

CausalMixFT Structural Causal Models Enhance Robustness in Tabular Foundation Models

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: What is the impact of CausalMixFT's structural causal model generation on the robustness of tabular foundation models against distribution shifts compared to standard mixup techniques. 10 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.7/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Rethinking Distribution Shifts: Empirical Analysis and Modeling for Tabular Data. Research question: What is the impact of CausalMixFT's structural causal model generation on the robustness of tabular foundation models against distribution shifts compared to standard mixup techniques?.

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 3.7/10.

3 Results

12 papers retrieved. 10 claims extracted; 0 independently verified. Quality review score: 3.7/10.

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
Existing tabular benchmarks for distribution shifts (e.g., Adult, BRFSS, COMPAS, ACS Public Coverage, ACS Income) tend to	×	0.07
In the Adult dataset, treating white men as the source distribution and non-white women as the target distribution result	×	0.03
The small relative regret observed in widely-used benchmarks (Adult, BRFSS, COMPAS, ACS Public Coverage, ACS Income) indicates	×	0.02
The proposed benchmark incorporates distribution shifts in tabular datasets covering socioeconomic and physical systems	×	0.09
Figure 1 displays relative regret for seven settings designed in the new benchmark, including shifts such as ACS Income	×	0.03
In Table (p16), the mean $Y X$ -ratio for the 0-10% bin quantile is 0.793 with a standard deviation of 0.092.	×	0.02
In Table (p16), the mean $Y X$ -ratio for the 40-50% bin quantile is 0.610 with a standard deviation of 0.304.	×	0.01
DRO (Distributionally Robust Optimization) performance is closely tied to the performance of its base model.	×	0.11
The choice of base model class has a greater impact on DRO performance than the specific choice of ambiguity sets, with	×	0.10
Figure 5 reports overall performances on different source-target shift pairs in Setting 5, specifically including shifts	×	0.03

References

- <http://arxiv.org/abs/2601.04110v2>
- <http://arxiv.org/abs/2512.03307v1>
- <http://arxiv.org/abs/2307.05284v6>