Robustness of Tabular Foundation Models to Distribution Shifts via Synthetic Pretraining Scale

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: How does the scale of synthetic pretraining data correlate with the robustness of tabular foundation models against distribution shifts as measured by TabBench metrics. 13 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.1/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Robust Tabular Foundation Models. Research question: How does the scale of synthetic pretraining data correlate with the robustness of tabular foundation models against distribution shifts as measured by TabBench metrics?.

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 3.1/10.

3 Results

12 papers retrieved. 13 claims extracted; 0 independently verified. Quality review score: 3.1/10.

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
Tabular foundation models (TFMs) rely on in-context learning (ICL) for classification and regression tasks with structur	×	0.11
TFMs can produce high-quality predictions on new datasets in milliseconds when GPU-accelerated.	×	0.08
Training TFMs relies on generating diverse synthetic datasets constructed from structural causal models (SCMs).	×	0.13
All current publicly available, competitive TFMs have been pretrained on datasets generated from a fixed prior distribut	×	0.06
Fixed priors in TFM training underrepresent certain regions of the parameter space, potentially degrading performance on	×	0.05
State-of-the-art TFMs lag behind tree-based methods on some benchmarks.	×	0.05
The proposed RTFM algorithm applied to TabPFN V2 used only 90k additional training datasets.	×	0.11
Applying RTFM to TabPFN V2 significantly improved its ranking on several real-world tabular benchmarks.	×	0.11
In the maximization stage of the proposed method, the model gW is frozen to maximize the optimality gap.	×	0.04
The study used a black-box optimization algorithm to search the SCM parameter space for parameters with large optimality	×	0.02
The optimality gap estimation in the experiments used nds=20 generators and e=7 baseline estimators.	×	0.06
With nds=20, e=7, and sufficient CPU cores (ncores = nds * e), the estimated optimality gap could be computed in a matte	×	0.03
Table (p10) lists synthetic dataset generation parameters including feature counts ranging from 3 to 128 and activation	×	0.03

References

- http://arxiv.org/abs/2601.04110v2
- http://arxiv.org/abs/2206.02435v2
- http://arxiv.org/abs/2512.03307v1