# Tabular Foundation Models with High Sample Diversity Improve Few-Shot Generalization Across Domains

Assignee Research

June 8, 2026

**Abstract**

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: Do tabular foundation models trained with higher sample diversity demonstrate improved generalization across different domain gaps when evaluated on few-shot learning benchmarks like TabMWP or GTabNet. 8 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: Tabular Few-Shot Generalization Across Heterogeneous Feature Spaces. Research question: Do tabular foundation models trained with higher sample diversity demonstrate improved generalization across different domain gaps when evaluated on few-shot learning benchmarks like TabMWP or GTabNet?.

## 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 3.8/10.

## 3 Results

12 papers retrieved. 8 claims extracted; 0 independently verified. Quality review score: 3.8/10.

## 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 |
|---|---|---|
| The study uses 118 tabular classification datasets from the UCI Machine Learning Repository. | × | 0.07 |
| 65 of the 118 datasets have more than two prediction classes and were binarized for the study. | × | 0.03 |
| FLAT models are trained and tested using an N-fold evaluation procedure. | × | 0.03 |
| Feature columns are standardized to mean 0 and variance 1 during the training and testing process. | × | 0.04 |
| The study employs a randomized sampling procedure for imbalanced few-shot learning, with the number of positive examples | × | 0.05 |
| The study compares FLAT against several baselines including logistic regression (LR), k-nearest neighbors (KNN), support | × | 0.03 |
| FLAT and Iwata are meta-trained and tested using the same N-fold evaluation procedure. | × | 0.03 |
| The benchmark tables include numerical results for various models across different datasets. | × | 0.06 |

## References


- http://arxiv.org/abs/2311.10051v1
- http://arxiv.org/abs/2311.14544v1
- http://arxiv.org/abs/2504.20862v1