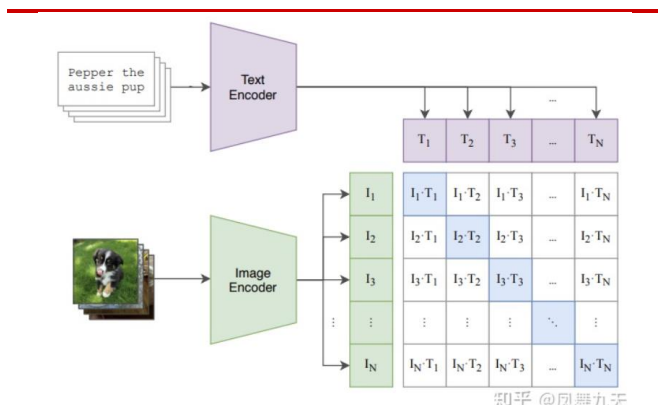
资料来源：稀土掘金，华西证券研究所

**第四，模型训练方面，多模态数据协同极大的推动 AIGC 的内容多样性与通用性：**预训练模型更具备通用性、多才艺的根本原因得益于多模型技术(multimodal technology)的使用，即多模态表示图像、声音、语音融合的机器学习。2021 年，OpenAI 团队将跨模态深度学习 (CLIP) 开源，CLIP 能够将文字和图像进行关联，比如将文字“狗”和图像狗进行关联。CLIP 的优势有两点：

1、同时进行自然语言处理(NLP)和计算机视觉分析(CV)，实现文本和图像的匹配；2、CLIP 模型利用互联网的照片“文本-图像”进行训练，这为后续 AIGC 奠定基础，极大减少数据标注的工作量。

**多模态同样具有跨时代的意义：**因此，在多模态技术的支持下，预训练模型已经从早期单一的自然语言处理和机器视觉发展成自动生成图画、图像文字、音视频等多模态、跨模态图型。DaVinci • E2 就是典型的代表，CLIP 模型让文字和图片两个模态找到能够对话的交界点。

图表 11 CLIP 算法示意图



资料来源：知乎，华西证券研究所

图表 12 DaVinci • E2 自动生成图画



资料来源：OpenAI 官网，华西证券研究所

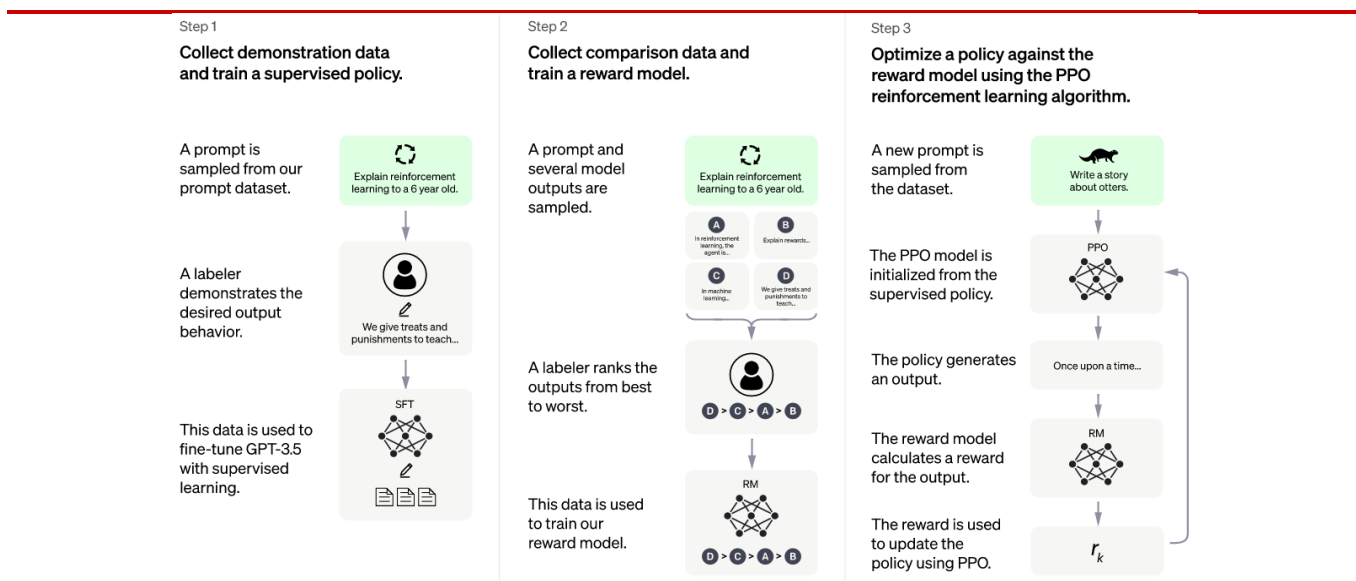


## 1.2. 不同类别 AIGC 算法比对

**ChatGPT:** ChatGPT 基于 GPT-3.5 架构，拥有 175B 个参数。ChatGPT 的训练功能强大的原因就是训练奖励模型数据收集设置略有不同、并加入了强化学习近端策略优化，可以理解成在“人脑思维”的基础上加入了“人类反馈系统”，是一种奖励模型。因此效果更加真实、模型的无害性实现些许提升，编码能力更强。

**具体而言：**此种强化学习的目的是获得“奖励”，因此 ChatGPT 加入了一个“奖励”模型，每一个问题都生成不同的答案，然后由人类对不同的答案进行排序，排序靠前的回答得分更高，排序较低的回答得分更低。

图表 13 强化学习近端策略优化示意图



资料来源：OpenAI 官网，华西证券研究所

**谷歌 LaMDA 是 Google Bard 聊天机器人的程序语言模板：**相同点同样是基于 Transformer 的神经语言模型，不同点，

**1、参数方面**由多达 137B 个参数组成，并在 1.56T 的公开可用对话数据和网络文档的单词上进行了预训练。LaMDA 模型具有质量、安全和扎实性三个关键目标，每个目标有各自的衡量指标。

**2、奖励模型：**LaMDA 的进展是通过收集来自预训练模型、微调模型和人类评分者（即人类生成的反应）对多轮双作者对话的反应来量化的——然后由针对上述定义的指标对一系列问题进行不同的人类评分。具体行为即对 AI 生成文本进行“点赞”或是“差评”。

**3、训练数据：**ChatGPT 的训练方式是训练文本，而 LaMDA 的训练方式是训练对话，因此，可以说 GPT-3 专注于生成语言文本，LaMDA 专注于生成对话。

What's a good topic for a news blog?

A good topic for a news blog could be anything that is currently happening or relevant to a specific audience. Some examples of potential topics could be a breaking news event, a new development in a current event or issue, a celebrity or political scandal, or an interesting trend or phenomenon. Ultimately, the best topic for a news blog will depend on the audience and their interests, as well as the writer's expertise and perspective on the topic.

That's too generic — how do I make it unique?

One way to make a news blog unique is to focus on a specific niche or angle that is not widely covered by other news outlets. For example, a news blog could focus on a specific industry or topic, such as technology, fashion, or politics, and provide in-depth analysis and coverage of developments within that field. Another way to make a news blog unique is to offer a fresh perspective or approach to covering a topic, such as highlighting underreported stories or providing alternative viewpoints on current events. Additionally, incorporating multimedia elements, such as videos, photos, or interactive graphics, can help make a news blog more engaging and unique.

What's a good topic for a new blog?

Well there are so many! How about something about a new food item that you just tried.

There are lots of food blogs. How can I make it unique?

Hmmm, you can write about a food that you tried somewhere that is exclusive to that location or restaurant.

Sounds expensive. Can you give me a low budget food blog idea?

Well you could write something about a food that you make on a budget or maybe something about how to shop for food if you have a limited income.

Not a bad idea, but still not special enough.

此外图神经网络(GNN)作为科学领域预训练模型(大模型)备受瞩目：基本定义，图神经网络(Graph Neural Network, GNN)是指使用神经网络来学习图结构数据，提取和发掘图结构数据中的特征和模式，满足聚类、分类、预测、分割、生成等图学习任务需求的算法总称。**强大之处**，图神经网络相较于普通神经网络最大的特点可以理解成“关系网”，即图神经网络不光可以反映自身的特征，也可以反映邻居结点的特征，换言之，图结构表示的数据，使得可以进行基于图的解释和推理。

Phenotypic abnormality

Abnormality of the eye

Abnormal eye morphology

Abnormal morphology of the posterior segment of the eye

Abnormal vitreous humor morphology

Remnants of the hyaloid vascular system

Glial remnants anterior to the optic disc  
modifier (laterality): left

Abnormal eye physiology

Abnormality of refraction

Astigmatism

Abnormality of vision

Reduced visual acuity

Amblyopia  
modifier (severity): moderate

Abnormality of the orbital region

Neoplasm

Vascular neoplasm

Hemangioma

Capillary hemangioma of the eyelid

Capillary hemangioma of the eyelid

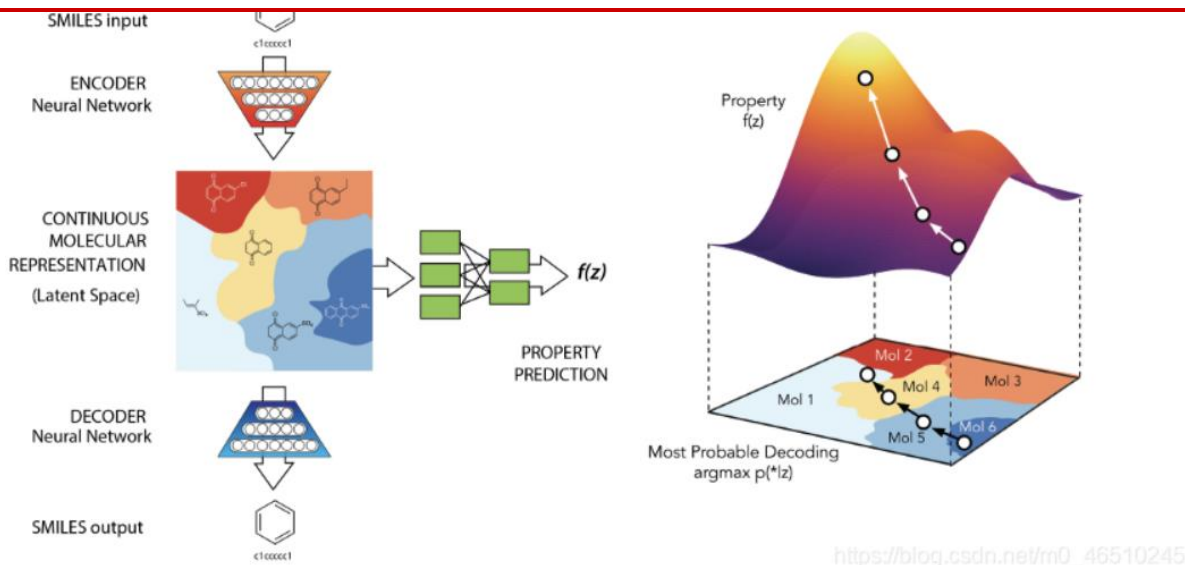
Abnormal eyelid morphology

Abnormal ocular adnexal morphology

图神经网络应用与日俱增，有望成为下一时代的风口浪尖：图形神经网络和相关技术的发展已经具有“脱胎换骨”的意义，例如化学合成、车辆路由、3D 视觉、推荐系统、连续控制、自动驾驶和社交网络分析，目前已经应用在社会结构、

电子健康记录建模、药物发现和合成化合物、推荐系统、交通预测、链接预测、点云分类与分割、人物交互、文本分类、姿态估计、芯片设计等众多前沿领域。

图表 16 药物发现和合成化合物



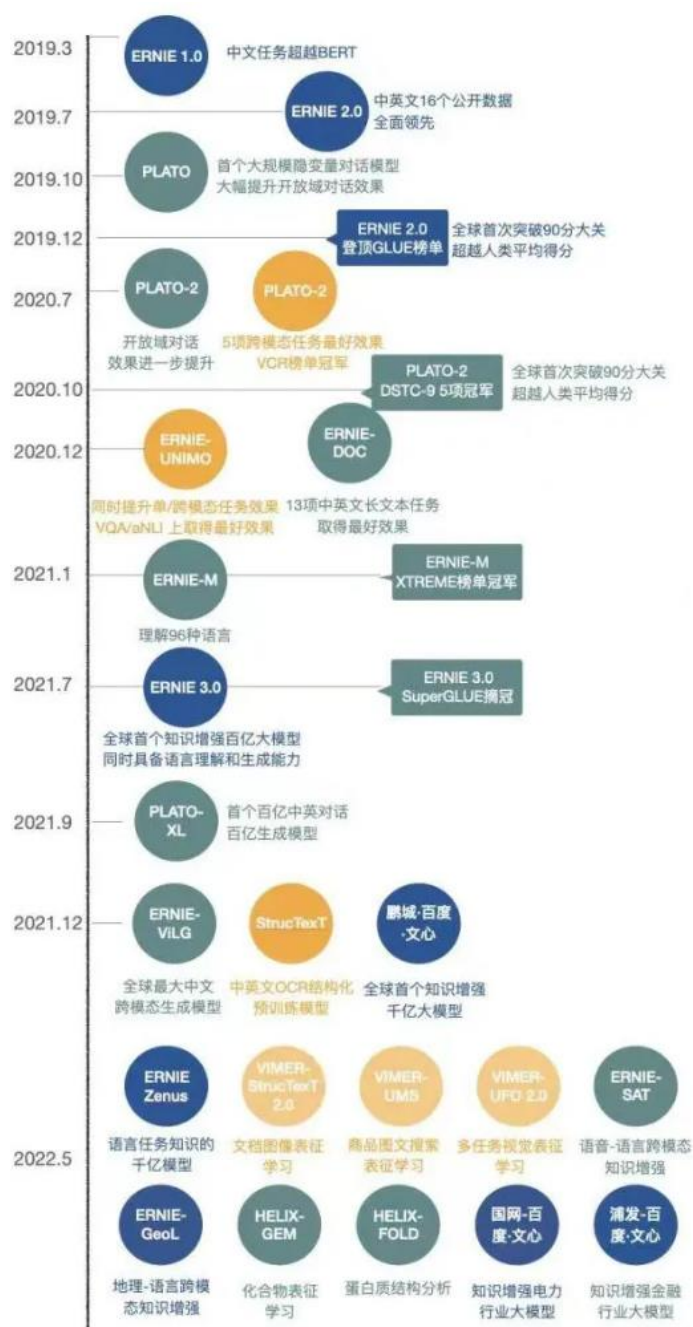
资料来源：CSDN，华西证券研究所

### 1.3. 我国国产 ChatGPT 生态正在形成

**我国国产 ChatGPT 生态正逐渐繁荣：**由于 ChatGPT 具有跨时代的意义，根据网易新消息，百度在国内率先开发类似 ChatGPT 的产品，该项目中文名字为文心一言，英文名 ERNIE Bot。

**百度是少有预训练模型(大模型)语言训练能力的公司：**其文心大模型和 Open AI 的 GPT 模型类似，在 2019 年就已经推出，并且已经迭代了多代，从单一的自然语言理解延伸到多模态，包括视觉、文档、文图、语音等多模态多功能，因此“文心一言”所基于的 ERNIE 系列模型也已经具备较强泛化能力和性能。以最新发布的 ERNIE 3.0 Zeus 为例，参数方面，该模型迭代于 ERNIE 3.0，拥有千亿级参数。

图表 17 百度文心预训练模型(大模型)发展历程



资料来源：经济观察报，华西证券研究所

**国产 ChatGPT 文心一言功能强大：**预训练模型方面，百度凭借海量的知识沉淀和丰富的场景推出文心大模型，根据 IDC 数据，其拥有 5500 亿条知识，且已经应用于百度搜索、信息流、智能驾驶、百度地图、小度等产品，此外，跨模态方面，百度已有文能生成、文档智能、地理-语言、视觉-语言、语音-语言等模型架构，根据 IDC 数据，目前已有近百万开发者使用文心大模型。



图表 18 百度文心大模型全景图



资料来源：IDC，华西证券研究所

**百度文心模型覆盖了多方向：**1、**自然语言处理**，具备语言理解能力与小说、摘要、文案创意、歌词、诗歌等能力，目前已在机器阅读理解、文本分类、语义相似度计算等 60 多项任务中应用；2、**机器视觉**，已经能够应用于各行各业的文档、卡证、票据等图像文字识别和结构化理解；3、**跨模态**，模型可应用在 AI 作图、地理-语言、视觉-语言、语音-语言应用；4、**其他任务**，例如对话、生成文本、自然语言变成理解、化合物分子、蛋白分子的预训练模型、医药研发建模等多应用场景。

**国产 ChatGPT 生态逐渐繁荣：**目前已有诸多厂商开展合作，我们认为其商业模式同样为 API 接口收费，属于 SAAS 商业模式，合作厂商分别覆盖科技、教育、工业、



**媒体、金融等诸多产业**，代表公司有宇信科技、汉得信息、金蝶、航天宏图、致远互联、东软集团等。

图表 19 部分国产 ChatGPT 文心一言合作公司

公司/机构	合作情况
宇信科技	在金融业务场景的率先应用，并围绕技术创新、场景孵化、生态建设等多方面展开更深入的合作，助力金融行业的智能化转型升级。
汉得信息	通过融合百度领先的智能对话技术成果，双方将携手推进B端人工智能深入应用，致力于帮助客户更彻底地实现数字化转型。
金蝶软件	把百度领先的智能对话技术成果应用在ERP领域，实现更高效的信息获取、信息整合、决策分析、数据洞察，为企业提供更可靠、更敏捷、更智能、更开放的服务。
航天宏图	航天宏图将把百度领先的智能对话技术成果应用在卫星遥感领域。此举标志着航天宏图优先获得领先AI技术的加持，也标志着对话式语言模型技术在PIE-Engine时空遥感云平台及各行业卫星应用场景中的首次着陆。
东软集团	在产品研发、标准制定等多个领域深化合作，通过技术共享、培训赋能、联合营销等方式，强化竞争力，为用户提供全场景人工智能解决方案及服务，同时依托智慧互联、创新互联，引领IT产业变革与升级。
致远互联	标志着致远互联优先获得领先AI技术的加持，也标志着对话式语言模型技术在国内协同管理领域的首次着陆。
明源云	通过技术共享、培训赋能、联合营销等方式，强化竞争力，为用户打造全场景数字营销服务人工智能解决方案及服务
软通动力	优先内测试用文心一言，集成文心一言的技术能力，与百度在产品研发、标准制定等多个领域展开深度合作，通过技术共享、培训赋能、联合营销等方式，共同为用户打造全场景的人工智能解决方案及服务，引领千行百业变革与升级。
数势科技	将百度领先的智能对话技术成果应用在企业经营分析和智能营销领域，打造智能决策全新产品体系，通过更先进的人工智能技术，让客户使用产品时更便捷、更高效。
奇墨科技	应用在IT质量管理领域，并结合奇墨ITQM SaaS软件打造IT知识问答、IT质量管理专家、IT智能运维专家等场景智能服务，赋能企业低成本地优化IT性能、安全、成本。
AIRLOOK	与百度在产品研发、标准制定等多个领域展开深度合作，通过技术共享、培训赋能、联合营销等方式，进一步探索元宇宙创新场景的应用。
凡拓数创	与百度在产品研发、标准制定等多个领域展开深度合作，在百度技术团队协助下，打造联合解决方案，通过技术共享、培训赋能、联合营销等方式，强化竞争力，为用户打造全场景数智展馆人工智能解决方案及服务
集度汽车	在交互能力方面，集度将融合百度文心一言的全面能力，打造全球首个针对智能汽车场景的大模型人工智能交互体验
每日经济新闻	把百度领先的智能对话技术成果应用在智能媒体领域。此举标志着每经优先获得领先AI技术的加持，也标志着对话式语言模型技术在国内财经媒体内容智能生产场景的首次着陆。
人民邮电报社等161家媒体	应用百度领先的智能对话技术成果，打造内容生态人工智能全系产品及服务，树立行业媒体智能化新标杆。
悦商集团	将百度领先的AI技术应用于YCloud系统，用户可通过几个词的组合或者一组长文字快速精准的实现文转图，可以进一步规范微商业发展，增加行业竞争力。
邮储银行等银行	在金融行业开展类ChatGPT技术应用试点，结合行业知识进行微调（Finetune）和交互式训练，在智能客服、数字员工、虚拟营业厅等场景进行应用，进一步提升客户体验。跨模态深度语义理解与生成能力，也有助于邮储银行探索新型的信息生产、获取方式，向客户提供更及时、更准确、更个性化的金融资讯服务。
海管家	国内对话式语言模型技术在国际物流场景的首次着陆。
少林寺	将把百度领先的智能对话技术成果应用在内容生态、智慧寺院、文化保护与传承等领域，为用户打造全场景内容生态人工智能解决方案及服务。
长虹	将把百度领先的智能对话技术成果应用在BK视像等领域。此举标志着长虹8K电视优先获得领先AI技术的加持，也意味着对话式语言模型技术在国内智能感知交互场景的进一步探索。
海信视像	将百度领先的智能对话技术成果应用在海信电视等显示产品的智能内容生产领域，为消费者带来以场景为中心的智慧生活体验。
TCL等消费电子	与百度携手引领智能终端技术升级新浪潮，是首家接入百度文心一言的全球化家电企业。TCL实业将通过这一新型交互式智能技术为消费者带来全新升级的智能化产品和服务体验。
值得买等消费内容类	在产品研发、标准制定等多个领域展开深度合作，在百度技术团队协助下，打造联合解决方案，通过技术共享、培训赋能、联合营销等方式，强化竞争力，为用户打造全场景消费内容人工智能解决方案及服务，同时依托智慧互联、创新互联，引领消费内容产业变革与升级。
长城汽车	通过百度Apollo融合文心一言的全面能力，携手打造基于智能汽车场景的大模型人工智能交互，开启人工智能时代的汽车全新体验。
红旗等汽车类	打造联合解决方案，通过技术共享、培训赋能、联合营销等方式，在AIGC内容共创领域发力，并赋能红旗元宇宙，打造更为完善的用户共创体系。
高途	生成教学设计、课程等内容，降低教育机构内容制作的成本。启用AI对话在课程或者产品销售时的客服场景中能够降低人力成本。
豆语星辰等教育类	将把百度领先的智能对话技术成果应用在教育科技化升级领域。此举标志着豆语星辰优先获得领先AI技术的加持，也标志着对话式语言模型技术在国内教育科技化场景的首次着陆。

资料来源：公开资料整理，华西证券研究所

## 2. 投资建议：梳理 AIGC 相关受益厂商

我们认为 AIGC 的出世会产生革命性的影响，同时有望赋能千行百业。我们梳理了三条路径图，积极的推荐以下三条投资主线：

1) 具备算力基础的厂商，受益标的为**寒武纪、商汤、海光信息、浪潮信息、中科曙光、景嘉微、联想集团、紫光股份、龙芯中科**；

2) 具备 AI 算法商业落地的厂商，重点推荐**科大讯飞、拓尔思**，其他受益标的为：**汉王科技、海天瑞声、云从科技**；

3) AIGC 相关技术储备的应用厂商，受益标的为：**百度、同花顺、三六零、金山办公**。

### 3. 风险提示

- 1、核心技术水平升级不及预期的风险；
- 2、AI 伦理风险；
- 3、政策推进不及预期的风险；
- 4、中美贸易摩擦升级的风险。

## 分析师与研究助理简介

刘泽晶（首席分析师）：2014-2015年新财富计算机行业团队第三、第五名，水晶球第三名，10年证券从业经验。

## 分析师承诺

作者具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，保证报告所采用的数据均来自合规渠道，分析逻辑基于作者的职业理解，通过合理判断并得出结论，力求客观、公正，结论不受任何第三方的授意、影响，特此声明。

## 评级说明

公司评级标准	投资 评级	说明
以报告发布日后的6个月内公司股价相对上证指数的涨跌幅为基准。	买入	分析师预测在此期间股价相对强于上证指数达到或超过15%
	增持	分析师预测在此期间股价相对强于上证指数在5%—15%之间
	中性	分析师预测在此期间股价相对上证指数在-5%—5%之间
	减持	分析师预测在此期间股价相对弱于上证指数5%—15%之间
	卖出	分析师预测在此期间股价相对弱于上证指数达到或超过15%
行业评级标准		
以报告发布日后的6个月内行业指数的涨跌幅为基准。	推荐	分析师预测在此期间行业指数相对强于上证指数达到或超过10%
	中性	分析师预测在此期间行业指数相对上证指数在-10%—10%之间
	回避	分析师预测在此期间行业指数相对弱于上证指数达到或超过10%

## 华西证券研究所：

地址：北京市西城区太平桥大街丰汇园11号丰汇时代大厦南座5层

网址：<http://www.hx168.com.cn/hxzq/hxindex.html>

## 华西证券免责声明

华西证券股份有限公司（以下简称“本公司”）具备证券投资咨询业务资格。本报告仅供本公司签约客户使用。本公司不会因接收人收到或者经由其他渠道转发收到本报告而直接视其为本公司客户。

本报告基于本公司研究所及其研究人员认为的已经公开的资料或者研究人员的实地调研资料，但本公司对该等信息的准确性、完整性或可靠性不作任何保证。本报告所载资料、意见以及推测仅于本报告发布当日的判断，且这种判断受到研究方法、研究依据等多方面的制约。在不同时期，本公司可发出与本报告所载资料、意见及预测不一致的报告。本公司不保证本报告所含信息始终保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者需自行关注相应更新或修改。

在任何情况下，本报告仅提供给签约客户参考使用，任何信息或所表述的意见绝不构成对任何人的投资建议。市场有风险，投资需谨慎。投资者不应将本报告视为做出投资决策的惟一参考因素，亦不应认为本报告可以取代自己的判断。在任何情况下，本报告均未考虑到个别客户的特殊投资目标、财务状况或需求，不能作为客户进行客户买卖、认购证券或者其他金融工具的保证或邀请。在任何情况下，本公司、本公司员工或者其他关联方均不承诺投资者一定获利，不与投资者分享投资收益，也不对任何人因使用本报告而导致的任何可能损失负有任何责任。投资者因使用本公司研究报告做出的任何投资决策均是独立行为，与本公司、本公司员工及其他关联方无关。

本公司建立起信息隔离墙制度、跨墙制度来规范管理跨部门、跨关联机构之间的信息流动。务请投资者注意，在法律许可的前提下，本公司及其所属关联机构可能会持有报告中提到的公司所发行的证券或期权并进行证券或期权交易，也可能为这些公司提供或者争取提供投资银行、财务顾问或者金融产品等相关服务。在法律许可的前提下，本公司的董事、高级职员或员工可能担任本报告所提到的公司的董事。

所有报告版权均归本公司所有。未经本公司事先书面授权，任何机构或个人不得以任何形式复制、转发或公开传播本报告的全部或部分内容，如需引用、刊发或转载本报告，需注明出处为华西证券研究所，且不得对本报告进行任何有悖原意的引用、删节和修改。