CausalMixFT Synthetic Sample Ratios and Robustness in Tabular Foundation Models

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: How does the synthetic sample ratio in CausalMixFT impact the robustness of fine-tuned tabular foundation models (TFMs) against distribution shifts, as measured by accuracy on out-of-distribution. 14 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.5/10. 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: How does the synthetic sample ratio in CausalMixFT impact the robustness of fine-tuned tabular foundation models (TFMs) against distribution shifts, as measured by accuracy on out-of-distribution benchmarks like TabShift or TabOD?.

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 4.5/10.

3 Results

11 papers retrieved. 14 claims extracted; 1 independently verified. Quality review score: 4.5/10.

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 median improvement of (+0.10 $\pm$0.98) over the pre-trained model on the same datasets.	×	0.09
Purely synthetic augmentation methods (CT-GAN, SCM, TabEBM, TableAugment, Mixed-Model) show negative median improvements o	×	0.08
CausalMixFT has a variability of $\pm$0.63, which is lower than the default fine-tuning variability of $\pm$0.98.	×	0.06
CausalMixFT ranks first overall in average ranks across datasets, followed by the default fine-tuning baseline.	×	0.06
Early stopping based on limited validation data leads to significant validation set overfitting depending on the fine-tu	✓	0.15
The normalization strategy used to compare performance across different data generators is based on Gorishniy et al. [12	×	0.05
CausalMixFT extends the fine-tuning framework of Bhler et al. [5] by mixing real and causally grounded synthetic sample	×	0.10
SCM-Based Synthetic Augmentation (CausalMixFT) uses SCMs fitted to the target dataset to generate synthetic data.	×	0.13
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 in CausalMixFT.	×	0.03
DoWhy's SCM framework with additive noise models is used to sample and fit DAGs in CausalMixFT.	×	0.02
Numerical features in CausalMixFT are modeled with regressors, and categorical features with classifiers.	×	0.04
Synthetic samples in CausalMixFT 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/2507.07829v1
- http://arxiv.org/abs/2512.03307v1