

LEARN Lab 团队简介

师宇哲

LEARN Lab

2020 年 9 月 20 日

摘要

LEARN Lab(Learners' Engineering and Research Network Lab) 是一个以认知机器学习为主要研究方向的学生自主科研创新团队，校方指导老师是华中科技大学的金海教授。团队的目标是通过将机器学习和认知科学互促结合研究，探索走向通用人工智能 (AGI) 的道路；并且为拥有多元化专业背景的大学生打造一个科研创新创业的交流平台。团队愿意为华中科技大学的本科生提供科研实习和科学交流的平台。

目录

第一章 引言	2
1.1 背景	2
1.2 LEARN Lab	3
第二章 LEARN Lab 团队情况综述	5
2.1 校方导师（荣誉导师）	5
2.2 导师及研究方向	5
2.2.1 认知科学与机器学习方向	5
2.2.2 理论机器学习方向	5
2.2.3 交互设计与认知心理学方向	5
2.3 团队负责人	6
第三章 团队项目	7
3.1 项目 1：开放世界反演绎谓词发明	7
3.2 项目 2：“等车问题”的再探究：从认知心理学和机器学习的角度	7
3.3 项目 3：研究知识驱动的神经网络	8
3.4 筹备中的项目：抽象与推理挑战	8
第四章 团队架构	9
第五章 团队运作规划	10
5.1 团队内部科研运作	10
5.1.1 关于例行讨论班的细则	10
5.1.2 关于以小组单位制开展科研的细则	11
5.2 团队平台运作	11
第六章 团队发展规划	12
6.1 短期规划	12
6.2 长期规划	12
第七章 其它说明	13

第一章 引言

1.1 背景

随着计算能力的提升和一些重要算法的提出，近十年来人工智能的应用在学术界取得了巨大的突破，在工业界的落地项目进展迅速。因此，人工智能毫无疑问地成为了新世纪人类科学技术的重要赛道。2018年7月8日，国务院印发了《新一代人工智能发展规划》，将人工智能提升到国家战略高度，要求有前瞻性地布局高级机器学习，重点突破自适应学习、自主学习等理论方法，实现具备高可解释性、强泛化能力的人工智能。作为大学生，在积极顺应发展大势的同时独立思考探索，走出属于自己的创新之路，为更好地帮助人类发展奠定基础势在必行。

机器学习领域按出现时间先后可大致分为逻辑理论主义，连接主义，符号学习主义和统计学习主义，这些思路在历史上分别经历过起伏。统计学习和以深度学习为代表的连接主义拥有以下优势：1) 可以直接处理大量“生数据”，即非结构数据识别出隐含在其中的规律和模式；2) 模型具有梯度，这就意味着统计学习可以同时优化经验风险和结构风险，能够比较高效地从假设空间中学得目标假设。符号学习和逻辑理论拥有以下优势：1) 可以直接利用人类背景知识，甚至是人类常识进行推理；2) 由于无论是演绎还是归纳、溯因，都可以用规则或一阶逻辑的形式表示，可以被人类理解，因此推理结果可靠。

然而遗憾的是，两种思路各自存在致命缺陷。一部分统计机器学习和深度学习存在以下缺陷：1) 一部分统计机器学习，尤其是深度学习模型从数据学得的知识不能够被人类理解，这导致了统计学习方法的不可靠性，这是因为人类没有办法通过理解模型学得的知识从而判断出模型学习误差的来源，相反地，只能通过它的输入和输出来判断学习效果的好坏，因此是“黑盒模型”。在一些极端追求结果可靠性的领域，如自动驾驶，医疗诊断，罪犯识别，文物鉴定等方面的应用出现问题；2) 另一方面，统计学习的应用需要在大量特定场景数据的监督下不断调整超参数，以尝试取得在该任务上的最好效果。这意味着对于每一个特定的任务，都要耗费大量人力、物力进行针对该领域的数据收集和标注。据统计数据，目前人工智能工业产品和服务的成本有超过 60% 来自数据准备。逻辑推理存在以下缺陷：1) 模型总是离散表示，不存在梯度，这意味着无法将经验风险和结构风险相结合，无法高效从假设空间搜索到目标假设；2) 无法直接处理“生数据”，而由于许多领域知识难以总结，形成了“知识工程瓶颈”，使逻辑推理的应用范围大受限制。此外，两种思路的理论基础都基于闭合世界假设 (Close-World Assumption)，即非正样例是负样例，然而真实世界中往往存在的是开放世界识别问题 (Open-World Recognition)，即测试集中存在属于训练时未见的类别 (novel class) 的样例，这种情况下贝叶斯理论不能直接使用，意味着需要全新的学习流程。而且，现有的方法无法直接对新类进行语义标注，需要人工完成。

随着今年疫情来袭，我们观察了人工智能技术在疫情防控方面的应用。尽管取得了巨大的成效，但上述的问题：统计学习模型缺乏很好可靠性，机器学习无法自动为识别出的未知类标记语义标签等问题仍然暴露了出来。例如，无法使用机器学习模型判断染病状态，因为一旦出现错误，就会造成极其严重

的后果，而且研究者也无从得知是什么导致了判断的错误；随着病毒变种越来越多的出现，需要大量人工来标注新出现的病毒数据，而无法依靠背景知识（如病毒的自身结构）使机器自助进行识别。

因此，将统计与逻辑相结合的机器学习模型成为了解决以上问题的一个潜在方案。这是一种混合模型 (hybrid model)，它拥有这样的潜力：充分利用统计学习和逻辑推理各自的优点，并相互弥补各自的缺陷。在国内外多位计算机科学家（如 Joshua B. Tenenbaum，周志华，John E. Hopcroft 等人）近期的谈话中均指出这类具有良好可理解性、可以处理感知之上更高级的认知任务的机器学习模型是未来十年机器学习的一个潜在发展方向。

面对以上的挑战和机遇，对人工智能方向感兴趣的大学生响应国家号召投入人工智能的创新势在必行，而在本科阶段就接触相关领域的研究，为之后的创新事业打下坚实基础尤为重要。然而遗憾的是，由于研究认知机器学习混合模型的研究领域是一个相对新的领域，从事相关问题研究的团队目前较少，能够为本科生提供的科研训练项目更是少之又少，华中科技大学目前也没有具有相关研究经验和具有交叉学科背景团队。由此，我们推出了 LEARN Lab (Learners' Engineering and Research Network Lab)，试图填补这个空白。

1.2 LEARN Lab

LEARN Lab 是一个面向本科生的研究团队，团队研究方向主要为统计与逻辑相结合的机器学习模型，目标找到通向人类水平智能 (Human-Level Intelligence) 的通用人工智能 (Artificial General Intelligence, AGI) 的道路。对于 AGI 的研究不能囿于传统的机器学习理论研究方式，而是要从更加高的视角——认知科学来考虑问题。认知科学是一个多领域交叉学科，包括人工智能、认知心理学、认知神经科学、哲学等方面，在 LEARN Lab 中，我们希望将机器学习的研究与认知心理学的研究相结合。简单来说，通过观察人类在一些任务上的认知学习方式，包括“人类发明新概念的过程是什么？”，“人类通过 demo 进行模仿学习时注意力在哪里？”，“人类通过什么机制认识到自己的错误？”，“为什么人非常善于发明抽象新概念，而机器则不行？”等等，帮助我们进行机器学习的创新，以此设计出能够解决高级认知问题的机器学习模型；同时，在机器学习实验过程中发现的一些现象，尝试通过开展相同任务上的人类行为实验，类比找出人类认知过程中相似的理论，以此补充和完善认知科学。总之，这是一个机器学习与认知心理学互促的研究机制。

由于这种交叉学科的研究需要大量的合作，因此为不同专业背景、不同研究兴趣的学生提供一个交流和沟通的平台尤其重要。LEARN Lab 将致力于打造满足此需求的平台 LEARN，这个概念直观地体现在团队和平台名称的“Network”（网络）中，我们的团队就像一张稠密社交网，将学生们连接起来，一起进行不平凡的工作。综上所述，LEARN Lab 目前的功能主要有两方面：

1. 通过机器学习与认知心理学互促的研究机制，寻找通向通用人工智能的道路。
2. 为来自不同专业背景，尤其是信息技术类和人文社科类专业同学间提供科研社交及创新创业平台。

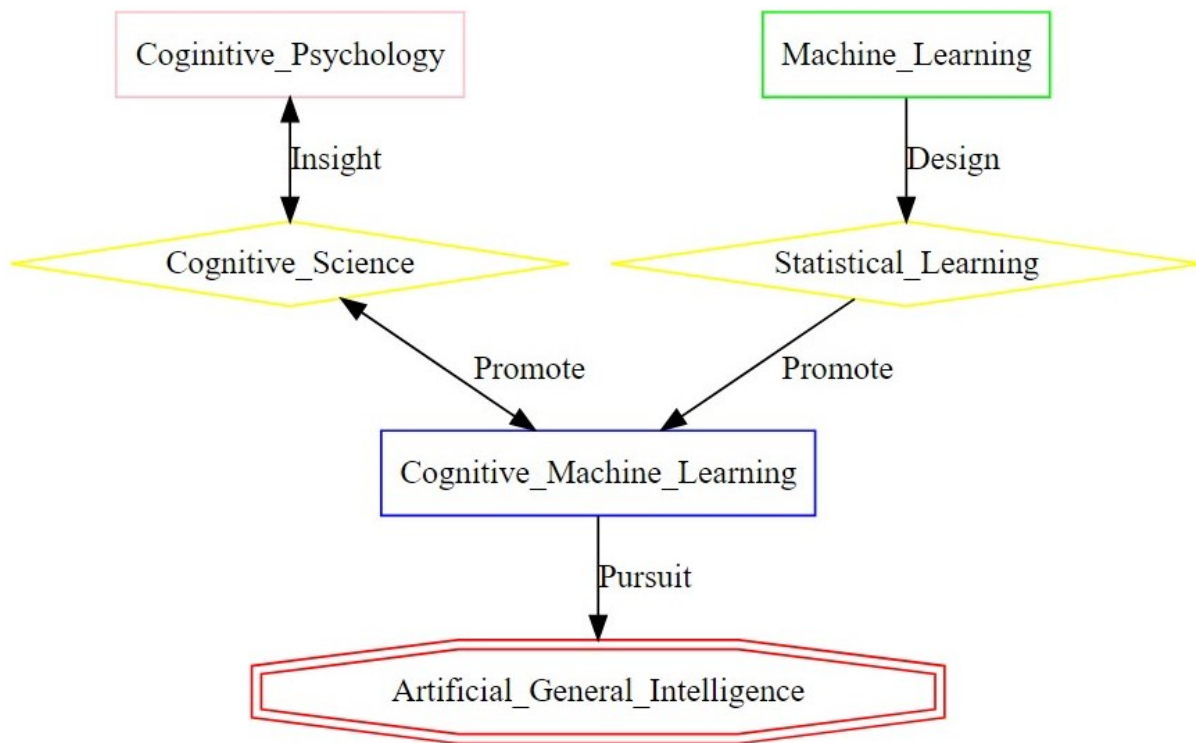


图 1.1: **LEARN Lab** 团队研究方向架构。在我们的蓝图中，认知机器学习将和认知科学的研究达到互促机制，共同帮助我们探索走向通用人工智能的道路；而理论机器学习方向则主要帮助成员进行理论学习，另外如果有成员参与相关研究也有机会为我们提供更好的统计学习模型。

第二章 LEARN Lab 团队情况综述

2.1 校方导师（荣誉导师）

华中科技大学计算机学院教授金海博士，是国际知名计算机科学家，电气电子工程师学会会士 (IEEE Fellow)，“长江学者”特聘教授，原华中科技大学计算机学院院长，华中科技大学“服务计算技术与系统”教育部重点实验室 & “集群与网格计算”湖北省重点实验室主任。他对于计算机相关行业有着深刻洞见，为团队发展道路的规划起到决定性帮助。

2.2 导师及研究方向

团队目前拥有导师三人，他们拥有不同的专业背景，分别负责指导团队不同的研究方向。这些研究方向正在逐渐构成团队规划中的机器学习与认知心理学互促研究的架构。

2.2.1 认知科学与机器学习方向

帝国理工学院 (Imperial College London) 副研究员戴望州博士 (Dr.Wang-Zhou Dai)，是一位机器学习领域优秀年轻学者，他尝试使用一阶逻辑表示的背景知识帮助进行机器学习并逐渐在混合模型这个较新的研究领域开辟出了一条道路。

戴望州博士将指导团队开展认知科学与机器学习方面的研究，试图通过理解人类认知方式帮助构建由数据与知识共同驱动的、可靠的、拥有良好的可理解性的机器学习模型。

此外，华中科技大学计算机学院教授何琨博士出于对此研究方向的兴趣，也将给予一定指导。

2.2.2 理论机器学习方向

华中科技大学计算机学院助理教授张腾博士，是一位机器学习领域优秀年轻学者，他基于“最优间隔分布学习机” (Optimal Margin Distribution Machine) 发展出了一个算法族。

张腾博士将引导团队成员进行机器学习理论学习，并指导团队开展统计机器学习方面的研究，并帮助混合模型获得更适合的统计感知模型。

2.2.3 交互设计与认知心理学方向

华中科技大学机械学院助理教授甘艳博士，是一位设计战略领域优秀年轻学者，曾参与东京奥林匹克运动会城市通用设计等项目。

甘艳博士将指导团队开展基于人类心理实验以及人机交互实验的交互心理学、认知心理学等方向的

研究，试图从认知心理学角度从更高层次理解设计学工作，并且试图获得由洞察力的认知心理学结论支持认知科学与机器学习方向组的研究。

2.3 团队负责人

团队负责人师宇哲是华中科技大学计算机系 18 级 ACM 班本科生，从 2019 年上半年就开始学习机器学习相关知识，并在 2019 年下半年开始与戴望州博士合作。师宇哲以探索走向通用人工智能的道路作为长期目标，希望能够首先在学术层面进行理论和实验的研究，在取得进展后使用成果出手解决产业界问题，为各行各业进行通用人工智能赋能。在深刻认识到个人力量的有限，研究认知机器学习对于学科交叉的迫切需要后，师宇哲于 2019 年底发起了 LEARN Lab。

第三章 团队项目

3.1 项目 1：开放世界反演绎谓词发明

目前本项目正在戴望州老师的指导下进行“在背景知识的帮助下进行未见类视觉感知并自动标记”的科研项目。

本科科研项目的启发来自于 1934 年皮尔斯 (Pierce) 提出的认知心理学过程“反演绎” (Abductive)。这个过程是，人类在观察某些事实时发现了“令人惊讶的事实” (Surprising Fact)，通过人类当前的背景知识，可能只能推出与观察事实相反的现象，而无法推出观察事实，即异常性 (Anomaly)；也可能既无法推出观察事实，也无法推出与观察事实相反的现象，即未见性 (Novelty)——人类必须为观察事实寻找一个解释 (Interpretation)，使得背景知识和解释能够推出观察事实。戴望州博士 19 年的工作 [Dai et al, 19] 提出利用诱导逻辑程序 (Abductive Logic Programming) 通过领域知识学习二进制加法规则，并对视觉感知模型的预测标签进行修改，使用修改后的标签作为监督信息重训视觉感知模型，最终实现对于手写二进制加法等式图片序列的学习。这项工作回应了反演绎过程的异常性。受到 [Dai et al, 19] 的启发，现在我的工作着重于两个方面：

- 处理反演绎学习的未见性。问题设定为我们在离线训练视觉感知模型的阶段无法得到全部类别的信息，而在在线测试过程中有可能遇到未见类别样本；这要求我们通过领域知识能够发现这些未见类别样本，并且通过领域知识推理出其符合观察事实的语义标签，为其标注后重新训练。为解决“自主学习问题”提供了潜在的解决方案。
- 将“逻辑反演绎 (Logic Abduction)”扩展至“反演绎 (Abduction)”，除了使用诱导逻辑程序进行反演绎学习的方法，我们尝试使用程序合成 (Program Synthesis) 和基于概率论的方法进行反译。

这项工作预计将在 2020 年下半年以科研论文的形式发表。

3.2 项目 2：“等车问题”的再探究：从认知心理学和机器学习的角度

目前本项目正在甘艳老师的指导下进行。

“等车问题”是一个经典的认知科学问题，它的最初形式由法尔克 (R.Falk) 提出，研究了不同情境设定下人在等车时心理概率期望 (Subjective Probability Expectation) 的上界和下界。项目组提出，在“校园校车”这种封闭的微观交通体系下，与外界公交车等待问题具有一系列区别，包括“选择代价低”、“拒绝代价高”等特征。另外，项目组分别在信息不完全/信息完全的情况下对比研究人类直接处理自己的先验估计和处理通过数据挖掘取得的后验信息对于等车心理决策判断的异同，以此获得关于人类等待心理、信息认知与处理方面的一些洞见。

3.3 项目 3：研究知识驱动的神经网络

基于连接主义和统计学习理论帮助下发展起来的深度学习一直是“数据驱动”机器学习模型的代表，只要数据足够“大”，也就是数据中实例足够多元化，标记信息足够精细，神经网络可以完成任何低阶感知任务。然而，“数据驱动”的本质和“黑箱模型”的自身缺陷使得神经网络无法进行高阶认知任务，我们也无法彻底了解其内在作用机理——这两个问题是相辅相成的。因此，作为神经网络研究中最 为高层次、也最为基础的问题之一，我们计划使用“知识驱动数据输入”，即拥有严格数学定义的视觉对象作为输入训练神经网络，以此观察它“究竟学到了什么”。另外，我们还将在“输出层”和全连接隐层之间引入逻辑背景知识，以此观察神经网络做归纳 (Induction) 和做演绎 (Deduction) 的能力，并且对比我们人为规定了语义的神经元和通过误差反向传播算法自动归纳语义的神经元权重的区别，看一看“机器学习表征”和“人类学习表征”以及“人类认为机器学习的表征”之间究竟有什么差别。

3.4 筹备中的项目：抽象与推理挑战

一项正在筹备中的科研项目是基于抽象与推理挑战 (Abstract Reasoning Challenge) 的研究。这一挑战内容为设计一种机器学习模型，能够使用三个简单样例就模仿学习出一些复杂类似“图片找规律”的任务。这项任务的复杂程度远超一般的程序合成任务，因此现有方法效果都非常差。我们认为解决这类问题应与认知心理学的研究紧密结合，通过探究人类处理类似任务时的表现，例如“人对于 demo 的注意力在哪里”，“人在模仿过程中的决策机制是如何做的”，以此对设计机器学习算法带来启发。

第四章 团队架构

团队总体分为科研部门和工程部门，这两者并不是完全分离的。科研部门为有志于从事科学研究的同学提供平台，工程部门则针对热爱开发的同学。工程部门将为科研部门提供技术支持，以及可能的科研成果转化，例如：

- 一方面，科研中需要的实验平台搭建。目前，认知机器学习组开始考虑使用交互游戏作为实验平台，通过设计游戏任务让机器学习“代理人”来解决，以其获得一些洞见。Minecraft 是一个非常适宜的实验环境，它拥有三维开放世界，并且可以任意地基于其上设计实验引擎。
- 另一方面，在科研中提出的算法需要更多的应用空间，工程组也将完成应用任务以验证科研成果的产业可扩展性。

各个项目组是科研部门和工程部门之上的虚拟组织结构，因此两个部门的任何成员都可以选择加入感兴趣的项目组。

团队目前项目 1 组有两人参与，一人负责科研一人负责工程；项目 2 组有 4 人参与；项目 3 组有两人参与，一人负责科研一人负责工程；团队目前还有后备成员两人，正在学习相关基础中。

团队设有经理一名，负责包括绩效考核、管理宣传系统、策划活动、招新在内的一切事务性工作。

第五章 团队运作规划

5.1 团队内部科研运作

在我们的规划中，在团队达到一定规模后，采取扁平化管理的机制，甚至将打破前述几个小组的界限，而是以项目为单位进行动态分组。

科研项目或工程项目的立项方式主要为：大方向由指导老师提出，由学生思考并提出相对具体的研究问题，再与老师讨论决定是否立项，并对研究内容进行把关。随后，提出具体研究问题的学生为项目组长成立项目组。项目组的选人遵循“双向选择”策略，由组长根据制定的研究计划提出能力要求并将能力要求和项目简介发布为所有人可见的信息，所有团队成员可以根据自己的兴趣、能力特点和时间规划酌情选择项目。对于项目组的要求如下：

- 每个项目组人数不宜超过三人；
- 组长负责对组员的管理，从招揽组员，到工作指派、进度控制等；
- 组长有权在合理合情的情况下提出更换或裁撤组员；
- 组长为研究项目负责，包括因各种原因导致的项目推迟或失败。

团队定期（初定为每周）召开交流研讨会，内容包括两方面：

- 所有成员轮流为大家讲解老师指定阅读的论文。
- 所有研究组介绍工作进展。

5.1.1 关于例行讨论班的细则

1. 开展频率每周一次，根据考试情况可在学期末阶段酌情调整。
2. 需要在 seminar 时脱稿情况下说出所学习论文的创新点及核心算法，其他内容依据论文来转述。具体标准是：使用幻灯片或思维导图等，讲述给具有基础能力的成员（在“2”中规定）听，能否给这一类（一个）听众讲明白，或者能否展示清楚，就是衡量的标准（在“3”中规定）。
3. 将在 seminar 之前一天决定参与监察的听众，“具有基础能力的成员”可以灵活决定，这个决定基于论文涉及的知识面，“听众”至少需要对所讲的基础知识有广泛的了解，否则难以辨别讲述内容。
4. “标准”是指某一次衡量，但不是检验成员成果的决定性因素，在此设定为：最多容忍 2 次不达标次数。

5. 在每个演讲的成员讲完之后，将接受“听众”提问，问题将从算法完备性、成果应用层面的展望等类似方面提出，不需要演讲者给出所有相应的准确回答（且可能不存在确切答案），但需要在回答中体现自己科研学习或者思考的过程。
6. 鼓励非监察听众热烈参与讨论，力所能及地提出见解和问题。

5.1.2 关于以小组单位制开展科研的细则

1. 优先选任有开发或者研究经验的成员担任组长，主要以项目为单位分组。
2. 每个组长与导师讨论后定题，然后在群里或者联系团队经理在 LEARN Lab 网站上贴出摘要，其他成员报名即可。
3. 每个项目组人数不宜超过三人（含组长），人满即调剂。
4. 组长负责对组员的管理，从招揽组员，到工作指派、进度控制等。
5. 组长有权在合理合情的情况下提出更换或裁撤组员。
6. 组长为研究项目负责，包括因各种原因导致的项目推迟或失败。
7. 在例行讨论班中，由于时间原因没有完成但需要及时完成的演讲任务，可以通过其他小组组长进行检查。

5.2 团队平台运作

团队建立官方社区 LEARN(Learners' Engineering and Research Network)，为不同专业背景之间的学生提供学术交流平台。初步计划暂时设立两个社区板块，分别为创新和创业：

- 创新：主要在科研创新方面。“学科交叉”主要以如下两种形式体现，出现在进行课题研究、论文撰写等场景：
 - 利用本专业方向知识和方法论出手解决对方专业问题。例如，用人机交互实验进行一些社会心理学问题的研究（人机交互出手解决社会学问题）；使用某些哲学方法论研究某种人工智能工程的道德可靠性（哲学出手解决人工智能问题）。
 - 在本专业方向研究中利用了对对方专业知识，寻求理论保障。例如，在反演绎学习的研究中寻找“诱导”和“科学知识论”相关概念的理论保障（机器学习利用哲学概念）。
- 创业：主要在创业思路和创业落地两方面，出现在组队参加互联网+、微软创新杯 (Imagine Cup) 等创业比赛场景，主要体现为商业计划的制定和落地都需要双方专业背景共同合作完成，即专业背景具有互补性。

第六章 团队发展规划

6.1 短期规划

团队在 2020 年初成立，还没有经历过 9 月份新生入学的招新高峰期。因此，今年 9 月份招新季必须认真把握。在宣传上，我们认为传统的“扫楼”式宣传效率很低，对应地，应充分利用线上下活动相结合的方式开展宣传。例如，可以通过介绍一些有趣的认知科学事实、认知心理学小实验、程序 demo 引起学生对于认知机器学习的兴趣；线上线下的科普讲座和学术研讨会则可以充分发挥团队的优势，进一步加深学生对团队的印象；此外，如果目前进行的科研项目能够有结果产出，还可以以此为基础开展学术交流会和 Tutorial。作为一个新领域的科研兴趣团队，宣传的目的从来不是为了招到更多的学生，而是使更多的学生了解我们的团队，从而有更大概率招到不同专业背景的优秀学生。

6.2 长期规划

作为“Network”特性的团队，我们致力于在较长的时间后能够建立一个跨学科领域、跨学校的科研社交网络。虽然在早期只能借助指导老师的资源，但长期来看 LEARN Lab 招入的优秀学生毕业后通常会选择深造（攻读博士学位），在这种情况下，他们就成长为可以定期的为团队成员进行教学的资源；当他们毕业后则成为团队的学术界或产业界资源。

在科研方面，尽管不一定团队每个成员都能得到理想的科研产出，但我们希望通过“以点带面”的方式，通过表现突出的个人带动团队进步；此外，特别注重科研方法、规范和创新创业思维能力的培养，争取将 LEARN Lab 打造出品牌效应，学术界和产业界人士会对 LEARN Lab 有这样的认识：“他们能够做出引领潮流的工作”；还会对 LEARN Lab 团队有这样的认识：“那里出来的学生不仅科研基础训练非常到位，而且很有想法”。我相信通过长时间的持续努力，LEARN Lab 能够发展成华中科技大学的一张“靓丽名片”，也能够发展成认知机器学习领域的“黄埔军校”。

第七章 其它说明

在本部分集中说明一些可能存在的、容易误解的概念。

- 团队性质：LEARN Lab 是一个学生自主管理的科研创新组织，以学生自主探究与导师指导相结合开展科研活动。目前团队活动资金来源是依托系列课题申请的大学生创新项目的资金支持。
- 导师与团队的关系：值得指出的是，四位导师中除了作为荣誉导师的金海老师，其余三位导师实际指导科研工作。他们与团队是合作关系，导师之间互不干涉，导师不过问团队的运营情况和事务性工作，导师不指派学生做与指导课题无关的工作，只是为成员探究课题提供建议和帮助。
- 团队成果产出的贡献划分：对于合作完成的科研论文/项目，均已贡献程度最大的成员所在单位作为第一单位，对于论文而言，只要第一作者是 LEARN Lab 成员，论文第一单位就是 LEARN Lab，此外对论文工作有一定贡献的成员都将按贡献由大到小的顺序挂名，指导老师可以挂通讯作者；对于项目而言，除了考虑贡献外，还要考虑支持经费主要来源，通常支持经费主要来源单位也会作为第一单位。