

Potential-Based vs. State-Based Rewards in Balancing Response Length and Accuracy for 7B and 70B Models

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 14 peer-reviewed papers addressing the following research question: What is the comparative effect of potential-based and state-based reward functions on the trade-off between response length and accuracy scores for 7B and 70B models evaluated on the CommonsenseQA. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 6.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: SLM-Bench: A Comprehensive Benchmark of Small Language Models on Environmental Impacts–Extended Version. Research question: What is the comparative effect of potential-based and state-based reward functions on the trade-off between response length and accuracy scores for 7B and 70B models evaluated on the CommonsenseQA dataset?.

2 Methodology

Systematic literature search across multiple databases yielded 14 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 6.5/10.

3 Results

14 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 6.5/10.

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.

References

- <http://arxiv.org/abs/2410.13187v3>
- <http://arxiv.org/abs/2508.15478v2>
- <http://arxiv.org/abs/2601.08844v1>