Instruction Fine-Tuning Effects on Language Model Mathematical Problem-Solving Accuracy

Assignee Research

June 6, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: What is the effect of instruction fine-tuning on language model mathematical problem-solving accuracy v15. 10 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Improving Large Language Model Fine-tuning for Solving Math Problems. Research question: What is the effect of instruction fine-tuning on language model mathematical problem-solving accuracy v15.

2 Methodology

Systematic literature search across multiple databases yielded 16 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.2/10.

3 Results

16 papers retrieved. 10 claims extracted; 0 independently verified. Quality review score: 4.2/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
The MATH dataset is used for experiments, with 4.5K original test examples for training and validation, and 500 test exa	×	0.05
Two sources of correct step-by-step solutions are used for model training: original human-written explanations in the MA	×	0.13
The automatic grading script provided by Lightman et al. (2023) checks the mathematical equivalence of generated solutio	×	0.02
Greedy decoding is used for Pass@1 performance evaluation, and nucleus sampling is used for majority voting performance	×	0.11
The sampling temperature is set to 0.6, and the top-p value is set to 0.95 for nucleus sampling.	×	0.03
Fine-tuning PaLM 2-S* and PaLM 2-L on step-by-step solutions with the MLE training objective improves performance compar	×	0.13
Models fine-tuned on PRM800K solutions achieve significantly better performance than those fine-tuned on original MATH s	×	0.09
The original solutions in the MATH dataset are more abstract, while the solutions generated by GPT-4 are more fine-grain	×	0.08
Solution-cluster re-ranking is evaluated using two loss functions: Lcls-margin (Eq. 8) and Lcls-xent (Eq. 9).	×	0.07
Two re-ranking strategies are used: re-ranking all candidate solutions (RR.All) and re-ranking all solutions in the top-	×	0.12

References

- `http://arxiv.org/abs/2312.10793v3`
- `http://arxiv.org/abs/2509.25160v1`

- http://arxiv.org/abs/2310.10047v1