15. 结课项目

本文档附录中的得分只用于计算现场汇报顺序，不是大作业最终得分。大作业得分还是参照网络学堂公告中说明的算法。

请负责本章节的同学，重点整理所有的结课报告，汇总成本章内容

唐杰、杜晋华、郑问迪、赖瀚宇、陈俊杰

# 前言

本章精选并汇编了2024年春季学期清华大学“高级机器学习”课程的结课项目报告，旨在全面展现机器学习领域，特别是深度学习及大语言模型（LLM）的最新研究动态、技术突破及广泛应用。这些报告由课程学生分组完成，不仅体现了他们对课程内容的深刻理解和掌握，更展现了他们在探索机器学习前沿问题时的创新思维和实践能力。

本章内容涵盖了从**语言模型与推理** 的理论创新到**图像、视频与多模态模型** 的技术突破，从**优化与推荐系统** 的算法改进到**教育、科学与医疗** 领域的深度融合，再到**商业与应用** 领域的广泛实践。特别值得注意的是，学生们在**大语言模型（LLM）**的应用方面进行了深入的探索，不仅将其作为强大的工具应用于各个领域，更在提升其性能、可解释性和可靠性方面进行了积极的尝试。例如，在语言模型方面，学生们提出了 RM-PoT 框架来增强 LLM 的数学推理能力，利用 LLM 进行因果推理和发现，以及构建基于 LLM 的智能体来参与特定游戏等；在多模态领域，学生们尝试通过双编码器增强图像字幕，利用 LLM 进行多模态对话系统的意图识别，以及进行场景分割长视频以实现更好的理解；在应用层面，学生们探索了 LLM 在智能教育、法律庭审模拟、个人投资助手、药物靶标相互作用预测等方面的应用。

本章的另一大亮点是**多智能体系统** 在多个领域的应用探索，包括多目标用户体验优化、法律庭审模拟、股票市场预测、智能仓库订单履行、交通信号灯控制以及机器翻译质量评估等。学生们通过构建基于 LLM 的多智能体框架，模拟复杂场景中的智能体交互和协作，为解决实际问题提供了新的思路和方法。

为了方便读者阅读和理解，我们将报告按照**语言模型与推理、图像、视频与多模态模型、优化与推荐系统、教育、科学与医疗、商业与应用** 六大主题进行分类。每个主题下，我们都对每篇报告的核心内容、创新点和实验结果进行了简明扼要的介绍，力求展现每篇报告的精华所在。

本章内容不仅是对“高级机器学习”课程教学成果的一次集中展示，更是对机器学习领域未来发展趋势的一次有益探索。我们希望这些报告能够激发读者的思考，为相关领域的研究和应用提供新的灵感和启示。同时，我们也期待这些年轻的学子能够在未来的学习和研究中继续探索，为人工智能领域的发展贡献更多力量。

# **第一篇 语言模型与推理**

## **通过外部记忆和自适应推理来编排大型语言模型的多智能体对齐 (Jiaxiang Liu, Wanlan Ren)**

**亮点：**这篇论文的重点和亮点在于提出了一个新颖的多智能体协作框架，通过智能体间的协作实现推理过程的迭代优化。系统采用 PPO 算法和 LoRA 适配器确保训练效率，同时实现了推理过程的透明化和可解释性。实验结果证明该方法在提升大语言模型推理能力方面具有显著优势，为相关研究提供了重要参考。

**介绍：**这篇论文提出了一个创新的多智能体协作框架，旨在提升大语言模型的推理能力。该框架包含推理、评估和重生成三个智能体，通过迭代优化的方式提升推理质量。论文采用 PPO 算法进行训练，并使用 LoRA 适配器实现高效的参数更新。在 HotpotQA 数据集上的实验结果表明，该方法显著提升了模型在复杂推理任务上的表现。这项工作为增强大语言模型的推理能力提供了新的研究思路。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文主要试图解决大语言模型在复杂推理任务中存在的局限性问题。具体来说，传统的大语言模型在处理需要多步推理、长期依赖和跨领域信息整合的任务时往往表现不佳。这些模型虽然能够处理表层的分类或翻译任务，但在需要深度推理的场景中常常难以保持逻辑一致性，且缺乏推理过程的透明度和可解释性。此外，现有模型缺乏有效的自我纠错机制，一旦在推理过程中出现错误，就难以及时发现和修正。论文提出通过多智能体系统来增强大语言模型的推理能力，使其能够更好地处理复杂的推理任务。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题，增强大语言模型的推理能力一直是自然语言处理领域的重要研究方向。然而，论文提出的通过多智能体协作框架来解决这一问题的方法是比较新颖的。传统方法主要关注单一模型的优化，而本文提出将推理过程分解为生成、评估和优化三个阶段，由不同的智能体协作完成，这种方法为解决大语言模型推理能力的局限性提供了新的思路。
  + 这篇文章要验证一个什么科学假设？
    - 论文要验证的科学假设是：通过构建多智能体协作系统，结合外部记忆模块和自适应推理机制，可以显著提升大语言模型的推理能力，使其能够更好地处理复杂的多步推理任务，同时提高推理过程的透明度和可靠性。具体而言，论文假设通过推理智能体、评估智能体和重生成智能体的协作，可以实现推理过程的迭代优化，从而产生更准确和连贯的推理结果。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：大语言模型推理能力研究、强化学习方法、多智能体系统研究。在大语言模型推理方面，GPT 系列模型的研究者以及一些专注于模型推理能力增强的研究团队值得关注。在强化学习领域，John Schulman 等人提出的 PPO 算法的研究工作很重要。在多智能体系统方面，Alex Graves 等人在混合计算和外部记忆网络方面的工作具有重要影响。这些研究为本文提供了重要的理论和技术基础。
  + 论文中提到的解决方案之关键是什么？
    - 论文提出的解决方案的关键在于三个方面：首先是多智能体协作框架的设计，通过推理、评估和重生成三个智能体的配合，实现推理过程的迭代优化；其次是采用 PPO 算法进行训练，确保模型参数更新的稳定性；最后是使用 LoRA 适配器进行高效的参数更新。这些关键设计使得系统能够在保持计算效率的同时，不断提升推理质量。
  + 论文中的实验是如何设计的？
    - 实验主要基于 HotpotQA 数据集进行，通过评估模型在多跳问答任务上的表现来验证系统的有效性。实验设计包括与基线模型的性能比较、错误分析以及不同组件的消融实验。具体的评估指标包括准确率、F1 分数和完全匹配分数。实验还专门分析了模型在处理不同类型推理任务时的表现。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验主要使用了 HotpotQA 数据集和 O1-OPEN/OpenO1-SFT 数据集。HotpotQA 是一个专门用于评估多跳推理能力的问答数据集，包含需要从多个文档中获取信息才能回答的问题。论文中没有明确提到代码是否开源。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果较好地支持了论文的科学假设。结果显示，与传统方法相比，该系统在准确率、F1 分数和完全匹配分数上都取得了显著提升，特别是在处理复杂的多步推理任务时表现出明显优势。错误分析也帮助确定了系统的优势和局限性。
  + 这篇论文到底有什么贡献？
    - 论文的主要贡献包括：提出了一个新的多智能体协作框架用于增强大语言模型的推理能力；设计了基于 PPO 和 LoRA 的高效训练方法；实现了推理过程的透明化和可解释性；在实验中证明了该方法的有效性。这些贡献为提升大语言模型的推理能力提供了新的解决方案。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的研究方向包括：改进智能体之间的协作机制；扩展评估指标的多样性；增强系统处理罕见实体和复杂逻辑的能力；将系统应用于更多实际场景如教育、研究和工业应用；探索多语言支持等功能扩展。

## **面向机器学习研究智能体的系统化思路精炼框架 (Zijun Liu, Cheng Gao, Hanxi Zhu)**

**亮点：**这篇论文的重点是提出了一个能够系统性探索和优化机器学习研究思路的多代理框架。其主要亮点在于：1) 创新地将重点放在研究思路的生成和优化上；2) 设计了支持推理计算扩展的代理协作机制；3) 通过实验证明了该方法的有效性。这些工作为提升机器学习研究自动化水平提供了新的范式。

**介绍：**这篇论文提出了一个名为 Baby-AIGS-MLer 的多代理框架，旨在提升机器学习研究的自动化水平。该框架包含四个专门的代理，通过形成多层次反馈循环来系统性地探索和优化研究思路。实验表明，该方法能显著提升现有代理的性能，虽然仍不及人类专家，但为机器学习研究自动化提供了新的思路。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文试图解决大语言模型在自动进行机器学习研究时存在的局限性问题。当前的 LLM 代理虽然能生成语法正确的代码，但往往局限于使用预设或自决定的方法论，无法像人类研究者那样探索多种解决方案并根据实验结果进行迭代优化。这导致它们生成的解决方案虽然在技术上可行，但性能往往不够理想。论文提出需要一个系统来使代理能够系统性地探索和优化研究思路，以产生更优的解决方案。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题。在机器学习自动化研究领域，已有多项工作关注如何使用 LLM 进行代码生成和模型设计。然而，现有研究主要聚焦于提高代码的有效性，而较少关注如何系统性地探索和优化研究思路。论文创新地将重点放在了研究思路的生成和迭代优化上，这个角度是相对新颖的。
  + 这篇文章要验证一个什么科学假设？
    - 这篇论文要验证的科学假设是：通过多代理协作系统实现研究思路的系统性探索和优化，可以显著提升机器学习研究代理的性能。具体而言，论文假设通过让多个专门的代理协同工作，在提案、实现、实验和评审环节形成反馈循环，可以帮助系统探索更多样的解决方案并朝更优的方向演进。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：第一类是机器学习系统开发代理研究，如利用多代理工作流、树搜索等方法提高代码正确性；第二类是研究思路生成相关工作，主要关注如何利用 LLM 的知识生成研究思路；第三类是 AI 驱动的自动化研究系统，探索如何实现研究过程的自动化。值得关注的研究者包括 Chan 等人 (提出 MLE-bench)、Schmidt 等人 (提出 AIDE) 等。
  + 论文中提到的解决方案之关键是什么？
    - 论文提出的解决方案的关键是 Baby-AIGS-MLer 框架，它包含四个核心代理：负责提出研究思路的 Proposal Agent、负责将思路转化为代码的 Code Generation Agent、负责执行实验的 Experiment Conductor 以及负责评估结果并提供反馈的 Review Agent。这些代理通过多层次反馈形成迭代优化循环，使系统能够探索多样的解决方案并持续改进。同时，框架支持推理计算扩展，可以并行采样多个思路进行探索。
  + 论文中的实验是如何设计的？
    - 论文的实验设计主要包括三个维度的对比：与基础模型的对比、与代码生成代理的对比、与人类开发者的对比。实验选取了两个 Kaggle 竞赛数据集，分别代表简单和复杂的机器学习任务。所有方法使用相同的 LLaMA-3.1-70B-Instruct 作为基础模型，并限制了推理时间。对于人类专家解决方案，邀请了三位有顶会发表经验的研究者估算开发时间。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验使用了两个 Kaggle 竞赛数据集：Titanic-Spaceship (二分类问题) 和 LMSYS-Chatbot Arena Human Preference Predictions (人类偏好预测)。评估指标采用 MLE-Bench 中的设置：Titanic 数据集使用准确率，LMSYS 数据集使用对数损失。论文未提及代码开源情况。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果较好地支持了论文的科学假设。在两个数据集上，Baby-AIGS-MLer 都显著优于基础模型 (提升 11.6% 和 126.8%) 和代码生成代理 (提升 4.8% 和 7.5%)，证明了系统性思路优化的有效性。虽然仍不及人类专家，但实验结果表明该方法确实能帮助代理生成更优的解决方案。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献有：1) 提出了一个新的多代理框架 Baby-AIGS-MLer，实现了机器学习研究思路的系统性探索和优化；2) 设计了支持推理计算扩展的代理协作机制；3) 通过实验证明了该方法的有效性，为提升机器学习研究自动化水平提供了新思路。
  + 下一步呢？有什么工作可以继续深入？
    - 未来可以从以下方向深入：1) 扩大实验规模，测试更多真实世界的机器学习问题；2) 研究如何将人类创造力转化为 LLM 驱动的思路生成；3) 进行更详细的消融实验，分析各组件的贡献；4) 尝试使用不同的基础模型，验证方法的普适性。

## **基于开源大语言模型增强文本到SQL转换的研究 (Ruilin Hu, Lu Fan, Yizhe Chen)**

**亮点：**这篇论文的重点和亮点在于它成功地提升了开源语言模型在 text-to-SQL 任务上的性能，使其接近闭源大模型的水平。特别值得注意的是其创新的全局 - 局部模式链接方法和多样化 SQL 生成策略，这些设计有效地克服了开源模型的固有限制。论文不仅在技术上有创新，还在实用性方面做出了重要贡献，为企业提供了一个既经济又安全的 text-to-SQL 解决方案。

**介绍：**这篇论文提出了 SageSQL，一个基于开源语言模型的多智能体 text-to-SQL 框架。该框架通过三个创新的阶段来提升性能：鲁棒的模式链接、多样化的 SQL 生成和基于自一致性的后处理。实验结果表明，SageSQL 在多个基准数据集上都取得了优秀的性能，证明了开源模型在合适框架下可以有效处理复杂的 text-to-SQL 任务，为实际应用提供了一个具有成本效益和隐私保护的解决方案。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文致力于解决开源语言模型在 text-to-SQL 任务中的性能问题。目前，虽然基于 GPT-4 等闭源大模型的方法在 text-to-SQL 任务上取得了很好的效果，但这类方法存在使用成本高、数据隐私风险大、部署灵活性差等问题。而开源模型虽然可以解决这些问题，但其性能与闭源模型相比还有较大差距。论文提出了一个名为 SageSQL 的多智能体框架，旨在提升开源语言模型在 text-to-SQL 任务上的性能，使其能够作为一个实用的替代方案。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题。text-to-SQL 任务已经有很多研究工作，包括基于传统机器学习方法、基于预训练语言模型的方法，以及最近基于大语言模型的方法。但是，如何有效利用开源语言模型来解决 text-to-SQL 任务，以及如何缩小开源模型与闭源模型之间的性能差距，这个具体问题还没有得到很好的解决。论文从这个角度切入，提出了新的解决方案。
  + 这篇文章要验证一个什么科学假设？
    - 这篇论文要验证的科学假设是：通过精心设计的多智能体框架，开源语言模型也能在 text-to-SQL 任务上达到接近闭源大模型的性能。具体来说，论文假设通过鲁棒的模式链接、多样化的 SQL 生成以及基于自一致性的后处理等技术的组合，可以有效克服开源模型在处理复杂 text-to-SQL 任务时的固有限制。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：第一类是基于开源模型的 text-to-SQL 方法，如 RAT-SQL、LGE-SQL 等；第二类是基于闭源大模型的 text-to-SQL 方法，如 GPT-4、DIN-SQL、DAIL-SQL 等；第三类是代码语言模型相关研究，如 StarCoder、CodeLlama 等。在这个领域内，值得关注的研究人员包括来自清华大学、微软研究院等机构的学者，他们在 text-to-SQL 任务和代码语言模型方面都有重要贡献。
  + 论文中提到的解决方案之关键是什么？
    - 论文提出的解决方案的关键在于其三阶段的多智能体框架设计。首先是鲁棒的模式链接阶段，采用全局 - 局部方法避免遗漏重要的表和列；其次是多样化的 SQL 生成阶段，通过集成采样策略生成多样化的 SQL 查询；最后是基于自一致性的后处理阶段，从多个候选 SQL 中选择最优结果。其中，模式链接阶段的创新设计对提升整体性能起到了关键作用。
  + 论文中的实验是如何设计的？
    - 论文的实验设计非常全面。首先在 Spider 和 BIRD 这两个公认的 text-to-SQL 基准数据集上进行了主要性能评估；然后通过消融实验分析了各个模块的贡献；最后还专门评估了鲁棒模式链接模块的效果。实验采用执行准确率 (EX) 作为主要评估指标，并与多个最新的基线方法进行了对比。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 论文使用了 Spider 和 BIRD 这两个数据集进行评估。Spider 是一个包含 200 个不同数据库的通用基准数据集，而 BIRD 是一个更具挑战性的真实场景数据集。关于代码开源情况，论文中没有明确提到，但考虑到这是一个学术研究工作，代码很可能会在论文发表后开源。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 是的，论文的实验结果很好地支持了其科学假设。在实验中，SageSQL-32B 在 Spider 验证集上达到了 88.5% 的执行准确率，在 BIRD 验证集上达到了 70.2% 的执行准确率，超过了现有的开源方法，并接近使用 GPT-4 的方法的性能。消融实验也证明了各个模块的必要性，特别是模式链接模块的重要作用。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献有三个：首先，提出了一个创新的多智能体框架来增强开源语言模型在 text-to-SQL 任务上的性能；其次，设计了鲁棒的模式链接方法，有效解决了开源模型在处理大规模数据库模式时的问题；最后，通过实验证明了开源模型在合适框架下可以达到接近闭源大模型的性能。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的工作可以从几个方向深入：一是进一步优化多智能体框架的可扩展性，使其能更好地适应不同的数据库系统；二是探索外部知识库的集成方法，提升模型对专业领域知识的理解能力；三是改进后处理阶段的选择策略，从多个候选 SQL 中更准确地识别最优结果。

## **像人类一样的自我反思：可编辑大语言模型是你所需要的 (Yanchen Wu, Gangxin Xu, Dongchen Zou)**

**亮点：**论文的重点和亮点在于提出了一种全新的可编辑大语言模型范式，使模型能够像人类一样通过反思和修改来提升输出质量。通过创新的模型架构设计和训练方法，成功实现了模型的自我修正能力，这是传统大语言模型所不具备的。实验结果也证实了该方法的有效性，为未来的研究提供了新的思路。

**介绍：**这篇论文提出了一种创新的可编辑大语言模型 (E-LLM)，它能够在生成过程中实时反思并修改输出内容。通过引入添加、删除和修改 token 的操作，以及基于强化学习的训练方法，使模型具备了类似人类的反思和修正能力。实验结果表明，这种方法能够显著提升模型输出的质量。这项工作为大语言模型研究开辟了新的方向，对提升 AI 系统的智能水平具有重要意义。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文试图解决传统大语言模型 (LLM) 在生成文本时缺乏自我修正和迭代改进能力的问题。传统 LLM 在生成过程中只能按顺序添加 token，无法对已生成的内容进行修改和优化，这导致错误会不断累积。论文提出了一种可编辑的大语言模型 (E-LLM)，它能够在生成过程中实时反思并修改输出内容，通过添加、删除和修正 token 来逐步提升输出质量。
  + 这是否是一个新的问题？
    - 这是一个相对新颖的问题。虽然已有研究探索了如何提升 LLM 输出质量，如使用人类反馈的强化学习 (RLHF) 方法，但大多数方法关注的是如何从多个输出中选择最优结果，而不是如何让模型具备自我修正能力。论文提出的 E-LLM 开创了一种新的研究范式，使模型能够像人类一样通过反思和修改来逐步完善输出。
  + 这篇文章要验证一个什么科学假设？
    - 论文要验证的科学假设是：通过引入可编辑机制和基于强化学习的训练方法，可以使大语言模型具备自我修正和迭代改进的能力，从而生成更高质量的输出。具体而言，论文假设通过让模型能够添加、删除和修改已生成的 token，并使用奖励信号来指导这些编辑操作，可以实现类似人类思维的反思和修正过程。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为以下几类:传统大语言模型研究,如GPT系列、BERT等;基于人类反馈的强化学习(RLHF)研究;思维链(Chain-of-Thought)推理相关研究。在这一领域,值得关注的研究人员包括OpenAI的Alec Radford和Ilya Sutskever,以及Google的Jacob Devlin等。
  + 论文中提到的解决方案之关键是什么？
    - 论文提出的解决方案的关键在于:设计了一个可编辑的模型架构,使模型能够执行添加、删除和修改token的操作;引入了预奖励模型来评估中间状态的潜在改进空间;采用PPO算法来优化编辑策略。这些设计使模型能够像人类一样通过多轮修改来提升输出质量。
  + 论文中的实验是如何设计的？
    - 实验设计包括以下步骤:使用GPT-2生成初始草稿;使用不同的信号来训练和优化模型的编辑能力;与基线模型(如GPT-2、GLM-4等)进行对比;使用Monte Carlo方法进行多次并行实验以确保结果的统计显著性。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验使用的数据集包括 LAMBADA、MMLU 以及多个中文语料库。论文没有明确提到代码是否开源。数据集的选择涵盖了不同类型的语言理解和生成任务，有助于全面评估模型性能。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果很好地支持了论文的科学假设。实验表明，随着训练的进行，模型输出的质量不断提升并最终趋于稳定。不同信号策略的对比实验也验证了更详细的中间过程奖励有助于提升输出质量。与基线模型的对比也显示了 E-LLM 的优越性。
  + 这篇论文到底有什么贡献？
    - 论文的主要贡献包括:提出了一种新型的可编辑大语言模型架构;设计了基于强化学习的训练方法来实现模型的自我修正能力;通过实验验证了该方法的有效性。这些工作为提升大语言模型的输出质量提供了新的研究方向。
  + 下一步呢？有什么工作可以继续深入？
    - 未来可以从以下方面深入研究:扩展到多模态任务;提升计算效率;探索对抗训练策略;应用更多强化学习方法如深度Q学习;训练更复杂的奖励模型。

## **通过外部记忆和自适应推理来编排大型语言模型的多智能体对齐 (Jiaxiang Liu, Wanlan Ren)**

**重点:** 这篇论文的重点和亮点在于提出了一个新颖的多智能体协作框架，通过智能体间的协作实现推理过程的迭代优化。系统采用 PPO 算法和 LoRA 适配器确保训练效率，同时实现了推理过程的透明化和可解释性。实验结果证明该方法在提升大语言模型推理能力方面具有显著优势，为相关研究提供了重要参考。

**亮点:** 这篇论文提出了一个创新的多智能体协作框架，旨在提升大语言模型的推理能力。该框架包含推理、评估和重生成三个智能体，通过迭代优化的方式提升推理质量。论文采用 PPO 算法进行训练，并使用 LoRA 适配器实现高效的参数更新。在 HotpotQA 数据集上的实验结果表明，该方法显著提升了模型在复杂推理任务上的表现。这项工作为增强大语言模型的推理能力提供了新的研究思路。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文主要试图解决大语言模型在复杂推理任务中存在的局限性问题。具体来说，传统的大语言模型在处理需要多步推理、长期依赖和跨领域信息整合的任务时往往表现不佳。这些模型虽然能够处理表层的分类或翻译任务，但在需要深度推理的场景中常常难以保持逻辑一致性，且缺乏推理过程的透明度和可解释性。此外，现有模型缺乏有效的自我纠错机制，一旦在推理过程中出现错误，就难以及时发现和修正。论文提出通过多智能体系统来增强大语言模型的推理能力，使其能够更好地处理复杂的推理任务。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题，增强大语言模型的推理能力一直是自然语言处理领域的重要研究方向。然而，论文提出的通过多智能体协作框架来解决这一问题的方法是比较新颖的。传统方法主要关注单一模型的优化，而本文提出将推理过程分解为生成、评估和优化三个阶段，由不同的智能体协作完成，这种方法为解决大语言模型推理能力的局限性提供了新的思路。
  + 这篇文章要验证一个什么科学假设？
    - 论文要验证的科学假设是：通过构建多智能体协作系统，结合外部记忆模块和自适应推理机制，可以显著提升大语言模型的推理能力，使其能够更好地处理复杂的多步推理任务，同时提高推理过程的透明度和可靠性。具体而言，论文假设通过推理智能体、评估智能体和重生成智能体的协作，可以实现推理过程的迭代优化，从而产生更准确和连贯的推理结果。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：大语言模型推理能力研究、强化学习方法、多智能体系统研究。在大语言模型推理方面，GPT 系列模型的研究者以及一些专注于模型推理能力增强的研究团队值得关注。在强化学习领域，John Schulman 等人提出的 PPO 算法的研究工作很重要。在多智能体系统方面，Alex Graves 等人在混合计算和外部记忆网络方面的工作具有重要影响。这些研究为本文提供了重要的理论和技术基础。
  + 论文中提到的解决方案之关键是什么？
    - 论文提出的解决方案的关键在于三个方面：首先是多智能体协作框架的设计，通过推理、评估和重生成三个智能体的配合，实现推理过程的迭代优化；其次是采用 PPO 算法进行训练，确保模型参数更新的稳定性；最后是使用 LoRA 适配器进行高效的参数更新。这些关键设计使得系统能够在保持计算效率的同时，不断提升推理质量。
  + 论文中的实验是如何设计的？
    - 实验主要基于 HotpotQA 数据集进行，通过评估模型在多跳问答任务上的表现来验证系统的有效性。实验设计包括与基线模型的性能比较、错误分析以及不同组件的消融实验。具体的评估指标包括准确率、F1 分数和完全匹配分数。实验还专门分析了模型在处理不同类型推理任务时的表现。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验主要使用了 HotpotQA 数据集和 O1-OPEN/OpenO1-SFT 数据集。HotpotQA 是一个专门用于评估多跳推理能力的问答数据集，包含需要从多个文档中获取信息才能回答的问题。论文中没有明确提到代码是否开源。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果较好地支持了论文的科学假设。结果显示，与传统方法相比，该系统在准确率、F1 分数和完全匹配分数上都取得了显著提升，特别是在处理复杂的多步推理任务时表现出明显优势。错误分析也帮助确定了系统的优势和局限性。
  + 这篇论文到底有什么贡献？
    - 论文的主要贡献包括：提出了一个新的多智能体协作框架用于增强大语言模型的推理能力；设计了基于 PPO 和 LoRA 的高效训练方法；实现了推理过程的透明化和可解释性；在实验中证明了该方法的有效性。这些贡献为提升大语言模型的推理能力提供了新的解决方案。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的研究方向包括：改进智能体之间的协作机制；扩展评估指标的多样性；增强系统处理罕见实体和复杂逻辑的能力；将系统应用于更多实际场景如教育、研究和工业应用；探索多语言支持等功能扩展。

# **第二篇 图像、视频与多模态模型**

## **CleveReward: 基于对比学习的不同基准数据集视频奖励训练方法 (Yuanming Yang, Xiaoqian Liu, Jian Chen)**

**亮点：**这篇论文的重点和亮点在于提出了新颖的对比学习方法来处理不同评分标准的数据集整合问题，这是之前研究中的一个难点。通过构建专门的 VideoCross 数据集和设计对比学习训练方案，论文为视频生成模型的评估提供了一个可行的自动化解决方案，具有重要的实践价值。

**介绍：**这篇论文提出了一个名为 CleveReward 的对比学习视频奖励模型，旨在解决 T2V 生成模型中内容一致性的自动评估问题。论文通过构建 VideoCross 数据集和采用对比学习方法，成功实现了跨数据集的模型训练，为视频生成质量评估提供了新的解决方案。虽然模型仍在开发中，但初步结果显示其有潜力超过现有的视频奖励模型。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文主要试图解决文本到视频 (T2V) 生成模型中指令遵循能力的自动评估问题。随着 T2V 技术的发展，生成视频的质量不断提升，但生成内容与输入文本的一致性仍然是一个重要挑战。目前评估这种一致性主要依赖人工评估，这种方式成本高昂且难以规模化。论文提出通过构建端到端的奖励模型来自动评估视频生成模型的指令遵循能力。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题。在 T2V 领域，评估生成内容质量一直是研究重点之一。但本文提出的针对跨数据集训练的对比学习方法是创新的。以往的研究往往因为不同数据集的评分标准不一致而难以整合利用，本文提出的方法很好地解决了这个问题。
  + 这篇文章要验证一个什么科学假设？
    - 这篇论文要验证的科学假设是：通过对比学习的方法，可以有效地整合来自不同评分标准的数据集来训练视频奖励模型，从而提高模型评估视频生成质量的能力。论文认为，即使不同数据集使用不同的评分标准，只要将评分数据转换为配对格式，就可以通过对比学习来训练统一的奖励模型。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：基于 GAN 的视频生成方法、基于自回归 Transformer 的方法、以及基于扩散模型的方法。在视频质量评估方面，研究主要集中在技术层面 (如压缩效果、传输质量等) 和内容质量层面 (如模糊度、运动稳定性等)。值得关注的研究人员包括来自 OpenAI 的 Sora 团队、Google 的 Lumiere 团队等。
  + 论文中提到的解决方案之关键是什么？
    - 论文提出的解决方案的关键在于两点：首先是提出了 VideoCross 数据集，这个数据集专门设计用于对比学习，整合了不同标准的维度，减少了冗余并提高了一致性；其次是采用对比学习的训练方法，将不同来源的视频评分数据转换为配对格式，通过最大化相似样本对的相似度来训练模型。
  + 论文中的实验是如何设计的？
    - 论文的实验设计主要包括两个方面：一是将传统监督式微调与对比学习方法进行对比，验证对比学习在跨数据集训练中的优势；二是通过消融实验研究不同组件对模型性能的影响。实验使用了 16 个 A800 GPU 进行训练，选择了 Qwen2-vl-7b-chat 和 cogvlm2-video 作为基础模型。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 论文使用了多个数据集，包括 T2VQA、videofeedback 等，并提出了新的 VideoCross 数据集。论文提到 VideoCross 是开源的，但在论文中没有明确给出代码链接。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 论文的实验结果部分相对简略，虽然展示了一些初步结果，但缺乏详细的定量分析和与现有方法的系统性对比。因此实验对科学假设的支持还不够充分。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献在于：提出了首个基于对比学习的视频奖励模型 CleveReward；构建了专门用于对比学习的 VideoCross 数据集；提出了一种有效的跨数据集训练方法。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的工作可以从以下几个方面深入：改进模型架构以提高性能；扩大数据集规模；加强与文本描述的对齐评估；探索在更多视频生成场景中的应用。

## **基于场景切分的长视频理解方法 (Ziyu Zhao, Jin Wang, Jinsong Xiao)**

**亮点：**这篇论文的重点是提出了一种新的长视频理解框架，最大的亮点在于其动态场景分割算法，该算法无需预设片段数量，能够自适应地根据视频内容进行分割。同时，论文提出的多层次特征提取策略也很好地平衡了局部细节和全局语境的理解，使得模型能够更好地处理长视频内容。

**介绍：**这篇论文提出了一种新的长视频理解方法，核心创新在于动态场景分割算法和多层次特征提取策略。通过将视频动态分割成语义连贯的片段，并采用层次化的特征提取方法，成功解决了长视频理解中的关键挑战。实验结果表明，该方法在多个长视频理解任务上都取得了优于现有方法的性能。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文主要致力于解决长视频理解的挑战。随着各类平台上视频内容的激增，有效理解长视频的需求日益突出。现有的视频语言模型主要针对短视频设计，在处理长视频时面临着严重的限制，包括 LLM 的上下文长度限制和 GPU 内存消耗过大等问题。此外，现有方法要么使用固定数量的采样帧，导致信息不足；要么采用记忆增强机制，可能会导致信息质量随时间逐渐降低；要么使用预定义的分组大小，缺乏灵活性。论文试图提出一种新的方法来更好地理解和处理长视频内容。
  + 这是否是一个新的问题？
    - 长视频理解并不是一个全新的问题，已经有不少研究工作在这个方向上进行探索。然而，现有解决方案都存在各自的局限性。这篇论文提出了一种新颖的场景分割方法来处理长视频，创新性地解决了现有方法中的一些关键问题，特别是在动态场景分割和特征提取方面提出了新的解决思路。
  + 这篇文章要验证一个什么科学假设？
    - 这篇论文要验证的科学假设是：通过基于场景分布的动态视频分割，结合多层次的特征提取策略，可以更有效地理解长视频内容。具体来说，论文假设：1) 自适应的场景分割比固定帧数的分割更适合保持语义一致性；2) 分层的特征提取策略能更好地捕获视频的局部和全局信息；3) 在长视频数据上微调预训练的短视频模型可以提升其长视频理解能力。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为两类：传统的视频语言模型 (Video-LLMs) 和专门针对长视频的模型 (Long-term Video-LLMs)。在传统 Video-LLMs 方面，Video-Llama、Video-ChatGPT、VideoChat 等工作主要关注短视频的理解。在长视频处理方面，MovieChat 提出了基于记忆的机制，Chat-UniVi 提出了统一的图像和视频处理方法，TimeChat 和 LVCHAT 则采用了分组聚合的策略。这个领域内值得关注的研究人员包括来自各大科技公司和研究机构的团队，他们在视频理解和多模态学习方面都有深入的研究。
  + 论文中提到的解决方案之关键是什么？
    - 论文解决方案的关键在于提出了一种新的动态场景分割算法 (Scene-Clipping) 和多层次的特征提取策略。Scene-Clipping 算法基于信息熵来动态划分视频片段，无需预先指定片段数量，确保了每个片段内的语义一致性。在特征提取方面，使用 Clip Q-former 处理局部时空特征，再用 Video Q-former 整合全局信息，形成了一个层次化的特征提取框架。
  + 论文中的实验是如何设计的？
    - 实验设计主要包括两个部分：首先，在长于 30 秒的视频数据集上训练模型，包括 VideoChat2、ShareGPT4Video 和 ActivityNetQA 等数据集；其次，在 MLVU 基准测试集上进行零样本评估，测试模型在多个长视频理解任务上的表现。实验还包括消融研究，验证了各个模块的有效性。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 论文使用的主要评估数据集是 MLVU (Multi-task Long Video Understanding Benchmark)，这是一个综合性的长视频理解基准测试集，包含多种类型的视频和多样的评估任务。论文中没有明确提到代码是否开源。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果强有力地支持了论文的科学假设。在 MLVU 基准测试上，该方法在多个任务中都优于现有方法，特别是在 M-Avg 指标上达到了 39.5% 的性能。消融实验进一步证明了 Scene-Clipping 算法和特征提取策略的有效性。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献包括：提出了一种新的动态场景分割算法，开发了多层次的特征提取策略，并证明了在长视频数据上微调预训练模型的有效性。这些贡献共同推进了长视频理解领域的发展。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的研究方向可以包括：改进场景分割算法的效率，探索更优的特征融合策略，研究如何处理更长时间跨度的视频，以及如何更好地保持长期依赖关系。同时，也可以考虑将该方法扩展到更多的应用场景。

## **基于双编码器的增强图像描述方法 (Guangjie Xu, Yangchi Gao)**

**亮点：**这篇论文的重点和亮点在于其创新的双编码器架构设计，成功地将不同特点的编码器（CLIP 负责整体特征，SAM 负责细节特征）结合起来，实现了更全面的图像理解。通过精心设计的混合语义预训练策略，模型能够有效平衡整体语义和局部细节的描述，在通用和医疗图像描述任务上都取得了优秀的表现。

**介绍：**这篇论文提出了一个基于双编码器的图像描述生成模型，通过结合 CLIP 和 SAM 编码器的优势，同时捕获图像的整体语义和局部细节信息。论文设计了创新的预训练策略，并通过大量实验验证了模型的有效性。实验结果表明，该模型在多个评估指标上都优于现有方法，特别是在医疗图像描述生成任务上取得了显著提升。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文主要试图解决图像描述生成中的一个关键问题：如何同时有效捕获图像的整体语义信息和局部细节信息。传统的图像描述模型往往倾向于只关注其中一个方面，要么擅长捕获整体场景信息，要么专注于局部细节，难以在两者之间取得平衡。该研究提出了一个基于双编码器的架构来解决这个问题，通过结合 CLIP 编码器和 SAM 编码器的优势，实现更全面和准确的图像理解和描述生成。
  + 这是否是一个新的问题？
    - 这并不是一个全新的问题，因为图像描述生成一直是计算机视觉和自然语言处理交叉领域的重要研究方向。但论文提出的解决方案是新颖的，特别是在使用双编码器架构来平衡全局和局部特征方面的创新。此前虽然有许多研究致力于改进图像描述生成的质量，但很少有工作专门关注如何有效结合不同层次的图像特征信息。
  + 这篇文章要验证一个什么科学假设？
    - 这篇文章要验证的科学假设是：通过结合两个不同特点的图像编码器（CLIP 用于整体特征，SAM 用于细节特征），并采用混合语义预训练策略，可以显著提升图像描述生成的质量和准确性。具体而言，假设 CLIP 编码器能够 better 捕获图像的整体语义信息，而 SAM 编码器能够补充更细致的局部特征信息。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：第一类是传统的图像 - 文本多模态预训练模型，如 CLIP 和 ALBEF；第二类是专注于图像描述生成的模型，如 BLIP 和 BLIP-2；第三类是针对特定领域（如医疗图像）的描述生成研究。在这个领域中，值得关注的研究人员包括开发 CLIP 的 Alec Radford 团队，以及提出 BLIP 和 BLIP-2 的 Junnan Li 团队。
  + 论文中提到的解决方案之关键是什么？
    - 论文解决方案的关键在于其创新的双编码器架构设计和混合语义预训练策略。具体包括：使用 CLIP 编码器获取整体特征，使用 SAM 编码器获取细节特征，通过两个独立的 Q-Former 处理这些特征，最后将处理后的特征输入到大型语言模型中生成描述。同时，采用了针对性的预训练策略，对两个编码器分别使用不同的数据集进行训练。
  + 论文中的实验是如何设计的？
    - 实验设计主要包括两个阶段：首先是混合语义预训练阶段，分别对两个 Q-Former 进行训练，使用了不同的数据集组合；其次是描述生成阶段，将预训练好的双编码器架构与冻结的大型语言模型结合。实验比较了不同训练策略和模型配置的效果，包括单独使用 CLIP 或 SAM 编码器的基线实验。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验使用了多个数据集，包括通用数据集 COCO 和医疗图像数据集 ROCO、MedICaT。评估指标使用了 BLEU 和 METEOR。遗憾的是，论文中没有提到代码是否开源。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果很好地支持了论文的科学假设。实验数据显示，使用双编码器的模型在各项评估指标上都显著优于单编码器基线模型。特别是在最优配置下（CLIP 训练用通用数据集，SAM 训练用混合数据集），模型取得了最好的性能，验证了双编码器架构的有效性。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献包括：提出了一个新的双编码器架构来改进图像描述生成；设计了有效的混合语义预训练策略；证明了结合不同特点的编码器可以提升模型性能；为医疗图像描述生成提供了新的解决方案。
  + 下一步呢？有什么工作可以继续深入？
    - 未来工作可以从以下几个方向展开：使用更大规模的语言模型来提升生成质量；探索更专业的数据集来改善特定领域的表现；将该架构应用到其他任务如图像修复等；优化模型架构以提高效率；探索更多编码器的组合可能性。

## **通过对齐推理解码加速多模态大语言模型推理 (Chaoqun Yang, Yuanda Zhang, Xiying Huang)**

**亮点：**这篇论文的重点和亮点在于提出了一种既实用又创新的 MLLMs 推理加速方案。它不同于传统的模型压缩或架构优化方法，而是通过巧妙利用目标模型的 KV Cache 并设计专门的视觉信息压缩机制，实现了高效的推测解码。论文的实验设计全面且具有说服力，为多模态模型的效率优化提供了新的研究思路。

**介绍：**这篇论文提出了一种名为 AASD 的框架，用于加速多模态大语言模型的推理过程。该方法的核心思想是利用目标模型的 KV Cache 来指导草稿模型生成更准确的预测，同时通过特别设计的 Vision KV Projector 来压缩处理视觉信息，并采用 Target-Draft Attention 机制确保训练和推理的对齐。实验表明，该方法在保持模型性能的同时，能够实现约 2 倍的推理速度提升。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文主要致力于解决多模态大语言模型 (MLLMs) 在推理阶段的速度问题。目前 MLLMs 在视觉理解、问答等任务中展现出强大能力，但其自回归解码的特性导致推理速度较慢，限制了实际应用。现有的模型压缩、轻量架构等加速方法往往会影响输出质量，而从语言模型迁移的加速技术在处理视觉特征时也面临挑战。论文提出需要一种既能保持模型性能又能有效提升推理速度的解决方案。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题，但在多模态领域中仍然具有重要的研究价值。虽然推理加速在语言模型领域已有大量研究，但将这些技术迁移到多模态场景时会遇到独特的挑战，比如视觉 token 序列长、多模态信息对齐等问题。现有工作主要关注单模态加速或简单的技术迁移，缺乏针对多模态特点的系统性解决方案。
  + 这篇文章要验证一个什么科学假设？
    - 这篇论文试图验证以下科学假设：通过合理利用目标模型的 KV Cache 并设计特定的对齐机制，可以在不损失精度的情况下显著提升 MLLMs 的推理速度。具体而言，论文认为目标模型的 KV Cache 包含了重要的上下文信息，如果能够有效压缩和利用这些信息，可以帮助草稿模型生成更准确的预测结果；同时，设计专门的训练对齐机制可以进一步提升推测解码的效率。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：1) 传统模型压缩技术，如蒸馏、量化、剪枝等；2) 轻量架构设计，包括视觉 token 压缩等；3) 语言模型加速技术向多模态场景的迁移。在这个领域，值得关注的研究人员包括提出 Speculative Decoding 的 Yaniv Leviathan 团队，以及在视觉 - 语言模型方面有重要贡献的 Haotian Liu 等人。
  + 论文中提到的解决方案之关键是什么？
    - 论文提出的解决方案关键在于两个创新点：1) 利用目标模型的 KV Cache 并通过 KV Projector 进行有效压缩，使草稿模型能够获取更多相关信息；2) 设计 Target-Draft Attention 机制，实现训练和推理阶段的有效对齐。这种方案既保留了重要的上下文信息，又避免了处理冗长视觉序列带来的计算负担。
  + 论文中的实验是如何设计的？
    - 实验设计主要包括三个方面：1) 与现有方法的全面对比，包括不同类型的草稿模型；2) 消融实验，验证 Vision KV Projector 等关键模块的作用；3) 在不同规模模型和多种任务上的测试，验证方法的普适性。实验还特别关注了视觉信息的重要性，通过选择性禁用图像或文本 KV cache 来分析不同模态信息的影响。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验使用了多个数据集，包括 LLaVA-Bench In-the-Wild 用于通用任务评估、COCO 数据集用于图像描述任务、Science QA 数据集用于链式推理任务。评估指标包括 walltime speedup、acceptance rate、block efficiency 和 decoding speed。论文提到代码将开源，但目前仅给出了匿名链接。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果很好地支持了论文的科学假设。在多个数据集上，该方法都实现了显著的速度提升 (约 2 倍)，同时保持了输出质量。消融实验也证实了各个模块的必要性，特别是 Vision KV Projector 的作用。实验结果的一致性和全面性为论文的主张提供了有力支持。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献包括：1) 提出了一种轻量级的 MLLMs 推理加速框架；2) 设计了创新的多模态对齐策略；3) 提供了大量实验验证，证明该方法在实际应用中的有效性。这些贡献不仅解决了当前的实际问题，也为未来研究提供了新思路。
  + 下一步呢？有什么工作可以继续深入？
    - 未来可以从以下方向继续深入：1) 探索更高效的视觉信息压缩方法；2) 研究如何自适应调整推测步数；3) 将该方法扩展到更多模态的场景；4) 探索在更大规模模型上的应用效果。

## **预测材料特性的Transformer模型(Juncheng Yu, Kaiwei Zhang, Haonan Li)**

**重点:** 论文的重点和亮点在于其创新性地将 transformer 架构引入材料科学领域，通过精心设计的 tokenizer 和预训练策略，成功实现了高效且可解释的材料性质预测。特别值得注意的是，模型学习到的嵌入展现出与化学原理高度一致的模式，这证明了该方法不仅在性能上优秀，还能捕捉到材料的本质特征。

**亮点:** 这篇论文提出了一个基于 transformer 的统一框架用于材料性质预测，通过创新的 tokenizer 设计和预训练策略，成功将 transformer 架构应用于材料科学领域。研究表明，该方法不仅能够有效捕捉材料结构中的全局依赖关系，还能学习到有意义的材料表示，在多个下游任务中展现出优秀性能。这项工作为材料科学中的机器学习应用开辟了新的方向。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文试图解决材料科学领域中材料性质预测的关键挑战。传统上，材料性质预测主要依赖密度泛函理论 (DFT) 计算，但这种方法计算成本高昂且耗时。虽然图神经网络 (GNNs) 在这个领域取得了一定进展，但仍存在一些局限性，如难以捕捉长程全局相互作用，缺乏明确的可扩展性证据。论文提出要开发一个基于 transformer 的统一框架来解决这些问题，使材料性质预测更加高效、准确和可解释。
  + 这是否是一个新的问题？
    - 材料性质预测本身并不是一个新问题，但论文提出的解决方案具有新颖性。过去的研究主要集中在使用 GNN 等方法，而本文首次系统地将 transformer 架构应用于材料科学领域，并提出了创新的 tokenizer 和 3D 位置编码方案。这种方法使得模型能够更好地处理材料的空间信息，实现自回归预训练，这在材料科学领域是一个新的尝试。
  + 这篇文章要验证一个什么科学假设？
    - 这篇论文要验证的科学假设是:transformer 架构能够比现有的 GNN 模型更好地捕捉材料结构中的全局依赖关系，并且通过适当的预训练策略，可以学习到更有意义的材料表示。具体来说，论文假设：1) 通过设计专门的 tokenizer 和 3D 位置编码，可以有效地将晶体结构转换为 transformer 可处理的形式；2) BERT 式和 GPT 式的预训练策略可以帮助模型学习到有意义的材料表示；3) 预训练得到的模型可以在多个下游任务中取得与专门模型相当或更好的性能。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：1) 基于 DFT 的传统方法，这是材料性质预测的金标准；2) 基于 GNN 的方法，如 SchNet、CGCNN、MEGNet 和 ALIGNN 等，这些方法在过去几年取得了显著进展；3) 早期将 transformer 应用于材料科学的尝试。在这个领域，Jeffrey Grossman (MIT)、Shyue Ping Ong (UCSD) 等研究员做出了重要贡献。
  + 论文中提到的解决方案之关键是什么？
    - 论文解决方案的关键在于其创新的 tokenizer 设计和预训练策略。tokenizer 将晶体结构划分为网格，并通过巧妙的编码方式保留了空间信息。同时采用 BERT 式和 GPT 式两种预训练策略，使模型能够学习到材料的内在特性。此外，论文还引入了对称性预测任务，这对于准确预测材料性质至关重要。
  + 论文中的实验是如何设计的？
    - 实验设计包括三个主要部分：1) 预训练实验，分别使用 BERT 式和 GPT 式策略；2) 预训练嵌入的可解释性分析，验证模型是否捕捉到了化学知识；3) 下游任务微调实验，评估模型在实际应用中的性能。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 数据来自三个主要来源：Alexandria、GNoME 和 Materials Project，总计约 500 万个材料结构。其中包括 11 种不同类型的材料性质数据。论文没有明确提到代码是否开源。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果较好地支持了论文的科学假设。预训练模型展现出了良好的性能，嵌入分析显示模型确实捕捉到了有意义的化学知识，下游任务的结果也证实了模型的有效性。
  + 这篇论文到底有什么贡献？
    - 论文的主要贡献包括：1) 提出了新的材料结构 tokenizer 和 3D 位置编码方案；2) 首次系统地将 transformer 架构应用于材料科学；3) 证明了预训练策略在材料科学中的有效性；4) 展示了模型的可解释性和在多个下游任务中的优秀表现。
  + 下一步呢？有什么工作可以继续深入？
    - 未来工作可以从几个方向深入：1) 改进 tokenizer 以更好地处理原子位置的微小变化；2) 探索将更多领域知识整合到预训练过程中；3) 扩展模型以处理更多类型的材料和性质；4) 开发更高效的预训练策略。

# **第三篇 优化与推荐系统**

## **T'-ARC：针对ARC-AGI挑战中转导性Transformer模型的测试时训练方法 (Shaoting Zhu, Shuangyue Geng, Un Lok Chen)**

**亮点：**这篇论文的重点是探索如何通过测试时训练来提升转导式变换器模型在 ARC 任务上的表现。最大的亮点在于证明了一个小型专用模型配合适当的 TTT 策略可以达到与大型预训练语言模型相当的性能，这一发现具有重要的实践意义，因为它为解决抽象推理任务提供了一个更经济高效的方案。同时，论文通过系统的实验研究，揭示了影响模型性能的关键因素，为后续研究提供了有价值的见解。

**介绍：**这篇论文提出了一个结合基础模型训练、测试时训练 (TTT) 和主动推理的三步框架来解决 ARC 基准测试中的几何推理任务。研究的独特之处在于，它不仅探索了预训练语言模型的应用，还专门设计了一个小型转导式变换器模型。通过系统的实验研究，论文证明了配备 TTT 的专用模型能够达到与大型预训练模型相当的性能，同时揭示了单任务 TTT 战略的优势以及专门架构设计的重要性。这项工作为解决抽象推理任务提供了一个更轻量级的可选方案。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文主要尝试解决抽象推理语料库 (ARC) 这一人工智能基准测试中的几何推理任务。ARC 是由 François Chollet 提出的一个衡量 AI 系统智能水平的基准测试，要求 AI 系统通过观察几个输入 - 输出示例对，推断出测试输入应该对应的输出网格。这个任务特别强调技能的泛化能力而不是技能本身的水平，因为每个任务都是全新的，需要模型能够快速适应和学习新的模式。论文特别关注如何通过测试时训练 (Test-Time Training， TTT) 来提升转导式变换器模型在这类任务上的表现。
  + 这是否是一个新的问题？
    - ARC 基准测试本身并不是一个新问题，它在 2019 年就已经提出。但是本文探索的方向 - 将 TTT 应用于专门设计的转导式变换器模型来解决 ARC 任务，以及研究不同模型架构、训练策略和数据规模对性能的影响，这是一个相对新颖的研究角度。特别是在以往的研究中，大多数工作都集中在使用预训练语言模型，而本文证明了小型专用模型配合 TTT 也能达到相当的效果，这是一个值得关注的新发现。
  + 这篇文章要验证一个什么科学假设？
    - 这篇文章主要验证以下科学假设：1) 专门设计的小型转导式变换器模型配合 TTT 策略能够达到与大型预训练语言模型相当的 ARC 任务表现；2) 单任务 TTT (为每个评估任务单独训练一个模型) 比多任务 TTT (用混合数据集训练一个模型) 更有效；3) 网格位置编码等专门的架构设计能够显著提升模型在几何推理任务上的性能。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为两类：一类是归纳式方法，如 Omni-ARC 和 BARC 等工作，它们首先学习显式的规则 (如程序代码) 再应用于测试输入；另一类是转导式方法，直接从输入生成输出。在最近的研究中，测试时训练 (TTT) 成为提升预训练语言模型推理能力的主要技术，这在 2024 年 ARC 竞赛的前两名解决方案中得到了验证。在该领域内，François Chollet 作为 ARC 基准测试的提出者，以及 2024 年 ARC 竞赛的冠亚军团队 (the Architects 和 Akyürek 等人) 都是值得关注的研究者。
  + 论文中提到的解决方案之关键是什么？
    - 论文解决方案的关键在于三个主要组成部分的有效结合：1) 基础模型训练：包括预训练语言模型的微调和从头训练专用转导式变换器模型；2) 测试时训练：采用单任务 TTT 策略，为每个评估任务单独训练一个模型；3) 主动推理：通过数据增强和多数投票来提升推理的稳定性。特别值得注意的是，论文证明了专门设计的小型模型配合 TTT 能够达到与大型预训练模型相当的效果。
  + 论文中的实验是如何设计的？
    - 论文的实验主要从以下几个方面进行设计：1) 比较自定义转导式变换器与预训练语言模型的性能；2) 研究不同 TTT 策略 (单任务 vs 多任务) 和主动推理的效果；3) 探究数据规模和模型大小对性能的影响；4) 验证网格编码等专门设计的有效性。实验使用了从评估集中选择的 80 个任务，包括不同难度级别的样本。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 论文使用的主要是 ARC 数据集，包括 400 个公开训练任务、400 个公开评估任务和 100 个私有评估任务。此外还使用了 BARC (40 万个样本) 和 Re-ARC (10 万个样本) 等扩展数据集。虽然论文提到了参考 Nikola Hu 的代码实现，但没有明确指出本研究的代码是否开源。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果很好地支持了论文的科学假设。实验表明，配备 TTT 的小型专用模型确实能达到与预训练语言模型相当的性能 (37/80 vs 38/80); 单任务 TTT 的效果显著优于多任务 TTT; 专门设计的网格编码确实能大幅提升模型性能 (与传统的正弦位置编码相比从 14/80 提升到 37/80)。这些结果都提供了有力的实验证据。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献包括：1) 证明了小型专用模型配合 TTT 可以达到与大型预训练模型相当的 ARC 任务性能，为解决抽象推理任务提供了一个更轻量级的可选方案；2) 系统地研究了不同 TTT 策略、数据规模、模型大小等因素对性能的影响；3) 验证了专门的架构设计 (如网格编码) 对几何推理任务的重要性。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的研究方向包括：1) 探索在推理阶段引入思维链 (Chain of Thought) 推理来识别输出之间的关系和矛盾，从而提高准确性；2) 进一步研究元学习和 DPO 等方法在某些情况下无效的原因；3) 测试最新发布的 2024 年 ARC 竞赛顶级方案，以获取更多见解。

## **BP-LLM：一种基于大语言模型的精确自适应带宽预测方法 (Zheng jiang, Iat Long long, Xin Chen)**

**亮点：**这篇论文的重点和亮点在于首次将大语言模型应用于带宽预测任务，提出了一个创新的多模态融合框架。通过将网络状态信息转换为文本描述，充分利用 LLM 的强大能力，实现了更准确的带宽预测和更好的用户体验。特别值得注意的是，该方法在处理长期依赖关系和多模态数据融合方面取得了显著突破，为网络性能优化提供了新的研究思路。

**介绍：**这篇论文提出了一种名为 BP-LLM 的创新方法，通过利用大语言模型的强大能力来解决带宽预测中的关键挑战。该方法首次将 LLM 引入带宽预测领域，通过设计新颖的网络上下文信息嵌入机制和跨模态对齐策略，实现了对网络状态信息的有效利用和准确预测。论文不仅构建了一个综合的评估基准，而且通过在实际应用场景中的验证，证明了该方法在提升用户体验和网络资源利用效率方面的显著优势。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文主要试图解决带宽预测领域中的三个关键挑战。传统的带宽预测算法在长期趋势分析、多模态输入处理和泛化适应性方面都存在明显不足。具体来说，现有算法难以有效捕捉网络流量的长期变化趋势，导致预测结果准确性受限；它们通常将带宽预测简化为单一时间序列任务，忽视了网络条件、用户位置、通信延迟等多模态输入的影响；同时，这些算法的简单建模方式限制了其在新任务和场景中的泛化能力。
  + 这是否是一个新的问题？
    - 带宽预测本身并不是一个新问题，但论文提出了一个新颖的解决方案。以往的研究主要集中在使用统计方法、机器学习技术和时间序列分析来建模预测网络流量模式。然而，将大语言模型 (LLM) 应用于带宽预测是一个创新性的尝试。论文提出的 BP-LLM 方法首次将 LLM 的强大序列建模能力引入带宽预测领域，这种将自然语言处理技术与网络性能优化相结合的方法是新颖的。
  + 这篇文章要验证一个什么科学假设？
    - 这篇文章要验证的科学假设是：利用大语言模型的强大序列建模能力和跨模态学习特性，可以显著提升带宽预测的准确性和适应性。具体而言，论文假设通过将网络状态信息转换为文本描述，结合 LLM 的预训练知识，可以更好地理解和预测网络带宽变化，同时提供更好的泛化能力和下游任务支持。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为两大类：传统带宽预测方法和基于 LLM 的时间序列预测。在传统方法中，包括了基于线性模型的方法（如 TiDE、N-Hits 和 Dlinear）和基于 Transformer 的方法（如 PatchTST、FEDformer、Pyraformer 等）。在 LLM 时间序列预测方面，值得关注的研究人员包括来自 OpenAI 的 Alec Radford 团队、斯坦福大学的 Percy Liang 团队等。这些研究为将 LLM 应用于时间序列预测奠定了重要基础。
  + 论文中提到的解决方案之关键是什么？
    - 论文提出的解决方案的关键在于三个核心组件：网络上下文信息嵌入引导、文本和时间模态对齐、以及参数高效的大语言模型微调。其中，最关键的创新点是通过将网络状态信息转换为文本描述，利用 LLM 强大的文本理解能力来增强带宽预测的效果，同时通过跨注意力机制实现了文本和时间序列数据的有效融合。
  + 论文中的实验是如何设计的？
    - 论文的实验设计主要包括两个部分：带宽预测性能评估和下游任务验证。在带宽预测实验中，将 BP-LLM 与多个最新的基线方法进行了比较，使用了不同的时间窗口（4-4、16-16、64-64）进行测试。在下游任务验证中，主要关注自适应比特率流媒体（ABR）和用户体验质量（QoE）改进两个方面，通过与 BBA、Pensieve 等传统算法的对比来验证 BP-LLM 的实际应用效果。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 论文构建了一个包含超过 500 万条记录的综合基准数据集，涵盖了视频点播（VoD）、直播流媒体和实时通信（RTC）三个应用场景。数据集包含了带宽数据以及网络类型、设备信息等元数据。然而，论文中没有明确提到代码是否开源。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 是的，论文中的实验结果很好地支持了其科学假设。在带宽预测任务中，BP-LLM 在多个评估指标上都优于现有方法，特别是在较长时间窗口（64-64）的预测中表现突出。在下游任务验证中，BP-LLM 在 ABR 和 QoE 改进方面都取得了最好的效果，充分证明了该方法在实际应用中的优势。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献包括：构建了一个全面的带宽预测基准数据集；首次将 LLM 引入带宽预测领域；提出了一个创新的多模态融合框架；在实际应用场景中验证了方法的有效性。特别是在处理长期依赖关系和多模态数据融合方面取得了重要突破。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的研究方向可以包括：将 BP-LLM 与强化学习或图神经网络等先进机器学习技术相结合；探索在 5G 等新兴技术场景中的应用；优化模型结构以提高计算效率；扩展到更多的网络应用场景；研究模型的可解释性等。

## **基于在线评测反馈的强化学习代码合成增强方法 (Zihan Wang, Jiajun Xu, Lei Wu)**

**亮点：**这篇论文的重点是提出了一个结合 SFT 和 PPO 的两阶段训练策略，并通过高效的分布式评估系统解决了执行反馈延迟的问题。论文的亮点在于：首次证明了非 RL 优化的基础模型可以作为代码生成任务的有效起点；开发的分布式评估系统将反馈延迟降低到秒级；设计的层次化奖励机制有效平衡了代码的多个质量维度。这些创新使得强化学习能够有效应用于复杂的算法编程任务。

**介绍：**这篇论文提出了一个名为 RLOJF 的强化学习框架，旨在提升大语言模型在复杂算法编程任务中的表现。通过创新的两阶段训练策略，结合监督微调和近端策略优化，以及高效的分布式评估系统，该框架成功将模型的程序生成能力提升到新的水平。实验结果表明，该方法在 CodeContest 数据集上将 pass@1 率从 48% 提升到 81%，验证了方法的有效性。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文致力于解决大语言模型在复杂算法编程任务中的表现问题。尽管现有的大语言模型在标准编程任务中表现出色，但在竞争性编程环境 (如信息学奥林匹克竞赛) 中仍面临重大挑战，这类任务需要复杂的数学推理和算法优化。论文提出了一个名为 RLOJF 的框架，通过高性能分布式评估系统实现快速执行反馈，使模型能够从实践中学习和改进。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题，但论文提出了新的解决方案。代码生成一直是自然语言处理领域的重要研究方向，特别是在大语言模型兴起后，已有多项研究致力于提升模型的编程能力。然而，之前的研究主要关注模型规模扩大和训练数据扩充，较少考虑从执行反馈中学习这一关键方面。论文创新性地将强化学习应用于代码生成任务，并解决了执行反馈周期长的技术难题。
  + 这篇文章要验证一个什么科学假设？
    - 论文要验证的科学假设是：通过结合监督微调 (SFT) 和近端策略优化 (PPO) 的两阶段训练策略，并利用在线评测系统的实时反馈，可以显著提升大语言模型在复杂算法编程任务中的表现。具体而言，假设 SFT 阶段可以提升代码结构和文档质量，而 PPO 阶段可以优化运行时性能和执行成功率。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：代码生成的大语言模型研究 (如 CodeLlama、StarCoder 等)、强化学习在代码生成中的应用研究、以及针对奥赛级编程任务的专门模型研究 (如 O1-Coder、DeepSeek-R1-Lite 等)。值得关注的研究者包括来自 Meta AI 的 Jonas Gehring 团队 (提出了 RLEF 框架)，以及来自 DeepMind 的 Yujia Li 团队 (开发了 AlphaCode)。
  + 论文中提到的解决方案之关键是什么？
    - 论文解决方案的关键在于其创新的两阶段训练策略和分布式评估系统。首先，通过 SFT 建立基础能力，再通过 PPO 优化性能，这种策略既保证了代码质量，又提升了执行效率。其次，基于 Kubernetes 的分布式评估系统将反馈延迟从分钟级降低到秒级，使得强化学习训练变得可行。此外，论文设计的层次化奖励机制平衡了代码正确性、效率和质量多个维度。
  + 论文中的实验是如何设计的？
    - 实验设计包含多个方面：首先在 CodeContest 数据集上进行基准测试，比较不同训练策略 (SFT、PPO、SFT+PPO) 的效果；其次通过消融实验分析各组件的贡献；最后通过案例分析展示代码生成质量的演进过程。实验还特别关注了模型在不同难度问题上的表现，以及训练过程中的稳定性。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 论文主要使用了 CodeContest 数据集，包含 1，360 个编程问题，并保留了 4，000 个问题用于评估。此外还使用了一个源自中学信息学奥林匹克竞赛的数据集。关于代码开源情况，论文中没有明确提及，这是一个需要改进的地方。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果很好地支持了论文的科学假设。通过实验数据显示，结合 SFT 和 PPO 的方法使模型的 pass@1 率从 48% 提升到 81%，且在不同难度的问题上都有显著改善。消融实验也证实了两阶段训练策略的必要性，验证了 SFT 在提升代码结构方面的作用和 PPO 在优化执行效率方面的贡献。
  + 这篇论文到底有什么贡献？
    - 论文的主要贡献包括：提出了一个新的强化学习框架 RLOJF；开发了高效的分布式评估系统；设计了有效的两阶段训练策略；证明了非 RL 优化的基础模型可以作为代码生成任务的有效起点。这些贡献推动了复杂算法编程任务的自动化进程。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的研究方向包括：改进多轮反馈机制，更好地模拟真实编程过程；扩展代码质量评估的范围，包含更多维度如可读性和模块化；增强数学推理能力；开发更高效的训练范式；探索在更广泛的 AI 辅助编程场景中的应用。

## **基于开源大语言模型增强文本到SQL转换的研究 (Ruilin Hu, Lu Fan, Yizhe Chen)**

**亮点：**这篇论文的重点和亮点在于它成功地提升了开源语言模型在 text-to-SQL 任务上的性能，使其接近闭源大模型的水平。特别值得注意的是其创新的全局 - 局部模式链接方法和多样化 SQL 生成策略，这些设计有效地克服了开源模型的固有限制。论文不仅在技术上有创新，还在实用性方面做出了重要贡献，为企业提供了一个既经济又安全的 text-to-SQL 解决方案。

**介绍：**这篇论文提出了 SageSQL，一个基于开源语言模型的多智能体 text-to-SQL 框架。该框架通过三个创新的阶段来提升性能：鲁棒的模式链接、多样化的 SQL 生成和基于自一致性的后处理。实验结果表明，SageSQL 在多个基准数据集上都取得了优秀的性能，证明了开源模型在合适框架下可以有效处理复杂的 text-to-SQL 任务，为实际应用提供了一个具有成本效益和隐私保护的解决方案。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文致力于解决开源语言模型在 text-to-SQL 任务中的性能问题。目前，虽然基于 GPT-4 等闭源大模型的方法在 text-to-SQL 任务上取得了很好的效果，但这类方法存在使用成本高、数据隐私风险大、部署灵活性差等问题。而开源模型虽然可以解决这些问题，但其性能与闭源模型相比还有较大差距。论文提出了一个名为 SageSQL 的多智能体框架，旨在提升开源语言模型在 text-to-SQL 任务上的性能，使其能够作为一个实用的替代方案。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题。text-to-SQL 任务已经有很多研究工作，包括基于传统机器学习方法、基于预训练语言模型的方法，以及最近基于大语言模型的方法。但是，如何有效利用开源语言模型来解决 text-to-SQL 任务，以及如何缩小开源模型与闭源模型之间的性能差距，这个具体问题还没有得到很好的解决。论文从这个角度切入，提出了新的解决方案。
  + 这篇文章要验证一个什么科学假设？
    - 这篇论文要验证的科学假设是：通过精心设计的多智能体框架，开源语言模型也能在 text-to-SQL 任务上达到接近闭源大模型的性能。具体来说，论文假设通过鲁棒的模式链接、多样化的 SQL 生成以及基于自一致性的后处理等技术的组合，可以有效克服开源模型在处理复杂 text-to-SQL 任务时的固有限制。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：第一类是基于开源模型的 text-to-SQL 方法，如 RAT-SQL、LGE-SQL 等；第二类是基于闭源大模型的 text-to-SQL 方法，如 GPT-4、DIN-SQL、DAIL-SQL 等；第三类是代码语言模型相关研究，如 StarCoder、CodeLlama 等。在这个领域内，值得关注的研究人员包括来自清华大学、微软研究院等机构的学者，他们在 text-to-SQL 任务和代码语言模型方面都有重要贡献。
  + 论文中提到的解决方案之关键是什么？
    - 论文提出的解决方案的关键在于其三阶段的多智能体框架设计。首先是鲁棒的模式链接阶段，采用全局 - 局部方法避免遗漏重要的表和列；其次是多样化的 SQL 生成阶段，通过集成采样策略生成多样化的 SQL 查询；最后是基于自一致性的后处理阶段，从多个候选 SQL 中选择最优结果。其中，模式链接阶段的创新设计对提升整体性能起到了关键作用。
  + 论文中的实验是如何设计的？
    - 论文的实验设计非常全面。首先在 Spider 和 BIRD 这两个公认的 text-to-SQL 基准数据集上进行了主要性能评估；然后通过消融实验分析了各个模块的贡献；最后还专门评估了鲁棒模式链接模块的效果。实验采用执行准确率 (EX) 作为主要评估指标，并与多个最新的基线方法进行了对比。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 论文使用了 Spider 和 BIRD 这两个数据集进行评估。Spider 是一个包含 200 个不同数据库的通用基准数据集，而 BIRD 是一个更具挑战性的真实场景数据集。关于代码开源情况，论文中没有明确提到，但考虑到这是一个学术研究工作，代码很可能会在论文发表后开源。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 是的，论文的实验结果很好地支持了其科学假设。在实验中，SageSQL-32B 在 Spider 验证集上达到了 88.5% 的执行准确率，在 BIRD 验证集上达到了 70.2% 的执行准确率，超过了现有的开源方法，并接近使用 GPT-4 的方法的性能。消融实验也证明了各个模块的必要性，特别是模式链接模块的重要作用。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献有三个：首先，提出了一个创新的多智能体框架来增强开源语言模型在 text-to-SQL 任务上的性能；其次，设计了鲁棒的模式链接方法，有效解决了开源模型在处理大规模数据库模式时的问题；最后，通过实验证明了开源模型在合适框架下可以达到接近闭源大模型的性能。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的工作可以从几个方向深入：一是进一步优化多智能体框架的可扩展性，使其能更好地适应不同的数据库系统；二是探索外部知识库的集成方法，提升模型对专业领域知识的理解能力；三是改进后处理阶段的选择策略，从多个候选 SQL 中更准确地识别最优结果。

## **S3PE：一种简单可扩展的S型位置编码方法 (Zhu Zhang, Tianxing Yang, Zihan Yan)**

**亮点：**这篇论文的重点和亮点在于首次提出了统一的理论框架来分析位置编码，并基于此设计出了性能更优的 S³PE 方案。特别值得注意的是，论文通过理论分析和实验验证，揭示了位置编码插值因子的最优性质，这对未来的研究具有重要的指导意义。

**介绍：**这篇论文主要解决了大语言模型长文本处理中的位置编码问题。通过对现有位置编码方案的理论分析，提出了统一的数学框架，并在此基础上设计了新的位置编码方案 S³PE。通过在不同规模模型上的实验验证，证明了 S³PE 的优越性。这项工作不仅提供了实用的解决方案，也深化了对位置编码本质的理解。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文主要试图解决大语言模型在长文本处理中的位置编码问题。具体来说，当模型需要处理超出预训练长度的文本序列时 (例如从 4k 扩展到 32k)，现有的位置编码方案如 PI、NTK-aware、ABF、YaRN 和 LongRoPE 等都存在一定的局限性。论文试图找到一个统一的理论框架来分析这些位置编码方案，并在此基础上提出更优的解决方案。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题。长文本处理中的位置编码问题一直是大语言模型研究中的重要课题。此前已有多种解决方案被提出，如 PI 通过线性插值扩展位置编码，NTK-aware 引入非线性插值，ABF 调整基频等。但这篇论文首次从理论角度统一分析了现有方案，发现它们都可以用同一个数学框架来描述，这种统一的视角是新颖的。
  + 这篇文章要验证一个什么科学假设？
    - 论文要验证的核学假设是：现有的位置编码方案都可以统一到一个数学框架下，并且在这个框架指导下，存在一个比现有方案更优的位置编码方案。具体而言，论文提出三个假设：1) 位置编码的非线性插值因子应随维度增加而增大；2) 对高频分量的修改应该较小；3) 对低频分量的修改可以显著大于上下文窗口扩展因子。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为两类：一类是基于 RoPE 的位置编码改进方案，包括 PI、NTK-aware、ABF、YaRN 和 LongRoPE；另一类是对位置编码理论性质的研究。在这个领域，值得关注的研究者包括提出 RoPE 的 Jianlin Su，以及在位置编码扩展方面做出重要贡献的 Hugo Touvron 等人。
  + 论文中提到的解决方案之关键是什么？
    - 论文解决方案的关键在于提出了 S³PE (Simple Scalable Sigmoid-Style Position Encoding)。这个方案基于对现有位置编码方案的统一理论分析，采用了 sigmoid 函数变换得到的插值因子形状。相比现有方案，S³PE 在设计上更加优雅简单，且在实验中表现出更好的性能。
  + 论文中的实验是如何设计的？
    - 实验设计包括两个主要部分：首先，在不同规模的模型上 (从 0.031B 到 0.849B) 进行持续预训练，对比不同位置编码方案的效果；其次，使用 RULER 的 NIAH 评估指标，在不同长度区间评估模型性能。实验中特别关注了插值因子的最大值 (8× 和 50× 两种设置) 对性能的影响。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验使用了多个数据集的混合，包括 25% CommonCrawl Chinese、25% Code Pretrain、24% Dolma、15% C4、8% Pile 等。评估使用了 RULER 的 NIAH 任务。论文未提及代码开源情况。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果很好地支持了论文的科学假设。实验表明 S³PE 在各种模型规模和评估指标上都优于现有方案，且 50× 插值因子的设置普遍优于 8×，这验证了论文提出的理论分析。
  + 这篇论文到底有什么贡献？
    - 论文的主要贡献有三点：首先，提出了一个统一的理论框架来分析现有位置编码方案；其次，基于理论分析提出了新的位置编码方案 S³PE；最后，通过大量实验验证了 S³PE 的优越性。
  + 下一步呢？有什么工作可以继续深入？
    - 未来可以从几个方向深入：研究位置编码在更大规模模型上的表现；探索位置编码与模型架构的关系；研究位置编码在特定任务上的优化等。

## **多功率定律预测不同学习率调度下的损失曲线 (Kairong Luo, Zhixuan Pan, Chumeng Jiang)**

**亮点：**论文的重点和亮点在于其提出了首个能够准确预测任意学习率调度下完整训练曲线的定量模型。这不仅为理解深度学习中的学习率调度提供了新的视角，还提供了一个实用的工具来优化训练过程。特别值得注意的是，论文通过严谨的实验设计和分析，揭示了学习率调度的内在机制，并将理论发现转化为了实际的性能提升。

**介绍：**这篇论文提出了一个创新的多幂次定律，用于预测大语言模型在不同学习率调度下的训练损失曲线。通过将训练损失分解为基础幂律项和损失降低项，该模型能够准确捕捉学习率变化对训练过程的影响。论文采用 "自下而上" 的方法，通过大量实验 gradually 构建和验证了理论模型，并最终将其应用于实际的学习率优化中，获得了优于现有方案的结果。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文试图解决大语言模型训练过程中学习率调度的预测和优化问题。具体来说，论文提出了一个多幂次定律 (Multi-Power Law) 来预测不同学习率调度下的整个训练损失曲线。这个问题的难点在于学习率调度是一个高维的连续函数空间，传统的 scaling law 只能预测最终损失而无法预测整个训练过程中的损失变化。通过准确预测损失曲线，可以帮助研究人员在不进行昂贵的完整训练的情况下，就能评估和优化学习率调度方案。
  + 这是否是一个新的问题？
    - 这是一个相对新颖的问题。虽然之前已经有许多关于学习率调度优化的研究，但大多数工作要么是基于经验的启发式方法，要么是在特定函数族 (如余弦衰减) 内进行参数优化。而本文首次提出了一个可以预测任意学习率调度下完整训练损失曲线的定量模型，这种方法可以帮助我们更系统地理解和优化学习率调度。
  + 这篇文章要验证一个什么科学假设？
    - 这篇文章主要验证以下科学假设：训练损失曲线可以被分解为两个主要部分：(1) 一个基于学习率累加和的幂律项，代表了在恒定学习率下的训练效果；(2) 一个损失降低项，捕捉了学习率衰减带来的额外收益。这种分解形式不仅能够准确预测各种学习率调度下的训练曲线，还能揭示学习率调度优化的内在机制。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：1) 学习率调度的启发式设计，如 Smith (2017) 的循环学习率、Loshchilov & Hutter (2017) 的余弦学习率等；2) 基于理论分析的学习率优化，如 Li & Arora (2019) 的指数衰减方案；3) 神经网络训练的 scaling law 研究，如 Kaplan 等人 (2020) 和 Hoffmann 等人 (2022) 的工作。在这个领域中，值得关注的研究人员包括 Jared Kaplan、Sam McCandlish 等人，他们在 neural scaling law 方面做出了开创性的工作。
  + 论文中提到的解决方案之关键是什么？
    - 论文解决方案的关键是提出了一个新的多幂次定律，将训练损失分解为两个可解释的组成部分，并通过仔细的实验设计和分析，逐步构建和验证了这个定律的形式。特别是，论文采用了一种 "自下而上" 的方法，从最简单的两阶段学习率调度开始，通过大量消融实验揭示了损失曲线与学习率调度之间的定量关系。
  + 论文中的实验是如何设计的？
    - 论文的实验设计分为三个主要阶段：1) 通过两阶段学习率调度的实验，研究损失降低项的基本性质；2) 扩展到多阶段调度，验证和完善理论模型；3) 在实际训练中验证模型的预测准确性和泛化性。实验覆盖了不同模型大小 (25M-400M 参数)、不同训练长度 (24k-72k 步) 和各种学习率调度方案。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验主要基于 Llama-2 模型架构进行，使用了不同大小的模型变体 (25M-1B 参数)。评估数据采用了标准的语言建模验证集。虽然论文没有明确提到代码开源情况，但实验方法和设置都有详细描述，便于复现。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 论文的实验结果很好地支持了其科学假设。首先，实验表明多幂次定律能够准确预测各种学习率调度下的训练曲线，包括未见过的调度方案和更长的训练时间。其次，基于该定律优化得到的学习率调度方案确实能够改善训练效果，这进一步验证了理论模型的有效性。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献有三个：1) 提出了一个新的多幂次定律，能够准确预测不同学习率调度下的训练损失曲线；2) 通过实验揭示了学习率调度对训练过程的影响机制；3) 基于该定律设计出了性能超过现有方案的新型学习率调度。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的研究方向包括：1) 探索理论模型的数学基础，解释为什么多幂次定律能够如此准确地描述训练动态；2) 将模型扩展到更大规模的训练场景；3) 研究学习率与其他超参数 (如批量大小、模型结构) 之间的相互作用。

# **第四篇 教育、科学与医疗**

## **基于大语言模型的可靠一致性药物-靶标相互作用预测的多智能体框架 (Bowen Gao, Peidong Zhang, Huajun Bai)**

**亮点：**这篇论文的重点和亮点在于通过模仿人类专家团队的工作方式，提出了一个创新的多智能体协作框架来提高药物 - 靶点相互作用预测的可靠性。论文的主要亮点是将复杂任务分解为标准化的子任务，并通过多智能体协作来完成这些任务，这种方法不仅提高了预测的准确性和一致性，而且为解决 AI 幻觉问题提供了一个新的思路。虽然在计算效率等方面还有待改进，但这种结构化的方法对于提高 AI 在科学发现领域的可靠性具有重要的启发意义。

**介绍：**这篇论文提出了一个创新的多智能体框架来解决药物 - 靶点相互作用预测中的可靠性问题。通过将复杂任务分解为多个子任务，并让专门的智能体协作完成这些任务，该框架显著提高了预测的准确性和一致性。论文使用 BindingDB 数据集进行了详细的实验验证，结果表明该框架在多个指标上优于单一模型方法，特别是在预测一致性方面取得了显著改善。尽管在混合专家模型和计算效率方面还有改进空间，但这项工作为解决 AI 在科学发现领域的可靠性问题提供了一个有价值的新思路。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文主要试图解决在使用大语言模型进行药物 - 靶点相互作用 (DTI) 预测时存在的两个关键问题。首先是预测的不一致性问题，即同一个模型在不同时间对同一个输入可能会给出不同的预测结果；其次是 AI 幻觉问题，即模型可能会产生与事实不符的预测。这些问题严重影响了大语言模型在药物研发这样的高精度科学场景中的实际应用。论文提出，这些问题的根源在于直接使用单一大语言模型来处理如此复杂的科学任务是不合适的，需要一个更结构化和可靠的解决方案。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题。药物 - 靶点相互作用预测是药物研发领域的一个经典问题，传统上主要依赖实验室实验和计算机模拟。随着大语言模型的发展，研究人员开始尝试将其应用于 DTI 预测，但很快发现了预测不一致和 AI 幻觉等问题。虽然这些问题已被认识到，但目前还缺乏有效的解决方案。本文提出的多智能体协作框架是一种新颖的解决思路，通过模仿人类专家团队的协作方式来提高预测的可靠性。
  + 这篇文章要验证一个什么科学假设？
    - 这篇文章要验证的核心科学假设是：通过将复杂的药物 - 靶点相互作用预测任务分解为多个子任务，并让专门的智能体协作完成这些子任务，可以显著提高预测的准确性和一致性。具体来说，论文假设通过模仿人类专家团队的工作方式，即将任务分解为分子分析、蛋白质序列分析和结合亲和力预测等子任务，并通过辩论式集成方法整合多个预测结果，可以有效减少 AI 幻觉并提高预测可靠性。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为两大类：第一类是传统的 DTI 预测方法研究，包括实验方法和计算方法；第二类是将大语言模型应用于科学发现的研究。在第一类中，主要包括使用机器学习模型预测结合亲和力的工作。在第二类中，值得关注的工作包括 Luo 等人 (2022) 的 BioGPT、Singh 等人 (2023) 的蛋白质语言空间对比学习方法，以及 Singhal 等人 (2023) 在医学知识编码方面的工作。而在多智能体协作方面，Hong 等人 (2023) 的 MetaGPT 和 Inoue 等人 (2024) 的 DrugAgent 都提供了重要的参考。
  + 论文中提到的解决方案之关键是什么？
    - 论文提出的解决方案的关键在于三个方面：首先是任务分解，将复杂的 DTI 预测任务分解为多个标准化的子任务；其次是多智能体协作，为每个子任务配备专门的智能体，并建立协作机制；最后是混合专家模型，通过多个独立的预测组和判断智能体的辩论式集成来提高预测的可靠性。这种设计模仿了人类专家团队的工作方式，每个智能体都专注于自己擅长的领域，通过协作来完成复杂任务。
  + 论文中的实验是如何设计的？
    - 实验设计主要包括以下几个方面：首先是基准测试，使用单一大语言模型 (GLM-4-FlashX 和 GPT-4o) 作为基准；然后是多智能体框架测试，评估单组多智能体的性能；最后是混合专家模型测试，评估多组智能体加判断智能体的性能。实验使用多个指标进行评估，包括 RMSE、MAE 等精度指标，以及方差等一致性指标。特别值得注意的是，实验还评估了计算时间成本。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 论文使用了 BindingDB 数据集进行评估，这是一个包含实验测定的蛋白质 - 配体结合亲和力的综合数据集。实验选取了三种结合亲和力类型（IC50、Ki 和 Kd）的数据，每种类型随机选择 100 个样本作为测试集。遗憾的是，论文中没有提到代码是否开源。这是一个限制，因为开源代码对于研究的可重复性和进一步改进都很重要。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果部分地支持了论文的科学假设。多智能体协作框架确实在多个指标上优于单一模型，特别是在预测一致性方面取得了显著改善（方差减少高达 40%）。然而，混合专家模型部分的结果不够理想，虽然在相关性指标上有所改善，但在其他指标上表现不佳。此外，计算时间成本显著增加（是单一模型的 5-6 倍），这也是一个需要解决的问题。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献可以概括为三点：首先，提出了一个领域知识指导的多智能体框架，通过任务分解和协作来提高 DTI 预测的可靠性；其次，设计了一个基于辩论的集成方法，虽然效果还需改进，但提供了一个新的思路；最后，通过实验证明了多智能体协作可以显著提高预测的一致性，为解决 AI 幻觉问题提供了可能的方向。
  + 下一步呢？有什么工作可以继续深入？
    - 论文在未来工作方面提出了几个重要方向：首先是优化任务分解方案，引入更多专门的角色和领域知识；其次是改进智能体之间的协作机制，使其更加高效和动态；再次是扩大数据集测试范围，验证框架的普适性；最后是集成其他技术，如检索增强生成和上下文学习等。此外，解决计算时间成本高的问题也是一个重要方向。

## **增强蛋白质功能预测：在几何感知图神经网络中整合预训练和微调 (Chenxi Hu, Fei Long, Renrui Tian)**

**亮点：**这篇论文的重点和亮点在于提出了一个集成几何感知和预训练的图神经网络框架，创新性地解决了蛋白质功能预测中的关键问题。特别是通过显式建模蛋白质的几何特征，以及充分利用大规模未标记数据的策略，使得模型能够更好地理解和预测蛋白质功能。这些创新不仅推进了蛋白质功能预测的研究，也为其他蛋白质相关任务提供了新的思路。

**介绍：**这篇论文提出了一个创新的蛋白质功能预测框架，通过几何感知图神经网络显式建模蛋白质的结构特征，并结合大规模预训练策略来提高预测准确性。该方法在处理蛋白质序列和结构信息时，特别注重空间几何关系的建模，通过多个特征编码器捕捉距离、方向等关键特征。同时，论文设计了有效的预训练任务，充分利用了大量未标记的结构数据。实验结果表明，该方法在多个评估指标上都优于现有方法。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文主要致力于解决蛋白质功能预测的问题。随着高通量测序技术的发展，未标注功能的蛋白质序列数量呈指数级增长，而实验方法成本高昂且耗时，难以跟上序列测定的步伐。传统的基于序列同源性的方法在处理低相似度蛋白质和新型蛋白质家族时表现不佳。虽然近期基于深度学习的方法展现出 promising 的效果，但现有方法往往忽视了蛋白质结构中的关键几何特征，且未能充分利用大量未标记的结构数据。论文提出的方法旨在通过整合几何感知的图神经网络和预训练策略来提高蛋白质功能预测的准确性。
  + 这是否是一个新的问题？
    - 蛋白质功能预测并非一个全新的问题，这是生物信息学领域长期关注的基础性问题。然而，论文提出的解决方案具有新颖性，特别是在如何更好地建模蛋白质的几何结构特征，以及如何利用大规模未标记数据方面提出了创新性的方法。随着 AlphaFold 等结构预测工具的出现，这个问题在当下具有新的研究价值和现实意义。
  + 这篇文章要验证一个什么科学假设？
    - 这篇论文主要验证以下科学假设：通过显式建模蛋白质结构的几何特征（如距离和方向），并结合大规模预训练策略，可以提高蛋白质功能预测的准确性。具体来说，论文假设：1）在图神经网络中引入几何感知机制能够更好地捕捉蛋白质结构特征；2）利用自监督预训练可以从大量未标记的结构数据中学习到通用的结构表示。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：1）传统的基于序列同源性的方法，如 BLAST、PSSM 等；2）基于深度学习的序列方法，如 DeepGO、DeepGO-Plus 等；3）基于结构的方法，如 DeepFRI、HEAL 等。在这个领域，值得关注的研究人员包括开发 DeepGO 的 Maxat Kulmanov、开发 DeepFRI 的 Vladimir Gligorijević，以及开发 HEAL 的 Zhonghui Gu 等。这些研究逐步推进了蛋白质功能预测的发展。
  + 论文中提到的解决方案之关键是什么？
    - 论文解决方案的关键在于提出了一个几何感知的图神经网络框架，该框架具有两个主要创新点：1）通过引入距离编码器、方向编码器等多个几何特征编码器，显式地对蛋白质结构中的空间关系进行建模；2）设计了一个基于掩码的预训练任务，能够从大规模未标记数据中学习通用的结构表示。此外，论文还采用了对比学习策略来提高模型的表示学习能力。
  + 论文中的实验是如何设计的？
    - 实验设计主要包括三个部分：1）在大规模未标记数据集上进行自监督预训练；2）在带标签的数据集上进行模型训练和微调；3）与现有方法进行性能对比。实验采用了 AUPR 作为评估指标，在 MF、BP、CC 三个 Gene Ontology 领域进行了全面的评估。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验使用了两个主要数据集：1）预训练数据集来自 AlphaFold 蛋白质结构数据库，包含 40 万个蛋白质结构；2）下游任务数据集包含 36，404 个来自 PDB 的蛋白质结构，按 8:1:1 的比例划分为训练集、验证集和测试集。论文中没有明确提到代码是否开源。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果较好地支持了论文的科学假设。在三个 GO 领域的评估中，提出的方法都取得了优于现有方法的性能。特别是，预训练策略的引入进一步提升了模型性能，验证了利用未标记数据的有效性。
  + 这篇论文到底有什么贡献？
    - 论文的主要贡献包括：1）提出了一个新的几何感知图神经网络框架，更好地建模蛋白质结构特征；2）设计了有效的预训练策略，充分利用大规模未标记数据；3）在多个基准数据集上验证了方法的有效性，推进了蛋白质功能预测的研究。
  + 下一步呢？有什么工作可以继续深入？
    - 未来可以从以下方面继续深入：1）探索更多的生化特性在模型中的整合；2）改进预训练策略，提高模型的泛化能力；3）将方法扩展到更多的蛋白质相关任务中；4）研究模型的可解释性，以更好地理解预测结果的生物学意义。

## **针对可自验证问题的大语言模型强化学习技术研究 (Guanglei He, Yingxin Li, Jiuyang Zhou)**

**亮点：**这篇论文的重点和亮点在于提出了一个全新的无人类先验知识干预的 AI 自主学习范式。通过让模型完全依靠问题反馈和历史记录进行迭代优化，探索了突破人类认知天花板的可能性。虽然目前的实验结果还有待改进，但这种完全自主的学习方法为 AI 系统能力提升开辟了新的研究方向。

**介绍：**这篇论文提出了一个创新的研究思路：在可自验证问题领域，让大语言模型完全依靠问题反馈和历史记录进行自主学习，不引入人类先验知识。论文设计了一个三阶段的迭代优化框架，包括思维链提示、监督微调和强化学习，并在编程竞赛题目上进行了验证。虽然实验结果还不够完善，但为探索 AI 系统自主进化的可能性提供了有价值的参考。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文试图解决大语言模型在可自验证问题 (如编程竞赛题目) 上的能力提升问题。具体来说，论文提出了一个问题：在没有人类先验知识指导的情况下，仅通过迭代尝试和性能反馈，大语言模型能否通过自我学习和优化最终超越人类水平的解题能力。这个问题的关键在于利用问题的自验证特性，让模型能够获得客观的成功 / 失败反馈，从而形成闭环的自主学习过程。
  + 这是否是一个新的问题？
    - 让 AI 系统通过自主学习提升能力并不是一个全新的问题，但该论文提出的无人类先验知识干预的完全自主学习方法具有新颖性。传统方法往往依赖人工设计的测试用例、搜索策略等先验知识来指导模型学习，而本文提出让模型完全依靠自身的尝试历史和反馈来进行迭代优化，这种完全自主的学习范式在该领域具有创新性。
  + 这篇文章要验证一个什么科学假设？
    - 这篇论文要验证的科学假设是：对于编程竞赛等可自验证的问题，如果让大语言模型仅依靠问题反馈和历史记录进行自主学习，不引入人类先验知识，模型是否能够通过持续迭代优化突破人类认知天花板，最终达到或超越人类专家水平的解题能力。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为两类：一类是针对可自验证任务的研究，如 SELF-REFINE 等工作探索了模型在无外部监督下通过自我迭代改进输出质量的方法；另一类是编程领域的大模型研究，包括 StarCoder、CodeLlama 等系列工作。在这一领域，值得关注的研究人员包括 OpenAI 和 Anthropic 的研究团队，他们在 GPT-4 和 Claude 等模型上取得了重要进展。
  + 论文中提到的解决方案之关键是什么？
    - 论文提出的解决方案的关键是建立了一个三阶段的迭代优化框架：首先使用思维链提示让模型进行多轮尝试并收集历史数据，然后通过监督微调对齐模型能力，最后使用直接偏好优化 (DPO) 进行强化学习。整个过程不依赖人工设计的测试用例或搜索策略，完全基于模型自身的尝试历史进行优化。
  + 论文中的实验是如何设计的？
    - 论文的实验设计包括三个阶段：第一阶段使用思维链技术进行问题求解并收集历史数据；第二阶段使用收集到的正确答案进行监督微调；第三阶段使用正确和错误答案进行 DPO 优化。实验评估了模型在每个阶段后的性能变化，并与基线模型进行了对比。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验使用了编程竞赛题目作为评估数据集，包括 Interview@any 和 Competition@any 两个测试集。论文没有明确提到代码是否开源。实验主要基于 Codegeex4-9B 模型进行，通过在线评测系统 (OJ) 获取解题反馈。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果部分支持了论文的科学假设。结果显示加入自反馈策略后模型性能有所提升，但在 DPO 阶段出现了性能下降，这可能是由于缺少 SFT 步骤导致的。论文坦诚承认了这一问题，并提出了可能的改进方向。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献在于提出了一个新的无人类先验知识干预的完全自主学习框架，探索了大语言模型仅依靠问题反馈进行能力提升的可能性。虽然实验结果还不够完善，但为该研究方向提供了有价值的思路。
  + 下一步呢？有什么工作可以继续深入？
    - 下一步工作可以从以下几个方面展开：补充完整 SFT 步骤研究其对 DPO 的影响；扩大实验规模收集更多数据；探索更有效的历史数据压缩和总结方法；研究如何更好地平衡探索与利用。

## **RM-PoT: 通过思维程序重构数学问题和求解 (Changsong Lei, Yu Zhang)**

**亮点：**这篇论文的重点是提出将问题重构与程序化思维相结合的新方法，其最大亮点在于首次系统地研究了问题表述形式对模型性能的影响，并提供了一个简单但有效的解决方案。实验结果表明，即使是简单的问题重构也能带来显著的性能提升，这为提升大语言模型的数学推理能力提供了一个新的研究方向。

**介绍：**这篇论文提出了一个称为 RM-PoT 的新框架来提升大语言模型解决数学问题的能力。该框架首先通过问题重构生成多个不同表述形式的问题版本，然后使用程序化思维方法生成 Python 代码来解决这些重构后的问题。通过在多个数学问题数据集上的实验，作者证明了这种方法能显著提升模型的解题准确率，特别是在处理复杂的推理问题时表现出色。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文试图解决大语言模型在数学问题求解中的局限性问题。具体来说，现有的大语言模型在处理需要精确计算和逻辑推理的数学问题时往往表现不佳。作者观察到，即使是同一个数学问题的不同表述形式，都可能导致模型的解答准确率产生显著差异。这反映出大语言模型在处理复杂推理问题时的鲁棒性不足，过分依赖问题的表面形式而非深层逻辑结构。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题。在此之前已经有研究关注大语言模型在数学推理方面的能力，也有学者提出了各种改进方法如 Chain-of-Thought 和 Program of Thoughts 等。但本文独特之处在于首次系统地研究了问题重构对模型性能的影响，并提出将问题重构与程序化思维相结合的新框架。
  + 这篇文章要验证一个什么科学假设？
    - 这篇文章要验证的核心科学假设是：通过对数学问题进行多样化的重构，并结合程序化思维方法，可以提高大语言模型解决数学问题的准确性和鲁棒性。具体而言，作者假设多样化的问题表述形式有助于模型更好地理解问题的本质结构，而程序化思维则能够将计算过程与推理过程分离，从而提高求解的准确性。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：一是针对大语言模型数学推理能力的研究，如 Chain-of-Thought 提示方法；二是专门面向数学任务的模型开发，如 MathGPT 和 GeoSolver；三是问题重构在数学问题求解中的应用研究。在这个领域中，Jason Wei、Wang 等人的 Chain-of-Thought 相关工作，以及 Chen 等人的 Program of Thoughts 工作都值得关注。
  + 论文中提到的解决方案之关键是什么？
    - 论文提出的解决方案关键在于 RM-PoT 框架，它包含两个核心组成部分：一是问题重构 (RM)，通过提示语言模型生成多个不同表述形式的问题版本；二是程序化思维 (PoT)，通过生成中间 Python 代码来表达推理步骤。这两个部分的结合既提高了模型对问题本质的理解，又保证了计算的准确性。
  + 论文中的实验是如何设计的？
    - 实验设计主要包括以下几个方面：首先验证了问题重构的有效性，通过比较原始问题和重构后问题的求解成功率；然后将 RM-PoT 与现有方法 (CoT、SC、PoT) 进行了对比实验；最后通过消融实验研究了不同重构数量 (K) 和重构方式 (朴素重构 vs 上下文重构) 对性能的影响。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验使用了三个公开数据集：GSM8K (8.5K 个小学数学题)、AQuA (100K 个代数应用题) 和 SVAMP (1K 个算术应用题)。论文中使用的是 GLM-4-9B 作为基础模型。遗憾的是，文中没有明确提到代码是否开源。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果较好地支持了论文的科学假设。实验表明 RM-PoT 在所有测试数据集上都优于基线方法，验证了问题重构和程序化思维结合的有效性。同时，消融实验也证实了增加重构数量和使用上下文重构能带来性能提升。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献有三点：提出了结合问题重构和程序化思维的 RM-PoT 框架；通过大量实验证明了该方法的有效性；深入分析了问题重构对语言模型数学推理能力的影响。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的研究方向包括：深入探究为什么简单的问题重构能提升模型性能；研究更有效的问题重构方法；探索将微调方法整合到当前框架中以进一步提升性能。

## **基于大语言模型的智能教育：自动问题生成技术研究 (Kehan Zheng, Yida Lu, Wenjing Wu)**

**亮点：**这篇论文的重点是探索 AI 自动生成高等教育专业课程练习题的方法，其主要亮点在于：创新性地将多层检索机制与反思增强的监督微调相结合，显著提升了生成题目的质量；建立了系统的评估框架；通过实验证明了方法的有效性。这些工作为 AI 在教育领域的应用提供了重要参考，具有很强的实用价值。

**介绍：**这篇论文针对高等教育中缺乏个性化练习题的问题，提出了一个基于 ChatGLM 的智能题目生成系统。系统通过多层检索增强生成机制和反思增强的监督微调方法，能够根据具体需求自动生成高质量的专业课程练习题。论文使用来自清华大学物理和化工课程的数据进行了实验，并建立了多维度的评估框架。实验结果表明该方法能够显著提升生成题目的质量，特别是在知识点相关性和问题正确性方面表现突出，为 AI 辅助教育提供了有价值的解决方案。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文主要试图解决高校教育中个性化练习题生成的问题。具体来说，当前大学生在学习物理和工程等复杂课程时，往往缺乏足够的个性化练习机会，这限制了他们对复杂材料的深入理解和学术表现的提升。虽然现有的 AI 教育工具可以生成练习题，但这些系统往往只能覆盖浅层知识，无法处理高等教育中的逻辑推理知识，也缺乏学科专业性。论文提出基于 ChatGLM 的 AI 教学助手，通过检索增强生成 (RAG) 和监督微调 (SFT) 来自动生成符合个性化需求的专业练习题。
  + 这是否是一个新的问题？
    - 使用 AI 生成教育练习题并不是一个全新的问题。已有多个商业和研究项目在探索这一方向，如匡优 AI 和豆包题目生成等工具。但本文针对的是高等教育中专业课程的练习题生成这一具体场景，这在现有研究中较少涉及。此外，本文提出将 RAG 和 SFT 结合应用于题目生成，这种方法也具有一定新颖性。
  + 这篇文章要验证一个什么科学假设？
    - 这篇文章主要验证以下科学假设：通过将检索增强生成 (RAG) 和监督微调 (SFT) 应用于大语言模型，可以显著提升自动生成的专业课程练习题的质量，使其更好地满足教学需求。具体而言，文章假设多层检索机制可以帮助模型获取更相关的知识来生成题目，而反思机制增强的微调可以提升生成题目的质量。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：第一类是大语言模型在文本生成方面的基础研究，如 GPT 系列、GLM 等模型的工作；第二类是针对特定任务的生成方法研究，包括提示工程、监督微调、检索增强生成等技术；第三类是问题生成在教育等领域的应用研究。在这一领域，值得关注的研究者包括清华大学的 Tang Jie 团队 (GLM 系列模型的开发者) 以及 Du Xinya 等人 (在教育问题生成方面有重要贡献)。
  + 论文中提到的解决方案之关键是什么？
    - 论文提出的解决方案的关键在于多层检索增强生成机制和反思增强的监督微调方法。在 RAG 方面，系统首先使用 M3E 嵌入模型将知识点和题目文本转换为向量，然后通过两级检索机制选择最相关的参考材料。在 SFT 方面，除了基本的监督微调外，还引入了评分标准指导和反思机制来提升生成质量。这些方法的结合使得系统能够生成更加相关、准确和有深度的练习题。
  + 论文中的实验是如何设计的？
    - 论文的实验设计主要包括两个部分：首先是数据准备，从清华大学 AI 教学助手项目中选取大学物理和化学工程热力学两门课程的练习题，经过清洗和处理后用于训练和测试。其次是模型评估，设计了包括知识点相关性、问题正确性、完整性和难度在内的多维度评估框架，并将 baseline、单层 RAG、多层 RAG、基础 SFT、带评分的 SFT 和带反思的 SFT 等六种方法进行了对比实验。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验使用的数据集来自清华大学 AI 教学助手项目，包括 2124 道大学物理练习题和 2942 道化学工程热力学练习题。经过清洗后分别保留了 1937 道和 2344 道纯文本题目。测试集为每门课程随机选取的 100 道题目。由于知识产权原因，具体题目内容未公开。论文没有提到代码是否开源。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果较好地支持了论文的科学假设。结果显示多层 RAG 方法在所有评估指标上都优于 baseline，证实了检索机制对提升题目质量的积极作用。带反思机制的 SFT 也取得了优于 baseline 的效果，但不及 RAG 方法，这表明反思机制确实有帮助但效果有限。这些发现与论文的初始假设基本一致，同时也揭示了当前方法在生成高难度题目方面的局限性。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献在于：首次系统地探索了将 RAG 和 SFT 应用于高等教育专业课程练习题生成的方法；提出了创新的多层检索机制和反思增强的监督微调方法；建立了针对练习题质量的多维度评估框架；通过实验证明了所提方法的有效性，为 AI 辅助教育提供了实用的解决方案。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的研究方向主要包括：通过引入更强大的推理能力来提升生成题目的深度和复杂度；借鉴 O1 系列模型的思路，构建专门用于题目生成的显式思维链；扩展到更多学科和课程；改进评估框架使其更全面和客观；探索如何更好地利用人类反馈来持续改进系统。

# **第五篇 商业与应用**

## **LLM-IR: 利用大语言模型进行多模态对话系统的意图识别 (Junyi Wang, Yuanpei Sui, Tao Liu)**

**亮点：**这篇论文的重点是提出了一种新的多模态意图识别方法，其亮点在于创新性地将大语言模型与多种技术手段相结合，特别是通过 LoRA 微调和 OCR 技术的结合，显著提升了模型在电商场景下的性能。该研究不仅在技术上具有创新性，而且在实际应用中具有重要的商业价值。

**介绍：**这篇论文提出了一种基于大语言模型的多模态对话系统意图识别方法，特别针对电商场景进行了优化。通过结合 LoRA 微调、OCR 技术、数据增强等多种技术手段，有效提升了模型在特定领域的表现。实验结果表明，该方法相比基线模型提升了 5.35 个百分点的性能，验证了该方法的有效性。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文主要试图解决多模态对话系统中意图识别的问题，特别是在电商场景下。随着电商平台用户表达方式的多样化，用户可能通过文本、语音或图像等多种方式来表达需求，这给精确理解用户意图带来了巨大挑战。现有的大型通用模型虽然在开放域任务中表现出色，但在特定领域任务中仍存在意图识别不准确、信息遗漏严重、上下文关联性差等问题，这不仅影响用户体验，还会降低电商平台的客户满意度和销售转化率。
  + 这是否是一个新的问题？
    - 多模态对话系统中的意图识别并不是一个全新的问题，但本文提出了一个新颖的解决方案。传统方法主要依赖于规则基础的方法或简单的深度学习模型，而本文首次系统地将大语言模型与多模态数据处理相结合，通过 LoRA 微调、数据增强等技术来提高模型在特定领域的表现。这种将通用大模型适应于特定领域任务的方法在电商场景下的应用是具有创新性的。
  + 这篇文章要验证一个什么科学假设？
    - 这篇论文主要验证以下科学假设：通过对大语言模型进行特定领域的微调，结合多模态数据增强和外部知识集成，可以显著提升模型在电商场景下的意图识别准确性。具体而言，论文假设通过 LoRA 微调、OCR 技术、知识蒸馏和检索增强生成等方法的组合，可以有效解决现有模型在特定领域多模态对话中面临的各种挑战。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：多模态对话系统研究、大语言模型微调技术研究、以及知识增强方法研究。Chen 等人在对话系统研究方面做出了重要贡献；Hinton 在知识蒸馏方面的工作具有开创性；Lewis 等人在检索增强生成 (RAG) 方面的研究也值得关注。这些研究为本文的方法提供了重要的理论基础。
  + 论文中提到的解决方案之关键是什么？
    - 论文解决方案的关键在于将大语言模型 (如 Qwen2-VL) 通过 LoRA 进行高效微调，同时整合了多种技术手段：包括 OCR 文本提取、多模态数据增强、知识蒸馏和检索增强生成等。这种多技术融合的方法能够有效提升模型在电商场景下的意图识别能力。特别是 LoRA 参数优化和 OCR 推理的结合，使得模型性能相比基线提升了 5.35 个百分点。
  + 论文中的实验是如何设计的？
    - 实验设计主要包括三个配置：1) 引入 OCR 训练但不使用 OCR 推理；2) 同时引入 OCR 训练和推理；3) 调整 LoRA 参数 (秩 = 16，缩放因子 = 32) 并引入 OCR 推理。实验在 A100 和 H100 GPU 上进行，使用 Llama-Factory 框架进行推理。通过这些不同配置的对比实验，系统地评估了各个组件对模型性能的影响。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验使用的数据集包括 1，000 个带标签的训练对话样本，10，000 个用于初步测试的无标签样本，以及额外的 10，000 个最终测试样本。评估指标采用加权 F1 分数。但遗憾的是，论文没有提到代码是否开源。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果很好地支持了论文的科学假设。通过详细的实验数据显示，结合 OCR 推理和 LoRA 参数优化后，模型在意图识别和图像场景理解两个方面都取得了稳定的提升，最终达到 84.17 的最佳性能，相比基线提升了 5.35 个百分点。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献在于提出了一个面向电商领域的多模态意图识别方法，通过微调和数据增强等策略有效解决了通用模型在特定领域的不足。特别是在 OCR 技术与 LoRA 参数优化的结合方面取得了突破性进展。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的研究方向包括优化 OCR 结果过滤机制、探索 RAG 和思维链推理的集成，以及进一步提升多模态对话系统的智能性和鲁棒性。特别是在处理更复杂的对话场景和提供更精准的个性化服务方面还有很大的提升空间。

## **利用大语言模型进行因果推理和发现 (Zhuofan Sun, Qingyi Li)**

**亮点：**这篇论文的重点是提出了一个将 LLMs 与传统因果发现方法相结合的新框架，其主要亮点在于设计了基于三元组的提示策略来避免因果图中的循环问题，并证明了将 LLMs 推断的因果顺序作为先验知识能够有效提高因果发现的性能。这种方法不仅解决了直接使用 LLMs 进行因果发现的问题，也为因果发现领域提供了新的研究方向。

**介绍：**这篇论文提出了一个创新的框架，将 LLMs 作为虚拟领域专家应用于因果发现任务。通过设计基于三元组的提示策略，该方法能够有效地从 LLMs 中提取因果顺序信息，并将这些信息作为先验知识融入传统的因果发现算法中。实验结果表明，这种混合方法能够显著提高因果发现的准确性和稳定性。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文试图解决在因果发现和因果推断中如何有效利用大语言模型 (LLMs) 的问题。传统的因果发现方法主要依赖于专家知识或统计数据分析，而本文提出将 LLMs 作为虚拟领域专家来辅助因果发现。具体来说，文章关注如何通过 LLMs 来推断变量之间的因果顺序，并将这种顺序信息与传统的因果发现算法相结合，从而提高因果图的发现准确率。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题，因为已经有一些研究尝试将 LLMs 应用于因果发现。但本文提出的基于三元组的提示策略，以及将 LLMs 推断的因果顺序作为先验知识融入传统因果发现算法的方法是创新的。这种方法避免了直接通过 LLMs 推断完整因果图可能带来的问题，如循环依赖等。
  + 这篇文章要验证一个什么科学假设？
    - 这篇文章主要验证两个科学假设：第一，基于三元组的提示策略比基于配对的提示策略能更好地避免因果图中的循环；第二，将 LLMs 推断的因果顺序作为先验知识融入传统因果发现算法能提高因果发现的准确性。这些假设建立在 LLMs 具有丰富的领域知识基础之上。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：传统的基于数据的因果发现方法 (如 PC 算法、CaMML 等)、基于 LLMs 的因果推断方法 (如 Kıcıman 等人的工作)、以及混合方法。在这个领域中，值得关注的研究人员包括 Pearl (因果推断的理论基础)、Kıcıman (LLMs 在因果推断中的应用) 等。
  + 论文中提到的解决方案之关键是什么？
    - 论文解决方案的关键在于提出了一个三步走的方法：首先，使用基于三元组的提示策略从 LLMs 中获取局部因果关系；然后，通过多数投票机制整合这些局部关系得到全局因果顺序；最后，将这个顺序作为先验知识输入到传统的因果发现算法中。其中，三元组提示策略的设计是一个重要创新点。
  + 论文中的实验是如何设计的？
    - 实验设计包括两个主要部分：首先，通过比较基于配对和基于三元组两种提示策略在不同节点数量下产生的循环数量，验证三元组策略的优势；其次，通过在多个基准数据集上比较纯 LLM 方法、纯 CAM 算法和结合方法的表现，验证混合方法的有效性。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验使用了多个标准的因果发现基准数据集，包括 Earthquake、Cancer、Survey、Asia、Asia-M、Child 等数据集，以及一个用于疼痛诊断的 Neuropathic 数据集。论文中没有明确提到代码是否开源。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果较好地支持了论文的科学假设。实验表明，基于三元组的提示策略确实显著减少了因果图中的循环数量，而且结合方法在大多数数据集上都取得了最好的性能，特别是在处理复杂数据集时表现出较好的稳定性。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献包括：提出了一个新的基于三元组的 LLM 提示策略，开发了将 LLM 推断的因果顺序融入传统因果发现算法的方法，并通过实验验证了这些方法的有效性。这些工作为因果发现领域提供了新的研究思路。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的研究方向可以包括：改进 LLM 的提示策略以进一步提高准确率，探索更好的方法来处理 LLM 输出中的不确定性，研究如何处理有潜在混杂因素的情况，以及将该方法扩展到更大规模的因果图中。

## **Sim-Court: 基于大语言模型智能体的法庭审判模拟系统 (Kaiyuan Zhang, Zhaoxi Li, Xuancheng Li)**

**亮点：**这篇论文的重点是构建了一个完整的法庭模拟系统，其亮点在于：1) 系统设计符合现实审判流程；2) 评估方法全面且具体；3) 实验发现了不同模型在角色扮演时的特点和共同问题；4) 开源了相关数据集。这些工作为缓解司法资源紧张问题提供了一个可能的解决方案，也为法律教育提供了新的工具。

**介绍：**这篇论文提出了一个名为 Sim-Court 的法庭模拟系统，通过基于大语言模型的智能体来模拟完整的刑事审判流程。系统包含法官、检察官、辩护律师等五个角色，模拟从审前准备到最终判决的全过程。论文详细设计了各个角色的任务指令和评估方法，并通过实验比较了不同大语言模型在角色扮演中的表现差异。研究结果表明这种模拟方式具有可行性，同时也发现了需要改进的方向。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文试图解决法律资源稀缺与案件数量庞大之间的矛盾问题。论文指出，在 2023 年中国有 2580 万个法律案件，但全国仅有 10145 个法院在运转。这种情况不仅造成司法资源紧张，也影响了法律教育的开展。为此，论文提出建立一个可解释、便携和高效的法庭模拟系统 Sim-Court，通过基于大语言模型的智能体来模拟完整的刑事审判流程。
  + 这是否是一个新的问题？
    - 利用人工智能处理法律相关任务并非新问题，但这篇论文提出的完整法庭模拟系统是一个相对新颖的尝试。此前已有研究使用 AI 进行法律判决预测、法律问答等任务，也有像 AgentCourt 这样的简化法庭模拟系统，但都没有实现对完整审判流程的细致模拟。该论文首次提出了一个包含审前准备、法庭调查、证据展示、法庭辩论和被告人陈述等完整环节的模拟系统。
  + 这篇文章要验证一个什么科学假设？
    - 这篇论文主要验证以下科学假设：基于大语言模型的智能体是否能够模拟真实法庭审判中不同角色的行为，包括法官、检察官、辩护律师、被告人和书记员等；这种模拟是否能够产生合理的审判过程和判决结果；不同大语言模型在扮演不同角色时的表现是否存在差异。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为两类：一类是基于大语言模型的社会模拟系统研究，如 Stanford Town (虚拟村庄)、RecAgent (推荐系统)、AgentHospital (医疗流程) 等；另一类是人工智能在法律领域的应用研究，包括法律判决预测、法律问答、法律文书理解等。值得关注的研究包括 Chen 等人的 AgentCourt 工作，以及 He 等人关于法律判决预测的研究。Jin 等人在社会模拟方面的工作也值得关注。
  + 论文中提到的解决方案之关键是什么？
    - 论文提出的解决方案关键在于：1) 设计了完整且符合现实的审判流程框架；2) 为每个智能体角色设计了详细的角色描述和任务指令；3) 建立了多维度的评估方法，包括任务完成质量、内容表达和现场表现等方面。这些设计使得模拟系统能够更好地还原真实法庭审判过程。
  + 论文中的实验是如何设计的？
    - 实验设计主要包括两个部分：首先收集了 100 个来自 2018 年江苏省的真实判决文书，使用 ERNIE-Speed-128K 模型提取相关信息；然后选取了 7 种不同的大语言模型 (包括 ChatGLM-4-air、Claude3-sonnet、GPT-4 等) 在前 10 个案例上进行测试，比较它们在扮演不同角色时的表现。评估指标包括判决准确性、角色一致性、表达逻辑性等多个维度。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验使用了 100 个真实判决文书组成的数据集，这些案例来自 2018 年江苏省。论文提到相关数据集和法庭模拟记录已在 GitHub 上开源 (<https://github.com/Miracle-2001/Sim-Court>)。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果较好地支持了研究假设。结果显示不同模型在扮演不同角色时确实表现出差异，如 GPT 系列在扮演法官时倾向于给出更重的判决，而 ChatGLM 和 Claude 则相对温和。实验也发现了一些共同问题，如在长对话情境下容易出现角色不一致等，这些发现为进一步改进提供了方向。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献包括：1) 提出了一个易用的法庭模拟框架，为法律场景下的大语言模型应用提供了新的基准；2) 相比直接预测判决结果的方法，提供了更具可解释性的审判过程；3) 生成并开源了审判信息和法庭记录数据集，为相关研究提供了资源。
  + 下一步呢？有什么工作可以继续深入？
    - 未来工作方向包括：1) 改进多轮对话和长文本检索能力；2) 通过真实法庭数据的微调来提升角色扮演的稳定性；3) 扩展到民事案件等更多类型，增加更多角色。

## **基于大语言模型的宝可梦对战智能体 (Zihan Lv, Qihang Cen)**

**亮点：**这篇论文的重点是提出了一个创新的多技术融合框架来增强 LLM 在游戏环境中的决策能力，其亮点在于系统地整合了多种先进技术，特别是在知识增强和决策一致性方面的创新设计。虽然实验还在进行中，但研究思路清晰，为游戏 AI 领域提供了新的研究方向。

**介绍：**这篇论文专注于开发一个基于大语言模型的宝可梦对战 AI 代理系统。研究团队提出了一个综合性框架，结合了上下文强化学习、知识增强生成、一致性动作生成和监督微调四种技术，旨在提升 AI 代理在复杂对战环境中的决策能力。尽管实验结果还在完善中，但初步结果显示这种方法对提升模型性能是有效的。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文试图解决如何利用大语言模型 (LLMs) 来开发一个能在宝可梦对战游戏中做出有效决策的智能代理系统的问题。具体来说，论文旨在提升 LLM 在复杂、动态的游戏环境中的决策能力和策略执行能力。这个问题的难点在于宝可梦对战涉及复杂的类型相克关系、技能效果和战术选择，需要智能代理具备深入的上下文理解能力和战略思维。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题。在游戏 AI 领域，已经有许多研究致力于开发能在各类游戏中表现出色的 AI 代理。特别是在近年来，随着 LLM 的发展，已经有研究探索将 LLM 应用于冒险游戏、竞技游戏、通信游戏和合作游戏等多个方向。不过，这篇论文针对宝可梦这种特定的回合制策略游戏场景，提出了一套综合性的解决方案，这个角度是相对新颖的。
  + 这篇文章要验证一个什么科学假设？
    - 这篇论文要验证的科学假设是：通过结合上下文强化学习、知识增强生成、一致性动作生成和监督微调这四种技术，可以显著提升 LLM 在宝可梦对战中的决策能力和战斗表现。论文假设这种多方法结合的方式能够让 AI 代理在复杂的对战环境中做出更合理的决策。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为四类：一是冒险游戏领域的 LLM 应用研究，如 Red Dead Redemption 2 和 Minecraft 中的 AI 代理开发；二是竞技游戏相关研究，包括星际争霸 II 和宝可梦对战；三是通信游戏研究，如狼人杀和阿瓦隆；四是合作游戏研究，如 Overcooked。在这个领域内，值得关注的研究包括 Voyager 项目的团队、CRADLE 项目的研究者等。
  + 论文中提到的解决方案之关键是什么？
    - 论文提出的解决方案的关键在于四个核心技术的协同：首先是基于上下文的强化学习 (ICRL)，通过文本反馈来优化决策；其次是知识增强生成 (KAG)，引入额外的宝可梦相关知识；第三是一致性动作生成，运用多种提示工程技术；最后是监督微调 (SFT)，利用高水平玩家的对战数据进行模型优化。这些技术的结合使得系统能够在复杂的对战环境中做出更合理的决策。
  + 论文中的实验是如何设计的？
    - 论文的实验设计主要围绕与启发式机器人的对战进行。实验分为两个主要部分：首先测试应用 ICRL、KAG 和一致性动作生成后的模型表现，每种方法进行 100 场对战；其次探索 SFT 对模型性能的影响。不过论文目前似乎尚未完成全部实验，结果部分仍在更新中。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 论文使用的是宝可梦对战模拟器 Pokémon Showdown 上的随机对战模式作为测试环境。关于数据集和代码开源情况，论文中没有明确提及，这可能是因为研究工作还在进行中。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 由于论文的实验部分尚未完全完成，目前难以全面评估实验结果是否充分支持了科学假设。已有的初步结果显示，优化后的模型相比基准确实有所提升，但具体改进程度还需要更多数据支持。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献在于提出了一个综合性的框架，将多种技术方法结合起来提升 LLM 在宝可梦对战中的表现。特别是在 ICRL、KAG、一致性动作生成和 SFT 的整合方面提供了新的思路。
  + 下一步呢？有什么工作可以继续深入？
    - 下一步的工作可以从几个方向展开：完善实验评估体系，增加与人类玩家的对战测试；深入研究各种技术方法的组合效果；探索更多的知识增强方式；开发更精细的战术决策机制；以及将该框架推广到其他类型的游戏中。

## **探索大语言模型在文字游戏中的应用：谁是卧底 (Chentian Wei, Jiewei Chen, Jinzhu Xu)**

**亮点：**这篇论文的重点是提出了一个无需训练的 CoT 框架来提升大语言模型在文字游戏中的表现。最大的亮点是通过设计三个互补的 CoT 模块，成功地让模型在 "谁是卧底" 游戏中展现出了接近人类的推理能力，特别是在关键词描述和身份推理方面取得了显著进展，平民胜率提升了 74 个百分点，这一成果为大语言模型在更广泛的游戏场景中的应用提供了有价值的参考。

**介绍：**这篇论文探索了如何让大语言模型在文字游戏 "谁是卧底" 中展现出接近人类的推理能力和角色扮演能力。作者提出了一个基于 Chain-of-Thought 的无训练框架，包含 Judge CoT、Describe CoT 和 Spy CoT 三个核心模块，分别负责推理判断、关键词描述和卧底伪装。通过在 100 组数据上的实验，作者证明了框架的有效性，使模型的表现得到显著提升。虽然在卧底伪装方面还有改进空间，但整体而言，这项工作为大语言模型在文字游戏中的应用提供了一个可行的解决方案。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文试图解决大语言模型在文字游戏 "谁是卧底" 中的应用问题。具体来说，作者希望构建一个无需训练的框架，使大语言模型能够在该游戏中表现出良好的推理能力和角色扮演能力。"谁是卧底" 是一个需要玩家通过对话和推理来识别隐藏在群体中的卧底的派对游戏，这个游戏对语言模型的逻辑推理、角色扮演和社交互动能力都提出了很高的要求。论文希望通过设计合适的 prompt 和框架，让语言模型能够准确描述关键词、进行有效的推理判断，并在扮演卧底时学会隐藏自己的身份。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题。在此之前已经有研究者探索过大语言模型在社交推理游戏中的应用，比如 "狼人杀" 等游戏。但是，这篇论文专注于 "谁是卧底" 这样的文字游戏，这个游戏相比 "狼人杀" 更加依赖于语言描述和逻辑推理，而不是复杂的社交互动。同时，作者提出的基于 Chain-of-Thought 的无训练框架也是一个创新点，因为之前的研究往往需要对模型进行专门的训练或微调。
  + 这篇文章要验证一个什么科学假设？
    - 这篇论文要验证的科学假设是：通过设计合适的 Chain-of-Thought 提示框架，大语言模型能够在不经过专门训练的情况下，在 "谁是卧底" 这样的文字游戏中展现出接近人类的推理能力和角色扮演能力。具体来说，作者希望验证：1) 模型能否准确描述关键词而不暴露身份；2) 模型能否通过逻辑推理正确判断卧底身份；3) 作为卧底时，模型能否学会合理地隐藏自己。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为三类：第一类是大语言模型在游戏中的应用研究，如 ChessGPT 在国际象棋中的应用；第二类是探索大语言模型推理能力的研究，特别是 Chain-of-Thought 提示方法的研究；第三类是社交推理游戏中的多智能体交互研究。在这些领域中，Jason Wei、Tom Brown、Shunyu Yao 等研究者的工作都值得关注。特别是在 Chain-of-Thought 方面，Jason Wei 的工作对本文有重要影响。
  + 论文中提到的解决方案之关键是什么？
    - 论文解决方案的关键是设计了三个核心 CoT 模块：Judge CoT、Describe CoT 和 Spy CoT。Judge CoT 引导模型进行逐步推理来判断卧底身份；Describe CoT 帮助模型更准确地描述关键词，避免产生幻觉；Spy CoT 则指导扮演卧底的模型如何通过模仿其他玩家的描述来隐藏自己。这些模块共同构成了一个完整的框架，使模型能够在游戏中扮演不同角色并展现出合理的行为。
  + 论文中的实验是如何设计的？
    - 论文的实验设计包括三个主要部分：首先是基线测试，只使用简单的提示让模型参与游戏；然后是逐步添加不同 CoT 模块的测试，包括只使用 Judge CoT、使用 Judge CoT 和 Describe CoT 的组合、使用所有三个 CoT 模块的完整框架；最后是针对具体案例的分析，考察模型在不同情况下的表现。实验评估指标包括平民胜率、平民投票错误率等量化指标，以及对模型行为的定性分析。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验使用了 100 组不同的词组数据集，这些词组来自 OpenAI 的 o1 模型进行词语相似度判断得到。实验中使用了 Zhipu AI 的 GLM4-9b-Flash 作为基线模型。根据论文信息，代码已在 GitHub 上开源，但具体链接在论文中未给出。实验记录了每场游戏中的角色分配、描述内容、投票结果等详细数据。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果较好地支持了论文的科学假设。实验显示，添加 CoT 模块后，模型的表现显著提升：平民胜率从基线的 7% 提升到 81%，平民投票错误率从 66.7% 降低到 16.7%。案例分析也表明，模型能够进行合理的推理和描述。不过，Spy CoT 模块的效果不够理想，这部分假设得到的支持相对较弱。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献有三个：第一，提出了一个无需训练的框架，使大语言模型能够在文字游戏中展现出良好的推理能力；第二，设计了三个相互配合的 CoT 模块，分别解决了描述、推理和角色扮演等关键问题；第三，通过大规模实验验证了框架的有效性，并提供了详细的案例分析。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的工作可以从几个方面深入：首先是优化 Spy CoT 模块的设计，提高卧底角色的表现；其次是探索如何使模型在多轮交互中保持更好的一致性；再次是将框架推广到其他类型的文字游戏中；最后是研究如何让模型学会更复杂的策略，比如在适当的时机改变描述策略等。

## **基于大语言模型的开源情报智能体系统:集成记忆、知识整合、工具应用与自反馈机制 (Zhijie Shen, Qian Wu, Keyu Shen)**

**亮点：**这篇论文的重点是提出了一个创新的 OSINT 系统架构，将 LLM 作为核心，通过模块化设计和多技术协同实现高效的情报分析。论文的亮点在于系统地整合了多项先进技术，包括知识图谱、RAG、多源数据采集等，并通过实验验证了系统在实际应用中的可行性。特别值得注意的是论文详细分析了系统的局限性并提出了具有针对性的改进方向，为该领域的未来研究提供了有价值的思路。

**介绍：**这篇论文提出了一个基于大语言模型的开源情报系统，通过整合知识图谱、动态查询机制和多源数据采集等技术，构建了一个完整的 OSINT 解决方案。系统采用模块化设计，包含数据收集、知识管理和智能分析三个主要子系统，能够有效处理多源异构数据并生成有价值的情报分析结果。实验表明系统在多个实际任务中表现良好，尤其在事件总结和人物分析方面取得了较高的准确率。然而，系统在处理多模态数据、自动化评估等方面仍有改进空间。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文试图解决开源情报 (OSINT) 收集和分析中的系统性挑战。传统的情报收集方法如使馆咨询、贸易展览和人工调查等存在数据源静态、工作流程分散、信息检索缓慢以及运营成本高等问题。虽然已有一些 AI 驱动的 OSINT 方法，但它们往往过度依赖单一数据源 (如 Twitter)，且无法有效对齐完整的情报周期，特别是在信息分析、上下文化和及时传递给利益相关者等方面存在不足。论文提出了一个基于大语言模型的 OSINT 智能体系统，通过整合记忆模块、领域特定知识图谱、自动化工具和自反馈机制，来实现高效的数据收集、处理和分析。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题。开源情报分析一直是情报和商业领域的重要课题。但随着数据源的激增和市场环境的快速变化，传统方法已难以应对。虽然已有研究将 AI 技术应用于 OSINT，但大多集中在情报周期的处理和分析阶段，对规划和评估阶段关注不足。同时，现有解决方案在处理多源数据、实时分析、上下文理解等方面仍有明显短板。本文提出将大语言模型作为核心，结合多种先进技术构建完整的 OSINT 解决方案，这种系统性的整合方案是具有创新性的。
  + 这篇文章要验证一个什么科学假设？
    - 这篇论文试图验证基于大语言模型的 OSINT 系统能否有效解决当前开源情报分析面临的主要挑战。具体而言，论文假设通过整合 LLM、知识图谱、多源数据采集和自反馈机制，可以构建一个能够实现高效、准确和实时情报分析的系统。这个假设涉及系统在人物分析、组织结构提取、事件总结和实体关系映射等多个具体任务上的表现。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为两类：传统 OSINT 工具 (如 Spiderfoot、Maltego 和 Sublist3r) 和基于 AI 的 OSINT 方法。在 AI 方法中，主要使用 SVM、朴素贝叶斯、随机森林以及 CNN、LSTM 等神经网络。具体应用包括使用 SVM 进行假新闻检测、CNN 进行社交媒体图像分析、混合模型进行恶意域名检测等。论文没有特别突出提到该领域的核心研究者，这可能是因为 OSINT 研究较为分散，或者论文更侧重于工程实现而非学术创新。
  + 论文中提到的解决方案之关键是什么？
    - 该解决方案的关键在于以 LLM 为核心，构建了一个模块化的 OSINT 系统架构。系统包含三个关键子系统：数据收集子系统 (支持多源数据采集和预处理)、知识图谱和动态查询机制 (基于 Neo4j 实现)、以及基于 LLM 的交互式 OSINT 智能体。智能体通过语义理解、记忆模块、知识集成、工具自动化和自反馈机制等功能模块的协同工作，实现从原始数据到可用情报的转化。特别是系统采用了基于提示的属性提取表和检索增强生成 (RAG) 技术，提高了领域知识提取的准确性和适应性。
  + 论文中的实验是如何设计的？
    - 实验设计主要围绕四类预定义的 OSINT 任务：人物档案分析、组织结构提取、事件总结和实体关系分析。这些任务被选择是因为它们能够全面测试系统的数据处理和分析能力。实验采用人工评估方式，由 5 位领域专家对系统输出进行评估，使用二元评分系统 (完全正确为 1，含有任何错误为 0)，总共评估了 200 个结果。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 论文没有明确说明使用了什么特定的数据集，也没有提到代码是否开源。实验似乎使用的是从各种公开渠道收集的真实数据，但缺乏标准化的测试集。这是论文的一个明显不足，影响了研究的可重复性和结果的可比性。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果部分支持了论文的科学假设。系统在不同任务上展现出不同程度的性能：事件总结达到 95% 的准确率，人物档案分析达到 82%，实体关系 71%，组织结构 51%。然而，实验评估过于依赖人工判断，缺乏客观的评估指标和基准对比，这在某种程度上削弱了结果的说服力。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献在于提出了一个完整的基于 LLM 的 OSINT 系统框架，整合了多项先进技术，包括知识图谱、动态查询机制、RAG 等。系统的模块化设计允许灵活扩展和适应新的领域需求。同时，论文详细分析了系统在实际应用中的优势和局限性，为未来研究提供了有价值的参考。
  + 下一步呢？有什么工作可以继续深入？
    - 未来工作主要可以从以下几个方面深入：改进多模态信息处理能力，开发更自动化的评估机制，增强生成内容的准确性以对抗偏见和幻觉，提升系统的自学习能力，以及探索多智能体协作机制。特别是在处理 konflicting 信息、集成多模态数据、验证事实准确性等方面仍有很大的改进空间。

## **LegalAgentBench: 法律领域大语言模型智能体评测基准 (Junjie Chen, Ruowen Zhao, Zhiyuan Feng)**

**亮点：**这篇论文的重点是构建了首个专门用于评估法律领域 LLM 代理能力的基准数据集 LegalAgentBench。其亮点在于：一是采用真实法律场景构建评估环境，确保了评估的实用性；二是设计了可扩展的任务构建框架，支持多样化的任务类型和难度级别；三是引入了细粒度评估指标，能够更全面地评估模型能力。这些创新为推动 LLM 在法律领域的应用提供了重要支持。

**介绍：**这篇论文提出了 LegalAgentBench，这是一个专门设计用于评估法律领域 LLM 代理能力的综合基准。论文首先构建了包含 17 个专业语料库和 37 个工具的评估环境，然后提出了一个可扩展的任务构建框架，最终构建了 300 个覆盖不同类型和难度的任务。通过对八个知名 LLM 的实验评估，论文不仅验证了该基准的有效性，还发现了当前模型在法律领域应用中的一些局限性，为未来研究提供了重要参考。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文试图解决大语言模型 (LLM) 作为智能代理在法律领域应用时缺乏标准化评估基准的问题。随着 LLM 在法律领域的应用日益广泛，现有的通用领域评估基准无法充分捕捉真实世界司法认知和决策的复杂性与微妙之处。虽然已有一些法律领域的数据集，但它们主要聚焦于法律案例检索或判决预测等基础任务，而实际法律实践涉及更深入的案例分析、法律推理以及基于大量法律和判例的综合判断。因此，作者提出了 LegalAgentBench，这是一个专门设计用于评估法律领域 LLM 代理能力的综合基准。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题，但是作者的解决方案具有创新性。先前已有研究者开发了用于评估 LLM 代理能力的基准，如 AgentBench 和 ToolBench，但这些基准主要针对通用领域。同时，也有研究者开发了法律领域的数据集，但主要集中在基础任务上。本文首次将 LLM 代理能力评估与法律领域特定需求相结合，提出了一个全面的评估框架，这种结合方式是新颖的。
  + 这篇文章要验证一个什么科学假设？
    - 这篇文章要验证的科学假设是：通过构建一个包含真实法律场景、多样化任务类型和难度级别、细粒度评估指标的综合评估基准，能够更好地评估和理解 LLM 代理在法律领域的能力边界，从而促进其在实际法律应用中的发展。这个假设包含三个层面：首先，真实法律场景的重要性；其次，任务多样性对全面评估的必要性；最后，细粒度评估对于深入理解模型能力的价值。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要可以分为两类：第一类是关于 LLM 代理的研究，包括 Chain-of-Thought、ReAct 等推理策略的发展，以及 HuggingGPT、LLM+P 等与外部工具集成的方法。第二类是关于 LLM 评估基准的研究，包括 AgentBench、AgentBoard、ToolQA、T-Eval 等。在这个领域内，值得关注的研究人员包括开发 ReAct 的 Shunyu Yao 团队，以及开发 AgentBench 的 Xiao Liu 团队。
  + 论文中提到的解决方案之关键是什么？
    - 论文解决方案的关键在于其提出的可扩展任务构建框架，该框架包括：1) 基于语料库和工具构建规划树；2) 通过分层采样和最大覆盖策略选择任务；3) 使用 GPT-4 改写问题使其更贴近实际应用；4) 通过中间步骤关键词标注实现细粒度评估。此外，论文还提供了 17 个专业语料库和 37 个工具，使 LLM 代理能够更好地处理法律任务。
  + 论文中的实验是如何设计的？
    - 论文的实验设计主要包括三个方面：首先，评估了八个知名 LLM（包括 GLM-4、GPT-4o 等）在 LegalAgentBench 上的表现；其次，比较了三种不同的方法（Plan-and-Solve、Plan-and-Execute、ReAct）的效果；最后，通过不同类型查询（从 1-hop 到 5-hop 以及写作任务）的测试来全面评估模型性能。实验采用了成功率、进度率和 BERT-Score 等多个评估指标。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 用于定量评估的数据集是 LegalAgentBench，包含 300 个不同类型和难度的任务，涵盖 17 个专业语料库和 37 个工具。这些任务包括多跳推理和写作任务。数据集的代码已在 GitHub 上开源，地址为 <https://github.com/cjj826/LegalAgentBench>。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 论文中的实验结果很好地支持了其科学假设。实验表明，LegalAgentBench 确实能够有效区分不同 LLM 的能力，反映出模型在工具使用和逻辑推理方面的差异。实验结果也显示，随着任务难度增加，所有基线模型的性能都会下降，证实了该基准覆盖了不同难度级别的任务。此外，细粒度评估指标（如进度率）确实提供了更深入的模型能力洞察。
  + 这篇论文到底有什么贡献？
    - 这篇论文的主要贡献有三个：首先，提出了首个专门用于评估法律领域 LLM 代理能力的基准数据集；其次，设计了一个可扩展的任务构建框架，支持多样化任务类型和难度级别；最后，引入了细粒度评估指标，能够更全面地评估模型能力。这些贡献为推动 LLM 在法律领域的应用提供了重要支持。
  + 下一步呢？有什么工作可以继续深入？
    - 未来的工作可以从以下几个方面继续深入：第一，扩展数据集支持英语等多语言，增强其国际适用性；第二，丰富任务类型，纳入更多国家和地区的法律场景；第三，改进对特定法律知识的评估方法；第四，开发更有效的法律信息检索和推理策略。

## **基于检索增强生成的个人投资助手 (Xue Zeng, Yinuo Li)**

**亮点：**论文的重点是提出了一个新的个人投资助手系统框架，其亮点在于创新性地将图数据库应用于 RAG 系统，很好地解决了传统向量数据库在保留文档关系方面的不足。通过严格的实验评估，证明了该方法在实际应用中的有效性，为 AI 辅助投资决策开辟了新的研究方向。

**介绍：**这篇论文提出了一个创新的个人投资助手系统 RBPA，通过将 RAG 技术与图数据库相结合，并对基础模型进行微调，解决了传统 LLM 在金融市场分析中的局限性。系统通过构建专业知识库和实时数据检索，提供个性化的投资建议。实验结果表明，该方法在投资回报率和推理质量方面都取得了显著改善。

* 论文十问
  + 论文试图解决什么问题？
    - 这篇论文试图解决个人投资助手系统中的关键问题。传统的大语言模型 (LLM) 在金融市场分析中存在局限性，无法准确把握市场的动态性和复杂性。论文提出了一个基于检索增强生成 (RAG) 的个人投资助手系统 RBPA，通过整合专业知识库和实时数据检索来提供个性化的投资建议。
  + 这是否是一个新的问题？
    - 这不是一个全新的问题，但论文提出了新颖的解决方案。虽然 RAG 技术和 AI 投资分析已有相关研究，但将图数据库与 RAG 结合用于个人投资建议是创新性的。这种方法可以更好地保留文档间的长距离关系，捕获更完整的信息。
  + 这篇文章要验证一个什么科学假设？
    - 论文要验证的科学假设是：通过将 RAG 技术与图数据库结合，并对基础模型进行针对性微调，可以构建一个比传统 LLM 更准确、更可靠的个人投资助手系统。这个系统应该能够提供更好的投资回报率，并具有更强的推理能力。
  + 有哪些相关研究？如何归类？谁是这一课题在领域内值得关注的研究员？
    - 相关研究主要分为两类：RAG 相关技术研究和 AI 投资领域研究。在 RAG 方面，重要工作包括 LangChain 框架、OpenAI 的 text-embedding-3 模型、Palm 和 Mistral 7B 等。在 AI 投资方面，主要包括 Kim 和 Nikolaev 的金融信息解释研究，以及 Citadel 等机构的实践应用。值得关注的研究人员包括 Chowdhery、Jiang 等人。
  + 论文中提到的解决方案之关键是什么？
    - 解决方案的关键在于两个创新点：首先是使用图数据库替代传统向量数据库来构建知识库，更好地保留文档间的关系；其次是通过 LoRA 对 GLM-4-Flash 模型进行微调，增强其在金融领域的分析能力。
  + 论文中的实验是如何设计的？
    - 实验设计包括两个主要部分：ROI 评估和推理质量评估。在 ROI 评估中，通过预测股票价格变动来构建投资组合，并计算长空组合收益等指标。在推理质量评估中，使用 Qwen2-7B 评估模型的准确性、深度和全面性。
  + 用于定量评估的数据集是什么？代码有没有开源？
    - 实验使用了多个数据集，包括 CSMAR 的金融新闻数据 (75k 条)、Uger 的股票价格和成交量数据 (7，498k 条)、以及 Choice 的分析师报告 (300 条)。这些数据涵盖 2018 年 1 月至 2024 年 12 月。论文中没有明确提到代码是否开源。
  + 论文中的实验及结果有没有很好地支持需要验证的科学假设？
    - 实验结果很好地支持了研究假设。微调后的模型在准确率上从 29% 提升到 42%，投资组合收益率从 - 45% 提升到 12%。在推理质量评估中，模型在准确性和深度方面都有显著提升。
  + 这篇论文到底有什么贡献？
    - 论文的主要贡献在于：提出了一个新的 RAG 与图数据库结合的框架；开发了一个全面的评估基准；证明了该方法在个人投资建议领域的有效性。
  + 下一步呢？有什么工作可以继续深入？
    - 未来工作可以从几个方向深入：扩大知识库覆盖范围；增强模型的逻辑推理能力；开发更复杂的投资策略；提高系统的实时性能。

# 附录

## 1. 实际汇报组

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 序号 | **组号** | **题目** | **组长** | **合并得分** |
| 1 | 8 | Efficient Inference for Large Multimodal Models | 杨超群 | 67.84285714 |
| 2 | 12 | Leveraging LLM-based Multi-Agent Collaboration to Enhance Embodied Agents’ Reasoning Capabilities for Solving Text-based Tasks in Human-populated Environments | 孙楠 | 67.11304348 |
| 3 | 19 | A Reinforcement Learning Technique for Large Language Model on Self-Verifiable Problems | 李颖欣 | 62.04210526 |
| 4 | 6 | A Multi-Power Law for Loss Curve Prediction Across Learning Rate Schedules | 江楚萌 | 57.75625 |
| 5 | 7 | LegalAgentBench: Evaluating LLM Agents in Legal Domain | 陈俊杰 | 57.712 |
| 6 | 18 | BenCourt: A Benchmark and Framework for Court Simulation using LLM-based Agents | 张开元 | 52.19310345 |
|  | 24 | Scene-Clipping Long Video For Better Understanding | 赵子钰 | 45.52222222 |
| 7 | 16 | ADSFT: Adaptive and Domain-Specific Fine-Tuning Paradigms for Generative Models | 王俊逸 | 42.1 |
|  | 2 | Exploring Large Language Models for Communication Games: Who is the Spy？ | 魏晨天 | 40.75454545 |
| 8 | 1 | Logic Extraction: Enhancing AI Generalization in Abstraction and Reasoning Corpus Tasks | 陈远乐 | 39.98333333 |
|  | 9 | A Comparative Study of RoPE-based Positional Encodings from A Scaling Perspective | 张柱 | 39.93333333 |
|  | 27 | RBPA: Retrieval Augmented Generation based Personal Investment Assistant | 李一诺 | 38.07142857 |
| 9 | 23 | Syn-QL: Prefernce Optimization with Synthetic Data for Text-to-SQL | 胡瑞麟 | 37.70526316 |
| 10 | 10 | RLOJF: Enhancing LLMs in Olympiad Programming with Online Judge Feedback | 汪子涵 | 37.17692308 |
|  | 22 | Agent for Oversea Marketing Support | 沈智杰 | 31.33684211 |
|  | 5 | VideoAlign: A Comprehensive Model for Evaluating Alignment Between Text and Generated Videos | 杨苑明 | 30.54615385 |
|  | 15 | A Large Language Model-based Bandwidth Prediction Algorithm | 蒋政 | 26.85384615 |
|  | 14 | Self-reflection like Humans, Editable-LLM (E-LLM) is All You Need | 吴彦辰 | 26.4 |
|  | 26 | Transformer Models for Predicting Material Properties | 张开维 | 26.36666667 |
|  | 3 | Systematic Idea Refinement for Machine Learning Research Agents | 刘子君 | 26 |
|  | 11 | Fine-Grained Reward Modeling in LLMs: An RL, PRM, and Memory-Augmented Approach for Advanced Reasoning | 任挽澜 | 24.424 |
|  | 17 | A Multi-Agent Framework for Reliable Drug-Target Interaction Prediction | 高博文 | 24.21 |
|  | 28 | Doubly Robust Estimation for Correcting Position Bias in LLM recommendations | 孙卓凡 | 21.93913043 |
|  | 13 | AI-empowered Intelligent Education: Question Generation based on LLMs | 郑克寒 | 21.85517241 |
|  | 25 | Pokémon Battle Agent based on LLMs | 岑奇航 | 21.8 |
|  | 21 | Enhanced medical image captioning via dual encoder | 胥广洁 | 21.52142857 |
|  | 4 | A Reinforcement Learning Framework for Protein Function Prediction | 胡晨曦 | 21.488 |
|  | 20 | MathLLaMA: A Specialized Language Model for Mathematical Reasoning and Problem-Solving | 雷长松 | 21.31428571 |
|  | 29 | Mini-Omni: Language Models Can Hear, Talk While Thinking in Streaming | 谢之非 | 0 |

## 2. 计算方式说明

1.互评得分计算规则：去掉大家在网络学堂互评中全部的最高分和最低分后取平均,最高为10，最低为0，占比20%

2.组内意向来自网络学堂问卷中“1单选:无论互评结果如何，你是否有意愿进行现场汇报(正确答案：有且强烈希望)”，无或者没有填写为0，可有可无为50，有为100，占比30%

3.助教意向：征集各个助教和老师的建议后给出的评分，有一个助教推荐为100，老师推荐为200，最高为500，最低为0，占比50%（目前这一部分应该是只收到两个助教的良性评分，老师和第三位助教对各组均表示同等态度，故这一部分实际应该最低为0，最高为200）

4.计算方式是：10\*10\*0.2+100\*0.3+500\*0.1=100\*0.2+100\*0.3+100\*0.5=100

### 2.1 互评中获得的详细分数

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **序号** | **标题** | **中文标题** | **得分列表** | **平均得分** |
| 2 | [AML&ML]Course Conclusion Report Guide | [AML&ML]课程结论报告指南 | 7, 10.0, 10.0, 10, 10, 9.0, 10.0, 8, 10.0, 10, 10.0, 10.0 | 9.5 |
| 3 | [AML] BP-LLM: A Large Language Model-based Approach for Accurate and Adaptable Bandwidth Prediction | [AML] BP-LLM: 基于大型语言模型的准确且可适应的带宽预测方法 | 8, 8.0, 8, 8.0, 8.6, 8.0, 9.0, 9, 8, 8.5, 9.0, 8.0, 8.0, 8.5, 9.0, 10.0, 9, 8.5, 8.0, 8.0, 8.0, 9, 9.0, 9.0, 7.5, 7.5 | 8.426923077 |
| 4 | RM-PoT: Reformulating Mathematical Problems and Solving via Program of Thoughts | RM-PoT: 通过思维程序重新表述并解决数学问题 | 8.0, 8, 8.0, 8.0, 9, 8.0, 9.0, 9.0, 9, 8, 8.0, 9.0, 8.0, 8.5, 9.0, 7.0, 8.0, 7, 8.4, 9.0, 8.0, 8.0, 9, 8.5, 8.0, 7.0, 7.0, 7.0 | 8.157142857 |
| 5 | [AML]Exploring Large Language Models for Word Games: Who is the Spy？ | [AML] 探索大型语言模型用于文字游戏：谁是间谍？ | 7, 8, 8.0, 8.0, 9, 8, 8.0, 7.5, 7.0, 7.5, 7.0, 8.0, 7.0, 8, 8.3, 7.5, 9, 7.5, 9.0, 7.0, 9.0, 8.0 | 7.877272727 |
| 6 | 【AML】Enhanced Image Captioning Approach via Dual Encoder | 【AML】通过双编码器增强图像文字描述方法 | 8.0, 8.0, 8, 7.0, 9.4, 8.5, 6.5, 8, 9.0, 7, 9.0, 8.5, 9.0, 7.0, 10.0, 8.0, 8.0, 6.0, 10.0, 7, 7.9, 8.0, 10.0, 9, 9.0, 9.0, 6.5, 10.0 | 8.260714286 |
| 7 | AASD: Accelerate Inference by Aligning Speculative Decoding in Multimodal Large Language Models | AASD: 通过多模态大型语言模型中的猜测解码加速推理 | 9, 9.0, 9.6, 9.0, 10.0, 9.0, 9.0, 9.0, 9.0, 9, 9.0, 8, 8.5, 10.0, 10.0, 8.0, 8.0, 7.5, 8.0, 9, 9.2, 8.0, 10.0, 9, 10.0, 9.0, 8.0, 9.0 | 8.921428571 |
| 8 | [AML] Enhancing Protein Function Prediction: Integrating Pretraining and Fine-Tuning within Geometric-Aware Graph Convolutional Networks | [AML] 增强蛋白质功能预测：在几何感知图卷积网络中集成预训练和微调 | 8, 9.0, 8, 7.0, 9.7, 8.0, 7.5, 8, 8.0, 8, 8.5, 8.9, 9.0, 8.0, 8.0, 9.0, 8, 8.5, 9.0, 8.0, 9, 8.0, 8.0, 7.0, 8.0 | 8.244 |
| 9 | LLM-IR: Leveraging Large Language Models for Intent Recognition in Multimodal Dialogue Systems | LLM-IR: 利用大型语言模型进行多模态对话系统中的意图识别 | 8, 9.0, 9.6, 10.0, 10.0, 8.5, 9.0, 8, 7, 9.0, 9.0, 7.0, 9.0, 8.0, 7, 7.6, 8.0, 8.0, 10.0, 9, 9.0, 9.0, 7.5, 9.0 | 8.55 |
| 10 | [AML] T$^5$-ARC: Test-Time Training for Transductive Transformer Models in ARC-AGI Challenge | [AML] T$^5$-ARC: 在ARC-AGI挑战中进行传道转化器模型的测试时间训练 | 7, 8.0, 8, 9.0, 9.1, 9.0, 8.0, 8, 8, 8.5, 9.0, 10.0, 9.0, 8.0, 9.0, 9.0, 8, 8.2, 9.0, 8.0, 9, 8.5, 8.0, 8.5 | 8.491666667 |
| 11 | [AML]Orchestrating Multi-Agent Alignment for Large Language Models Through External Memory and Adaptive Reasoning | [AML] 通过外部存储器和自适应推理协调多代理一起对齐大型语言模型 | 8, 9, 9.0, 9.0, 7.0, 9.0, 9.0, 7.0, 8, 8.0, 8.0, 7, 9.0, 9.0, 8.0, 8.0, 7.0, 7.0, 8, 8.3, 8.0, 9, 9.0, 9.0, 8.0 | 8.212 |
| 12 | [AML] A Multi-Power Law for Loss Curve Prediction Across Learning Rate Schedules | [AML] 跨学习率计划的损失曲线预测的多力法 | 8, 9, 8.0, 9.2, 7.0, 8.0, 9.5, 9.5, 9.0, 8, 9.5, 10.0, 7, 10.0, 10.0, 8.0, 9.0, 8.0, 10.0, 9.5, 10.0, 9.5, 10.0, 8.0, 10, 7.9, 8.0, 8, 10.0, 9.0, 9.5, 8.0 | 8.878125 |
| 13 | RLOJF: Reinforcement Learning for Enhanced Code Synthesis using Online Judge Feedback | RLOJF: 利用在线评判反馈增强代码生成的强化学习 | 8, 8.4, 9.0, 9, 9.0, 10.0, 8.0, 9.5, 9.0, 8, 8, 9.0, 9.0, 7.0, 8.5, 8.0, 10.0, 8, 8.9, 9.0, 8.0, 9, 8.5, 9.0, 7.5, 8.0 | 8.588461538 |
| 14 | [AML]Self-reflection like Humans, Editable-LLM (E-LLM) is All You Need | [AML] 像人类一样的自我反思，可编辑的大型语言模型（E-LLM）是你所需的一切 | 9.0, 8, 8.0, 8.8, 9.0, 9.0, 9.0, 8, 8.5, 9.5, 9.0, 7.0, 8.9, 7.0, 7.0, 7.0, 7, 8.1, 8.5, 9, 7.5, 9.0, 7.0, 8.0 | 8.2 |
| 15 | [AML]Enhancing Human-Involved Embodied Task Reasoning through LLM-Based Multi-Agent Collaboration | [AML] 通过基于LLM的多代理协作增强参与人类的具体任务推理 | 8, 8.5, 8, 8.7, 9.0, 9, 9.0, 9, 8.5, 9.5, 8.5, 8.0, 8.0, 9.0, 8.0, 8, 8.1, 8.5, 10.0, 9, 8.5, 9.0, 7.0 | 8.556521739 |
| 16 | RBPA: Retrieval-Augmented-Generation based Personal Investment Assistant | RBPA: 基于检索增强生成的个人投资助手 | 8, 7, 9.0, 8.0, 8.7, 7.0, 8.0, 7.0, 8, 8, 8.0, 8.0, 7.0, 7.0, 8.0, 7.0, 6.0, 7, 7.8, 7.0, 7.0, 7.0, 9, 8.0, 7.0, 7.5, 6.0, 8.0 | 7.535714286 |
| 17 | [AML]Systematic Idea Refinement for Machine Learning Research Agents | [AML] 机器学习研究代理的系统性理念提炼 | 8, 8.0, 8.0, 8.6, 9.0, 8.0, 8.0, 8, 7, 10.0, 10.0, 7.0, 10.0, 8.0, 10.0, 8.0, 9.0, 9, 8.4, 8.5, 8, 9.0, 9.0, 8.0, 8.0 | 8.5 |
| 18 | Sim-Court: A Simulation of Court using LLM-based Agents | 模拟法庭：使用基于LLM的代理来模拟法庭 | 8, 8.0, 8, 7.0, 7.0, 8.9, 8.0, 7.0, 8.0, 9.0, 8, 8.5, 9.0, 9.0, 9.5, 8.5, 9.5, 7.0, 8.0, 9.5, 7, 7.9, 7.5, 9, 7.5, 8.0, 9.5, 7.0, 6.0 | 8.096551724 |
| 19 | [AML] 22. OSINT Agent for Oversea Marketing Support with Memory, Knowledge, Tool-Using and Self-reflection | [AML] 具有记忆、知识、工具使用和自我反思功能的海外市场支持的OSINT代理 | 8, 8, 7.0, 8.9, 8.0, 6.5, 8.0, 8, 8, 8.0, 6.0, 8, 8.3, 7.0, 9, 8.0, 7.0, 8.0, 6.0 | 7.668421053 |
| 20 | [AML] A Multi-Agent Framework for Reliable and Consistent Drug-Target Interaction Prediction Using Large Language Models | [AML] 基于大型语言模型进行可靠且一致药物靶标相互作用预测的多代理框架 | 8, 8, 9.0, 9.0, 8, 8.0, 7, 9.0, 9.0, 8.0, 8.0, 7.0, 8, 8.1, 8.0, 9, 9.0, 8.0, 7.0, 7.0 | 8.105 |
| 21 | [AML] A Reinforcement Learning Technique for Large Language Model on Self-Verifiable Problems | [AML] 用于大型语言模型的自我验证问题的强化学习技术 | 8, 8.2, 8, 7.0, 8.7, 9.0, 8.0, 8, 8, 9.0, 8.0, 7.0, 8.0, 7, 8.0, 9, 8.5, 7.0, 8.0 | 8.021052632 |
| 22 | [AML] Unified Material Transformer as Scalable Material Property Predictor | [AML] 统一材料转化器作为可扩展的材料属性预测器 | 8, 9, 8.0, 8.7, 8.0, 7.0, 10.0, 8, 7, 9.0, 9.0, 9.8, 8.0, 7.0, 8.0, 9, 7.9, 8.0, 9, 8.5, 8.0, 8.5, 6.0, 7.0 | 8.183333333 |
| 23 | [AML]Pokémon Battle Agent based on LLMs | [AML] 基于大型语言模型的神奇宝贝对战代理 | 8, 7.0, 9.0, 8.6, 8.0, 7.0, 8.0, 7.0, 9, 9.0, 7, 8.0, 8.0, 8.0, 7.0, 9.0, 7.5, 8.0, 8.0, 7, 8.0, 7.5, 8.0, 9, 7.5, 8.0, 7.0, 7.0, 9.0 | 7.9 |
| 24 | [AML]Leveraging LLMs for Causal Inference and Discovery | [AML] 利用大型语言模型进行因果推断和发现 | 8, 8.0, 7, 8.0, 8.6, 8.0, 8.0, 8.0, 8, 7, 9.0, 8.0, 7.5, 8.0, 7.0, 7, 8.2, 8.0, 9.0, 9, 8.0, 8.0, 8.0 | 7.969565217 |
| 25 | [AML] S$^3$PE: A Simple Scalable Sigmoid-Style Position Encoding | [AML] S$^3$PE: 一种简单可扩展的Sigmoid风格位置编码 | 8, 8.0, 9, 9.0, 8.6, 9.0, 8.0, 8, 9, 9.5, 9.0, 8.0, 8, 9.2, 9.0, 8.0, 7.0, 8.5, 9.0, 8.0, 8.0 | 8.466666667 |
| 26 | [AML]LegalAgentBench: Evaluating LLM Agents in Legal Domain | [AML] LegalAgentBench：评估法律领域的大型语言模型代理 | 8.0, 8, 9.5, 8.0, 8.5, 8.0, 9, 10, 10.0, 10.0, 8.5, 9.5, 10.0, 9.5, 7.0, 9, 8.4, 8.0, 10.0, 8.0, 9, 10.0, 9.0, 9.5, 7.0 | 8.856 |
| 27 | Enhancing Text-to-SQL with Open-source LLMs | 利用开源大型语言模型增强Text-to-SQL | 8, 9.0, 8.3, 8.0, 9.0, 8, 8, 9.0, 9.0, 8.0, 7, 7.9, 10.0, 8.0, 8, 9.0, 9.0, 9.5, 6.0 | 8.352631579 |
| 28 | [AML] AI-empowered Intelligent Education: Question Generation based on LLMs | [AML] AI赋能的智能教育：基于大型语言模型的问题生成 | 8, 7.7, 9, 8.0, 8.3, 8.0, 7.5, 7.0, 8.0, 8, 7, 9.5, 9.5, 6.6, 7.0, 8.0, 7.0, 9.5, 9.0, 8, 7.8, 7.0, 7.5, 8.0, 8, 9.0, 8.0, 7.0, 7.0 | 7.927586207 |
| 29 | Scene-Clipping Long Video For Better Understanding | 场景剪辑长视频以更好地理解 | 8, 8, 8.0, 8.5, 8.5, 8, 7, 9.0, 9.0, 8.5, 9.0, 7, 8.2, 9.0, 8.0, 9, 8.0, 8.0 | 8.261111111 |
| 30 | [AML]CleveReward: Contrastive Learning-Engined Video Reward Training on Different Benchmark Datasets | [AML] CleveReward: 在不同基准数据集上进行对比学习引擎的视频奖励训练 | 8, 9.0, 8, 7.0, 8.7, 9.0, 9.0, 6.0, 7.0, 9, 9.0, 9, 9.0, 9.8, 6.0, 9.5, 7, 8.1, 8.0, 8.0, 10, 10.0, 9.0, 7.0, 7.0, 8.0 | 8.273076923 |

### 2.2 总评中的详细分数

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **组号** | **题目** | **组内意向** | **互评得分** | **助教意向** | **合并得分** | **组长** | **姓名** |
| 1 | Logic Extraction: Enhancing AI Generalization in Abstraction and Reasoning Corpus Tasks | 50 | 8.49166666666667 | 80 | 39.98333333 | 陈远乐 | 陈远乐 |
| 耿双越 |
| 朱少廷 |
| 2 | Exploring Large Language Models for Communication Games: Who is the Spy？ | 50 | 7.87727272727273 | 100 | 40.75454545 | 魏晨天 | 魏晨天 |
| 许金朱 |
| 陈杰伟 |
| 3 | Systematic Idea Refinement for Machine Learning Research Agents | 0 | 8.50000000000000 | 90 | 26 | 刘子君 | 刘子君 |
| 高骋 |
| 朱晗希 |
| 4 | A Reinforcement Learning Framework for Protein Function Prediction | 0 | 8.24400000000000 | 50 | 21.488 | 胡晨曦 | 田仁睿 |
| 龙飞 |
| 胡晨曦 |
| 5 | VideoAlign: A Comprehensive Model for Evaluating Alignment Between Text and Generated Videos | 0 | 8.27307692307692 | 140 | 30.54615385 | 杨苑明 | 杨苑明 |
| 刘晓倩 |
| 陈健 |
| 6 | A Multi-Power Law for Loss Curve Prediction Across Learning Rate Schedules | 100 | 8.87812500000000 | 100 | 57.75625 | 江楚萌 | 江楚萌 |
| 罗开荣 |
| 潘至璇 |
| 7 | LegalAgentBench: Evaluating LLM Agents in Legal Domain | 100 | 8.85600000000000 | 100 | 57.712 | 陈俊杰 | 陈俊杰 |
| 赵若雯 |
| 冯志远 |
| 8 | Efficient Inference for Large Multimodal Models | 100 | 8.92142857142857 | 200 | 67.84285714 | 杨超群 | 杨超群 |
| 张远达 |
| 黄茜瑛 |
| 9 | A Comparative Study of RoPE-based Positional Encodings from A Scaling Perspective | 50 | 8.46666666666666 | 80 | 39.93333333 | 张柱 | 杨天行 |
| 闫梓涵 |
| 张柱 |
| 10 | RLOJF: Enhancing LLMs in Olympiad Programming with Online Judge Feedback | 0 | 8.58846153846154 | 200 | 37.17692308 | 汪子涵 | 徐家骏 |
| 吴垒 |
| 汪子涵 |
| 11 | Fine-Grained Reward Modeling in LLMs: An RL, PRM, and Memory-Augmented Approach for Advanced Reasoning | 0 | 8.21200000000000 | 80 | 24.424 | 任挽澜 | 刘家祥 |
|
| 任挽澜 |
| 12 | Leveraging LLM-based Multi-Agent Collaboration to Enhance Embodied Agents’ Reasoning Capabilities for Solving Text-based Tasks in Human-populated Environments | 100 | 8.55652173913044 | 200 | 67.11304348 | 孙楠 | 孙楠 |
| 施铖铭 |
| 董毓文 |
| 13 | AI-empowered Intelligent Education: Question Generation based on LLMs | 0 | 7.92758620689655 | 60 | 21.85517241 | 郑克寒 | 路宜达 |
| 郑克寒 |
| 武文静 |
| 14 | Self-reflection like Humans, Editable-LLM (E-LLM) is All You Need | 0 | 8.20000000000000 | 100 | 26.4 | 吴彦辰 | 吴彦辰 |
| 徐港新 |
| 邹冬晨 |
| 15 | A Large Language Model-based Bandwidth Prediction Algorithm | 0 | 8.42692307692308 | 100 | 26.85384615 | 蒋政 | 陈鑫 |
| 容逸朗 |
| 蒋政 |
| 16 | ADSFT: Adaptive and Domain-Specific Fine-Tuning Paradigms for Generative Models | 50 | 8.55000000000000 | 100 | 42.1 | 王俊逸 | 王俊逸 |
| 隋元培 |
| 刘涛 |
| 17 | A Multi-Agent Framework for Reliable Drug-Target Interaction Prediction | 0 | 8.10500000000000 | 80 | 24.21 | 高博文 | 高博文 |
| 白华君 |
| 张沛东 |
| 18 | BenCourt: A Benchmark and Framework for Court Simulation using LLM-based Agents | 100 | 8.09655172413793 | 60 | 52.19310345 | 张开元 | 张开元 |
| 李兆曦 |
| 李宣成 |
| 19 | A Reinforcement Learning Technique for Large Language Model on Self-Verifiable Problems | 100 | 8.02105263157895 | 160 | 62.04210526 | 李颖欣 | 李颖欣 |
| 何光磊 |
| 周九阳 |
| 20 | MathLLaMA: A Specialized Language Model for Mathematical Reasoning and Problem-Solving | 0 | 8.15714285714286 | 50 | 21.31428571 | 雷长松 | 张宇 |
| 雷长松 |
| 21 | Enhanced medical image captioning via dual encoder | 0 | 8.26071428571429 | 50 | 21.52142857 | 胥广洁 | 胥广洁 |
| 高杨驰 |
| 22 | Agent for Oversea Marketing Support | 0 | 7.668421053 | 160 | 31.33684211 | 沈智杰 | 沈智杰 |
| 吴倩 |
| 沈珂宇 |
| 23 | Syn-QL: Prefernce Optimization with Synthetic Data for Text-to-SQL | 50 | 8.35263157894737 | 60 | 37.70526316 | 胡瑞麟 | 樊陆 |
| 胡瑞麟 |
| 陈奕哲 |
| 24 | Scene-Clipping Long Video For Better Understanding | 50 | 8.26111111111111 | 140 | 45.52222222 | 赵子钰 | 赵子钰 |
| 王锦 |
| 肖锦松 |
| 25 | Pokémon Battle Agent based on LLMs | 0 | 7.90000000000000 | 60 | 21.8 | 岑奇航 | 岑奇航 |
| 吕子涵 |
| 26 | Transformer Models for Predicting Material Properties | 0 | 8.18333333333333 | 100 | 26.36666667 | 张开维 | 于竣丞 |
| 张开维 |
| 李昊楠 |
| 27 | RBPA: Retrieval Augmented Generation based Personal Investment Assistant | 50 | 7.53571428571429 | 80 | 38.07142857 | 李一诺 | 曾雪 |
| 李一诺 |
| 28 | Doubly Robust Estimation for Correcting Position Bias in LLM recommendations | 0 | 7.96956521739130 | 60 | 21.93913043 | 孙卓凡 | 孙卓凡 |
| 李青仪 |
| 29 | Mini-Omni: Language Models Can Hear, Talk While Thinking in Streaming | 0 | 0.00000000000000 | 0 | 0 | 谢之非 | 谢之非 |

## 3. 助教点评

1。优化 创新点很solid 但 比较单一 时间有些短

2。具身智能 偏应用陈述 超时 1min

3。自反馈在解题方面的应用 应用为主：生成+微调

SFT单纯的效果其实就差不多的，所以实验需要验证DPO的必要性

人工标注数据不足，可以进一步优化

4。学习率的可控研究，从scaling law角度蛮有想法的

5。法律智能体评估框架。有点集成主义，可用性足够，但是更新感觉会很依赖人工和时间

6。场景模拟。多角色情况下，不同角色在prompt外的优化是必要的，建议增加角色agent的能力设计。而且prompt的自优化被认为也是必要的，建议考虑。在评估方面，考虑借鉴国外法律流程。（法律方面有标注的通用数据集是有很多的，可以尝试迁移评估）

7。意图识别。目前多轮意图识别是一个比较缺的研究方向，可进一步考虑。然后相关工作的对比有点欠缺。

8。ARC子方向的探索 蛮规范的工作 期待进一步创新

9。SQL语言自动生成 工程上挺扎实的 希望增加相关工作对比（不是说可支持测评的相关模型，还包括TextToSQL的相关开源方案比如 ），毕竟目前很多数据公司都在做这个

10。OJ上的大模型一次通过率提升 数据构造上可以展开说说

## 4. 唐老师建议

1。尝试颠覆传统的范式

2。尝试选择超越界限 比如达到100%，而不是99%

3。真正有意义的不是获得什么，而是做了什么