

Tabular Foundation Models: Synthetic vs. Real-World Pretraining on TabBench Under Domain Shift

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 8 peer-reviewed papers addressing the following research question: How does the generalization performance of tabular foundation models pretrained on synthetic data compare to those pretrained on real-world datasets when evaluated on TabBench under domain shift. 12 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.2/10. 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 generalization performance of tabular foundation models pretrained on synthetic data compare to those pretrained on real-world datasets when evaluated on TabBench under domain shift conditions?.

2 Methodology

Systematic literature search across multiple databases yielded 8 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.2/10.

3 Results

8 papers retrieved. 12 claims extracted; 0 independently verified. Quality review score: 3.2/10.

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
Tabular foundation models (TFMs) rely on in-context learning (ICL) for classification and regression tasks with structure	×	0.10
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).	×	0.12
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.	×	0.06
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 described methodology, the estimated optimality gap could be computed in a matter of seconds when parallelized with	×	0.04
The methodology utilized n_{cores} equal to the product of n_{ds} (number of generators) and e (number of baseline estimators)	×	0.02
The benchmark table includes synthetic dataset configurations with activation functions such as tanh, identity, elu, and	×	0.03
The benchmark table includes synthetic dataset configurations with distribution types including uniform, exponential, an	×	0.04

References

- <http://arxiv.org/abs/2601.04110v2>

- <http://arxiv.org/abs/2512.03307v1>
- <http://arxiv.org/abs/2510.12208v1>