Correlation of Novel Tabular Generative Metrics with Mode Collapse Detection in Imbalanced Datasets Versus Precision and Recall

Assignee Research

June 12, 2026

Abstract

Synthetic data is becoming an increasingly promising technology, and successful applications can improve privacy, fairness, and data democratization. While there are many methods for generating synthetic tabular data, the task remains non-trivial and unexplored for specific scenarios. One such scenario is survival data. Here, the key difficulty is censoring: for some instances, we are not aware of the time of event, or if one even occurred. Imbalances in censoring and time horizons cause generative models to experience three new failure modes specific to survival analysis: (1) generating too f

1 Introduction

This paper examines: SurvivalGAN: Generating Time-to-Event Data for Survival Analysis. Research question: How do novel tabular generative metrics correlate with mode collapse detection in imbalanced datasets compared to precision and recall across different synthetic data generators?.

2 Methodology

Systematic literature search across multiple databases yielded 2 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 8.1/10.

3 Results

2 papers retrieved. 7 claims extracted; 6 independently verified. Quality review score: 8.1/10.

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
Imbalances in censoring and time horizons cause generative models to experience three failure modes specific to survival	✓	0.45
The authors formalize three failure modes specific to survival analysis and provide three new generative metrics to quan	✓	0.29
SurvivalGAN is a generative model that addresses imbalance in censoring and event horizons and uses a dedicated mechanis	✓	0.32
SurvivalGAN was evaluated via extensive experiments on medical datasets.	×	0.13
SurvivalGAN outperforms multiple baselines at generating survival data.	✓	0.28
SurvivalGAN addresses the identified failure modes as measured by the new metrics.	✓	0.19
Using SurvivalGAN improves the downstream performance of survival models trained on the synthetic data.	✓	0.23

References

- https://doi.org/10.1007/s10462-025-11440-2
- https://doi.org/10.48550/arxiv.2302.12749