

深度模型构建

关键科学问题

本文研究内容

参数实现与优化

架构设计

数据与知识获取

微调成本高昂

多模型参数融合冲突难以调和



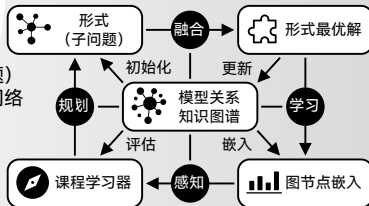
深度模型 | 参数

Q3

多模型场景下的
低损耗模型融合

多形式优化的模型融合

- 构建模型关系网络与表征
- 给出待探索形式（子融合问题）
- 求解子融合问题，更新关系网络
- 形成探索-优化闭环

基于知识迁移的
高效深度模型构建

跨域架构迁移

- 统一架构自编码器
- 神经架构表征学习
- 跨域表征映射



异构空间表征对齐

在表征空间
对齐异构架构

编码器

解码器

监督微调

无监督预训练



知识蒸馏

小样本微调



双重对比学习

提示学习



表征对齐，挖掘和引导预训练知识

小样本知识蒸馏

- 提示学习引导预训练知识
- 双重对比学习分步表征对齐
- 小样本、低资源下的知识迁移

基于小样本蒸馏得到
小规模微调模型

深度模型 | 参数



神经架构 | 计算图

架构设计验证低效

异构空间碎片化难利用

Q2

跨异构架构空间的
可迁移神经架构搜索

知识数据 | 嵌入

训练数据瓶颈

小规模模型难以训练

Q1

少样本场景下的
稳定知识蒸馏