CausalMixFT and Traditional Augmentation Methods in Tabular Foundation Model Generalization

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: What is the comparative impact of CausalMixFT versus traditional data augmentation methods (e.g., SMOTE, MixUp) on the generalization gap in fine-tuning tabular foundation models across low-resource. 12 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Causal Data Augmentation for Robust Fine-Tuning of Tabular Foundation Models. Research question: What is the comparative impact of CausalMixFT versus traditional data augmentation methods (e.g., SMOTE, MixUp) on the generalization gap in fine-tuning tabular foundation models across low-resource and high-resource settings, evaluated on datasets like Adult or Criteo?.

2 Methodology

Systematic literature search across multiple databases yielded 12 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.8/10.

3 Results

12 papers retrieved. 12 claims extracted; 0 independently verified. Quality review score: 3.8/10.

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
CausalMixFT achieves the highest median improvement of (+0.12$\pm$0.63) over the pre-trained model on 33 classification data	×	0.09
Default fine-tuning has a variability of $\pm$0.98, while CausalMixFT has a variability of $\pm$0.63.	×	0.07
CausalMixFT ranks first overall in average ranks across datasets, followed by the default fine-tuning baseline.	×	0.06
Purely synthetic generators, including CTGAN, SCM, TabEBM, TableAugment, and MixedModel, show negative median improvemen	×	0.08
The experiments involve 2,310 fine-tuning runs across 33 classification datasets with 10 folds each.	×	0.13
Model performance is reported as normalized ROC-AUC relative to the pre-trained model.	×	0.07
CausalMixFT combines real and causally generated samples for fine-tuning.	×	0.11
SCMs explicitly encode causal dependencies among features through a directed acyclic graph (DAG) and a set of structural	×	0.05
The PC and FCI algorithms are used to estimate the structural relations between the features.	×	0.03
DoWhy's SCM framework with additive noise models is used to sample and fit DAGs.	×	0.02
Numerical features are modeled with regressors, and categorical features with classifiers in the SCM framework.	×	0.04
Synthetic samples are generated by sampling exogenous noise and propagating it through the fitted SCM.	×	0.05

References

- http://arxiv.org/abs/2601.04110v2
- http://arxiv.org/abs/2408.13622v1
- http://arxiv.org/abs/2512.03307v1