Federated vs. Centralized Large Language Models: Robustness to Label Flipping on GSM8K

Assignee Research

May 31, 2026

Abstract

This report synthesises findings from 14 peer-reviewed papers addressing the following research question: How does the robustness of federated large language models against label flipping attacks compare to centralized training when evaluated on reasoning benchmarks like GSM8K. Data poisoning and leakage risks impede the massive deployment of federated learning in the real world. This chapter reveals the truths and pitfalls of understanding two dominating threats: {\em training data privacy intrusion} and {\em training data poisoning}. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Data Poisoning and Leakage Analysis in Federated Learning. Research question: How does the robustness of federated large language models against label flipping attacks compare to centralized training when evaluated on reasoning benchmarks like GSM8K?.

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 3.3/10.

3 Results

14 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/2312.17080v4
- http://arxiv.org/abs/2305.02022v2
- http://arxiv.org/abs/2409.13004v1