

七月集中学习安排

为使大家在工程能力、科研能力快速成长，根据现国际科研发展前沿和国家重大战略需求，将七月集体学习安排如下：

★学习形式：

研 0、研 1 同学按《安排表》必须按时参加（研 2 同学不做硬性要求，但希望能安排时间，也尽量按学习表参与到线下学习）。

学习以线下交流、分享开展，老师、教授、博士讲授为主，辅以线上参与；完成良好的进行奖励。

- 1、 每周一各项目组内部开会交流项目进展情况。
- 2、 周二、周四下午，分别邀请老师、项目组进行专题分享，拟计划安排如下

七月学习安排

周一	周二	周三	周四	周五
3 日 下午 14:30 小组分享	4 日 下午 14:30 《科研论文写作》	11 日	6 日 前沿研究分享《AI+脑科学+安全》 工程项目开发分享《视频压缩+AI》	7 日
10 日 下午 14:00 小组分享	11 日 下午 14:30 《科研申请》	12 日	13 日 前沿研究分享《下一代人工智能方法》 工程项目开发分享《AI 平台》	14 日
17 日 下午 14:30 小组分享	18 日 下午 14:30 《科研论文写作》	19 日	20 日 前沿研究分享 《数据驱动+知识驱动融合的人工智能、大模型》 工程项目开发分享《三维重建、VAPS 开发等》	21 日
24 日 下午 14:30 小组分享	25 日 下午 14:30 《科研学习成果汇报》	26 日	27 日 下午 14:30 《科研学习成果汇报》	28 日

★学习要求：

参与学习人员均需完成工程能力+科研能力

工程能力的培养（就项目完成度进行考核）：进入项目组开展工程能力的培养。在项目负责人、学长带领下，掌握一个具体开发技术，完成开发任务。

研 0、研 1 同学，请微信直接联系张老师，工程能力学习分小组进行。

科研能力的培养（7.25 日、7.27 日全体做汇报）：写一个相关领域研究进展报告、或者一个具体技术的报告（20 页以上），配合 ppt，作 5-10 分钟汇报（领域自行选择，见附件科研学习选题范围推荐）。

★奖励设置：

- 1) 开展的以线下活动为主，希望研 0 同学能在校参与。有特殊原因不能到校的，需要在项目负责人或导师保证其学习效果基础上，线上参加。
- 2) 科研报告的范围和要求：我们提供方向，大家在方向内自选（见附件科研学习选题范围推荐）；并按照格式要求完成科研报告；按照 PPT 模板要求完成制作并汇报科研能力（汇报或者科研报告），汇报评级达到优良以上的，团队给予补贴。每份 400。
- 3) 工程能力的培养达到目标要求的，团队基于研 0 同学以住房补助，800 元/月。
- 4) 期间撰写专利、论文的，撰写费用 400；受理、发表的相关奖励另行计算。

附件：

★科研学习选题范围推荐：

以总体科学目标为牵引，拟以培育项目的方式资助探索性强、选题新颖的申请项目，研究方向如下：

1. 深度学习的表示理论和泛化理论。

研究卷积神经网络（以及其它带对称性的网络）、图神经网络、transformer 网络、循环神经网络、低精度神经网络、动态神经网络、生成扩散模型等模型的泛化误差分析理论、鲁棒性和稳定性理论，并在实际数据集上进行检验；研究无监督表示学习、预训练-微调范式等方法的理论基础，发展新的泛化分析方法，指导深度学习模型和算法设计。

1. 深度学习的训练方法。

研究深度学习的损失景观，包括但不限于：临界点的分布及其嵌入结构、极小点的连通性等，深度学习中的非凸优化问题、优化算法的正则化理论和收敛行为，神经网络的过参数化和训练过程对于超参的依赖性问题、基于极大值原理的训练方法、训练时间复杂度等问题，循环神经网络记忆灾难问题、编码-解码方法与 Mori-Zwanzig 方法的关联特性，发展收敛速度更快、时间复杂度更低的训练算法及工具，建立卷积网络、Transformer 网络、扩散模型、混合专家模型等特定模型的优化理论及高效训练方法，深度学习优化过程对泛化性能的影响等。

2. 微分方程与机器学习。

研究求解微分方程正反问题及解算子逼近的概率机器学习方法；基于生成式扩散概率模型的物理场生成、模拟与补全框架；基于微分方程设计新的机器学习模型，设计和分析网络结构、加速模型的推理、分析神经网络的训练过程。

面向具有实际应用价值的反问题，研究机器学习求解微分方程的鲁棒算法；研究传统微分方程算法和机器学习方法的有效结合方法；研究高维微分方程的正则性理论与算法；研究微分方程解算子的逼近方法（如通过机器学习方法获得动理学方程、弹性力学方程、流体力学方程、Maxwell 方程以及其它常用微分方程的解算子）；融合机器学习方法处理科学计算的基础问题（求解线性方程组、特征值问题等）。

3. 隐私保护的机器学习方法。

针对主流机器学习问题，结合安全多方计算、全同态加密、零知识证明等方法构建具备实用性的可信机器学习环境。发展隐私保护协同训练和预测方法，发展加密和隐私计算环境的特征聚类、查询和多模型汇聚方法，发展加密跨域迁移学习方法，发展面向对抗样本、后门等分析、攻击、防御和修复方法，研究机器学习框架对模型干扰、破坏和控制方法，发展可控精度的隐私计算方法。

4. 脑科学启发的新一代人工智能方法。

发展对大脑信息整合与编码的定量数学刻画和计算方法，设计新一代脑启发的深度神经网络和循环神经网络，提高传统神经网络的表现性能；建立具有树突几何结构和计算功能的人工神经元数学模型，并用于发展包含生物神经元树突计算的深度神经网络和循环神经网络，提高传统神经网络的表现性能；发展包含多种生物神经元生理特征和生物神经网络结构特点的人工神经网络及其训练算法，解决图像识别、图像恢复、医学图像重构、地震波检测等应用问题。

5. 数据驱动与知识驱动融合的人工智能方法。//大模型

建立数据驱动的机器学习与知识驱动的符号计算相融合的新型人工智能理论和方法，突破神经网络模型不可解释的瓶颈；研究知识表示与推理框架、大规模隐式表达的知识获取、多源异构知识融合、知识融入的预训练模型、知识数据双驱动的决策推理等，解决不同场景的应用问题。

6. 科学计算领域的人工智能方法。

针对电子多体问题，建立薛定谔方程数值计算、第一性原理计算、增强采样、自由能计算、粗粒化分子动力学等的人工智能方法，探索人工智能方法在电池、电催化、合金、光伏等体系研究中的应用。

针对典型的物理、化学、材料、生物、燃烧等领域的跨尺度问题和动力学问题，通过融合物理模型与人工智能方法，探索复杂体系变量隐含物理关系的挖掘方法，建立构效关系的数学表达，构建具有通用性的跨尺度人工智能辅助计算理论和方法，解决典型复杂多尺度计算问题。)

7. 人工智能驱动的下一代微观科学计算平台。

发展基于人工智能的高精度、高效率的第一性原理方法；面向物理、化学、材料、生物等领域的实际复杂问题，建立多尺度模型，实现高精度、大尺度和高效率的分子动力学模拟方法；探索建立人工智能与科学计算双驱动的“软-硬件协同优化”方法和科学计算专用平台。

8. 开放环境下多智能体协作的智能感知理论与方法。

针对多模态信息融合中由于数据视角、维度、密度、采集和标注难易程度不同而造成的融合难题，研究基于深度学习的融合模型，实现模态一致性并减少融合过程中信息损失；研究轻量级的模态间在线时空对齐方法；研究能容忍模态间非对齐状态下的融合方法；研究用易采集、易标注模态数据引导的难采集、难标注模态数据的预训练与微调方法；研究大规模多任务、多模态学习的预训练方法，实现少样本/零样本迁移。

9. 可通用的专业领域人机交互方法。

针对多变输入信号，建立自动化多语种语言、图像、视频等多模态数据生成模型，发展可解释的多轮交互决策方法；建立机器学习和知识搜索的有效结合方法；探索新方法在不同专业领域场景中的应用。

10. 下一代多模态数据编程框架。//流计算的研究

发展面向超大规模多模态数据（文本、图像、视频、向量、时间序列、图等）的存储、索引、联合查询和分析方法。发展一体化的多模态数据编程框架，建立自动化数据生成、评估和筛选方法，实现自动知识发现和自动模型生成性能的突破，并完成超大规模、多模态数据集上的可靠性验证。

11. 支持下一代人工智能的开放型高质量科学数据库。

研究跨领域、多模态科学数据的主动发现、统一存储和统一管理方法。研究基于主动学习的科学数据、科技文献知识抽取与融合方法。研究跨学科、多尺度科学数据的知识对象标识化、语义化构建方法。研究融合领域知识的多模态预训练语言模型，开发通用新型数据挖掘方法。形成具有一定国际影响力的覆盖生命、化学、材料、遥感、空间科学等领域的高质量、通用型科学数据库，为人工智能驱动的科学新范式提供基础科学数据资源服务。

12. 高精度、可解释的谱学和影像数据分析方法。

发展光谱、质谱和各类影像数据处理的人工智能方法。建立融合模拟与实验数据的可解释“谱-构-效”模型，开发人工智能驱动的光谱实时解读与反演软件；基于 AlphaFold 等蛋白结构预测方法，建立高精度冷冻电镜蛋白结构反演算法等。

