



人工智能通识教程

（农林院校版）

<https://ai4ag.github.io>





第八章 大模型



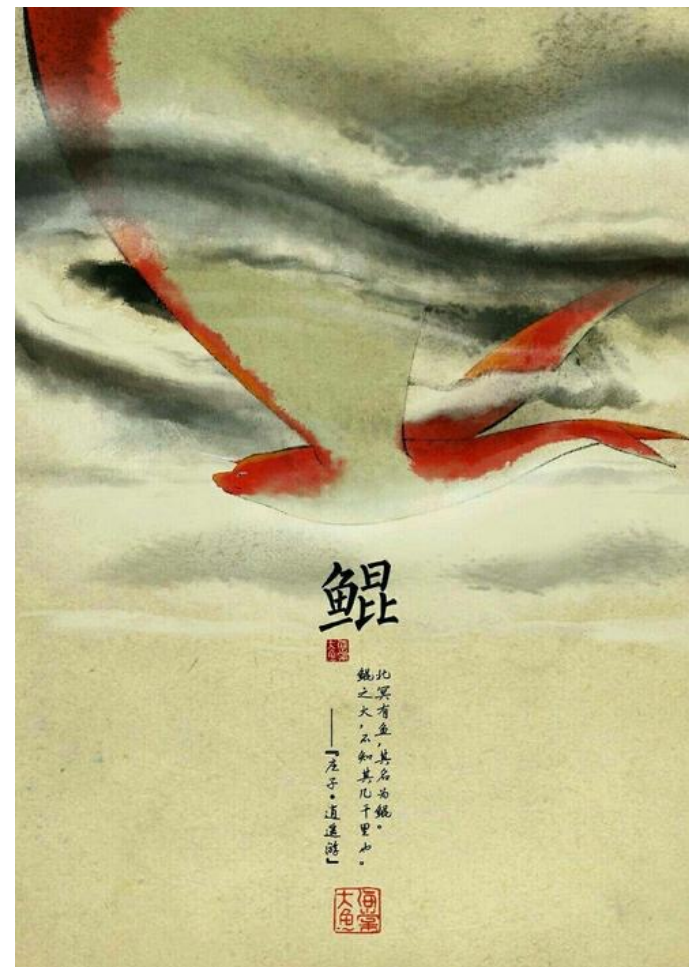
引子

北冥有鱼，其名为鲲。鲲之大，不知其几千里也。

——庄子《逍遥游》

或如鲲之巨，大而不可量度，当人工智能进入大模型时代，其所蕴藏的潜力与影响也难以估量。

在行业发展与大模型崛起相互交织的时代背景下，如何运用以大模型为代表的人工智能前沿技术，为各行业发展注入新的活力，为其科研、生产与管理方式带来新的变革，已成为当前各行业领域研究与实践的重要议题。



本节教学内容

- 介绍大模型章节授课内容

- 大模型概述

- 理解大模型的基本概念和发展历程。
- 了解国内外大模型。

- 大模型基本原理

- 理解下一个token预测、注意力机制和词嵌入。
- 初步了解大模型训练中的强化学习。



大模型章节授课内容



大模型章节授课内容

- 大模型概述
- 大模型基本原理
- 大模型应用开发
- 大模型安全与伦理



大模型概述





(一) 大模型简介



大模型简介

近年来，人工智能领域迎来了以大模型为核心的爆发式发展。这些模型的出现不仅为各行各业带来了革命性的应用前景，也引发了人们对AI未来发展的广泛讨论和思考。代表性成果包括OpenAI推出的ChatGPT和SORA，以及深度求索推出的DeepSeek。



大模型简介

□ 萌芽期（1950 年-2005年）

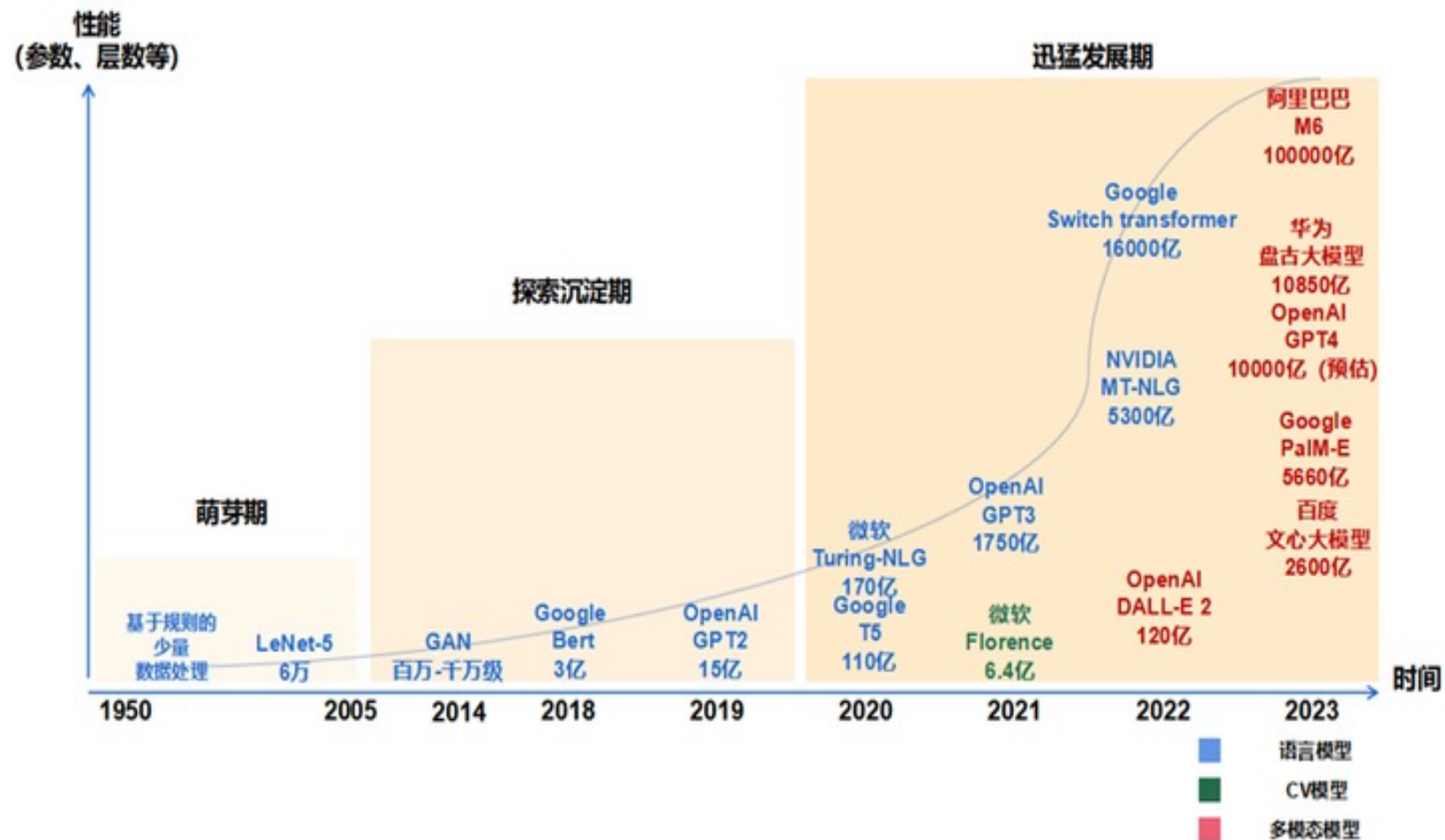
以CNN为代表的传统神经网络模型阶段

□ 探索沉淀期（2006年-2019年）

以Transformer为代表的全新神经网络模型阶段

□ 迅猛发展期（2020至今）

以GPT为代表的预训练大模型阶段



大模型简介

●GPT系列

通过参数规模扩展实现性能跃升

□ GPT-1（2018年）

1.17亿参数，基于 Transformer 架构的解码器设计，首次实现预训练+微调范式。

□ GPT-2（2019年）

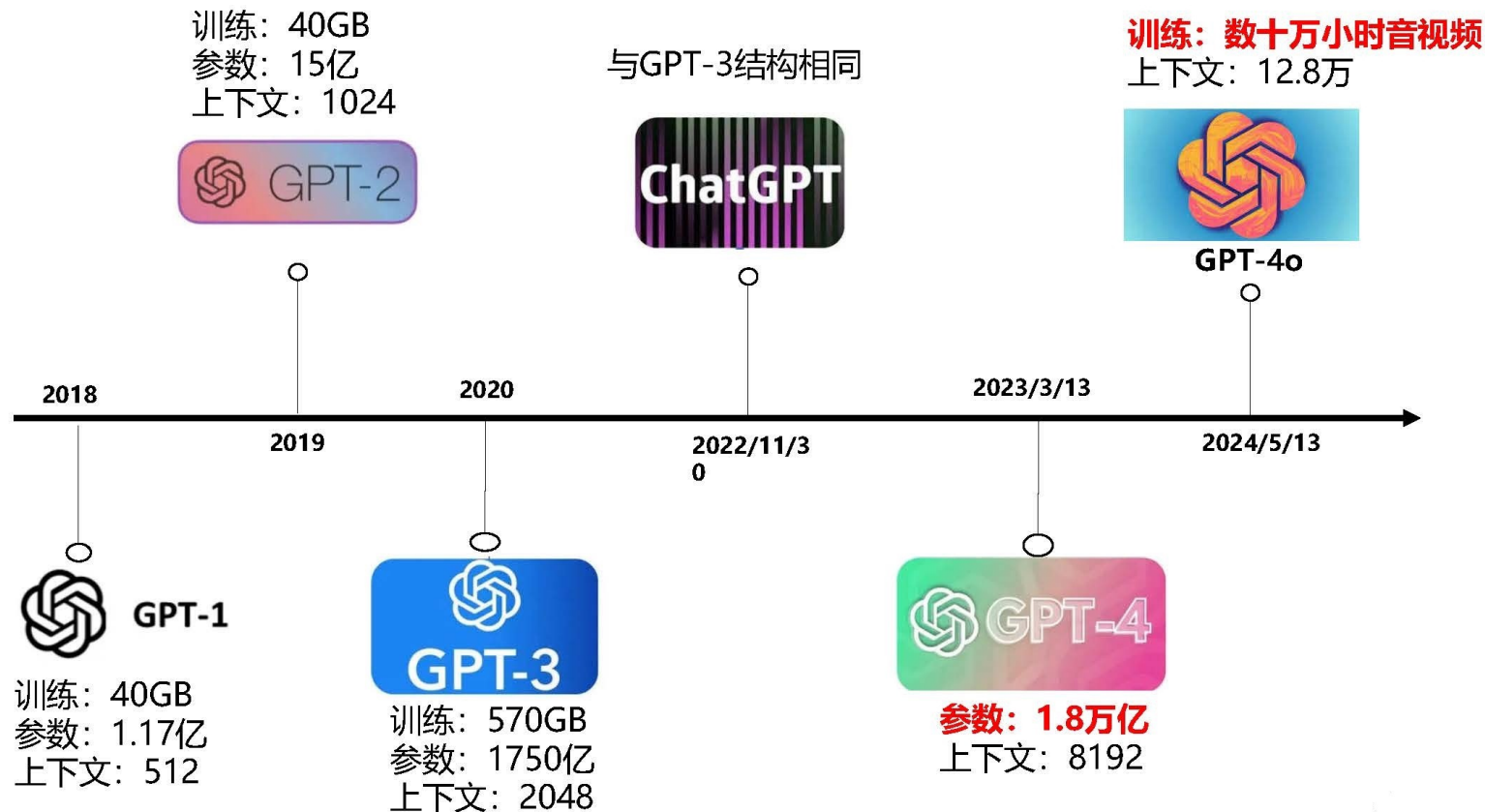
参数规模扩大至15亿，支持零样本学习，无需额外微调即可完成新任务。

□ GPT-3（2020年）

参数达1750亿，引入少样本学习，支持复杂任务处理。

□ GPT-4（2023年）

参数规模约1.8万亿，采用混合专家模型（MoE）提升推理能力。



大模型简介

●DeepSeek快速出圈

■ 2024年12月26日，DeepSeek推出对标OpenAI GPT-4o的 语言模型DeepSeek V3，随后在美国AI 行业内部引起轰动。

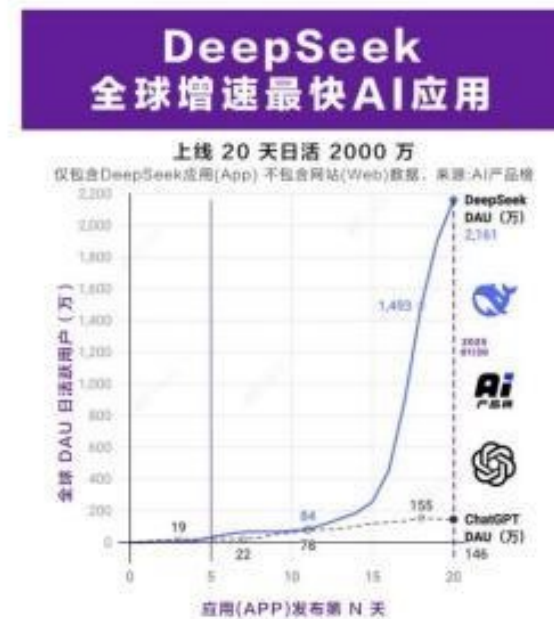
■ 2025年1月20日，DeepSeek发布对标OpenAI o1 的DeepSeek R1大语言模型，并于1月24日引起美国投资界KOL关注。

■ 2025年1月26日，关于DeepSeek颠覆了大模型的商业模式（堆算力、拼资本），引发英伟达股价大跌，DeepSeek首先在美国出圈，引发国际社会讨论。

■ 2025年1月底（春节前后），DeepSeek在中国出圈，并上升到中美竞争高度，同时纷纷接入DeepSeek，DeepSeek成为AI和大模型的代名词。



2025年1月20日下午，中共中央政治局常委、国务院总理李强主持召开专家、企业家和教科文卫体等领域代表座谈会，听取对《政府工作报告（征求意见稿）》的意见建议。DeepSeek公司创始人梁文峰作为企业家代表之一参加了此次座谈会。



大模型简介

●DeepSeek R1 对大模型行业的重大影响

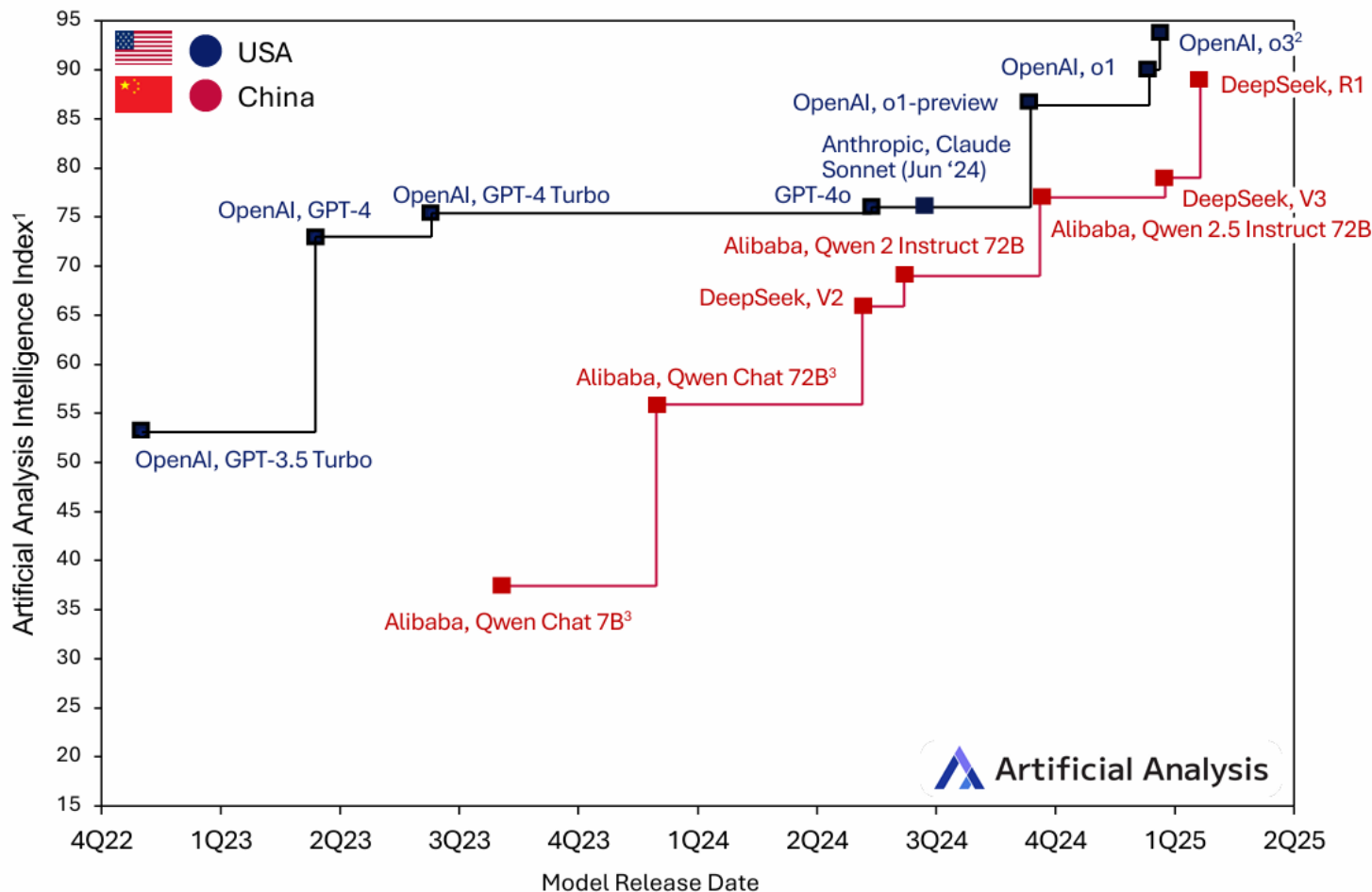
■打破垄断：DeepSeek-R1以低成本和开源特性打破以往头部企业巨头割据局面。

■价格下调：DeepSeek-R1的API定价仅为行业均价的1/10,推动了中小型企业低成本接入AI，对行业产生了积极影响。

■推动创新：DeepSeek-R1促使行业开始从“唯规模论”转向更加注重“性价比”和“高效能”方向。

DeepSeek R1模型的能力突破

US & China: Frontier Language Model Intelligence, Over Time¹





(二) 国内外大模型



国内外大模型

月份	国内 (TOP1)	国内 (TOP2)	国内 (TOP3)	海外Top3
2024年12月	DeepSeek-V3、SenseChat 5.5-latest、360gpt2-01	Doubao-pro-32k-241215、NebulaCoder-v5、Qwen-max-latest	Step-2-16k、GLM-4-Plus	01、01-preview、ChatGPT-4o-latest
2024年10月	GLM-4-Plus、SenseChat 5.5、AndesGPT 2.0、Qwen2.5-72B-Instruct (并列)	Hunyuan-Turbo、360gpt2-pro、Step-2、DeepSeek-V2.5、Doubao-pro	Baichuan4、山海大模型4.0、TeleChat2-Large	01-preview、Claude 3.5 Sonnet (20241022)、ChatGPT-4o-latest
2024年8月	Hunyuan-Turbo-Preview	AndesGPT-2.0、DeepSeek-V2-0628	Qwen2-72B-Instruct、SenseChat 5.5、Doubao_pro_preview	ChatGPT-4o-latest、GPT-4-Turbo-2024-04-09、Claude 3.5 Sonnet
2024年6月	Qwen2-72B	GLM-4-0520、DeepSeek-V2、SenseChat5.0	AndesGPT	GPT-4o、Claude-3.5-Sonnet-200k、GPT-4-Turbo-0409
2024年4月	Baichuan3	GLM-4、通义千问2.1	腾讯Hunyuan-pro、文心一言4.0	GPT-4-Turbo-0125、GPT-4-Turbo-0409、GPT-4
2024年2月	文心一言4.0	GLM-4	通义千问2.1	GPT4-Turbo-0125、GPT4 (网页)、Claude2
2023年12月	文心一言4.0	通义千问2.0	AndesGPT	GPT4-Turbo、GPT4(网页)、Claude2

国内大模型发展

文本

通用闭源	文心一言	通义千问	腾讯混元	商汤日日新 sensenova	BlueLM	360智脑	天工	MiLM	中科闻歌	紫东太初	润舟科技 langboat
	字节豆包	Kimi.ai	百川智能 BAICHUAN AI	MINIMAX	盘古大模型	云从科技 CLOUDWALK	DeepSeek	阶跃星辰	OpenBayes	Transn传神	
	智谱·AI	云知声	山海	零一万物		OPPO	AndesGPT	ZTE中兴	讯飞星火	天翼AI	Scietrain 西湖心辰
通用开源	Qwen2.5	deepseek coder	GLM-4	面壁小模型 MiniCPM	Yi	Baichuan2	RWKV-LM	TeleChat2-35B	书生·浦语		
推理	QWQ-32B-Preview	DeepSeek-R1-Lite	InternThinker	K0-math	Skywork o1	360gpt2-o1	LLaVA-CoT				

多模态

实时交互	星火极速	智谱清言	海螺AI	豆包	文小言	通义APP	日日新 sensenova	Kimi	语音合成/声音复刻	
文生视频	可灵AI	即梦AI	清影	Vidu	PixVerse	海螺AI	HiDream.ai	通义万相	Doubao-语音合成	百度TTS
视觉理解	腾讯混元	阶跃星辰	Qwen2-VL	Doubao-vision	SenseChat-Vision	海螺AI	GLM-4v	书生·万象	讯飞语音合成	CosyVoice
文生图	即梦AI	混元-DiT	快手可图	CogView	讯飞星火	meitu	通义万相	文心一格	Fish Audio	speech-01

行业

部分领域	医疗	汽车	教育	金融	工业	更多行业
	百度灵医	理想 MindGPT	MathGPT	蚂蚁金融大模型	奇智孔明 AInno-15B	营销: 探迹 SalesGPT
	医联 MedGPT	DriveGPT	作业帮	妙想金融大模型	华为盘古工业大模型	文化: 阅文集团 妙笔大模型
	百川AI全科医生	极氪 Kr大模型	子曰	轩辕大模型	SmartMore SMoore LrMo	法律: Chat Law
	讯飞晓医	易车大模型		HithinkGPT	羚羊工业大模型	AI4S: DP 深势分子大模型



大模型基本原理





(一) 下一个token预测



下一个token预测

●什么是token?

哥 憚 及 □ 达 晶

Tokenization

に 哥 ぬ に 憚 及 ぬ に □ ぬ に 达 ぬ に 晶 ぬ

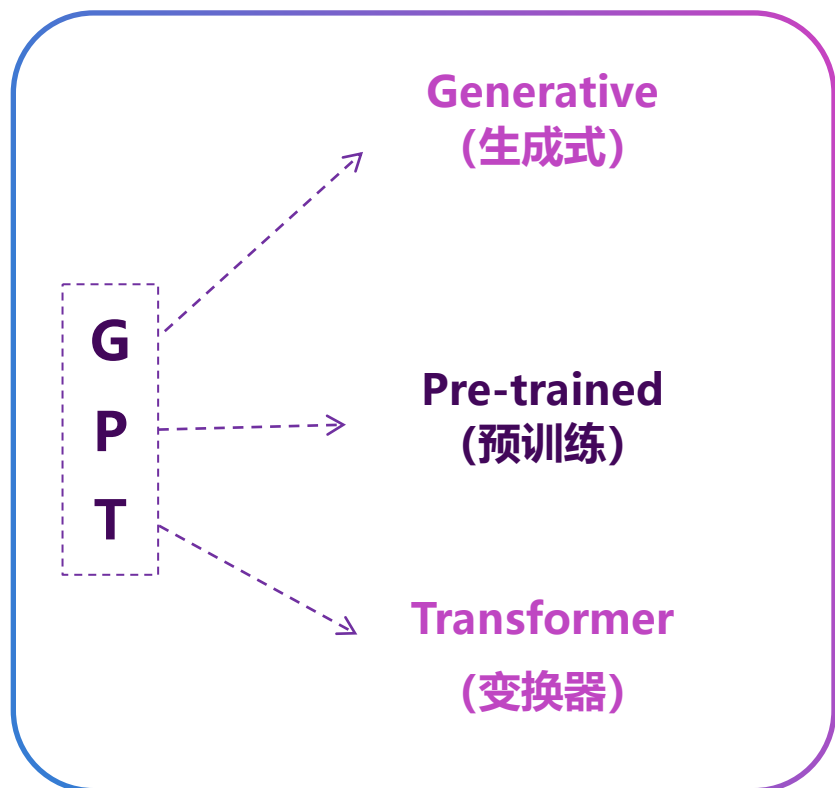
棕 褚 觥 济 悸 塘 宪 别 tokens □ 茗 猴 璇 □ □

token

在NLP中，token是文本的**基本单位**。token在大语言模型中的作用至关重要，因为它们 是模型理解和处理**自然语言的基础单位**。
token 可以是一个完整的词、词的一部分(子词)，甚至是单个字符。**Tokenization**，即分词过程，是将文本转换为token的**关键步骤**。

下一个token预测

● NTP (Next Token Prediction)



1. 收到提示词

示例: “今天天气不错, 我决定”

2. 将输入拆分为token (标记)

[“今”, “天”, “天”, “气”, “不”, “错”, “,”, “我”, “决”, “定”]

3. 采用Transformer架构处理token

- 理解token之间的关系
- 识别提示词的整体含义

4. 基于上下文预测下一个token

- 为可能的单词分配概率分数
- 示例: { “去” :0.7, “停”:0.2, “站”:0.1 }

5. 根据概率分数选择标记

示例: “去”

概率预测+文字接龙

重复步骤4和步骤5直到形成完整的句子

示例: 今天天气不错, 我决定去公园



(二) 注意力机制



注意力机制

大模型 在“续写”时，不会对所有词都一视同仁，它会“划重点”！

这项能力叫做 “注意力机制”：把注意力更多地放在重要的词语上。

清晨，阳光暖暖地洒在街道两旁，空气中弥漫着淡淡的花香。人们匆忙赶路，却不忘抬头看看明朗的天空。微风拂过脸庞，带来丝丝清凉，让人忍不住感叹：今天天气真不错啊！

注意力机制

通过注意力机制，大模型能够理解句子中词语之间的关系，从而更好地“续写”出通顺的内容。

清晨，**阳光暖暖地洒**在**街道**两旁，空气中弥漫着淡淡的花香。**人们**匆忙赶路，却不忘抬头看看**明朗的**天空。**微风**拂过脸庞，带来丝丝清凉，让人忍不住**感叹**：**今天天气**真不错啊！

清晨，**阳光**暖暖地洒在街道两旁，空气中弥漫着淡淡的花香。**人们**匆忙赶路，却不忘抬头看看明朗的天空。**微风拂过脸庞**，带来**丝丝清凉**，让人忍不住**感叹**：**今天**天气真不错啊！

清晨，**阳光**暖暖地洒在街道两旁，**空气中弥漫着淡淡的花香**。**人们**匆忙赶路，却不忘抬头看看明朗的天空。**微风**拂过脸庞，带来丝丝清凉，让人忍不住**感叹**：**今天**天气真不错啊！

清晨，**阳光**暖暖地洒在街道两旁，空气中弥漫着淡淡的花香。**人们**匆忙赶路，却不忘抬头看看明朗的天空。**微风**拂过脸庞，带来丝丝清凉，让人忍不住**感叹**：**今天天气真不错啊！**

清晨，**阳光**暖暖地洒在街道两旁，空气中弥漫着淡淡的花香。**人们匆忙赶路**，却不忘**抬头看看**明朗的**天空**。**微风**拂过脸庞，带来丝丝清凉，让人忍不住**感叹**：**今天**天气真不错啊！

清晨，**阳光**暖暖地洒在街道两旁，空气中弥漫着淡淡的花香。**人们**匆忙赶路，却不忘抬头看看明朗的天空。**微风**拂过脸庞，带来丝丝清凉，让人忍不住**感叹**：**今天天气真不错**啊！

但问题又来了：大模型是怎样去辨别，哪些内容彼此更相似呢？

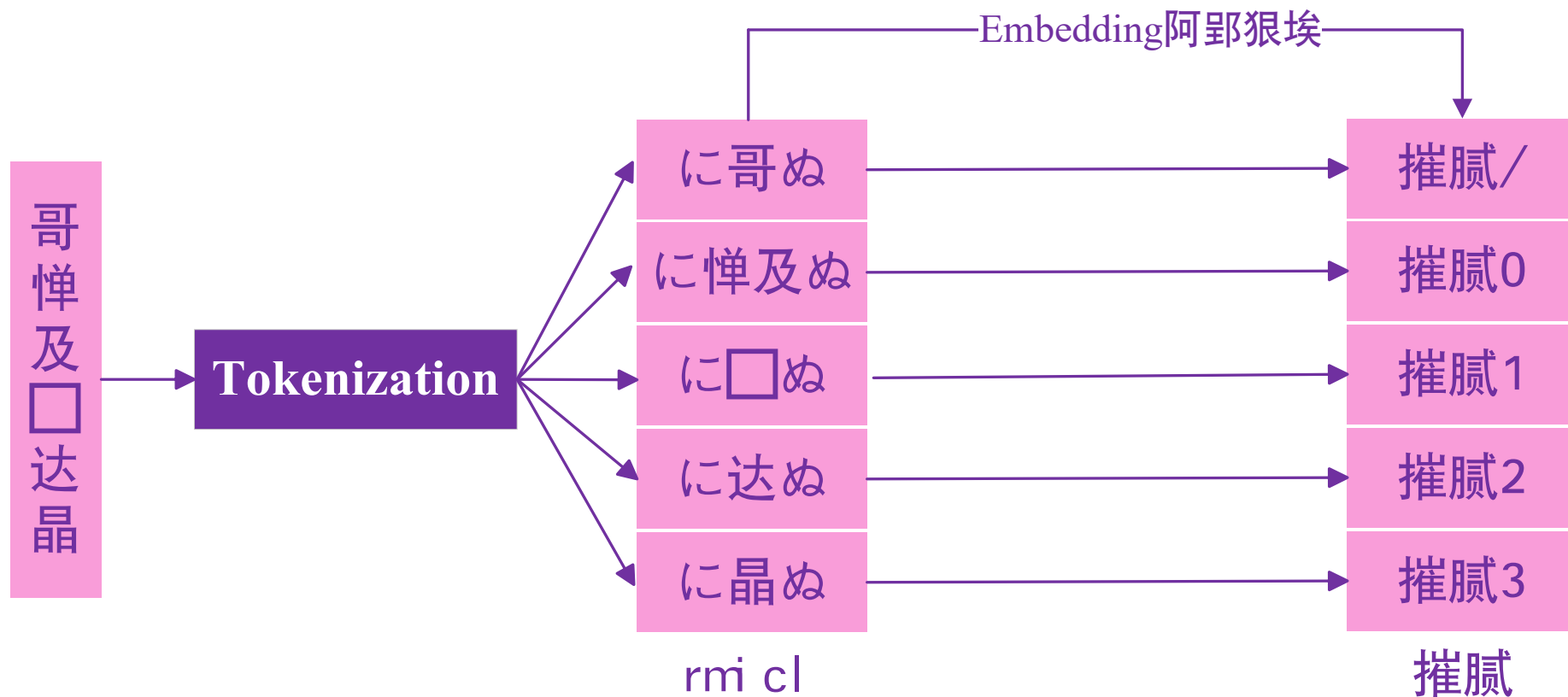


(三) 词嵌入



词嵌入

我们能看懂文字，但大模型看见的却是一串串数字。因为大模型会先把每个词变成一个数字编号（向量），这个过程叫做“词嵌入” (Word Embedding)。



词嵌入

举个例子：比如“猫”和“狗”这两个词，我们知道它们都是动物，也都很可爱，但在叫声上有所不同。而大模型却不懂这些，它看到的一切信息，都是被从不同维度标记的。假设标记数字的取值范围是 -1 到 1：

对于“猫”来说：“喵喵叫”这个特征可能对应的数字是 0.99，因为猫确实会喵喵叫；“可爱”这个特征对应的数字可能是 0.8，因为猫通常被认为是可爱的。

对于“狗”来说：“喵喵叫”这个特征对应的数字可能只有 0.3，因为它不太常发出这种叫声；但“可爱”这个特征也对应 0.8，和猫一样。

这像这样，大模型在一通数值比较后，就知道“猫”和“狗”一样可爱，但叫声不同。

于是，通过大量的“词嵌入”，大模型就能“理解”词语的含义，然后开始干活了。



（四）强化学习



强化学习

RLHF (Reinforcement Learning from Human Feedback): 即使用强化学习的方法, 利用人类反馈信号直接优化语言模型。

■ 多种策略产生样本并收集人类反馈

■ 训练奖励模型

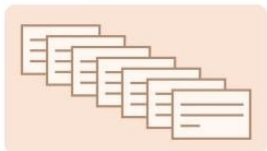
■ 训练强化学习策略, 微调语言模型

1 Collect human feedback

A Reddit post is sampled from the Reddit TL;DR dataset.



Various policies are used to sample a set of summaries.



Two summaries are selected for evaluation.



A human judges which is a better summary of the post.



"j is better than k"

2 Train reward model

One post with two summaries judged by a human are fed to the reward model.



The reward model calculates a reward r for each summary.



r_j

r_k

The loss is calculated based on the rewards and human label, and is used to update the reward model.

$$\text{loss} = \log(\sigma(r_j - r_k))$$

"j is better than k"

3 Train policy with PPO

A new post is sampled from the dataset.



The policy π generates a summary for the post.



The reward model calculates a reward for the summary.



The reward is used to update the policy via PPO.

r

《Learning to summarize from human feedback》

思考

- 回顾大模型的发展历程，其参数规模有着什么程度的扩展，支持其实现性能跃升。
- 在2024年底，我国的DeepSeek大模型在多项评测中表现出色，其性能已经可以与国际顶级模型相媲美。简述DeepSeek R1 对大模型行业的重大影响。
- 大模型本质上是“概率预测+文字接龙”，其续写机制称为下一token预测（Next Token Prediction, NTP），请问什么是token？以及举例简述NTP的过程。

谢谢观看

- 本课程所引用的一些素材为主讲老师多年的教学积累，来源于多种媒体及同事和同行的交流，难以一一注明出处，特此说明并表示感谢！