

TabPFN Performance in Few-Shot Tabular Reasoning with Non-Causal Foundation Models

Assignee Research

June 12, 2026

Abstract

Tabular data is a pervasive modality spanning a wide range of domains, and the inherent diversity poses a considerable challenge for deep learning. Recent advancements using transformer-based in-context learning have shown promise on smaller and less complex datasets, but have struggled to scale to larger and more complex ones. To address this limitation, we propose a combination of retrieval and fine-tuning: we can adapt the transformer to a local subset of the data by collecting nearest neighbours, and then perform task-specific fine-tuning with this retrieved set of neighbours in context. U

1 Introduction

This paper examines: Retrieval & Fine-Tuning for In-Context Tabular Models. Research question: How does the in-context learning accuracy of TabPFN compare to other foundation models pre-trained on non-causal synthetic datasets when evaluated on few-shot tabular reasoning tasks using metrics like F1-score or AUC?.

2 Methodology

Systematic literature search across multiple databases yielded 13 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 9.0/10.

3 Results

13 papers retrieved. 11 claims extracted; 11 independently verified. Quality review score: 9.0/10.

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 data is a pervasive modality spanning a wide range of domains.	✓	0.24
Recent advancements using transformer-based in-context learning have shown promise on smaller and less complex datasets.	✓	0.32
Transformer-based in-context learning has struggled to scale to larger and more complex datasets.	✓	0.26
The proposed approach combines retrieval and fine-tuning to adapt the transformer to a local subset of the data by colle	✓	0.32
The proposed approach performs task-specific fine-tuning with the retrieved set of neighbours in context.	✓	0.25
TabPFN is currently the best tabular in-context learner.	✓	0.20
Applying the retrieval and fine-tuning scheme on top of TabPFN results in a locally-calibrated PFN, or LoCalPFN.	✓	0.32
LoCalPFN establishes a new state-of-the-art on 95 datasets curated by TabZilla from OpenML.	✓	0.19
LoCalPFN outperforms tuned tree-based models on the 95 datasets curated by TabZilla from OpenML.	✓	0.22
LoCalPFN shows a significant boost in performance compared to the base in-context model.	✓	0.23
The proposed approach advances the frontier of deep learning in tabular data.	✓	0.15

References

- <https://doi.org/10.48550/arxiv.2406.05207>

- <https://doi.org/10.48550/arxiv.2407.00956>
- <https://doi.org/10.48550/arxiv.2406.08311>