What is the robustness of quantized SLMs (1B-10B parameters) across SLM-Bench domains when evaluated under fed

Assignee Research

June 10, 2026

Abstract

Small Language Models (SLMs) offer computational efficiency and accessibility, yet a systematic evaluation of their performance and environmental impact remains lacking. We introduce SLM-Bench, the first benchmark specifically designed to assess SLMs across multiple dimensions, including accuracy, computational efficiency, and sustainability metrics. SLM-Bench evaluates 15 SLMs on 9 NLP tasks using 23 datasets spanning 14 domains. The evaluation is conducted on 4 hardware configurations, providing a rigorous comparison of their effectiveness. Unlike prior benchmarks, SLM-Bench quantifies 11 me

1 Introduction

This paper examines: SLM-Bench: A Comprehensive Benchmark of Small Language Models on Environmental Impacts–Extended Version. Research question: What is the robustness of quantized SLMs (1B-10B parameters) across SLM-Bench domains when evaluated under federated learning setups with edge Kubernetes distributions?.

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 3.3/10.

3 Results

11 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 3.3/10.

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.

References

- http://arxiv.org/abs/2508.15478v2
- http://arxiv.org/abs/2505.08588v1
- http://arxiv.org/abs/2302.13473v2