

早上好!! 今天是 2025-05-06 上午 10 点 14 分, 星期二。接下来将为各位同事播报最多 5 篇 AI 领域相关论文, 大家边听论文播报边开始一天的工作吧!

第 1 篇论文:

论文的标题是: **特征对齐和表示转移在大语言模型知识蒸馏中的应用。**

论文的内容是: 知识蒸馏 (KD) 是一种从复杂的教师模型向较简单的学生模型转移知识的技术, 显著提高了模型的效率和准确性。它在诸多应用中展示了重要进展, 包括图像分类、目标检测、语言建模、文本分类和情感分析。最近在 KD 方法上的创新, 如基于注意力的方法、逐块对数蒸馏和解耦蒸馏, 显著提升了学生模型的表现。这些技术着重于刺激复杂性、注意力机制和全局信息捕获以优化知识转移。此外, KD 被证明在压缩大语言模型的同时保持准确性、减少计算负担和提高推理速度方面是有效的。这篇综述整合了最新的文献, 强调了关键发现、贡献和未来方向, 为研究人员和实践者提供了关于其在人工智能和机器学习中不断变化的角色的见解。

以下是我的思考是: 这篇论文对于我们在 LLM (大语言模型) 上进行训练微调的项目有重要的提升。知识蒸馏技术可以用于提升学生模型的效率和准确性, 这对于我们在税务领域构建高效的垂直 AI 认知模型是非常有帮助的。特别是, 论文中提到的基于注意力的方法和逐块对数蒸馏可以作为我们 fine-tuning (微调) 和 full-tuning (全量调优) 过程中优化模型性能的方法。以下是一个简洁的分步实施方法:

1. **\*\*研究知识蒸馏的最新技术\*\***: 深入理解基于注意力和逐块对数蒸馏的方法, 分析其在大语言模型中的应用潜力。
2. **\*\*模型选择与设计\*\***: 选择合适的教师模型和学生模型, 设计蒸馏过程, 确保在税务领域的应用中保持高效性和准确性。
3. **\*\*实验与验证\*\***: 在税务相关的数据集上进行实验, 验证蒸馏方法对模型性能的提升, 尤其是在减少计算负担和提高推理速度方面的效果。
4. **\*\*优化与调整\*\***: 根据实验结果, 进一步优化蒸馏过程, 调整模型参数, 以达到最佳效果。
5. **\*\*应用与部署\*\***: 将优化后的模型应用到实际的税务场景中, 评估其在实际业务中的表现, 并进行必要的调整。

第 2 篇论文:

论文的标题是: **d1: 通过强化学习扩展扩散大语言模型中的推理能力**

论文的内容是: 最近的大型语言模型 (LLMs) 展示了通过在线强化学习 (RL) 增强的强大推理能力。这些能力主要是在从左到右的自回归 (AR) 生成范式中得到了证明。相比之下, 基于扩散的非自回归范式以由粗到细的方式生成文本。虽然最近的基于扩散的大型语言模型 (dLLMs) 在语言建模性能上与 AR 模型相当, 但尚不清楚 dLLMs 是否也能利用 LLM 推理的最新进展。为此, 我们提出了 d1, 这是一种通过监督微调 (SFT) 和 RL 结合将预训练的掩码 dLLMs 适配为推理模型的框架。具体来说, 我们开发并扩展了技术以提高预训练 dLLMs 的推理能力: (a) 我们利用掩码 SFT 技术直接从现有数据集中提取知识并灌输自我改进行为, 以及 (b) 我们引入了一种称为 diffu-GRPO 的新颖无批评家、基于策略梯度的 RL 算法。通过实证研究, 我们调查了不同后训练方法在多个数学和逻辑推理基准上的表现。我们发现 d1 产生了最佳性能, 并显著提高了最

先进的 dLLM 的表现。

以下是我的思考是：该论文提出了一种将扩散大语言模型（dLLM）适配为推理模型的新框架 d1，通过结合监督微调 and 强化学习，特别是引入了一种新的 RL 算法 diffu-GRPO。这对于我们的项目在训练设计方法和训练优化方法方面有直接的提升意义。我们可以考虑以下步骤实施：

1. **模型选择与准备**：选择合适的基础 dLLM 进行实验，确保其具备良好的语言建模性能。
2. **掩码监督微调**：应用论文中提到的掩码 SFT 技术，从现有税务数据集中提取相关知识，增强模型的自我改进能力。
3. **强化学习适配**：实施论文中介绍的 diffu-GRPO 算法，进行在线强化学习，以进一步优化模型的推理能力。
4. **性能评估与调整**：在多个税务相关的数学和逻辑推理基准上评估模型性能，根据结果进行参数调整和方法优化。

综上，这篇论文的研究方法为我们的项目提供了一个新颖的思路，尤其是在训练优化和推理能力提升方面。

第 3 篇论文：

论文的标题是：**超越下一个标记的视角。**

论文的内容是：因果语言模型训练的结构假设每个标记都可以从先前的上下文中准确预测。这与人类自然的写作和推理过程形成对比，人类通常在明确论点或措辞之前就已知晓目标。虽然文献中已对这种不匹配进行了深入研究，但普遍的假设是需要架构上的改变来解决这种不匹配。我们认为，通过重新排列和处理训练数据序列，模型可以更准确地模仿真实的数据生成过程，而无需对架构或训练基础设施进行任何其他更改。我们证明了这种技术，称为 Trelawney，以及从中推导出的推理算法，使我们能够在涵盖规划、算法推理和故事生成任务的多个关键基准上提高性能。最后，我们的方法自然地支持在不增加额外成本的情况下生成长期目标。我们研究了如何利用模型的目标生成能力进一步改善规划和推理。此外，我们相信 Trelawney 可能会开启超越当前语言建模范式的新能力。以下是我的思考是：1. 考虑在税务领域的 AI 训练中应用 Trelawney 技术，通过重新排列和处理训练数据序列，使模型更好地模拟人类专家在税务规划和决策中使用的推理过程。2. 结合税务场景中的长期目标设定，设计数据集和训练流程，使 AI 能够自然生成并优化长期税务策略。3. 在微调过程中，逐步引入不同层次的规划和推理任务，以考察 Trelawney 技术在提高模型性能方面的具体效果。4. 通过实验验证，在特定税务任务（例如税务报告生成、税务合规性审查等）中引入该技术，评估其对任务准确性和效率的提升。

第 4 篇论文：

论文的标题是：**《语言建模在金融未来中的应用：衡量标准、任务及数据机会的定量调查》。**

论文的内容是：最近语言建模的进展引起了人们对将自然语言处理（NLP）技术应用于金融问题的浓厚兴趣，从而启用了新的分析和决策方法。为了系统地检验这一趋势，我们审查了 2017 年至 2024 年间在 38 个会议和研讨会上发表的 374 篇 NLP 研究论文，并重点分析了直接涉及金融相关任务的 221 篇论文。我们从 11 个定性和定量维度对这些论文进行了评估，确定了关键趋势，如通用语言模型的使用增加、情感分析和信息提取方面的稳步进展，以及围绕可解释性和隐私

保护方法的新兴努力。我们还讨论了评估指标的使用，强调了特定领域指标对标准机器学习指标的补充作用。我们的研究结果强调了更易于访问、适应性强的数据集的需求，并强调了将金融危机时期考虑在内以增强模型在现实条件下的稳健性的重要性。该调查为应用于金融的 NLP 研究提供了结构化的概述，并为在这一交叉领域工作的研究人员和从业者提供了实用见解。

以下是我的思考是：这篇论文对我们的项目有一定的提升。原因是该论文系统地分析了将 NLP 技术应用于金融领域的趋势，并探讨了如何使用特定领域的评估指标来增强模型在金融领域的适应性和稳健性。对于我们正在进行的税务领域的垂直 AI 认知训练项目，论文中提到的关于如何利用语言模型应对金融特定任务的经验和趋势分析，可以帮助我们更好地理解如何在税务场景中应用 NLP 技术。同时，关于数据集的适应性和可访问性以及在实际条件下增强模型稳健性的讨论，也为我们在构建和优化训练数据集时提供了重要的参考。

以下是一个简洁的分步实施方法：

1. **\*\*研究现有的 NLP 应用于金融领域的趋势\*\***：深入理解论文中提到的关键趋势及应用案例。
2. **\*\*分析特定领域评估指标的应用\*\***：探索如何将论文中提到的特定领域指标应用于税务 AI 认知的模型评估中。
3. **\*\*构建适应性强的数据集\*\***：参考论文中关于数据集建设的建议，构建和优化适用于税务场景的数据集。
4. **\*\*增强模型的稳健性\*\***：借鉴论文中关于在金融危机条件下增强模型稳健性的讨论，探索如何在税务场景中实现类似的稳健性。
5. **\*\*实践测试与评估\*\***：通过实际应用，测试和评估改进后的模型和数据集在税务场景中的表现。

以上是本次的论文播报，祝大家今天工作顺利！