

# 教育部司局函件

## 教育部科技司关于组织开展人工智能算法战略研究工作有关要求的通知

各有关高校：

为落实中央领导同志批示精神，充分发挥高校在人工智能算法方面的优势，我司拟布局若干“人工智能算法战略研究项目”，引导优势高校围绕关键核心算法开展研究。结合优势高校报送的项目建议和专家论证意见，初步拟定了项目方向和委托单位。现将具体事项和要求通知如下：

1.采取定向委托方式，拟委托南京大学、西安交通大学、北京大学、清华大学、西安电子科技大学、浙江大学、复旦大学、同济大学等8所高校牵头开展相关算法战略研究。

2.请各牵头高校务必高度重视，进一步组织吸纳全国高校优势力量参与攻关，并确定专人做好全过程管理工作。每一个项目必须明确研究团队具体人员和单位（包含责任人与参与人），同一人只能参加一个项目。

3.各项目团队要在开展关键核心算法研究的基础上，兼顾关键核心算法战略规划研究，并面向“科技创新—2030 重大项目”或“十四五”重点研发计划、基金项目等，提出关键核心算法相关的重大项目建议。

4.项目不以人才培养、论文产出等作为考核目标，重在考核关键核心算法研究和项目建议的凝练。同时，为证实算法的有效性，项目研究可开展必要的应用验证，但不作为研究重点。

请各牵头高校于9月22日前将修改完善后的任务书（一式一份，加盖学校公章）报送我司高新处。我司将于9月底组织开展答辩，通过答辩的予以立项支持（每项经费不超过250万元）。

附件：1.人工智能算法战略研究项目指南

2.人工智能算法战略研究项目任务清单

3.人工智能算法战略研究项目任务书（模板）

联系人：刘法磊 蔡固顺

联系电话：010-66096733 66092093



2019年9月17日

## 附件 1

# 人工智能算法战略研究项目指南

## 方向一：大数据分析的核心关键算法规划研究

在开放动态的实际大数据应用环境中，以往基于经典统计理论的大数据分析方法理论假设不再成立，面临海量实时、监督信息低质、学习要素多变、事件类别不平稳等新的挑战，独立的大数据分析过程也难以适应人机物交互频繁的趋势。该方向拟针对大规模、强鲁棒、弱监督、以及交互式等方面开展适于实际应用的大数据分析新算法。针对大规模数据，研究大数据在线学习，在样本快速积聚与更新的大数据应用中能高效利用样本提升泛化性能的机器学习理论，发展提升学习效率的在线分布式基础算法，在资源受限条件下有效利用大数据。针对强鲁棒性需求，研究鲁棒机器学习理论与方法，在数据分布等学习要素发生变化后仍能保持较稳定泛化性能，发展模型可演进的多目标学习方法。针对弱监督特性，研究弱监督机器学习理论与方法，在包括小样本等监督信息不足条件下仍能获得较强泛化性能，发展高效利用小样本、不完备标记、噪声标记数据的学习方法。针对交互式特性，研究能有效利用环境交互来提升与保持泛化性能的机器学习理论与方法，发展交互式强化学习、人机协作学习方法，增强模型可理解性、可信性、隐私保护与攻击对抗能力，保证学习过程的高效性、安全性与可靠性。

预期目标：针对开放动态环境大数据海量、低质、动态等特点，发展与之匹配的大数据分析理论与核心算法，提出一套大数据分析学习理论，构建不少于 4 种新型学习算法，推出两个以上大数据分析开源系统/工具，初步建设大数据分析平台。

建议优势团队：南京大学、西安交通大学、复旦大学、中山大学

## 方向二：机器学习自动化的核心算法规划研究

当前机器学习算法的成功应用依赖大量的人工工作，包括预订训练数据、预设深度网络结构、预制机器学习模型等一系列的手工选择，并且选择结果在不同的应用问题上难以复用，严重阻碍机器学习技术的落地。本方向拟从数据样本、学习过程（模型算法）、环境任务三个层面提出创新性与系统性的机器学习自动化方法与技术，面向大范围、跨任务的应用牵引，通过理论基础的界定机器学习自动化的理论边界，发展训练样本自选择、损失度量自调节、正则化项自设定、学习模型自构建、学习算法自设计、富经验自重用、跨环境自适应的新型机器学习方法，最终形成“机理启发、模型驱动、知识嵌入”的机器学习自动化新理论与新算法。

预期目标：在机器学习自动化理论与算法方面取得突破，研制面向“数据样本自生成/自选择、模型算法自构建/自设计、环境任务自适应/自转换”三类任务的不少于 5 种高效算法，支持盲图像去噪、鲁棒深度学习、弱监督/无监督目标检测等不少于 3 种

典型应用任务，并在智能通信或智慧医疗等实际领域尝试研制方法的示范应用。

建议优势团队：西安交通大学、南京大学、清华大学

### **方向三：深度神经网络的优化算法规划研究**

深度学习已经在计算机视觉、自然语言处理等应用领域取得显著进展。但是深度学习仍然在深度神经网络的构架设计和表达能力、深度学习算法、深度学习的泛化能力等基础理论方面存在很大不足，设计合适的网络结构和参数调整通常需要大量的人力试错，耗费大量的数据资源和计算资源；另外，深度神经网络通常被用作“黑盒”模型，模型和算法的可解释性弱。该方向拟通过充分利用数学工具（如概率论、随机理论、函数逼近、流形嵌入、微分方程等），开展深度学习的理论分析（如泛化性、稳定性、能量景观等），探究深度神经网络的模型结构与模型特性（如稳定性、泛化能力）之间的联系，指导设计具有良好理论性质（如全局最优、收敛性质好、通讯代价低等）的新型学习算法。

预期目标：结合深度神经网络的结构，通过模型和算法的融合，发掘问题的内在特点，在深度神经网络架构设计和高效学习算法方面取得突破，提出不少于3种具有良好理论性质的深度神经网络模型和算法。

建议优势团队：北京大学、复旦大学、清华大学、上海交通大学、武汉大学

### **方向四：深度概率建模及高效推理算法规划研究**

由于环境噪声、数据缺失、信息不完整、对抗攻击等因素的存在,人工智能系统需要具备强大的不确定性建模和推理的能力。贝叶斯方法提供了一套具有坚实理论基础和强大功能的建模语言,通过贝叶斯推理将知识与数据有机融合,在小样本学习、不确定性推理、决策等方面具有显著优势。该方向将融合贝叶斯方法与深度神经网络的互补优势,探索二者有机融合的机理,发展面向无监督学习、小样本学习、对抗学习等场景的深度概率建模理论框架及高效推理算法,研制支持深度概率模型的编程语言和高效算法库,支持人工智能的高效环境感知与决策等任务。

预期目标:在深度概率建模理论和算法方面取得突破,研制面向不确定性建模和推理的高效算法不少于3种和1个深度概率编程算法库,支持小样本学习、无监督学习、对抗鲁棒学习等不少于3种典型任务。

建议优势团队:清华大学(人工智能研究院)、西安交通大学、北京大学

## **方向五:演化认知深度学习的核心算法规划研究**

传统深度学习方法通常基于封闭环境与理想标注假设,其应用性能严重依赖决策场景与训练场景的一致性、专家对问题定义的合理性,以及所选择模型的适应性等。现实人工智能问题面临复杂未知的非结构化场景,环境数据在规模/性态/复杂度等方面快速变化,给深度学习技术的应用提出了挑战。深入研究人脑感知、认知、自学习等功能并以机器来实现,让计算机能像人一样



具有连续认知环境的能力，是发展下一代深度学习需要迫切解决的问题。该方向拟针对当前深度学习技术的局限，开展适于复杂未知环境应用的演化认知深度学习技术研究。借鉴生物大脑的信息高层感知机制和认知机理，以及自然演化过程中的进化学习等特点，围绕数据、模型与算法三个方面，发展“自主感知--稀疏认知--演化进化”为一体的演化认知深度学习新理论与新方法，探索自动化机器辅助标注、自动化模型架构搜索、自动化开集优化学习等新技术。

预期目标：针对复杂未知环境下的深度学习应用，发展与之匹配的自动化机器辅助标注、自动化模型架构搜索、自动化开集优化学习技术与核心算法，提出一套“自主感知--稀疏认知--演化进化”为一体的演化认知深度学习新理论，构建不少于 3 种新型学习算法，推出 2 个以上自动化深度学习开源工具。

建议优势团队：西安电子科技大学、华南理工大学、南京大学、南开大学

#### **方向六：脑启发机制的人工智能模型与算法规划研究**

为构筑我国人工智能发展的先发优势，积极响应党中央、国务院号召和部署，围绕新一代人工智能发展规划，借鉴脑神经结构及其信息处理机制，本方向旨在研究脑启发机理下人工智能新算法。研究神经突触可塑性模型和可计算类脑元学习神经网络模型，提升人工智能算法的知识抽取和泛化能力。针对当前“冯·诺依曼”传统计算架构对人工智能算法的限制，研究存储与计算深

度融合的近内存和存内计算等算法。针对通用处理器芯片与类脑人工智能模型失配的问题，研究基于新型神经形态器件的智能算法，大幅提升计算性能并降低能耗，同时提高集成度与灵活度，最终迈向通用人工智能。

预期目标：提出神经突触可塑性学习、存储与计算深度融合的近内存计算和存内计算、新型神经形态器件的类脑模型等算法；研发一整套相关算法和工具包，面向重大应用对算法验证，形成本方向核心算法规划建议。

建议优势团队：浙江大学、中科大、清华大学、北京大学、复旦大学

## **方向七：人机协同混合增强智能算法规划研究**

当前人工智能算法最大的挑战是在具有不确定性、动态性和开放性条件下仍然具有鲁棒性。传统人工智能中的机器学习算法过于依赖规则，可移植性和可扩展性较差，只能在约束条件紧、工作对象少的简单环境下工作，无法适应动态、非结构化和非完整信息处理的需求。尽管我们可以构建由人工神经网络、模糊推理和近似推理等方法组成混合计算智能系统，克服个体的局限性，但依然无法有效解决常识知识和随时间变化的复杂动态问题，也无法做到使用过去的经验来有效指导未来的决策。而大脑自主信息处理的信息多为动态、复杂非结构化数据。本方向拟通过借鉴人类大脑的工作机制，提出一种新型的基于认知计算的人机协同混合增强智能理论与方法，构建相应的可计算模型，来解决复杂、



动态、不确定性环境下的人工智能难点问题。本方向拟重点研究开放环境中不确定条件下的推理，开放环境中的智能认知与行为，建立新的人机混合增强智能理论与方法，提升记忆与交互、自主学习以及与环境/场景结合紧密的智能能力等。通过直觉和因果推理来实现非确定条件下自主智能；通过在线知识演化与推理方法实现场景适应多任务的人工智能；通过人在回路的多异构机器人协同认知计算实现人机多任务自适应重构协同智能。

预期目标：在面向不确定性和脆弱性环境下的基于认知计算的人-机协同混合智能理论与方法方面取得突破，研制面向开放环境中不确定条件下的直觉和因果推理，适应多任务多场景的在线知识演化、人在回路的增强智能学习等人机协同认知等三类任务的不少于四种高效算法，并在实际领域对算法进行验证，形成本方向核心算法规划建议。

建议优势团队：复旦大学、西安交通大学、中国科技大学、浙江大学、同济大学

#### **方向八：多智能体自主协同与进化算法规划研究**

自主智能无人系统作为新一代人工智能核心技术与基础，属于国家竞争的战略科技，我国在基础理论与核心算法等方面与国外仍有较大差距。为建立我国人工智能战略优势，围绕国家下一代人工智能发展规划，本方向旨在建立自主智能体在复杂环境和条件下发育、协同、进化的新理论与新算法。针对当前自主智能体面对复杂环境的脆弱性问题，研究自主智能体仿人感知认知基

基础理论和算法，建立复杂环境下智能体多任务长期自主行为进化算法，建立自主智能体的拟生命发育体系；针对当前多智能体面对突变环境协同优化能力不足的问题，研究自然群体系统中个体交互模型和表示，突破多自主智能系统传统规划决策模式，提出可表达、可计算、可解释的群体智能算法；针对自主智能体面临的安全性和防欺骗性问题，研究自主智能系统的协同进化与安全机制模型，建立复杂网络环境下自主智能系统的分布式协同稳健控制理论和算法，实现多智能体系统自主协同与稳健运行。

预期目标：研究进化、博弈、协同和控制等方面的算法，研发一整套相关无人系统算法和工具包，初步形成自主智能无人系统开源开放框架和平台，形成本方向核心算法规划建议。

建议优势团队：同济大学、北京理工大学、浙江大学

人工智能算法战略研究项目任务清单

方向	算法名称	预期创新点	预期关键指标
方向一	面向资源受限条件的大数据在线学习算法	突破以往在线学习算法适应范围窄、缺乏对适用性进行评判的局限，提出能够适应多种环境、并具备“自省”性的在线学习算法。	提出的新型算法能适应凹凸、强凸、指数凹等多种函数；能够识别超出自身适用性的数据，识别率高于90%。
	面向开放动态环境非独立分布数据的迁移学习算法	突破以往迁移学习依赖足够的目标环境数据的局限，提出能够识别环境变动、感知环境参数的算法，并通过环境参数迁移适配历史模型。	提出的新型算法对环境参数的识别，在10次交互感知的条件下误差小于10%，模型迁移适配的准确率不低于最优模型的90%。
	面向开放动态环境低质标记数据的主动学习算法	突破以往标记数据缺乏时训练数据的构建依赖大量人工标注的被动局面，提出高效主动学习算法，对少量不确定数据发起询问，以获得大幅度能提升。	提出的新型算法能够准确识别需询问的低质数据，识别正确率高于90%，相比全量数据标注，节省数据标注量超过60%。
	面向开放动态环境多步决策数据的高效强化学习算法	突破以往避免模型强化学习算法样本利用效率低下、而基于模型强化学习算法准确率低下，导致无法应用于实际问题的局限，提出高准确率的基于模型强化学习算法。	提出的新型算法能从真实业务数据中有效还原环境模型，还原准确率高于90%；在实际业务场景中验证有效性，性能超越以往监督学习算法。

方向二	基于自适应课程学习的样本选择算法	突破现有样本选择策略需要人工预制标准，人工预设超参的局限，借助人类接受教育/课程设置原理与人类认知模式启发，实现自动生成标准、自动设定超参的样本自选择策略。	能够针对各种数据偏差情形（如噪声标注、类不平衡等）实现有效样本自动选择，在3个公开数据集上达到最佳学习样本选择效果，学习精度提升10%以上。
	基于模型驱动的网络结构构建与优化算法	突破目前人工预设优化算法与启发式构建网络结构的局限，通过建立网络多层结构与算法多步迭代的本质关联，实现模型驱动的网络结构构建策略，进而实现针对测试数据可直接调用的优化算法自形成策略。	能够在3个以上计算机视觉的典型任务上超过现有国际最好算法的性能，达到state-of-the-art的计算效果，网络参数量比现有网络减少50%以上。
	基于数据环境引导的误差项与正则项自构建算法	突破当前优化目标函数需要人工假定误差项与正则项格式的局限，借助代表算法执行环境的数据与知识作为引导，采用模型参数与环境交互适应的学习策略，实现针对机器学习目标函数误差项与正则项的自构建算法。	针对2个以上典型模式识别与信号处理问题任务，在相比传统方法减少50%以上调参工作量的前提下，达到或超过已有最优模型的计算性能。
	适用于复杂非i.i.d.分布数据的自更新学习模型自更新算法	突破目前机器学习方法需要人工预设数据i.i.d.分布的局限，通过自动捕获环境中存在的各类相关性信息（如时空谱维等），实现概率模型自动感知并编码复杂非i.i.d.分布数据的鲁棒性学习模型自更新算法。	在两个以上的典型非i.i.d.数据分布场景实现鲁棒学习效果，精度超过现有最佳方法10%以上，在1个实际应用场景（如无线通信）尝试落地实用。
	基于自适应课程	突破现有样本选择策略需要人工预制标准，人工	能够针对各种数据偏差情形（如噪声标注

方向三	学习的样本自选择算法	预设超参的局限,借助人类接受教育/课程设置原理与人类认知模式启发,实现自动生成标准、自动设定超参的样本自选择策略。	注、类不平衡等)实现有效样本自动选择,在3个公开数据集上达到最佳学习样本选择效果,学习精度提升10%以上。
	基于微分方程的可解释学习算法	为分析神经网络模型提供数学工具,理清深度神经网络的机理。	深入研究深度神经网络中关键决策信息的传播路径和传导机制,建立有效的理论框架描述相关机制;构建神经网络的偏微分方程模型。提供不少于2个典型应用。
	数据驱动和机理驱动融合的算法	为实际问题提供数据和机理融合计算的指导思想 and 理论基础,在保证算法精度的同时提高算法的可解释性。	发展数据和机理的互补与融合的可计算建模方法论。提供不少于2个典型应用。
	高度非凸非光滑问题的算法	突破非凸、非光滑、维数高等带来的困难,发展新的准确、快速、实时高效的算法,建立算法收敛性条件。	构造高效一阶和二阶信息相结合的算法,在泛化性能保持的前提下,显著提高训练速度。
	通信有效的分布式机器学习算法	提出监督学习的分布式(统计)算法、深度网络的分布式算法和混合模型的异质学习算法,建立局部和全局的算法收敛性理论	开发1套自主版权的分布式计算基础软件包。
	贝叶斯深度学习高效算法及概率编程算法库	应对深度学习的复杂函数变换,结合变分优化的长处,设计高效的蒙特卡洛采样算法;融合传统概率编程和深度神经网络编程的优点,支持贝叶斯深度学习模型和算法的快速编程和实现。	显著随机梯度推断算法的收敛效率和精度;概率编程算法库集成支持不少于10种贝叶斯深度学习模型和算法。
方向四			



	基于贝叶斯深度学习 的持续学习 算法	利用脑启发机制，克服机器学习算法普遍存在的遗忘灾难问题，研制数据高效的持续学习算法。	在分类任务逐渐增加等持续学习场景下，显著优于已有算法。
	对抗鲁棒的深度 学习算法及对抗 攻防算法平台	设计有理论保证的对抗鲁棒学习算法；以及研制对抗攻防高效的对抗攻击算法。	在对抗噪声存在的情况下，比已有算法显著提升识别精度。
	深度学习的自适 应推理算法	从样本自适应、空间自适应和时间自适应等角度研究自适应推理算法，探索与之相适应的网络模型和推理机制。	将算法应用于图像分割、物体检测、视频分析等视觉任务，取得显著效果。
	场景自主感知算 法	面向场景变换，提出数据主动“感知-筛选-标注”的自主演化增强算法，从数据层面形成智能自主感知发育。	形成一套数据自主感知算法，实现无需人工标注的机器自主学习与目标任务集扩展，对不少于5类场景目标形成有效的数据自主感知。
方向五	开集渐进认知算 法	面向任务演变，提出“闭集任务--度量增量学习--开集任务”的开集渐进认知算法，实现变化任务下的自主学习。	形成一套深度神经网络开集学习算法，支持开放环境下的新目标自动聚类、分类、识别与连续学习，实现任务变化时深度神经网络的自主学习与认知。
	网络演化优化算 法	面向资源变化，提出基于多目标协同动态优化的深度网络结构自演化算法，使深度网络适配开放环境下的动态变化资源。	形成一套高效的深度学习模型自主演化优化算法，支持分类、识别等典型任务下的网络自动演化，支持大规模分布式实



			现，比优化前模型性能提升不低于 20%。
方向六	基于神经突触可塑性的元学习算法	探索脉冲时间相关的突触可塑性规则，研究突触可塑性调控机制，提出可计算类脑元学习神经网络算法。	实现一套基于突触可塑性的元学习算法工具包。
	近内存计算和内存计算算法	模拟人脑的信息处理方式，研究基于存储与计算深度融合的内存计算算法，提升人工智能系统的计算效率。	实现存储与计算深度融合的内存计算算法原型。
	支持新型神经形态器件的类脑算法	研究基于新型神经形态器件的类脑智能算法，算法自适应能力强。	提出和实现类脑芯片上运行的原型算法。
	直觉推理算法	通过在线决策过程构建直觉模型和直觉推理算法，有效识别场景中的物体及姿态。	相比现有算法，显著提升对动态场景的检测（10%以上），并提升高实时性。
方向七	因果推理算法	根据相关知识进行有序的、程序化的逻辑响应，建立因果推理算法，提取场景中物体间作用关系及语义结构，理解行为和意图等高层语义信息。	在 5 个以上的知名公共数据库上，提升对动态场景理解与推理性能。
	在线知识演化算法	结合新任务对已有知识进行推演，从而学习新知识的过 程，构建在线知识演化算法，提升模型泛化能力和多任务适应能力，并使机器人适应多任务。为解决学习新知识而对已有知识灾难性遗忘	能够在 5 个以上的大数据公开数据库上进行任务自适应。

		的问题，项目基于可微神经网络对内存进行管理，从而对复杂数据结构进行存储。	
	人在回路的增强智能学习算法	通过机器人的主动询问式学习，以人类专家的经验，通过模仿学习（连续学习、示教学习、迁移学习和强化学习等）帮助机器人获取经验和知识，实现智能体在线学习算法。	研制人在回路的增强智能学习方法，提升人工智能算法的鲁棒性与可迁移性。
	动态自适应人机协同算法	基于增强智能学习、示教学习等构建一种新的人在回路的在线学习算法，实现人-机器人在线自适应协同。并利用人对不确定因素的判断更新可计算模型中的状态和模型参数，从而达到任务分解、多异构机器人自适应任务重构与协作等目标。	在 2 种个以上的典型应用任务问题上达到在线人机协同增强智能效果。
方向八	基于主动感知的自主智能体行为进化算法	提出基于感知大数据的自主智能体仿人感知理论，建立基于类脑思维能力的知识模式自辨别、处理算法自优化、思维模型自迁移，形成拟生命发育体系	构建一套针对复杂环境的自主智能体长期自主行为进化算法，提升现有自主智能体在开放环境中的适应性
	自主智能系统的分布式优化与博弈算法	提出自主智能系统一体化任务规划架构，建立任务驱动层次化交互学习与优化模型，形成任务动态博弈分析、机制设计与学习能力	发展自主多智能体系统交互学习与优化理论；构建面向多任务的动态博弈理论框架。提供不少于2个典型应用。

	自主智能系统的自主协同与稳健控制算法	提出基于系统逻辑和监督控制的自主协同与稳健安全的分布式控制方法，实现多任务耦合多集群智能涌现，非合作集群演化的博弈与进化	搭建集群行为的协同进化平台，实现协同作业效率提高20%以上，并可检测、可抵御DoS攻击、欺骗攻击等网络攻击
--	--------------------	--	---

# 人工智能算法战略研究项目任务书

项目名称: \_\_\_\_\_

牵头单位: \_\_\_\_\_

负 责 人: \_\_\_\_\_

联系方式: \_\_\_\_\_

中华人民共和国教育部

2019 年 9 月制

# 参考大纲

- 一、 概述.....
- 二、 研究目标.....
- 三、 基础性问题.....
- 四、 项目总体规划和任务分解 .....
- 五、 研究成果及应用前景 .....
- 六、 研究周期、阶段划分和经费安排 .....
- 七、 项目负责人和研究团队 .....
- 八、 研究基础与保障条件 .....