Self-Supervised Pretext Tasks and Adversarial Robustness in Tabular Data Representations

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: How do different self-supervised learning pretext tasks in tabular data affect the adversarial robustness of representations when evaluated with the Robustness Benchmark for Tabular Data (RBT) on the. 18 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: A Survey on Self-Supervised Learning for Non-Sequential Tabular Data. Research question: How do different self-supervised learning pretext tasks in tabular data affect the adversarial robustness of representations when evaluated with the Robustness Benchmark for Tabular Data (RBT) on the Adult Income dataset?.

2 Methodology

Systematic literature search across multiple databases yielded 16 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.2/10.

3 Results

16 papers retrieved. 18 claims extracted; 0 independently verified. Quality review score: 4.2/10.

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
The OpenML-CC18 benchmark contains 72 datasets.	×	0.01
The OpenML-CC18 benchmark datasets contain between 500 and 92,000 samples.	×	0.01
The OpenML-CC18 benchmark datasets contain between 5 and 3,073 features.	×	0.03
The DLBench benchmark includes datasets for both classification (C) and regression (R) tasks.	×	0.03
The DLBench benchmark consists of 11 datasets.	×	0.01
The TabularBench benchmark consists of 45 datasets.	×	0.01
The TabZilla benchmark consists of 36 datasets intended for classification tasks.	×	0.03
The TP-BERTa benchmark includes 202 unlabeled datasets and 145 labeled datasets for classification and regression.	×	0.03
The OpenTabs dataset contains 2,000 unlabeled datasets.	×	0.04
The average number of samples in the OpenTabs dataset is 23,000.	×	0.04
The UniTabE dataset contains 283,000 unlabeled datasets.	×	0.04
Masking strategies completely remove targeted features, whereas perturbations leave partial information.	×	0.02
Levin et al. (2023) introduced a pseudo-feature approach to predict missing features in upstream data that are present i	×	0.04
Ye et al. (2023) pre-trained a Transformer encoder using 2,000 high-quality cross-table datasets.	×	0.04
Analyses indicate that pre-training provides more transferability over tree-based baselines.	×	0.08
The DoRA method (Du et al., 2023) utilizes an intra-sample pretext task based on domain knowledge in the financial domai	×	0.06
In the DoRA method, two real estates located in the same town are positioned closer together via inter-sample contrastiv	×	0.04
Tabular data has practical utility in diverse domains including medicine and finance.	×	0.09

References

- http://arxiv.org/abs/2405.07414v2
- http://arxiv.org/abs/1909.08072v2
- http://arxiv.org/abs/2402.01204v4