Multi-Objective Optimization Effects on Code Generation Accuracy in HumanEval-JavaScript

Assignee Research

May 31, 2026

Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: What is the impact of multi-objective optimization on code generation accuracy in HumanEval-JavaScript relative to standard PPO when training with diverse reward signals. The evolution of Large Language Models (LLMs) like ChatGPT and GPT-4 has sparked discussions on the advent of Artificial General Intelligence (AGI). However, replicating such advancements in open-source models has been challenging. 12 claims were extracted from source literature; 8 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 7.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: InternLM2 Technical Report. Research question: What is the impact of multi-objective optimization on code generation accuracy in HumanEval-JavaScript relative to standard PPO when training with diverse reward signals?.

2 Methodology

Systematic literature search across multiple databases yielded 11 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 7.3/10.

3 Results

11 papers retrieved. 12 claims extracted; 8 independently verified. Quality review score: 7.3/10.

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
InternLM2 is an open-source Large Language Model.	✓	0.17
InternLM2 outperforms its predecessors in comprehensive evaluations across 6 dimensions.	✓	0.18
InternLM2 outperforms its predecessors across 30 benchmarks.	×	0.13
InternLM2 was initially trained on a context length of 4k tokens.	×	0.13
InternLM2 advanced to a context length of 32k tokens during pre-training and fine-tuning stages.	✓	0.23
InternLM2 exhibited performance on the 200k 'Needle-in-a-Haystack' test.	✓	0.17
InternLM2 is aligned using Supervised Fine-Tuning (SFT).	✓	0.23
InternLM2 utilizes a novel Conditional Online Reinforcement Learning from Human Feedback (COOL RLHF) strategy.	✓	0.23
The COOL RLHF strategy addresses conflicting human preferences.	✓	0.21
The COOL RLHF strategy addresses reward hacking.	✓	0.17
InternLM2 models are released in different training stages.	×	0.15
InternLM2 models are released in different model sizes.	×	0.11

References

- `https://doi.org/10.48550/arxiv.2412.15115`
- `https://doi.org/10.4230/oasics.icpec.2025.4`

- https://doi.org/10.48550/arxiv.2403.17297