

附件一：2022 CCF-百度松果基金申报主题

目录

1. 开放课题.....	1
1.1 深度学习基础理论和技术.....	1
1.2 深度强化学习/进化学习技术.....	2
1.3 深度学习框架技术.....	2
1.4 自然语言处理.....	4
1.5 计算机视觉.....	4
1.6 智能语音.....	5
1.7 个性化推荐.....	6
1.8 AI+科学计算.....	7
2. 半开放课题.....	7
2.1 推理部署关键技术研究及开源实现.....	8
2.2 深度学习框架评估评价体系研究.....	8
2.3 千言 - 面向中文自然语言处理的开源数据集建设和方法研究.....	9
2.4 机器翻译关键技术.....	9
2.5 跨模态对话.....	10
3. 领域课题.....	10
3.1 小度人工智能助手中的复杂语义理解和对话.....	10

2022 CCF-百度松果基金申报主题

(以下主题均不限于给定的建议研究方向，可基于研究者背景及兴趣确定。)

1. 开放课题

1.1 深度学习基础理论和技术

本课题期望通过对深度学习基础技术、基础理论层面的研究，来指导解决当前深度学习研发和应用上的挑战性问题。侧重对实际问题更有直接指导意义的或者对未来技术发展有引领意义的基础研究。

建议研究方向：

- 1) 优化算法和收敛性研究：**研究能加速深度学习训练收敛的更好数值优化策略；研究大 batch 训练下收敛性的保证问题，以支持更大节点数的分布式训练；研究大规模异步并行训练的收敛性问题及对应新优化策略，提供对异步训练效果的更好理论保证，并拓展至更多场景。
- 2) 可解释性和鲁棒性研究：**研究深度学习模型的可解释性问题，更清晰定位数据、特征和网络结构对效果的影响，并指导模型设计和训练可视化工作；研究深度学习训练的鲁棒性问题，包括对抗样本攻防等，增强深度学习应用的安全性；研究语义分割、人脸识别、3D 检测、OCR 等具体领域的鲁棒性评测方法和检测标准。
- 3) 低精度训练和量化训练收敛性研究：**研究 FP8/FP16/BF16 低精度训练对收敛性的影响，指导低精度训练策略的应用；研究整型 INT8/INT16 量化训练的方法，实现训练加速。
- 4) 模型小型化技术研究：**研究包括稀疏化、自动剪枝、低精度量化、蒸馏技术以及模型结构自动搜索在内的模型压缩技术，保持模型精度的同时使模型体积更小计算更快。
- 5) 多模态、多任务学习研究：**研究深度学习表示和抽象能力的底层机理，在多模态、多任务学习上给出更好学习方式。
- 6) 深度学习和知识融合研究：**研究图神经网络，提升对结构化数据和知识的表达和处理能力；探索连接主义和符号主义结合的新方式，研究深度学习技术和知识融合的基本原理和新方法，以及利用知识提升深度学习模型的推断能力。
- 7) 通用预训练和迁移学习研究：**研究通用预训练模型的作用原理和优化提升策略，以及对应的迁移学习机理和新方法。
- 8) 无监督、弱监督学习研究：**降低对大规模标注数据的依赖，研究对弱标注数据更好的利用方式，研究自我监督学习的能力空间，探索新的无监督、弱监督学习方法，探索摆脱梯度反向传播的学习范式。

[返回目录](#)

1.2 深度强化学习/进化学习技术

强化学习指的是智能体通过和环境不断进行交互，根据反馈提升自己的决策能力，从而获得更高的奖励。强化学习自 2014 年与深度学习结合以来，已经在多个决策领域取得了显著的提升效果，包括围棋、机器人控制、多智能体协同、推荐系统优化等领域。本课题期望开展强化学习和实际应用场景相结合的研究，解决不同应用场景中的关键技术难题，以及降低强化学习应用门槛、提升实用性方面的共性问题研究。

建议研究方向：

- 1) **强化学习应用于交通信号灯控制：**研究如何设定合理的单步奖励、路网信息表征等以提升信号灯控制效率。
- 2) **强化学习应用于能源调度：**研究如何通过强化学习提升电网系统的调度稳定性和鲁棒性，以应对新能源逐步接入传统电网的趋势。
- 3) **离线强化学习：**研究通过离线交互数据学习智能体策略的新算法，降低强化学习的应用门槛。
- 4) **超大规模多智能体学习范式：**研究可以支撑上千智能体同时学习的高效训练机制，在同等计算资源下比已有算法提升 10% 以上的收敛效率。

[返回目录](#)

1.3 深度学习框架技术

本课题期望对深度学习框架技术进行系统、深入和前瞻的研究，来优化当前深度学习框架主流技术中的关键难点问题，并进一步研究下一代框架的设计实现和对应核心技术。

建议研究方向：

- 1) **统一中间表示：**研究深度学习统一中间表示和优化策略，以满足训练、分布式、编译器、压缩、推理等多种使用场景的需要。
- 2) **神经网络编译器技术及多硬件对接：**研究深度学习底层编译优化策略，包括运行时编译优化、压缩-编译协同设计、动态代码自动生成等，实现对多硬件的更好自适应支持、更高性能执行，研究编译器和手工优化结合的方法；研究和底层硬件协同的软硬一体整体优化方案，包括 AI 芯片和编译器的协同设计，在对应硬件上发挥出最佳性能。

- 3) **稀疏技术及在训练和推理中的应用**: 研究使用非结构化稀疏 (CSR/COO 等通用格式) 或者硬件结构化稀疏等技术实现训练推理加速; 研究稀疏算子加速技术; 研究稀疏模型在典型场景的应用。
- 4) **基于 CPU、GPU 对典型模型的性能优化技术**: 研究某类典型模型在通用 CPU、GPU 硬件上的性能优化, 实现性能超越当前最佳水平, 实现对 Transformer 类视觉模型在 Nvidia GPU 上的极致优化, 性能超越 Faster-Transformer。
- 5) **图的快速子图融合技术**: 研究项目希望能基于图提出快速融合的设计方案, 实现新增算子融合方案的快速实现; 基于图的快速子图匹配和融合, 以最少的代码完成新增算子融合的实现。
- 6) **深度学习模型异构部署技术**: 研究单模型在异构硬件上的调度方案, 实现模型推理的延迟或是吞吐的提升; 基于 Nvidia GPU Jetson 硬件的异构部署, 如 X86+NVIDIA GPU 的组合, Nvidia GPU Jetson 系列内部 CPU + GPU + DLA 的组合, 手机上 ARM CPU + GPU + DSP + NPU 的组合等。将单个模型进行异构部署, 实现效率的最大化提升。
- 7) **自动化深度学习技术**: 研究自动化深度学习 (AutoDL) 技术, 研究更通用的自动网络结构设计和模型超参数搜索技术, 特别是针对机器视觉多任务的通用骨干网络架构搜索; 研究 AutoDL 的高效实现方式, 包括搜索的分布式化、联邦学习化; 并设计 AutoDL 技术和深度学习框架的有机结合方案, 简化深度学习框架使用复杂度, 进一步提升开发效率。
- 8) **分布式训练技术**: 研究更高效的分布式训练策略, 包括训练效率的提升 (加速比提升)、扩展性的提升 (更多节点下保持加速比和收敛性); 研究超大规模模型训练技术, 如数据并行、张量模型并行、流水线并行等能更好的统一架构方案及协同方式, 研究自动化的硬件感知的并行策略选择; 研究图神经网络模型的高效大规模训练方法; 研究低配网络条件下的并行训练算法, 减小或甚至消除网络带宽限制对训练速度和收敛效果的影响; 研究异构硬件/集群下的新的分布式训练方案; 研究大规模参数服务器架构技术。
- 9) **端云协同训练技术**: 研究移动端的分布式训练, 以及端云协同的训练技术, 充分利用端侧算力并提升数据安全性。
- 10) **深度学习弹性调度技术**: 研究深度学习框架训练和推理的弹性调度部署, 提升资源利用率和任务调度效率。
- 11) **深度学习可视化**: 研究设计更好的数据可视化、模型可视化和训练可视化策略, 提升深度学习开发调优的便利度。
- 12) **深度学习框架扩展能力的研究**: 研究对深度强化学习、联邦学习、量子机器学习、图神经网络以及树类模型等传统机器学习技术的更高效全面的支持方案。
- 13) **新交互界面设计和下一代框架研究**: 研究命令式编程 (静态图) 和声明式编程 (动态图) 更好的结合方式, 做到灵活性和效率、部署兼顾; 探索新的编程范式、

用户交互方式以及对应的底层实现逻辑，进一步提升用户开发效率、简化底层架构复杂度；研究框架和平台层更好的结合融合方式，探索提升用户体验的全新产品形态。

[返回目录](#)

1.4 自然语言处理

自然语言处理研究能实现人与计算机之间用自然语言进行有效交互的各种理论和方法。本课题希望对深度学习技术在自然语言处理中的前沿问题和产业化实践中的实际难题进行分析和探索。在前沿问题探索中，着重方法的创新性和可推广性，在实际产业化难题解答中，针对实际场景着重方法的效果和效率。

建议研究方向：

- 1) **深度学习语义表示技术**，包括通用基础大模型及其微调范式（如 Prompt-Tuning 等）方式的研究、领域语义模型研究等。
- 2) **深度学习自然语言对话技术**，包括任务型对话和开放域对话相关的技术研究。
- 3) **基于知识图谱和 NLP 的智能系统研究**，比如医疗辅助诊断系统研究。
- 4) **机器翻译技术**，包括不限于低资源翻译（多语言、多领域等）、多模翻译（融合语音、图像等）等研究。

[返回目录](#)

1.5 计算机视觉

本课题希望对深度学习技术在计算视觉中的前沿问题和产业化实践中的实际难题进行分析和探索。在前沿问题探索中，着重方法的创新性、领先性和可推广性，鼓励探索基础模型的突破，填补视觉技术领域的空白。在实际产业化难题解答中，针对实际场景，着重方法的效果和效率，促进产业化的高效、高质量发展。

建议研究方向：

- 1) **工业质检/巡检场景关键技术**：在训练缺陷或隐患检测模型的过程中，针对训练数据量较少、检测目标区域面积较小等一系列问题，系统化研发自监督学习、对比学习、迁移学习、元学习等技术来提升实际应用效果。
- 2) **卫星遥感图像分析关键技术**：基于深度学习对图像的处理能力，对卫星遥感影像进行分析，实现包括土地、河流、湖泊、建筑物、云层等资源属性检测、分割、变化对比检测等功能。
- 3) **自动视频生成**：面向人脸、肢体、手势研发生成网络，实现表象建模、动作迁移和时序平滑等技术；基于少量图片样本、结合参数化 3D 建模等图形学技术对视频自动生成进行探索。

- 4) **视频分析关键技术**: 针对视频分类和视频检索, 探索图像、文本和语音多模态融合策略, 提升视频分类和视频检索的实际应用效果。
- 5) **模型小型化和自动化设计**: 针对模型训练中学习率、Batchsize 等超参以及训练数据增强策略对模型最终性能和训练速度的影响, 研发超参搜索和自动数据增强的方法, 解决效果和效率问题。
- 6) **OCR 关键技术**: 探索视觉信息和语义信息的深度融合方法, 研发如何利用语义信息辅助文字检测和文字识别新方法。
- 7) **人脸识别**: 在现实视频中, 针对尺度、遮挡、低质量等问题, 研发多属性结构化人脸检测技术; 在超大规模人脸图像下, 研发人脸特征学习和聚类等算法; 针对多样攻击, 研发多模态、高效、便捷的反欺诈技术。
- 8) **3D/点云**: 以机器人导航、混合现实、智能交通、无人驾驶等场景为基础, 研发三维点云识别、点云生成、点云分割、实例分割、3D 检测、3D 跟踪等相关领域领先技术。
- 9) **交互式图像分割关键技术**: 针对图像分割标注困难的问题, 系统化研发弱监督学习和主动学习等技术, 进一步提升图像分割标注效率。
- 10) **图像分类关键技术**: 针对服务器端和移动端不同设备, 在推理速度受限的情况下, 系统化研发更高精度的图像分类网络结构和预训练模型, 包括自监督等前沿技术。
- 11) **图像检索关键技术**: 针对人脸识别、行人再识别、商品识别和地标识别等类别多, 更新频率高, 样本获取难的场景, 系统化研发主体检测和度量学习等技术来提升实际应用效果。同时, 针对多样攻击, 研发多模态、高效、便捷的反欺诈技术。
- 12) **文档智能分析关键技术**: 面向文档图像智能分析, 尤其是中文场景, 系统化研发版面分析

[返回目录](#)

1.6 智能语音

本课题希望对深度学习技术在智能语音中的前沿问题和产业化实践中的实际难题进行分析和探索。在前沿问题探索中, 着重方法的创新性、领先性和可推广性, 鼓励探索基础模型的突破。在实际产业化难题解答中, 针对实际场景, 着重方法的效果和效率, 系统完整性和实用性, 促进产业化的高效、高质量发展。

建议研究方向:

- 1) **语音识别**: 在语音交互、训练数据筛选等场景中, 研究语音识别置信度量算法或模型; 研究基于模型的轻量、快速的语音后处理系统, 包括自动标点、反

正则、文本纠错、文本顺滑等模块；在中英混合场景下，探索端到端（流式/非流式）语音识别系统；针对语音识别跨领域和跨语言场景，研究底资源迁移学习方法；在会议场景下，研究提升语音识别模型准确率方法，并建设会议日志系统。

2) **语音合成**：研究基于模型的统一中文文本前端模块，支持文本正则、分词、拼音预测、韵律预测等；在少量录音数据的条件下，研究说话人自适应技术，为某一新说话人定制语音合成模型；研究语音合成风格化技术，包括语速、停顿、重音、情感等，并搭建风格化合成系统。

3) **声纹识别**：在现实音频中，针对跨信道、场景失配、时变、短语音、低质量等问题，研发鲁棒声纹识别技术；在语音会议等多说话人场景下，研发声纹特征学习、说话人分割、说话人聚类、流式说话人日志算法；在金融、电信等场景下针对多样攻击，研发准确、高效、便捷的鉴伪技术。

4) **语音预训练技术**：利用大规模无监督数据，研究语音预训练技术，并提升语音识别、语音翻译、音素识别、语音合成、关键词检测、声纹识别、说话人日志、语种识别、性别识别、年龄识别、语音分离、语音增强等下游任务或多任务混合学习系统性能。

5) **模型小型化和稀疏化**：针对语音识别和语音合成系统，研究在识别准确率和合成效果不损失的前提下，研究基于模型的降低模型参数量和计算量的方法，并做到端侧部署。

[返回目录](#)

1.7 个性化推荐

个性化推荐技术是在互联网信息爆炸式增长的时代背景下，帮助用户高效获得感兴趣信息的关键技术。本课题希望探索基于深度学习的个性化推荐前沿技术，并针对实际产业应用中的挑战性问题进行深入研究。

建议研究方向：

- 1) **研究基于深层语义交互的匹配技术**，实现用户兴趣和内容的深度匹配。
- 2) **研究基于多任务、序列建模技术**，支持信息流场景的高效排序。
- 3) **研究大规模推荐预训练模型技术**，可迁移应用与不同推荐任务场景。

[返回目录](#)

1.8 AI+科学计算

AI 科学计算是指使用人工智能方法、利用计算机再现、预测和发现客观世界运动规律和演化特征的全过程。通过 AI 学习自然规律、求解数学模型并应用于工程实践和科学探索，解决航空航天、船舶制造、生物计算、地球科学、能源勘探等领域的难题。科学计算中的物理规律，通常使用微分方程或其它数学模型来描述，求解一般面临维数高、计算时间长、计算量大、并行效率低等难题，通过神经网络模拟函数的方法，通常可以简化方程的求解。当前相关的工作包括但不限于物理信息约束神经网络 PINN 方法，数据驱动的傅里叶神经网络操作 FNO 方法，AI 算法与传统数值方法结合算法。本课题期望探索 AI 技术和科学计算任务相结合的创新方法，具备良好的实用性和可推广性，更高效解决各领域的科学计算问题。

建议研究方向：

1) 研究科学计算相关各领域的 AI 技术解决方案，超越传统方法，包括但不限于：

- a. **计算流体力学：**使用数值方法在计算机中对流体力学的控制方程进行求解，从而可预测流场的流动，用 AI 的方法更好地解决圆柱绕流问题、多孔介质多相流问题、湍流模拟、超高速流场仿真等。
- b. **分子动力学：**依靠计算机来模拟分子、原子体系的运动，从而以动态观点考察系统随时间演化的行为，用 AI 方法更好地解决材料属性预测、新材料发现、新能源材料等问题。
- c. **生物计算：**新药发现，蛋白质结构预测模型、分子属性预测、药物靶点亲和力和分子生成；疫苗设计；精准医疗等。
- d. **地球科学：**气象预报、海洋气候模拟、地壳运动、生态、对地遥感等。
- e. **能源勘探：**油藏模拟、石油开采、海洋采油、电网控制等。
- f. **工程计算：**研究 AI 在计算机辅助工程领域的应用，分析静态结构和动态结构，模拟电磁、流体、结构等物理场，应用于汽车、船舶、钢铁、土木工程等领域的产品研发和工程设计。

2) 研究 AI 方法与数值方法结合的科学计算问题解决方案：通过 AI 算法与数值方法结合解决以上或其它科学计算问题，打通 AI 方法和数值模拟软件全流程。

3) 研究 AI 科学计算中神经网络训练相关的关键技术：包括但不限于研究 BFGS 等高阶优化器在 AI 科学计算任务的作用，研究不同数据精度 FP16/BF16/FP32/FP64 对 AI 科学计算任务的影响。

[返回目录](#)

2. 半开放课题

2.1 推理部署关键技术与开源实现

■ 基于 Paddle Lite 实现对 Risc-V 架构的硬件支持；

Risc-V 是与 X86、ARM 齐平的开源硬件架构。随着国内外 Risc-V 产业联盟以及深度学习的发展，越来越多的硬件厂商设计了基于 Risc-V 架构的深度学习硬件。本研究项目希望能够完成 Paddle Lite 对 Risc-V 的支持，支撑应用在 Risc-V 架构的硬件上高效部署。

建议研究方向：

1) 实现 Paddle Lite 对 Risc-V 的支持，并且能够满足实际生产的应用需求。

■ 基于 Paddle Lite 实现对 RTOS 操作系统支持；

实时操作系统（Real-time operating system, RTOS），又称**即时操作系统**，在机器人、智慧生活等领域有着广泛的应用。随着深度学习的发展，逐渐开始有应用将深度学习的模型基于 RTOS 在 ARM CPU 或是 MCU 上进行部署。本研究项目希望能够完成 Paddle Lite 对 ARM CPU 在 RTOS 系统的支持。

建议研究方向：

1) 实现 Paddle Lite 对 RTOS 的支持，并且能够满足实际生产的应用需求。

■ 基于 Paddle Lite 实现 ARM CPU 代码自动生成；

随着深度学习模型的快速发展，模型种类越来越多，想要通过手工优化，实现对多样模型的高效部署，变得越来越困难。将手工优化经验，结合编译器的原理实现对人力的释放，已经成为各个推理框架的趋势。本研究项目希望能够完成 Paddle Lite 对 ARM CPU 自动代码生成，比如单点算子的融合，Conv+各类激活的融合等等。

建议研究方向：

1) 实现 Paddle Lite 对 ARM CPU 自动代码生成的支持，并且能够满足实际生产的应用需求。

[返回目录](#)

2.2 深度学习框架评估评价体系研究

■ 深度学习框架易用性评价体系研究

对深度学习框架的用户而言，除框架功能灵活性、丰富性以外，易用性为用户感知的重要因素。由于深度学习框架的用户多为开发者，对使用体验的目标和衡量方法

没有明确的定义，需要技术、产品多方面兼顾。如何在兼顾性能的基础上，满足各类用户的易用性体验，是提升框架能力的核心思路。故易用性评价体系研究对框架的长远发展十分重要。本课题旨在制定合理贴切的易用性评价体系，保障框架合理有效提升。

建议研究方向：

- 1) **易用性提升兼顾功能损失特性研究：**在满足易用性的前提下，框架的组网能力、支持的算法类型是否受影响，同等模型在易用性提升的情况下模型的性能和精度是否受影响。
- 2) **易用性评价维度体系研究：**从框架硬件使用、统一接口、语言支持，支持集群和云计算环境、模型参数、组网难易程度、数据读取和处理、输入输出工具、文档支持等方面探究框架的易用性。
- 3) **易用性定量评价指标体系：**结合外部市场调研和框架认知经验，设计全面且合理的可量化易用性评估维度，根据不同用户群体覆盖和使用偏好，设计合理的权重；针对2)中1-2个方向思考定量计算方法，实现长线动态监控效果。

[返回目录](#)

2.3 千言 - 面向中文自然语言处理的开源数据集建设和方法研究

“[千言](#)”是由百度联合中国计算机学会、中国中文信息学会共同发起的面向自然语言处理的中文开源数据集项目，目前已经得到了国内多家高校和企业的资源研发者的共同建设，覆盖了10多个自然语言处理的任务，包含了开放域对话、机器阅读理解、机器同传、文本生成、情感分析等。“千言”为研究者提供了一站式的数据集浏览、整理、下载以及评测体验，受到了越来越多研究者的关注和使用。近两三年，随着大模型技术的出现和发展，基于大模型的自然语言处理技术取得了长足的进步。在取得进步的同时，大模型也带来了新的技术挑战和新的技术机遇，包括了可信、跨模态、知识融合等。本课题旨在鼓励中文自然语言处理开源数据集的建设，并推动大模型时代自然语言处理技术的进步。

建议研究方向包括但不限于：

- 1) 面向跨模态任务的开源数据集建设；
- 2) 面向可信（可控、可解释、鲁棒性）自然语言处理的开源数据集建设；
- 3) 面向语言和知识融合任务的开源数据集建设；
- 4) 面向自然语言处理数据标注的工具建设和方法研究；
- 5) 面向自然语言处理任务的评估方法研究；
- 6) 基于千言已有的开源数据集开展相关研究，如机器阅读理解、情感分析、文本生成、文本相似度等。

2.4 机器翻译关键技术

机器翻译是自然语言处理领域的重要研究方向之一。近年来，神经网络机器翻译技术迅速发展，翻译质量显著提升，神经网络机器翻译成为主流的方法。本项目

希望针对机器翻译技术在产业化中的实际问题，比如低资源翻译、篇章翻译和跨模态翻译等，探索前沿技术，促进机器翻译技术的创新发展，创造更大价值。

建议研究方向：

- 1) 低资源语言的机器翻译技术；
- 2) 基于篇章信息的机器翻译技术；
- 3) 翻译领域自适应技术：从数据和模型方面探索缓解领域数据稀疏问题；
- 4) 多模态翻译技术：研究翻译模型融合文字、视觉和语音等多模态知识，提升翻译质量；
- 5) 机器同传技术：研究质量和时延的平衡、翻译一致性和语音识别错误修正等。

2.5 跨模态对话

理解和使用不同模态信息进行交互是人人对话的关键特征。当前对话系统大多局限于文本载体，跨模态对话系统缺少关键探索，面临数据严重缺乏，内容理解、回复生成等技术难度极大。本项目希望克服当前纯文本对话技术的缺陷，探索完全开放域的跨模态对话预训练技术，并在此基础上实现表情聊天、跨模态情感抚慰和人设一致性对话，实现跨模态的对话理解和交互，使对话系统在社交、智能屏、虚拟人等场景下实现广泛应用，与人类进行更自然聊天。

建议研究方向：

- 4) 开放域跨模态对话技术：研究开放域跨模态对话技术，使对话模型具备开放域图片理解并生成相关回复的能力；
- 5) 融合表情图片的回复生成技术：研究对话模型根据上文回复表情图片的对话技术；
- 6) 跨模态情感对话技术：研究结合用户脸部、肢体等视觉信息（例如面部表情）的跨模态情感对话技术；
- 7) 环境感知对话技术：研究结合环境视觉信息输入的开放域对话技术；
- 8) 可低成本定制的跨模态对话技术：研究通过少量跨模态对话数据低成本定制跨模态对话能力的技术。

[返回目录](#)

3. 领域课题

3.1 小度人工智能助手中的复杂语义理解和对话

小度作为国内最大最活跃的人工智能助手，在手机、智能音箱、电视、车等各个智能设备中都是人机交互的重要方式（2021年小度系智能音箱出货量中国第一，智能屏出货量全球第一）。小度助手的智能程度是产品体验的核心，其中影响智能程度的关键模块除了语音识别的效果外，还有小度的语义理解和对话能力。小度的语义理解和对话中面临的挑战很多，包括远场语音交互下的语音识别错误、个性化和多轮场景下的语义歧义问题、超大规模的理解表示体系等。本课题期望将深度学习、

增强学习、大规模预训练模型等技术方向的研究结合人工助手的业务场景问题，提升开放、复杂、远场语音交互下人工智能助手的理解和对话能力。

建议的研究方向：

- 1) **语义理解模型在语音识别错误下的模型鲁棒性研究：**比如通过对比学习的训练方法改进、预训练模型的任务设计、主动样本增强等技术方向和方法。
- 2) **个性化、上下文相关的大规模对话意图、槽位理解和追踪研究：**解决同样的表达不同人、不同上下文的语义需求不同。
- 3) **超大规模弱标签监督下的语义理解模型研究：**如何更充分有效的利用超大规模但是弱标签（准确率85%）的有监督数据构建深度学习的语义理解模型。
- 4) **交互式语义理解模型研究：**研究如何通过Human-in-the-Loop进行Interactive训练（更好的Sample策略、超大模型的增量训练）。
- 5) **基于增强学习的多领域对话管理研究：**在多领域的任务式对话中研究如何基于增强学习来获得一个更好的决定，什么时候可以直接给用户返回结果，什么时候改和用户对话交互进行澄清的对话策略。

[返回目录](#)