

Synthetic Data Diversity and Calibration Error in Tabular Foundation Models Under Distributional Drift

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: Does increasing the diversity parameter in synthetic data generators improve the calibration error of tabular foundation models when evaluated on domains with distributional drift. 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: Robust Tabular Foundation Models. Research question: Does increasing the diversity parameter in synthetic data generators improve the calibration error of tabular foundation models when evaluated on domains with distributional drift?.

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
Tabular foundation models (TFMs) rely on in-context learning (ICL) for classification and regression tasks with structured	×	0.12
TFMs can produce high-quality predictions on new datasets in milliseconds when GPU-accelerated.	×	0.07
Training TFMs relies on generating large amounts of diverse synthetic datasets constructed from structural causal models	×	0.06
All current publicly available, competitive TFMs have been pretrained on datasets generated from a fixed prior distribution	×	0.06
Fixed priors in TFM training underrepresent certain regions of the parameter space, potentially degrading performance on	×	0.06
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.	×	0.10
Applying RTFM to TabPFN V2 significantly improved its ranking on several real-world tabular benchmarks.	×	0.12
In the described methodology, the estimated optimality gap could be computed in a matter of seconds when parallelized with	×	0.04
The methodology utilized ncores equal to the product of nodes (20) and elements (7) for parallelization.	×	0.00

References

- <http://arxiv.org/abs/2411.15497v3>

- <http://arxiv.org/abs/2504.20900v1>
- <http://arxiv.org/abs/2512.03307v1>