

Synthetic Problem Quality Effects on Reinforcement Learning for Code Generation

Assignee Research

June 4, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: What is the impact of synthetic problem quality on the inference efficiency and convergence speed of reinforcement learning for code generation tasks. 5 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Explainable Machine Learning for Public Policy: Use Cases, Gaps, and Research Directions. Research question: What is the impact of synthetic problem quality on the inference efficiency and convergence speed of reinforcement learning for code generation tasks?.

2 Methodology

Systematic literature search across multiple databases yielded 15 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.8/10.

3 Results

15 papers retrieved. 5 claims extracted; 0 independently verified. Quality review score: 3.8/10.

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 |
|--------------------------------------------------------------------------------------------------------------------------|----------|------------|
| In public policy settings, selecting models that work well for precision in the top k% of predicted scores is preferred | × | 0.07 |
| Optimizing for accuracy or AUC-ROC is sub-optimal for public policy applications concerned with precision in the top k% | × | 0.06 |
| Developing a feature set for individuals in policy applications typically entails combining several heterogeneous data s | × | 0.02 |
| Combining heterogeneous data sources in policy applications often introduces complex correlation structures to the featu | × | 0.06 |
| The ML model in Bauman et al. [2018] combines data sources with strong spatiotemporal patterns. | × | 0.03 |

References

- <http://arxiv.org/abs/2605.22893v1>
- <http://arxiv.org/abs/2507.02910v1>
- <http://arxiv.org/abs/2010.14374v3>