

# CausalMixFT Accuracy on GLUE and SuperGLUE Under Differential Privacy Constraints

Assignee Research

June 9, 2026

## Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: How does the accuracy of CausalMixFT on GLUE and SuperGLUE benchmarks compare to DP-AdamW and DP-LAMB across epsilon values ranging from 0.5 to 8.0. 9 claims were extracted from source literature; 6 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: DP-AdamW: Investigating Decoupled Weight Decay and Bias Correction in Private Deep Learning. Research question: How does the accuracy of CausalMixFT on GLUE and SuperGLUE benchmarks compare to DP-AdamW and DP-LAMB across epsilon values ranging from 0.5 to 8.0?.

## 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 7.2/10.

## 3 Results

15 papers retrieved. 9 claims extracted; 6 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.

## 5 Extracted Claims

Claim	Verified	Confidence
DP-AdamBC achieved an improvement over DP-Adam across several different target privacy budgets on image, text and graph	✓	0.19
DP-AdamW and DP-AdamW-BC are proposed to improve the performance of DP-Adam and DP-AdamBC via adding weight decay.	✓	0.21
DP-AdamW and DP-AdamW-BC are evaluated using the same suite of tasks as in Tang et al., 2023.	×	0.10
DP-AdamW and DP-AdamW-BC carry the same privacy guarantees as DP-SGD and DP-AdamBC.	✓	0.24
DP-AdamW outperforms DP-Adam in all privacy budgets for CIFAR-10 image classification.	✓	0.18
DP-AdamW-BC outperforms DP-AdamBC across privacy budgets except for the large $\epsilon = 7$ .	✓	0.18
The most significant performance increase is in the $\epsilon = 3$ setting, where both DP-AdamW and DP-AdamW-BC outperform DP-Ada	✓	0.20
DP-AdamW achieves state-of-the-art accuracy among the evaluated DP-SGD and DP-Adam variants for CIFAR-10 image classific	×	0.14
DP-AdamW consistently outperformed DP-AdamW-BC across all tested privacy budgets, achieving statistically significantly	×	0.13

## References

- <http://arxiv.org/abs/2511.07843v1>

- <http://arxiv.org/abs/2204.07288v1>
- <http://arxiv.org/abs/2411.16121v3>