

Adversarial Noise Injection in Pretraining Enhances Robustness of Tabular Foundation Models

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: What is the impact of synthetic adversarial noise injection during pretraining on the robustness of tabular foundation models evaluated on TabMNAR. 10 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.2/10. 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: What is the impact of synthetic adversarial noise injection during pretraining on the robustness of tabular foundation models evaluated on TabMNAR?.

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 3.2/10.

3 Results

12 papers retrieved. 10 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 structured data.	×	0.12
TFMs can produce high-quality predictions on new datasets in milliseconds when GPU-accelerated.	×	0.07
Training TFMs relies on generating diverse synthetic datasets constructed from structural causal models (SCMs).	×	0.07
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 some benchmarks.	×	0.05
State-of-the-art TFMs still lag behind tree-based methods on some benchmarks.	×	0.06
The proposed method, RTFM, was applied to TabPFN V2.	×	0.11
Training TabPFN V2 with RTFM using only 90k additional training datasets significantly improved its ranking on several benchmarks.	×	0.11
The optimality gap estimation step in the proposed algorithm can be computed in a matter of seconds when parallelized with 140 cores.	×	0.02
In the described implementation, the optimality gap estimation used $n_{ds}=20$, $e=7$, and $n_{cores}=140$.	×	0.03

References

- <http://arxiv.org/abs/2402.04325v1>

- <http://arxiv.org/abs/2408.07579v1>
- <http://arxiv.org/abs/2512.03307v1>