

致理书院暑期基础学科交叉实践课程项目申请表

姓 名	刘知远	性别	男	院系	计算机系		
最高学位	博士	专业方向		人工智能	职称	长聘副教授	
主要科研 方向和成 果	刘知远，清华大学计算机系副教授、博士生导师。主要研究方向为自然语言处理、基础模型。2011 年获得清华大学博士学位，已在 ACL、EMNLP、IJCAI、AAAI 等人工智能领域的著名国际期刊和会议发表相关论文 200 余篇，Google Scholar 统计引用超过 4.5 万次。曾获教育部自然科学一等奖（第 2 完成人）、中国中文信息学会钱伟长中文信息处理科学技术奖一等奖（第 2 完成人）、中国中文信息学会汉王青年创新奖，入选国家青年人才项目、北京智源研究院青年科学家、2020-2022 连续三年 Elsevier 中国高被引学者、《麻省理工科技评论》中国区 35 岁以下科技创新 35 人榜单、中国科协青年人才托举工程。担任中文信息学会理事、社交媒体处理专委会副主任，期刊 AI Open 副主编，ACL、EMNLP、WWW、CIKM、COLING 等国际著名会议领域主席。						
课程题目	大模型技术与交叉应用						
教学团队 介绍 (请附上团 队每位老 师照片)	刘知远						

课程对学生的先修要求	建议考虑学生的交叉性，根据课程设计设置课程先修课要求或者面向不同学科的先修要求，如先修描述为 满足以下条件： 1. 具有编程能力，完成《计算机程序设计基础》或相关课程的学习。 2. 掌握微积分、线性代数、概率论等基础数学知识。 3. 有机器学习基础优先。
课程设计	包括但不限于以下方面： 容纳人数：10 人 教学资源或设备：高性能 GPU 服务器 课时安排：32 授课学时（1 周） + 72 实践学时（3 周） 目标和特色：让学生建立对深度学习和预训练模型的主要思想，掌握预训练模型的理论方法与实现细节，建立针对已有预训练模型进行局部改进和改造的能力，能够利用预训练模型解决相关领域数据的智能理解与处理，并在相关交叉学科领域（如生物医学、化学）进行实践。本课程特色包括：（1）立足前沿，将深度学习最前沿的预训练模型引入教学环节；（2）鼓励交叉，预训练模型在理论基础方面存在很多挑战性问题，面向交叉应用充满想象空间，是鼓励多学科背景学生开展跨学科合作的绝佳课题。

课程方案	建议但不限于以下格式，仅供参考 总体设计：课程将分为理论授课、实践指导、自由探索三个部分，以由浅入深，全程指引的方式，让学生掌握利用预训练模型开展跨学科研究。 课程实施步骤：（如课程专题项目引入导引，任务设计，教学不同环节等） （1）理论授课环节，由教师介绍预训练模型的基础理论、方法、应用和开放性挑战性问题，让学生掌握预训练模型的理论知识；考虑到学生的不同学科背景，将根据情况提供不同方面的自学材料与讲义。 （2）实践指导环节，由助教介绍预训练模型的计算平台与实现细节，并手把手指导学生完成指定实现任务，让学生掌握预训练模型的实践知识。 （3）自由探索环节，根据学生的学科特点，提出若干候选探索课题，也允许学生自由选题。由实验室有丰富科研经验的高年级博士生或博士后担任小导师，指导不同学科背景的学生讨论组队和选题，并利用 2 周时间合作完成探索课题任务，撰写成果报告，并鼓励原创性强的成果以代码开源或者学术论文形式公开。
------	---