

Causal Structure Complexity in SCMs and Generalization Performance of Fine-Tuned Transformers on Tabular Data

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: How does the causal structure complexity of the SCMs in CausalMixFT impact the generalization performance of fine-tuned TFMs across different tabular domains, measured by validation loss and. 12 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.0/10. 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 causal structure complexity of the SCMs in CausalMixFT impact the generalization performance of fine-tuned TFMs across different tabular domains, measured by validation loss and convergence speed?.

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 4.0/10.

3 Results

11 papers retrieved. 12 claims extracted; 0 independently verified. Quality review score: 4.0/10.

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 ± 0.98 , while CausalMixFT has a variability of ± 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 Mixed-Model, show negative median improvement	×	0.09
The experiments involved 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.	×	0.08
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	×	0.04
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.	×	0.03
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/2508.11281v3>
- <http://arxiv.org/abs/2504.20900v1>