

新学习范式： 自监督，迁移学习

(理工类)

东南大学 计算机科学与工程学院



课程要点

1

自监督学习和迁移学习基本概念

2

自监督学习和迁移学习代表性技术

3

自监督学习和迁移学习应用

课程安排

01

自监督学习概念、代表性技术、应用

02

迁移学习概念、代表性技术、应用

目录

1

为什么有自监督学习?

- 自监督学习的优势

2

自监督学习的基本概念

- 自监督学习的定义

3

自监督学习的代表性技术

- 自监督学习的原理

4

自监督学习的应用

- 自监督学习的应用与前景

5

迁移学习的历史背景

- 迁移学习最早背景

6

迁移学习的基本概念

- 迁移学习的问题求解

7

迁移学习的实际应用

- 迁移学习在不同学习场景中的应用

8

大语言模型为例

为什么有自监督学习?



为什么需要自监督学习?

有监督学习 (例如句子情感分类)

这听起来很不错!
我并不喜欢这个电影。

标签: 积极的
标签: 消极的

无监督学习

自动根据句子特征聚类, 可能会把积极的句子归为一类, 但是不会有显式的标签。

自监督学习

这[MASK]起来很不错!
我并不喜[MASK]这个电影。

遮住一个词并以之为标签, 根据上下文预测这个词

为什么需要自监督学习?

有监督学习

有监督学习对大量标注数据的依赖，且标注的数据很昂贵

无监督学习

无监督学习在提取数据结构时缺乏具体指导信号的局限性

自监督学习

自监督学习通过预设任务生成伪标签，从**未标注数据中提取特征**，减轻对大规模标注数据的依赖并提升模型泛化能力

自监督学习的基本概念

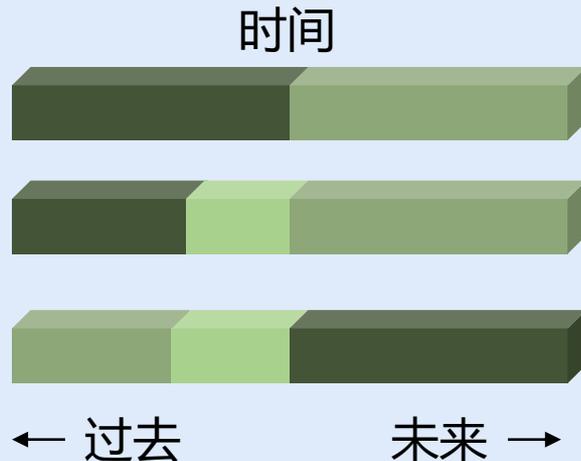


自监督学习的基本概念

在自监督学习中，模型通过设计**预设任务**从数据中提取有用的特征。这些预设任务通常要求模型根据部分可见的数据来预测或重建被故意隐藏的信息。这种方法**不需要外部标注数据**，使模型能够在没有人工干预的情况下，从大量未标注的数据中学习。

• 预测任意部分

- 从**过去**预测**未来**
- 从**最近的过去**预测**未来**
- 从**现在**预测**过去**



From LeCun

自监督学习的**核心任务**之一是假设输入中有一部分是未知的，然后通过模型来预测这一部分的内容。这种方法促使模型挖掘数据中的潜在模式和特征，提高其对未知场景的预测能力。

自监督学习的代表技术



自监督学习的代表技术

自监督学习算法

1

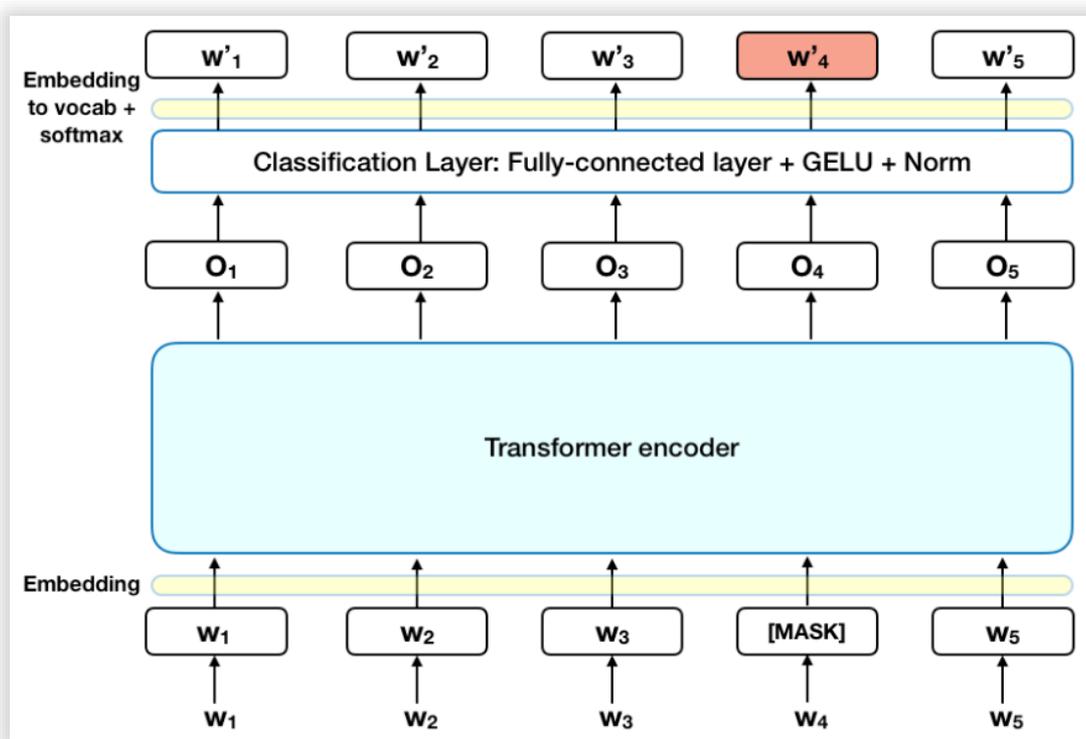
大语言模型

2

视觉模型

重构破坏的数据

大语言模型：完形填空



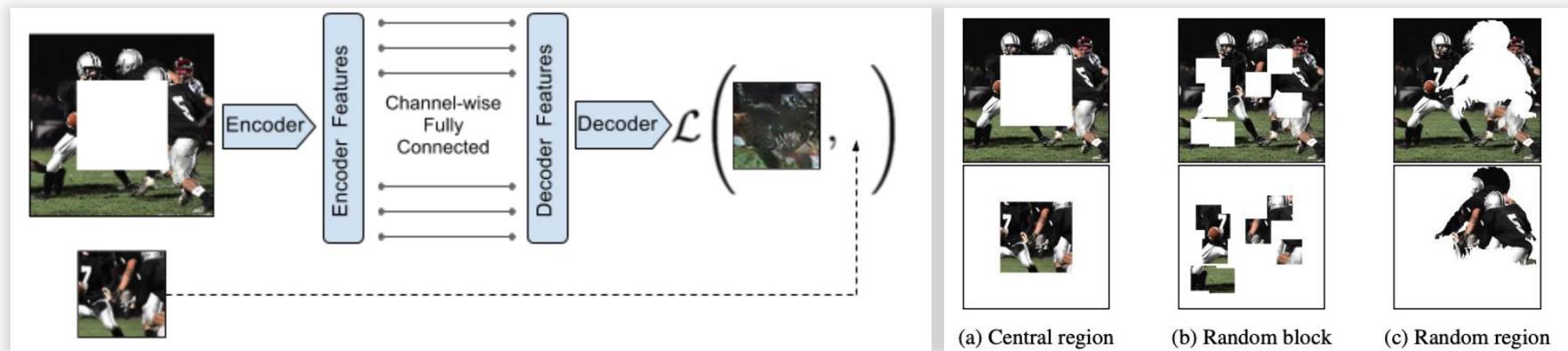
- 预测输入中15%的标记。
- 80%被替换为掩码标记
- 10%被替换为随机单词
- 10%保持不变

例子：

- 输入：The [mask] sat on the [mask].
- 输出：The cat sat on the mat.

重构破坏的数据

视觉模型：图像填充



从一张图片中移除一部分，然后将剩下的部分输入到一个学习器（编码器-解码器 encoder-decoder）架构的网络中，解码器的目标是重建原始图片。这种挖去图片部分的方法是多样的，可以挖去图片的中央区域、随机的块状区域，或者不规则的随机区域等。

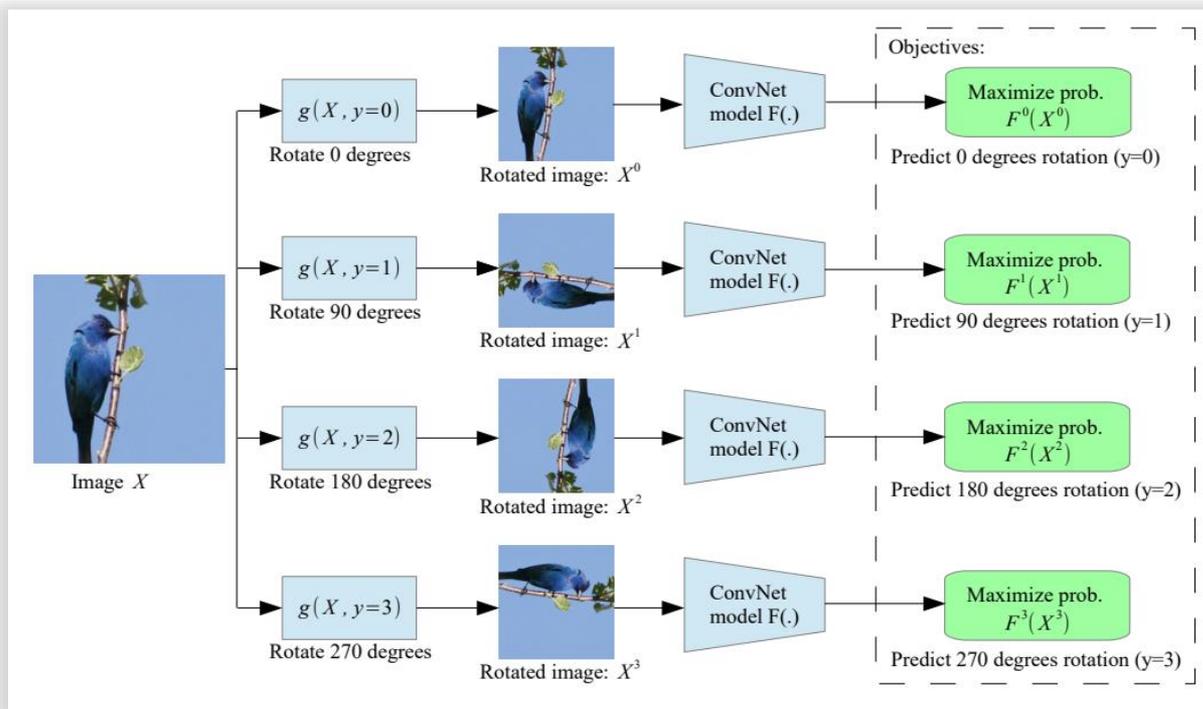
视觉常识

视觉模型：拼图游戏



利用**拼图游戏**来训练卷积神经网络的上下文理解能力，并学习对分类和检测有用的特征。左侧是原始图像中选定的拼图块（带有随机间隙），中间是随机打乱的拼图块，右侧是拼图的解决方案。

视觉模型：图像旋转

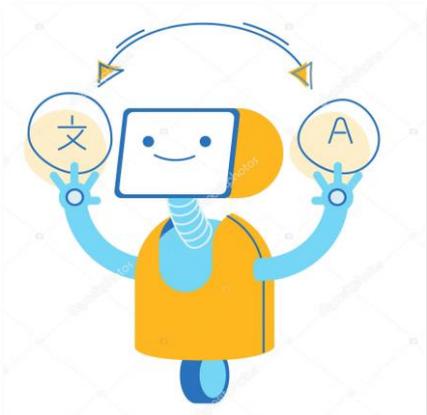


由于图片没有标签，因此模型的学习目标是给定一张图片后，模型需要预测这张图片被**旋转**了多少度，这样问题就简化为一个分类问题。

自监督学习的应用



自监督学习的应用



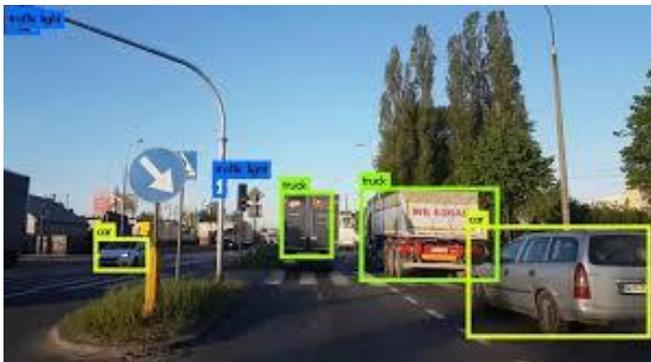
机器翻译



Chat-GPT



医疗图像分类



目标检测



图像分割

迁移学习的历史背景



迁移学习的历史背景

01 早期理论基础：认知科学与迁移

迁移学习的思想可以追溯到认知科学的早期研究。在20世纪初，心理学家如爱德华·桑代克 (Edward Thorndike) 和查尔斯·贾德 (Charles Judd) 研究了人类学习过程中的“学习迁移”现象。学习迁移是指一个人在某个任务上的学习可以影响另一个任务的学习过程。例如，学习骑自行车可能会有助于滑板等类似任务。这一早期研究奠定了迁移学习的基本思想，即知识和技能可以在不同但相关的任务之间迁移。

02 20世纪下半叶：人工智能的萌芽

在计算机科学和人工智能的早期，迁移学习并没有被系统地研究，但类似的想法在一些研究中有所体现。1970年代，随着人工智能的逐步发展，研究者开始探索如何在不同的任务中复用已有的知识。符号学习 (Symbolic Learning) 和专家系统 (Expert Systems) 是这一时期的主要研究方向，但这些系统的局限性使得迁移学习没有得到广泛应用。

迁移学习的历史背景

03 1990年代：迁移学习概念的正式提出

1990年代，迁移学习的概念逐渐被正式提出，特别是在强化学习（Reinforcement Learning）和多任务学习（Multi-task Learning）领域，相关研究开始逐步发展。**Reid Simmons** 等人（1992年）首次在机器人学习中提出了“迁移”概念，他们探索了如何将一个机器人在某个环境中学到的策略迁移到另一个环境。

04 21世纪：深度学习与迁移学习的崛起

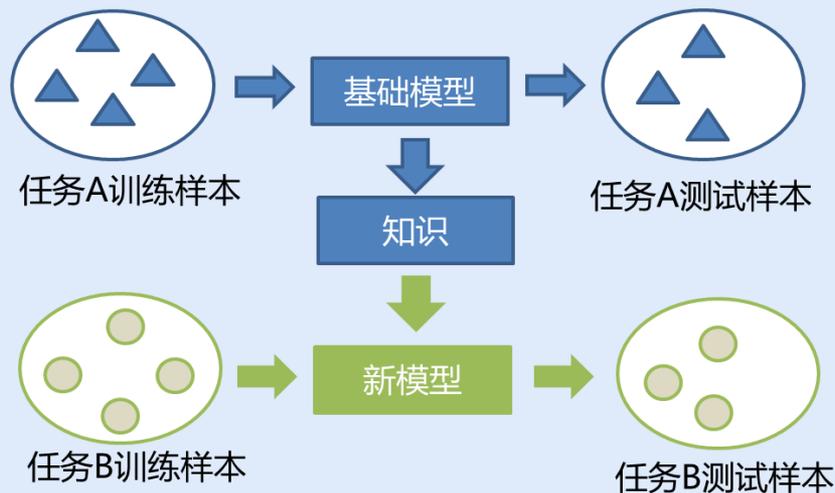
2000年以后，随着深度学习的兴起，迁移学习迎来了飞速发展。深度神经网络（DNN）在计算机视觉、自然语言处理等领域取得了巨大的成功，但深度学习的模型训练通常需要大量的计算资源和数据。因此，如何将预训练的模型应用到新的任务上成为一个核心问题。**2006年**，**Hinton**提出的深度置信网络（DBN）和深度自编码器为迁移学习提供了新的途径。**2014年**，随着AlexNet在ImageNet比赛中的成功，迁移学习在计算机视觉领域的应用开始迅速普及。利用在大型数据集上预训练的卷积神经网络（CNN），将其迁移到不同的视觉任务中。

迁移学习的基本概念



迁移学习的基本概念

- 迁移学习的基本概念是：**将一个模型在某个任务中学到的知识应用到另一个相关但不同的任务中**，以减少训练时间或提升性能。理解为把一个任务的学习成果迁移到新的任务上。
- 因此，在迁移学习的过程中，需要利用已知的任务训练一个基准的模型，利用基准模型学习到的先验知识来对于新的任务进行学习。**这个模型学到的“经验”去帮助新的任务更快、更好地学习。**



- 比如各位同学大一学习基础课，就是一个预训练过程，后面再学专业课，就是一个迁移学习过程。

迁移学习的概念

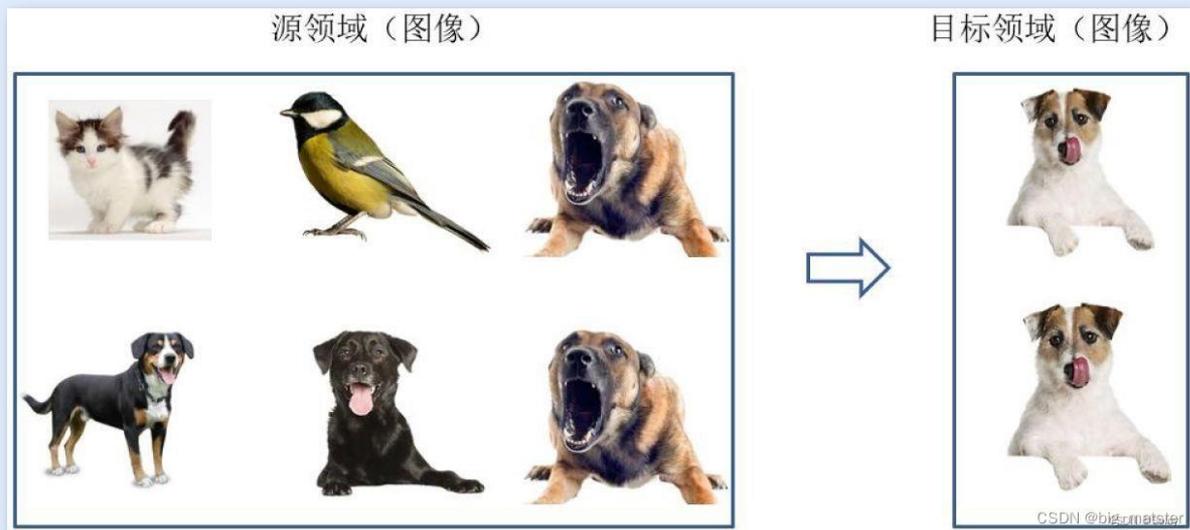
举一个例子，就是想象你学会了**骑自行车**，这就是你的“基准模型”。在你学会了骑自行车后，如果你再去学**骑摩托车**，你会发现它们有很多相似的地方，比如保持平衡、掌握方向。因为你已经有了骑自行车的经验，所以学摩托车的时候就不需要从头开始学习，只需要微调一下，比如学会用油门和刹车。因此，**骑自行车的经验帮助你更快地学会了骑摩托车。**



迁移学习的概念

对应到迁移学习：

- **基准模型**就是我们在**一个任务上训练好的模型**。比如我们在大量图片上训练了一个模型，让它学会识别猫和狗的特征。
- 当我们要处理**新任务**时，比如识别老虎和狮子，**我们不会从头训练模型**，而是直接用之前训练的模型。因为模型已经学会了识别动物的基本特征（比如毛发、眼睛、形状），我们只需要稍微调整一下，模型就能很好地完成新任务



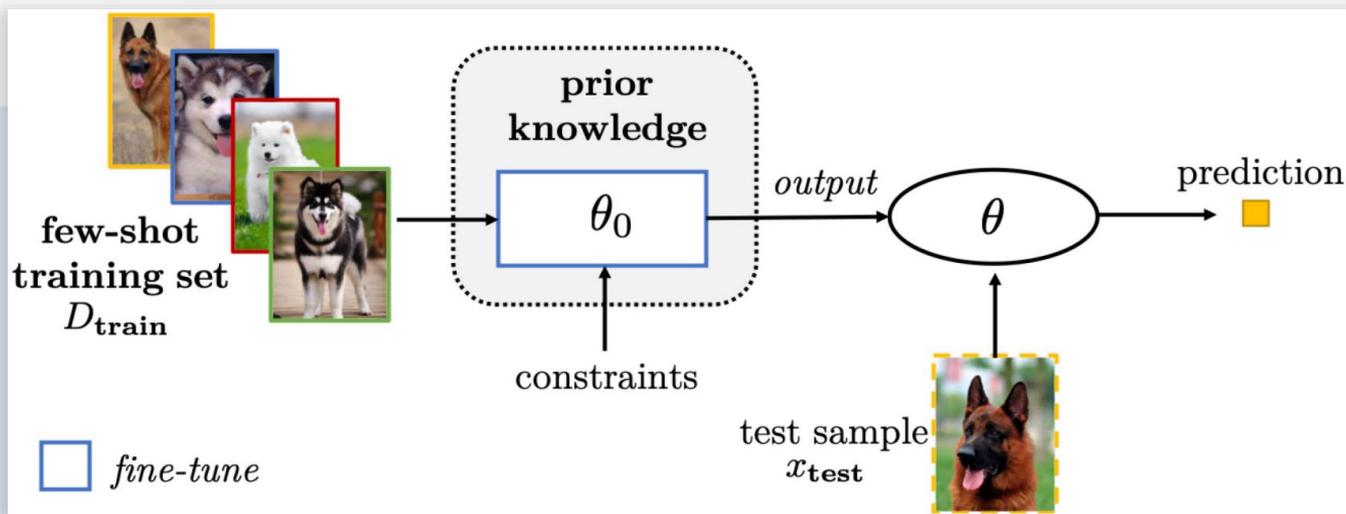
迁移学习的实际应用



迁移学习的主要应用1---小样本学习

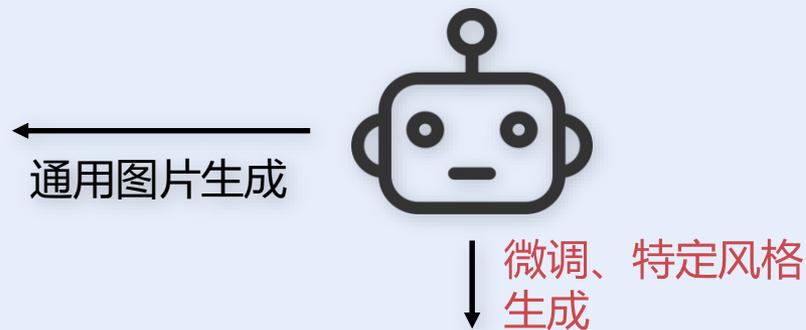
当我们想要给一种动物进行分类时，通常需要大量的图片数据来训练一个高精度的分类器。然而，现实中有一些稀有的动物，我们能找到的图片可能非常有限，可能只有几十甚至几百张。这时，我们该如何通过这些少量的图片，成功训练出一个能准确识别稀有动物的分类器呢？这就是小样本学习技术发挥作用的地方。

具体来说，小样本学习可以通过迁移学习等方式，来弥补数据不足。例如，它可以将从常见动物中学到的识别特征（如鸟类的翅膀形状、羽毛纹理）迁移到稀有鸟类的分类任务中。这样，即便我们手上只有少量的稀有动物图片，模型也能通过已有的知识，快速掌握如何识别这些稀有动物。



迁移学习的主要应用2---图像生成领域

迁移学习在图像生成领域同样发挥着作用。例如我们想要生成某个特定风格的图像（例如手绘风格），但是往往很难收集到大量特定风格的图片。因此，我们可以先训练一个具有通用图像生成能力的模型，然后对其进行微调，使其生成特定的风格。



大模型中的自监督学习和迁移学习



实例：大模型中的自监督学习和迁移学习-训练

大模型通过已知的文本数据，预测下一个词或填补缺失的词。在这个过程中，模型逐渐学习到关于语法、语义、常识知识等复杂的语言规律。

训练阶段

输入：

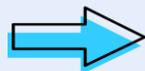
“ 爱因斯坦在物理学方面
有卓越贡献，曾提出
过相对论，是物理学的
[MASK] ”

...

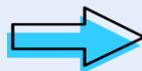
标签：

“ 爱因斯坦在物理学方面
有卓越贡献，曾提出
过相对论，是物理学的
奠基人 ”

...



大语言
模型



推理阶段

可能的输出

:

符合事实	{	奠基人 (p=0.73)
		学者 (p=0.82)
符合语法但 不符合事实	{	发明者 (p=0.12)
		实验物理学家 (p=0.08)
不符合语法	{	优秀 (0.0003)
		奔跑 (0.0002)

实例：大模型中的自监督学习和迁移学习

-Fine Tuning

Fine-tuning（微调）是迁移学习中的一个重要过程，指的是在一个已经通过大规模数据预训练的模型上，使用较小规模的特定任务数据进行进一步的训练，以提高模型在该任务上的表现。

在迁移学习中，模型首先在大规模的通用数据集上进行预训练（如 GPT），学习了语言的基本规律、语法结构、词汇关系等。

Fine-tuning:

- **定义**：微调是将预训练模型应用于某个特定任务,利用原有大模型的知识 and 少量特定任务领域的的数据将模型的能力迁移到某个特定任务上
- **目的**：通过微调，让模型不仅能保留原有的通用知识，还能学会与目标任务相关的特定知识。

例如，预训练的大模型可能已经学习了语言的语义和句法，但并不一定了解特定领域的专业知识。通过在特定领域的数据集上进行微调，可以让模型更适应该领域的需求，比如医学、法律等领域的文本生成或分类。

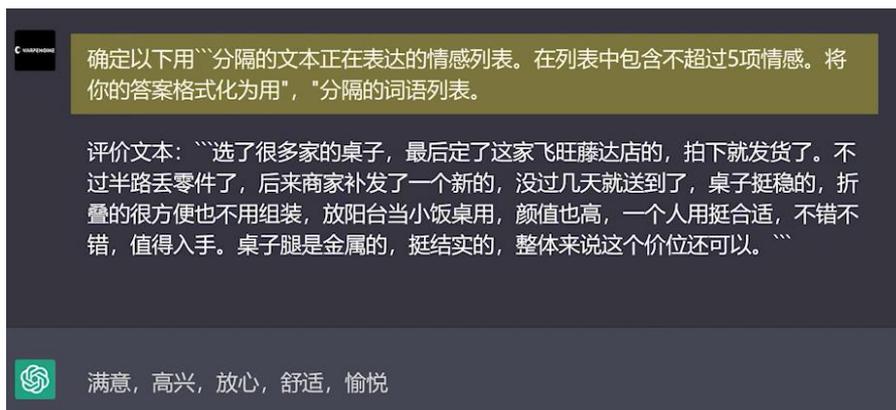
实例：大模型中的自监督学习和迁移学习

-Prompt Engineering

Prompt Engineering 类似于“对大模型进行引导”，这就像迁移学习中我们使用预训练模型。预训练的语言模型已经在大规模文本数据上学习了语法、语义和常识知识，而 Prompt Engineering 是利用这些已学习的知识，通过设计精细的提示（prompt），使模型在特定任务或新领域上表现得更好。

- **预训练模型**：大语言模型（如 GPT-3）经过了在大量文本数据上的预训练。
- **Prompt 微调**：用户通过设计和优化提示（prompt），让模型能够针对特定任务或目标领域提供合适的输出，而不需要重新训练模型。这类似于通过 prompt 来“微调”模型的输出。

右图就是利用 Prompt Engineering 来完成情感分类的例子



确定以下用“”分隔的文本正在表达的情感列表。在列表中包含不超过5项情感。将你的答案格式化为用“”分隔的词语列表。

评价文本：“”选了很多家的桌子，最后定了这家飞旺藤达店的，拍下就发货了。不过半路丢零件了，后来商家补发了一个新的，没过几天就送到了，桌子挺稳的，折叠的很方便也不用组装，放阳台当小饭桌用，颜值也高，一个人用挺合适，不错不错，值得入手。桌子腿是金属的，挺结实的，整体来说这个价位还可以。”

满意，高兴，放心，舒适，愉悦