

LEARN Lab 团队简介

师宇哲, LEARN Lab 负责人

2020 年 8 月 5 日

摘要

LEARN Lab(Learners' Engineering and Research Network) 是一个以认知机器学习为主要研究方向的研究团队, 由帝国理工学院 (Imperial College London) 的研究员戴望州博士 (Dr.Wang-Zhou Dai) 指导, 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心理学互促结合研究, 找到通向通用人工智能 (AGI) 的道路。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。

1 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相关的团队。由此，我们推出了 LEARN Lab (Learners' Engineering and Research Network)，试图填补这个空白。

1.2 LEARN Lab

LEARN Lab (Learners' Engineering and Research Network) 是一个面向本科生的研究团队，由帝国理工学院 (Imperial College London) 的研究员戴望州博士 (Dr. Wang-Zhou Dai) 指导。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？”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，类比如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 致力于打造这样的平台，这也体现在团队名称的“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：

- 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。

- 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平台。

2 LEARN 团队情况综述

2.1 导师

帝国理工学院 (Imperial College London) 的研究员戴望州博士 (Dr.Wang-Zhou Dai)，是一位机器学习领域优秀年轻学者，他在博士期间开始尝试使用一阶逻辑表示的背景知识帮助进行机器学习，在 AAAI, NIPS, IJCAI 等人工智能类顶会发表多篇论文，逐渐在混合模型这个较新的研究领域开辟出了一条道路。其目前的研究兴趣正是认知科学与机器学习。

2.2 校方导师

华中科技大学计算机学院的教授金海博士，是国际知名计算机科学家，电气电子工程师学会会士 (IEEE Fellow)，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。

2.3 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。他的理想是在未来找到通向 AGI 的道路，为此，他从 2019 年上半年就开始学习机器学习相关知识，并在 2019 年底与戴望州博士开始合作。目前正在戴博士的指导下进行“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视视觉感知模型，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。这为解决“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。

除此之外，另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。

2.4 团队成员

团队目前规模不大，但胜在精悍、分工明确。内部架构包括认知组、视觉组和工程组，心理组还在筹备中。目前认知组两人，视觉组两人，工程组两人，共有六名核心成员。此外，在暑期学习活动中吸纳了两位新成员，目前他们还在学习相关知识。

认知组是团队核心部门，负责认知机器学习的研究，目前在研项目一项；视觉组负责实验内容相关的底层视觉任务的研究，目前在研项目一项；工程组负责将团队研究内容进行实际应用的尝试，目前在参加的大创项目两项；心理组在筹备完成后将负责开展与认知组相对应的人类实验。

3 团队运作规划

3.1 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。

3.2 团队平台运作

团队将在校内有影响力的平台上建立官方社区，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。

4 团队发展规划

4.1 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。

4.2 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：“那里出来的学生科研基础训练非常到位”。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”。