# 项目基本介绍

**项目名称**：基于微调大模型分析课堂逻辑问答对授课效果影响

**团队成员**：刘承龙、陈强、张浩宇

**项目简介**：本项目通过微调大语言模型，构建针对课堂问答逻辑关系的分析系统，旨在识别教师提问与学生回答之间的逻辑关联，进而评估课堂互动效果，为教学优化提供数据支持。

**项目代码链接**：https://github.com/GuoDragon/MindSpore\_CCNU\_MindTrio

# 项目的意义和价值

## 项目要解决什么样的问题

传统课堂中，对师生在课堂中的问答对的深层逻辑关系语义发掘不足，无法通过此项任务来对提高教师授课水平；且目前多数针对 NLP 分类的任务都缺乏可解释性。

## 问题解决所带来的商业或社会价值

**教育质量提升**：通过分析课堂问答逻辑，帮助教师调整提问策略，提升教学效果。

**教育智能化**：为智慧课堂、在线教育平台提供底层能力支持，推动教育评估的自动化与智能化。

**社会价值**：促进教育公平与个性化教学发展，助力优质教育资源的科学利用。

## 项目方案介绍

### 现有解决方案的不足

当前已有的课堂问答分析方案多基于规则或浅层模型，存在以下问题：

1. 判断准确率低，难以应对多样化的自然语言表达；
2. 缺乏可解释性，无法给出判断依据；
3. 泛化能力弱，难以迁移到不同学科或教学场景。

### 项目整体方案pipeline介绍

**数据构建**：人工标注1万条课堂问答数据，涵盖5类逻辑关系；

**模型微调**：基于DeepSeek-R1-Distill-Qwen-1.5B模型，使用LoRA进行高效微调；

**向量库构建**：构建逻辑关系文本向量库，增强模型对语句逻辑的感知能力；

**推理与分析**：模型输出逻辑类别及判断原因，辅助教学效果评估。

### 项目创新点与技术亮点

**创新点**：

1. 引入逻辑关系向量库，增强模型对语句结构的理解；
2. 结合LoRA微调，提升模型在少样本场景下的表现；
3. 输出带解释的逻辑判断结果，增强可解释性。

**技术亮点**：

1. 使用MindSpore框架实现高效训练与推理；
2. 模型轻量化设计，支持边缘设备部署；
3. 系统具备良好的扩展性，可接入多种教育平台。

## 项目实现思路

我将从**原理解析**和**程序设计**两个方面来介绍本项目的实现思路，重点围绕如何通过微调大语言模型实现对课堂问答逻辑关系的识别与分析。

### 原理解析

#### 基于LoRA的高效微调机制

#### 传统的全参数微调大语言模型需要更新数十亿参数，计算和存储成本极高。本项目采用LoRA（Low-Rank Adaptation） 技术，通过引入低秩矩阵分解，在不改变原始模型参数的情况下，仅训练少量新增参数，即可实现对模型行为的精准控制。具体而言，我们在Transformer的Self-Attention层中插入可训练的低秩矩阵，使其在推理时能与原有权重合并，从而显著降低训练成本并提升微调效率。

#### 解释性输出

#### 模型不仅需要判断逻辑关系类型（如因果、对比、递进等），还需输出判断原因。为此，我们精心设计了完整的数据结构，并使用“理据占位符“，使模型既能完成分类任务，又能生成自然语言形式的解释，增强了结果的可信度和实用性。

# 程序设计



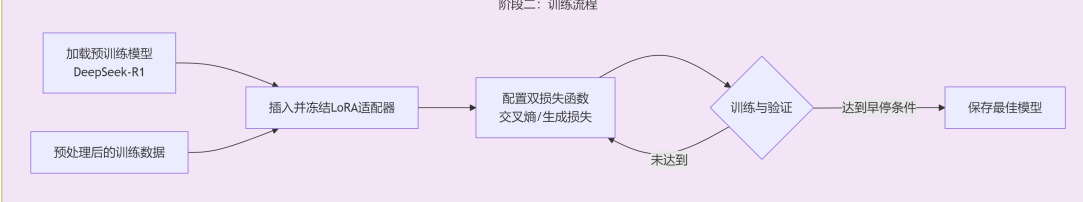
### 数据流设计与训练流程

项目的核心数据处理与推理流程如下：

1. 数据预处理：

* 对原始问答文本进行清洗、分词、去除停用词；
* 构建标签体系，包括5类逻辑关系；
* 按分类比例随机抽取2000条数据进行高质量的原因人工标注

1. 训练流程：



* 加载预训练模型（DeepSeek-R1-Distill-Qwen-1.5B）；
* 插入LoRA适配层，冻结原模型参数；
* 使用交叉熵损失函数训练分类任务，使用生成损失训练解释生成任务；
* 每轮训练后验证集评估，早停策略防止过拟合。

### 遇到的问题与解决方案

**问题1**：标注数据不一致

**解决**：制定细粒度标注规范，并进行多轮标注与仲裁，确保数据质量。

**问题2**：模型在少数类别上表现差  
**解决**：合并数据量过少标签，调整学习率等参数

# 成果展示

[昇思MindSpore创新训练营 MindTrio小组项目答辩.pptx](file:///D:\LLaMA-Factory\relationAI_APP\hw\MindSpore_CCNU_MindTrio\昇思MindSpore创新训练营%20MindTrio小组项目答辩.pptx)

# 项目总结

### 本项目通过使用MindSpore平台，利用LoRA方法微调大语言模型，成功构建了一个能够识别课堂问答逻辑关系的分析系统。系统不仅具备较高的准确率，还能输出判断原因，为教学效果评估提供了可解释的AI支持。未来将继续优化模型泛化能力，拓展更多逻辑类型，并探索在边缘设备上的部署应用。