Impact of Structural Causal Model Backbone Choice on Few-Shot Accuracy in CausalMixFT for Tabular Benchmarks

Assignee Research

June 11, 2026

Abstract

Fine-tuning tabular foundation models (TFMs) under data scarcity is challenging, as early stopping on even scarcer validation data often fails to capture true generalization performance. We propose CausalMixFT, a method that enhances fine-tuning robustness and downstream performance by generating structurally consistent synthetic samples using Structural Causal Models (SCMs) fitted on the target dataset. This approach augments limited real data with causally informed synthetic examples, preserving feature dependencies while expanding training diversity. Evaluated across 33 classification datas

1 Introduction

This paper examines: Causal Data Augmentation for Robust Fine-Tuning of Tabular Foundation Models. Research question: To what extent does the choice of Structural Causal Model (SCM) backbone (e.g., linear vs. nonlinear) in CausalMixFT affect few-shot accuracy on tabular benchmarks like TabBench or TabMIX?.

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 8.2/10.

3 Results

12 papers retrieved. 9 claims extracted; 8 independently verified. Quality review score: 8.2/10.

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
CausalMixFT achieves the highest median improvement of (+0.12$\pm$0.63) over the pretrained model on 33 classification data	✓	0.26
Default fine-tuning has a variability of $\pm$0.98, while CausalMixFT has a variability of $\pm$0.63, indicating greater instabi	✓	0.18
CausalMixFT ranks first overall in average ranks across datasets, followed by the default fine-tuning baseline, while pu	✓	0.29
The normalization strategy used to compare performance across different data generators is based on the zero-shot perfor	×	0.06
CausalMixFT extends the fine-tuning framework by mixing real and causally grounded synthetic samples into the fine-tunin	✓	0.23
SCM-Based Synthetic Augmentation (CausalMixFT) uses SCMs fitted to the target dataset to generate synthetic data that re	✓	0.20
The PC and FCI algorithms are used to estimate the structural relations between features in CausalMixFT.	✓	0.17
DoWhy's SCM framework with additive noise models is used to sample and fit DAGs in CausalMixFT.	✓	0.16
Numerical features in CausalMixFT are modeled with regressors, and categorical features with classifiers.	✓	0.15

References

- http://arxiv.org/abs/2311.14544v1

- http://arxiv.org/abs/2601.04110v2
- http://arxiv.org/abs/2008.09301v1