
基于 LoRA 和 Prompt 优化的大模型数学推理任务研究

李佳磊*

软件学院

复旦大学

22302010024@m.fudan.edu.cn

刘育麟†

软件学院

复旦大学

22302010046@m.fudan.edu.cn

赵亮哲‡

软件学院

复旦大学

22302010032@m.fudan.edu.cn

Abstract

本文针对大语言模型 (LLM) 在数学推理任务中的性能提升问题, 使用了一种基于低秩适配 (LoRA) 和 Prompt 优化的微调方法。预训练模型使用 Qwen/Qwen2.5-0.5B, 在 GSM8K 和 MATH 数据集的基础上, 通过数据增强与扩充优化训练样本, 同时采用 LoRA、混合精度训练和梯度累计技术有效减少显存消耗。在模型微调完成后, 通过设计针对性 Prompt 进一步提升模型对复杂数学问题的回答能力。实验结果表明, 与基线模型相比, 我们的方法在数学推理任务上取得了显著的性能提升, 验证了所提方法的有效性。

1 Introduction

近年来, 大语言模型在自然语言处理任务中表现出强大的能力。然而, 在数学推理等高复杂度任务中, 现有模型的性能仍有较大提升空间。为此, 本文聚焦于在资源受限的环境下, 通过 LoRA 技术优化微调过程, 并结合数据增强和 Prompt 设计, 提升模型的数学推理能力。我们以 GSM8K 和 MATH 数据集为实验基础, 探索低显存条件下的有效训练策略, 并验证了优化方法的可行性。

本文的贡献主要有以下几点:

*与赵亮哲和刘育麟同等贡献

†与李佳磊和赵亮哲同等贡献

‡与李佳磊和刘育麟同等贡献

- 首先，数据集选择与处理：使用了 GSM8K 和 MATH 数据集，涵盖了广泛的数学问题，确保了实验的代表性和挑战性。并且进行了数据增强与数据扩充的工作。
- 其次，模型微调减少显存消耗的应用：将 LoRA 方法应用于数学推理任务的微调，并且使用混合精度以及梯度累计进一步减少显存消耗，提高了模型应用于实践的可能性与实用性。
- 最后，Prompt 优化：通过设计优秀的适配于数学问题解答的 Prompt，进一步增强模型的数学能力，同时也为算力不足支持进行微调的实验提高新思路。

本研究不仅为大语言模型的高效微调提供了新思路，还为数学推理任务中的应用研究提供了有益的经验和方法，具有一定的学术和实践意义。

2 Approach

本研究采用了低秩适配 (LoRA) 技术对大语言模型 (LLM) 进行微调，以提高其在数学推理任务中的表现。具体而言，我们使用了 Qwen2.5-0.5B 模型，选择了 GSM8K 和 MATH 数据集进行实验。方法部分主要包括以下几个方面：数据集处理、模型微调的配置、Prompt 优化设计和评估方法。

本文采用的主要策略如下：

2.1 数据集处理

- 数据收集与预处理：从 GSM8K 和 MATH 数据集中提取训练样本，清洗数据并统一格式。
- 数据增强与扩充：利用大语言模型生成同义词替换和问题改写样本，增加数据多样性。

2.2 模型微调方法

- LoRA 微调：通过低秩适配技术减少显存需求，保留模型的原始权重并对新增参数进行训练。
- 混合精度训练：使用 FP16 与 FP32 混合精度训练以降低显存占用。
- 梯度累计：在显存有限的情况下，通过分批次累积梯度实现大批量训练。

2.3 Prompt 优化

- 针对数学推理任务设计特定 Prompt 模板，以引导模型生成更精确的答案。
- 对比不同 Prompt 的效果，选取最佳模板用于最终评估。

3 Experiments

本实验旨在评估 LoRA 微调方法对 Qwen2.5-0.5B 模型在数学推理任务中的表现。实验分为多个部分：实验细节、实验过程、数据集与预处理、Prompt 设置与优化、实验结果和性能分析。

3.1 实验细节

- 硬件设置：本实验在 Google Colab 环境下进行，使用了 40GB 显存的 GPU (A100 GPU) 进行训练。为了避免显存溢出，LoRA 微调提供了较为高效的训练方案，允许我们在有限资源下进行大规模实验。
- 超参数设置：在 LoRA 微调中，学习率被设置为 $5e-5$ ，批量大小为 22，训练周期为 20 个 epochs。具体超参数根据实验过程中的性能调优进行过多次调整。
- 训练时间：按照上述超参数进行训练，20 个 epoch 大概运行 27min。

3.2 实验过程

- 基线实验：采用 LoRA、混合精度训练和梯度累计技术进行优化对 Qwen2.5-0.5B 模型进行微调，并记录其在测试集上的表现。
- 优化实验：结合数据增强和 Prompt 设计进行性能提升。

3.3 数据集与预处理

本研究使用了两个数学推理数据集：GSM8K 和 MATH。

- GSM8K：该数据集包含 8500 个数学问题，主要涵盖基础算术问题，解决问题需要 2 至 8 步的计算。问题包括加法、减法、乘法、除法等基本运算。
- MATH：该数据集包括 12500 个数学问题，来源于高中数学竞赛，涵盖代数、几何、数论等多个领域。每个问题都带有完整的解题步骤，并标注了问题的难度等级（1 至 5）。

在实验中，我们将 GSM8K 数据集的 7500 个样本用于训练，1000 个样本用于测试；MATH 数据集的训练样本和测试样本分别为 10000 个和 2500 个。为了简化测试工作，对于 MATH 数据集我们采用 MATH500 作为验证集来进行测试。

数据预处理步骤包括：

- 将每个问题与答案拼接，并用 tokenizer 进行编码。
- 筛选有效的训练和测试样本，去除无效或格式错误的样本。

数据增强与数据扩充包括：

- 数据增强：使用大语言模型 GPT4o 对原有输入进行同义词替换与问题改写。调换语言描述的顺序。
- 数据扩充：使用 GPT4o 随机生成一些易于回答的数学题样本参与训练。

3.4 Prompt 设置与优化

本次实验应用的 Prompt 为

```
1 PROMPT_DICT = {  
2     "prompt_input": (  
3         f"{DEFAULT_BOS_TOKEN}system\n"  
4         "You are a logical reasoning expert.\n"  
5         "Analyze the following problem and provide a detailed solution.\n"  
6         f"{DEFAULT_EOS_TOKEN}\n"  
7         f"{DEFAULT_BOS_TOKEN}user\n"  
8         "Problem: {problem}\n"  
9         f"{DEFAULT_EOS_TOKEN}\n"  
10        f"{DEFAULT_BOS_TOKEN}assistant\n"  
11        "Here is the step-by-step solution:\n"  
12    ),  
13    "prompt_output": (  
14        "{answer}" + f"{DEFAULT_EOS_TOKEN}"  
15    )  
16 }  
17  
18
```

3.5 训练过程

训练过程分为以下几个步骤：

- 数据加载：首先加载训练集和测试集，并将它们预处理为适合模型输入的格式。训练集通过数据增强方法（如输入输出变换、同义词替换等）进行扩充。
- 模型初始化：加载 Qwen2.5-0.5B 模型并应用 LoRA 微调配置。模型在初始状态下冻结绝大部分参数，LoRA 微调仅更新低秩矩阵的参数。
- 训练与验证：在训练过程中，模型每 epoch 进行一次验证，评估模型在验证集上的性能。训练完成之后保存模型。
- 评估指标：主要评估模型在数学推理任务中的准确率。

3.6 实验结果

我们在 GSM8K 和 MATH 数据集上进行了实验，以下是基线模型和 LoRA 微调模型的性能对比。

3.6.1 基线模型结果

在 GSM8K 数据集上的基线模型表现如下：

测试准确率：31.28%

在 MATH 数据集上的基线模型表现如下：

测试准确率：8.71%

3.6.2 Prompt 优化模型结果

在 GSM8K 数据集上应用 LoRA 微调后，模型的表现如下：

测试准确率：35.37%

在 MATH 数据集上应用 LoRA 微调后，模型的表现如下：

测试准确率：10.00%

3.7 性能分析

从实验结果来看，设置 Prompt 并且优化后的模型在数学推理任务上相比基线方法表现出了较为明显的提升。特别是在 GSM8K 数据集上，Prompt 模型测试准确率提升%，说明 Prompt 设计与优化可以获得较高的推理准确性。

在 MATH 数据集上，设置 Prompt 并且优化后的模型虽然在准确率上相较基线模型有所提升，但提升幅度较小，1.69%。这可能与 MATH 数据集的问题难度和模型的推理能力有关，未来的研究可以尝试结合更多的训练策略（如强化学习）来进一步提升模型在高难度问题上的表现。

4 Analysis and Discussion

本研究通过在数学推理任务中使用 LoRA 微调技术和 Prompt 优化方法，验证了在有限计算资源条件下提升大语言模型性能的可行性。实验结果显示，LoRA 微调在 GSM8K 和 MATH 数据集上均取得了较好的表现，Prompt 优化进一步提升了模型的推理能力。然而，针对不同数据集和问题难度，模型的表现仍存在差异。

4.1 LoRA 微调的有效性

LoRA 技术显著降低了显存需求，同时保持了模型的推理能力。实验表明，在 GSM8K 数据集上，LoRA 微调后的模型测试准确率相对较高。这一结果验证了低秩适配方法在资源受限条件下的高效性。然而，在 MATH 数据集上的准确率较小，可能是由于

MATH 数据集问题的复杂性更高，模型在处理复杂推理任务时对显存和训练样本多样性有更高的需求。

4.2 Prompt 优化的作用

设计针对数学推理任务的 Prompt 模板，为模型提供了清晰的解题步骤引导，从而有效提升了推理准确性。特别是在 GSM8K 数据集上，Prompt 优化后的模型表现出更高的鲁棒性。尽管如此，Prompt 的改进对 MATH 数据集的提升效果有限，这可能是因为该数据集包含更复杂的问题类型和更高的推理难度，Prompt 优化无法完全弥补训练数据或模型能力的不足。

4.3 数据集特性对结果的影响

GSM8K 和 MATH 数据集在问题类型和难度分布上存在显著差异。GSM8K 主要包含基础算术问题，而 MATH 则涵盖了代数、几何、数论等多个领域。实验中，LoRA 微调和 Prompt 优化对 GSM8K 的提升更为明显，反映出当前方法对基础问题的适配性较好。然而，对于 MATH 数据集中难度等级较高的问题，模型表现仍有较大提升空间。这表明未来研究需要更关注复杂推理任务的优化策略。

4.4 研究局限性

本研究的主要局限性包括：

- 数据增强的质量与覆盖面：尽管使用了数据增强方法生成了更多样化的训练样本，但其覆盖范围可能不足以满足 MATH 数据集的复杂性需求。
- 计算资源限制：尽管 LoRA 技术降低了显存需求，但进一步提升模型性能可能需要更大规模的计算资源和更长的训练时间。
- 方法的通用性：本研究方法在数学推理任务中表现较好，但其在其他高复杂度任务中的通用性仍需进一步验证。

5 Conclusion

本文研究了基于 LoRA 和 Prompt 优化的大语言模型在数学推理任务中的应用。通过对 Qwen2.5-0.5B 模型进行微调，并结合 GSM8K 和 MATH 数据集的实验，验证了所提方法在资源受限条件下提升模型性能的有效性。

研究的主要贡献包括：

- LoRA 微调的成功应用：在显存受限的条件下，通过低秩适配技术实现了模型的高效训练。
- Prompt 优化的有效设计：针对数学推理任务设计了特定的 Prompt 模板，进一步提升了模型的推理能力。

- 数据增强与扩充：通过数据预处理和增强策略，提升了训练数据的多样性和质量。
- 实验结果表明，LoRA 微调和 Prompt 优化方法在 GSM8K 和 MATH 数据集上的测试准确率均有所提升，其中 GSM8K 的提升幅度更为显著。尽管如此，对于高难度数学问题的处理仍有较大优化空间。

未来的研究方向包括：

- 引入强化学习：通过强化学习方法提升模型在高难度数学问题上的推理能力。
- 改进数据增强策略：利用更多样化的数据生成技术扩展训练样本的覆盖范围。
- 探索跨任务迁移学习：验证 LoRA 和 Prompt 优化方法在其他复杂推理任务中的通用性。
- 总之，本研究为大语言模型在数学推理任务中的优化提供了新的思路和方法，为低资源环境下的模型训练实践提供了有益的参考。

References

- [1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017) Attention is all you need. In I. Guyon, et al. (eds.), *Advances in Neural Information Processing Systems 30*, pp. 5998-6008. Curran Associates, Inc.
- [2] Qin, G., Zhang, Y., Liu, Y., & Zhang, L. (2023) Qwen-7B: A general-purpose pre-trained language model. *arXiv preprint arXiv:2306.00360*.
- [3] Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., ... & Howard, W. (2021) LoRA: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*.
- [4] Kingma, D. P., & Ba, J. (2014) Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- [5] Hendrycks, D., & Gimpel, K. (2021) MATH: A dataset for evaluating mathematical problem solving. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pp. 4915-4931.
- [6] Cobbe, K., et al. (2021) GSM8K: A dataset for evaluating mathematical problem solving in language models. *arXiv preprint arXiv:2110.14168*.
- [7] Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., ... & Rush, A. M. (2019) Hugging Face's Transformers: State-of-the-art natural language processing. *arXiv preprint arXiv:1910.03771*.
- [8] Lhoest, Q., et al. (2021) Datasets: A community library for natural language processing. *arXiv preprint arXiv:2106.06690*.
- [9] Lhoest, Q., et al. (2022) Evaluate: A library for evaluating machine learning models. *arXiv preprint arXiv:2209.00288*.

- [10] Liu, P. J., et al. (2022) Parameter-efficient fine-tuning for natural language processing. *arXiv preprint* arXiv:2210.11416.
- [11] McKinney, W. (2012) *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. O'Reilly Media, Inc.
- [12] Van Rossum, G., & Drake, F. L. (2009) *Python 3 Reference Manual*. CreateSpace.

A Appendix / supplemental material

小组三位成员的具体分工如下:

李佳磊: 数据集收集、预处理以及数据增强、扩充, 进行论文写作与完成

刘育麟: 提出以及实现减少显存消耗, 使用 lora、混合精度训练、精度累计等

赵亮哲: 提供算力支持, 进行离线 lora 训练

需要额外强调的是, 所有小组成员在本次实践中的贡献程度相同, 并无多少高低之分。