大模型的有监督微调

吴振一
郑州大学 河南先进技术研究院

- **研究目标**

行业背景：

传统的人才评价方法包括业绩评估、360度反馈、能力测评、任职资格评价、工作观察、个人发展计划等。

随着科技的发展和人力资源管理理念的转变，越来越需要更加灵活和多样化的评价方式。

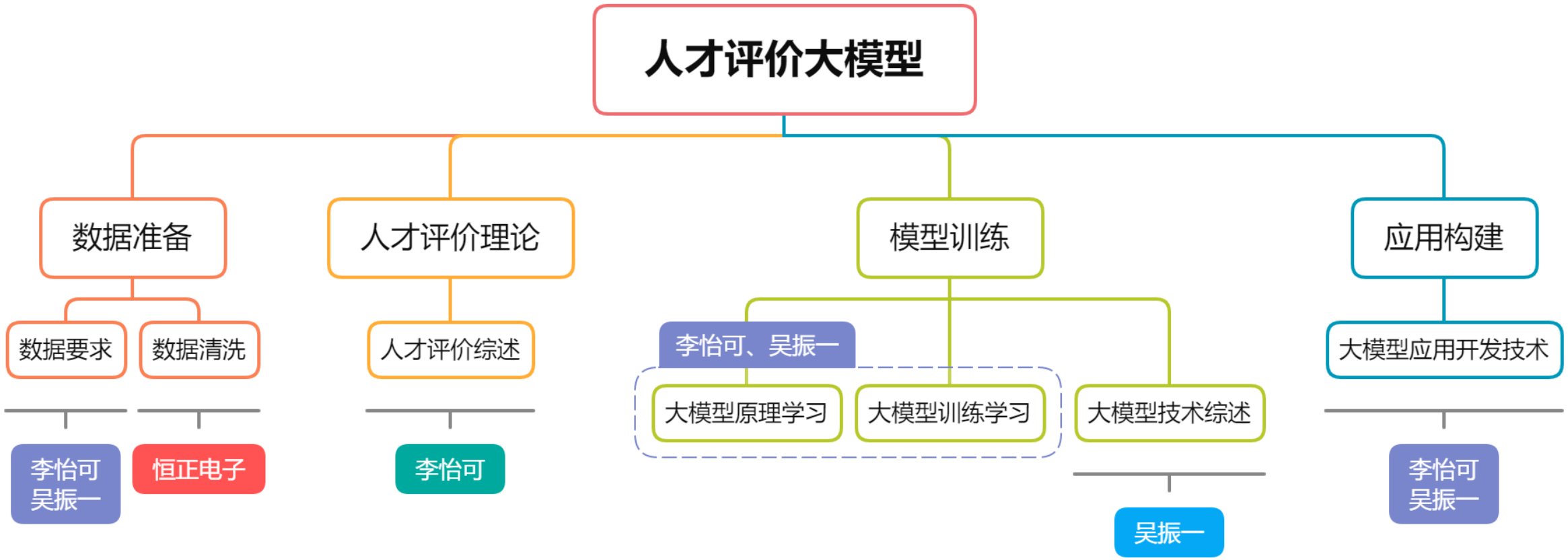
2022年以来，大语言模型领域涌现出惊人的技术力，2023~2024年期间，伴随着其技术成熟度的提升，超强的适应性、极高的自动化与效率、多维度的分析能力，以及可持续改进的特性，使大语言模型在各行各业中展现出巨大的应用潜力。

研究目标：

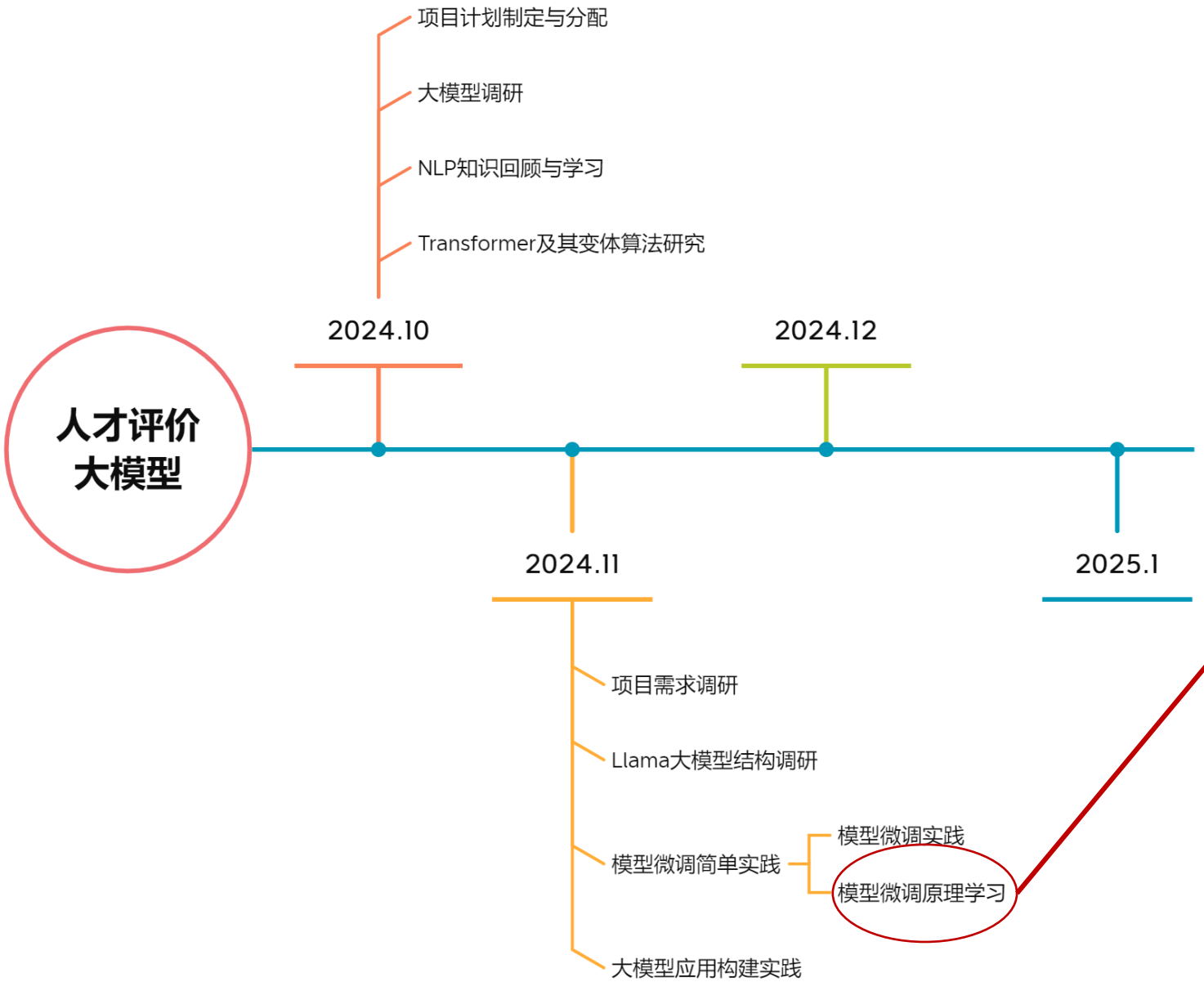
1.开发出拥有人事人才领域知识，能够进行人才评价的大语言模型。

2.构建基于人才评价大语言模型的智能人才评价系统。

- **研究内容**


人才评价大模型
数据准备
人才评价理论
模型训练
应用构建
数据要求
数据清洗
人才评价综述
李怡可、吴振一
大模型原理学习
大模型训练学习
大模型技术综述
大模型应用开发技术
李怡可
吴振一
恒正电子
李怡可
李怡可
吴振一
吴振一
李怡可
吴振一

- **研究进度**



汇报主题：模型微调原理学习

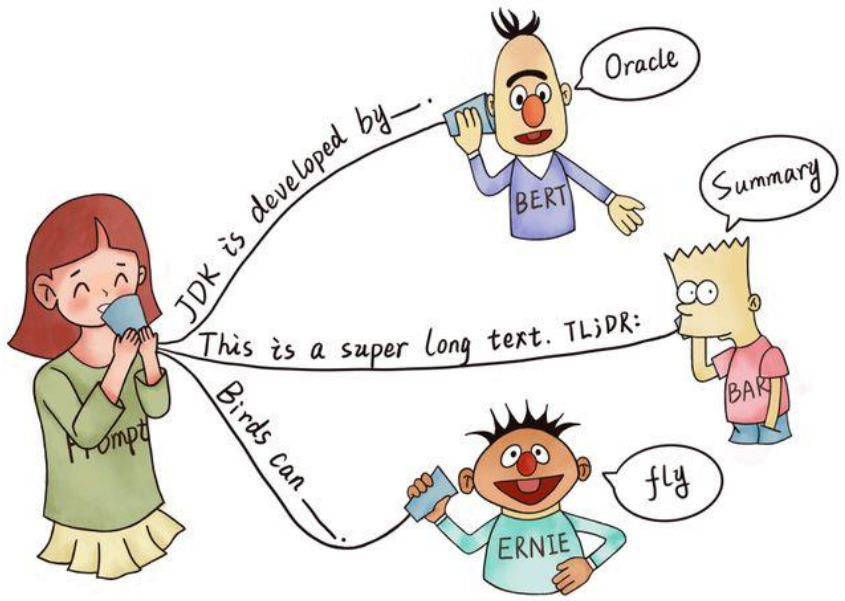
项目地址：https://github.com/ayispace/LLMwithTalentEvaluation

Outline

- Prompt-based Learning
- Incontext Learning, ICL
- Low-Rank Adaptation, LoRA
- QLoRA

- **Prompt-based Learning**

What is Prompt-based Learning?

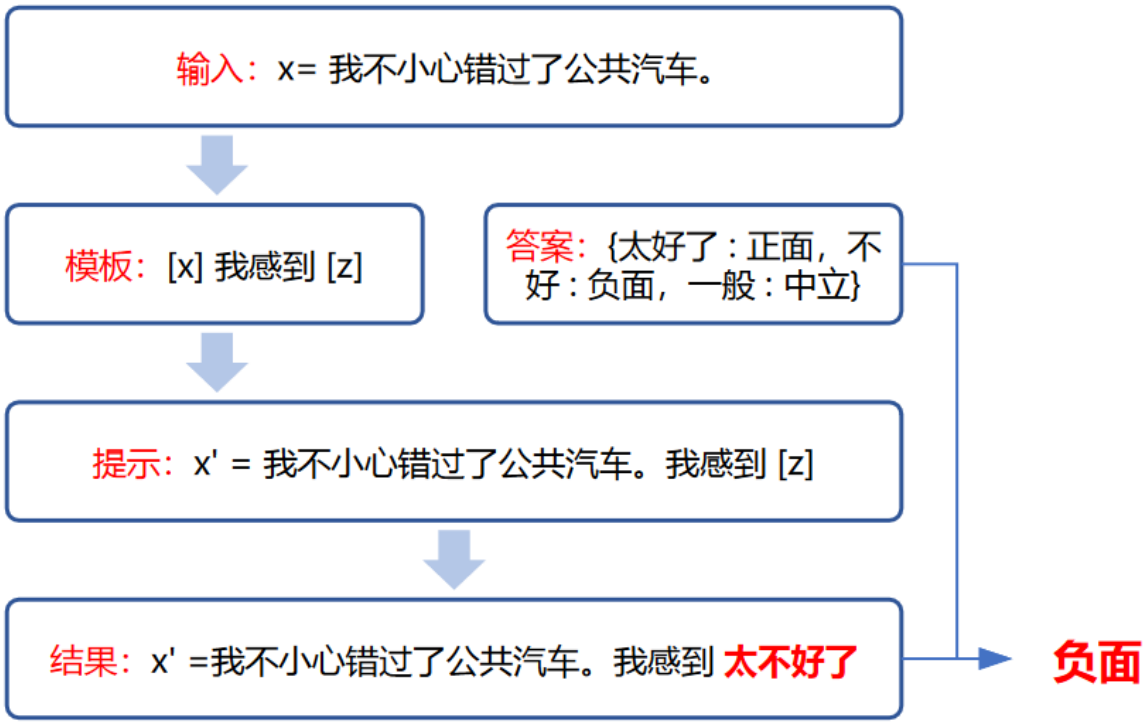


预训练模型的输入端添加特定的“提示”（prompt）内容，以引导模型在特定任务上的表现

(https://arxiv.org/pdf/2107.13586)

- **Prompt-based Learning**

What is Prompt-based Learning?



- **Prompt-based Learning**

Three steps about Prompt-based Learning:

1. 提示添加：

借助特定的模板，将原始的文本和额外添加的提示拼接起来，一并输入到语言模型中

Template： “[X] 我感到 [Z]”

$x =$ “我不小心错过了公共汽车。”

$x' =$ “我不小心错过了公共汽车。我感到$[Z]$”

- **Prompt-based Learning**

Three steps about Prompt-based Learning:

2. 答案搜索：

将构建好的提示整体输入语言模型后，需要找出语言模型对[Z] 处预测得分最高的文本

Z = {“太好了”，“好”，“一般”，“不好”，“糟糕”}

$$\hat{z} = \underset{z \in \mathbb{Z}}{\text{search}} P(f_{\text{fill}}(x', z); \theta)$$

- **Prompt-based Learning**

Three steps about Prompt-based Learning:

3. 答案映射：

得到的模型输出并不一定就是最终的标签。还需要将模型的输出与最终的标签做映射。

$$\begin{cases} \text{if } \hat{z} \in \{\text{“太好了”}, \text{“好”}\} & \hat{y} = \text{“正面”} \\ \text{if } \hat{z} \in \{\text{“不好”}, \text{“糟糕”}\} & \hat{y} = \text{“负面”} \\ \text{if } \hat{z} \in \{\text{“一般”}\} & \hat{y} = \text{“中立”} \end{cases}$$

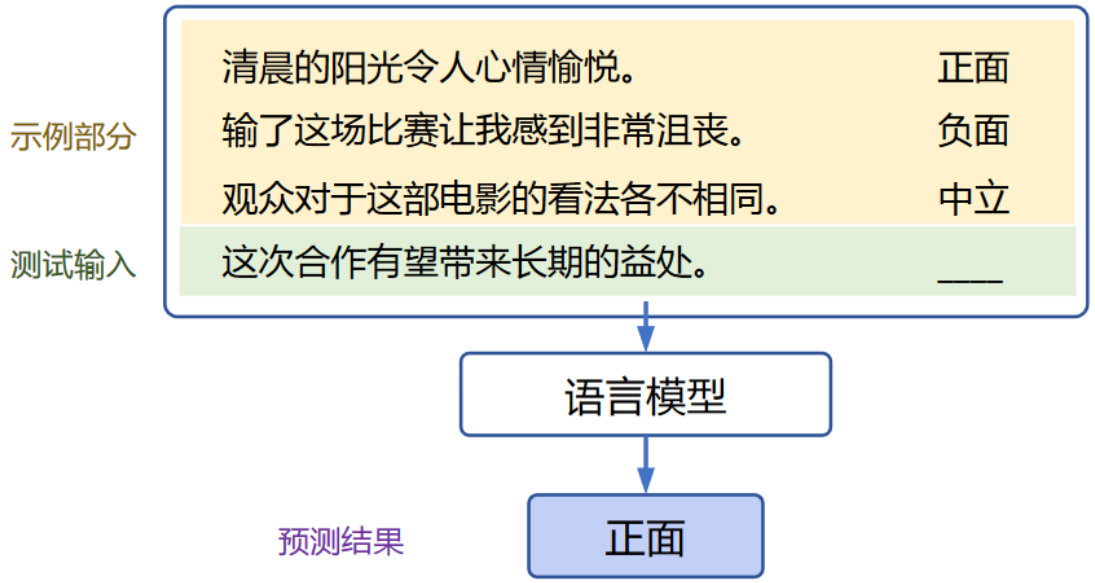
Outline

- Prompt-based Learning
- Incontext Learning, ICL
- Low-Rank Adaptation, LoRA
- QLoRA

- **ICL**

What is Incontext Learning?

向模型输入特定任务的一些具体例子（也称示例，Demonstration）以及要测试的样例，模型可以根据给定的示例续写出测试样例的答案



（并不需要对模型进行参数更新，仅执行前向的推理）

Outline

- Prompt-based Learning
- Incontext Learning, ICL
- Low-Rank Adaptation, LoRA
- QLoRA

- **LoRA**

What is Low-Rank Adaptation?

A method for Parameter Efficient Fine-tuning.

So what is Parameter Efficient Fine-tuning?

由于大语言模型参数量十分庞大，当将其应用到下游任务时，微调全部参数需要相当高的算力。为了节省成本，研究人员提出了多种参数高效（Parameter Efficient）的微调方法，旨在仅训练少量参数使模型适应到下游任务。

以 LoRA（Low-Rank Adaptation of Large Language Models）为例，其可以在缩减训练参数量和 GPU 显存占用的同时，使训练后的模型具有与全量微调相当的性能。

- **LoRA**

What is Low-Rank Adaptation?

有文献表明：语言模型针对特定任务微调之后，权重矩阵通常具有很低的本征秩（Intrinsic Rank）。研究人员认为参数更新量即便投影到较小的子空间中，也不会影响学习的有效性。

因此，提出固定预训练模型参数不变，在原本权重矩阵旁路添加低秩矩阵的乘积作为可训练参数，用以模拟参数的变化量。

文献：Intrinsic Dimensionality Explains the Effectiveness of Language Model Fine-Tuning[https://arxiv.org/abs/2012.13255]

- **LoRA**

Low-Rank Adaptation working mechanism.

假设我们有一个大模型的线性变换矩阵 $W \in \mathbb{R}^{d\times k}$ 参与到计算中。LoRA 的操作过程如下：

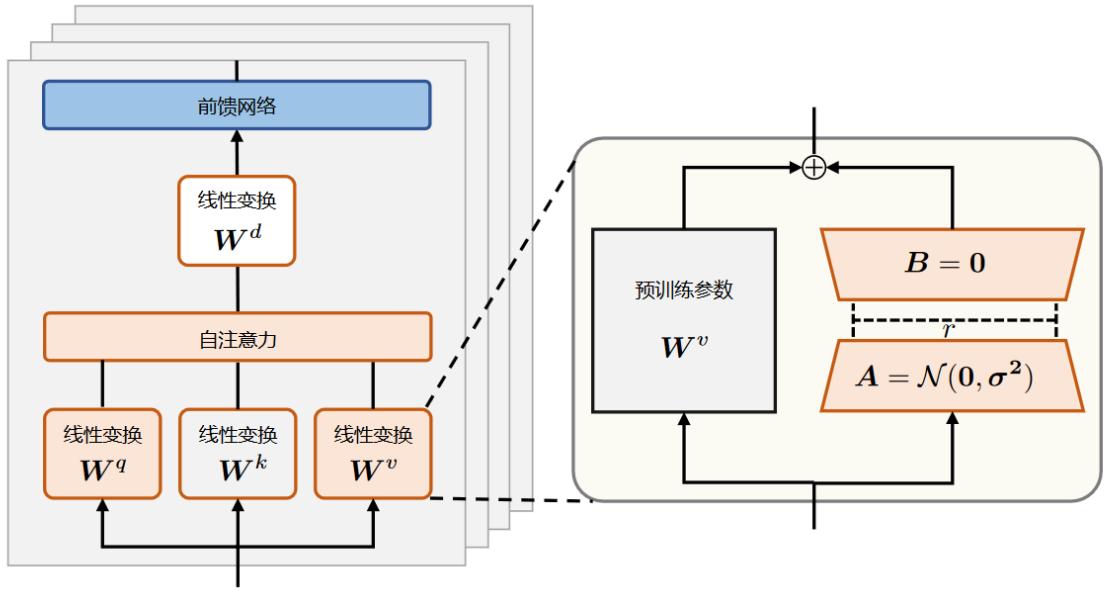
1.LoRA 引入两个较小的矩阵 $A \in \mathbb{R}^{d\times r}$ $B \in \mathbb{R}^{r\times k}$ 其中：$r \ll \min(d,k)$ r表示矩阵W的秩

2.使用 LoRA 进行微调时，模型计算公式变为：

$$Wx + \alpha \cdot ABx$$

其中 x 是输入，α 是一个缩放因子，通常用来调节新引入的调整矩阵对原始输出的影响。

- **LoRA**


前馈网络
线性变换
W^d
自注意力
线性变换
W^q
线性变换
W^k
线性变换
W^v
预训练参数
W^v
$B = 0$
r
$A = \mathcal{N}(0, \sigma^2)$

LoRA

Outline

- Prompt-based Learning
- Incontext Learning, ICL
- Low-Rank Adaptation, LoRA
- QLoRA

- QLoRA

What is QLoRA?

Q = Quantized. QLoRA = Quantized Low-Rank Adaptation

So what is Quantized?

量化（Quantization） 是一种将数值从较高精度表示转换为较低精度表示的方法。是一种常见的模型压缩方法。

- 定点量化：
通过一个缩放因子来将原始浮点数值映射到定点值

- 动态范围量化：
允许在模型运行时调整缩放因子，使得不同权重可以根据其值范围进行动态调整，从而减少精度损失

- 对称和非对称量化：
对称量化使用零点为中心，值在正负方向上对称地分布；非对称量化则允许更灵活的范围，特别适合数值范围不对称的分布

- QLoRA

What is QLoRA?

- 数据类型：
4-bit NormalFloat（NF4）

- 量化技术：
双重量化（Double Quantization）

- 内存优化：
分页优化器（Paged Optimizers）