Scaling Laws of Tabular Foundation Models: Efficiency and Few-Shot Adaptation Trade-offs

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: How does the computational efficiency (measured in FLOPS or inference latency) of tabular foundation models scale with increasing sample diversity during training, and what trade-offs exist between. 15 claims were extracted from source literature; 4 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 6.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: TabDPT: Scaling Tabular Foundation Models on Real Data. Research question: How does the computational efficiency (measured in FLOPS or inference latency) of tabular foundation models scale with increasing sample diversity during training, and what trade-offs exist between efficiency and few-shot adaptation performance on benchmark datasets?.

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 6.2/10.

3 Results

12 papers retrieved. 15 claims extracted; 4 independently verified. Quality review score: 6.2/10.

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
In-context learning (ICL) enables a model to adapt to new tasks by modifying the context without per-dataset fine-tuning	×	0.10
Repurposing large language models (LLMs) for in-context learning on tabular data faces obstacles due to inefficient toke	×	0.11
LLM performance on tabular data varies based on prompt format and is sensitive to the order of given examples.	×	0.05
TabDPT utilizes in-context learning retrieval combined with self-supervised learning based on column masking for pre-tra	×	0.13
Pre-training TabDPT on real data leads to faster convergence compared to training exclusively on synthetic data.	✓	0.16
Pre-training TabDPT on real data yields improved downstream accuracy on unseen datasets compared to training exclusively	✓	0.16
TabDPT achieves leading accuracy on new, unseen datasets for both classification and regression tasks without further tr	×	0.11
Increasing model size or pre-training data size (number of cells) leads to consistent improvements in TabDPT performance	✓	0.16
On the 'cmc' dataset with 10 labeled examples per class, TabDPT (semi) achieved an accuracy of 44.24.	×	0.02
On the 'karhunen' dataset with 10 labeled examples per class, TabDPT (semi) achieved an accuracy of 92.08.	×	0.02
On the 'optdigit' dataset with 10 labeled examples per class, TabDPT (semi) achieved an accuracy of 94.31.	×	0.02
TabDPT (semi) achieved an average few-shot accuracy of 77.56 across seven CC18 datasets.	×	0.03
STUNT achieved an average few-shot accuracy of 76.76 across seven CC18 datasets.	×	0.03
CACTUs achieved an average few-shot accuracy of 75.16 across seven CC18 datasets.	×	0.03
This work provides the first analysis of scaling laws for Tabular Foundation Models (TFMs) that are not restricted to an	✓	0.17

References

- http://arxiv.org/abs/2106.11959v5
- http://arxiv.org/abs/2410.18164v3
- http://arxiv.org/abs/2306.11113v2