CausalMixFT Scalability vs Traditional Fine-Tuning in Tabular Transformers

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: How does the scalability of CausalMixFT compare to traditional fine-tuning methods (e.g., AdamW, LAMB) in terms of convergence speed and downstream performance when fine-tuning large TFMs (e.g.,. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 7.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Differentially Private Fine-tuning of Language Models. Research question: How does the scalability of CausalMixFT compare to traditional fine-tuning methods (e.g., AdamW, LAMB) in terms of convergence speed and downstream performance when fine-tuning large TFMs (e.g., TabTransformer, DeepFM) on small to medium-sized tabular datasets (e.g., Adult Census, Yelp) using fixed computational budgets (e.g., training steps)?.

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 7.2/10.

3 Results

11 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 7.2/10.

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/2601.04110v2
- http://arxiv.org/abs/2606.01947v1
- http://arxiv.org/abs/2110.06500v2