# What is the impact of using GAN-generated versus diffusion-generated tabular data on the reasoning capabilitie

Assignee Research

June 10, 2026

**Abstract**

Deep generative models have made rapid progress in image, text, audio, and video generation, and are increasingly being applied to structured records. For tabular data, however, generative modeling remains difficult: a dataset may contain numerical and categorical attributes, missing values, sensitive fields, imbalanced categories, complex feature dependencies, and domain constraints. Earlier tabular data modeling methods based on GANs or VAEs have achieved useful results, but they can suffer from unstable training, mode collapse, weak modeling of multimodal distributions, and fragile handling

## 1 Introduction

This paper examines: Diffusion and Flow Matching Models for Tabular Data: A Survey. Research question: What is the impact of using GAN-generated versus diffusion-generated tabular data on the reasoning capabilities of multimodal models trained on mixed-type datasets?.

## 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 3.8/10.

## 3 Results

13 papers retrieved. 0 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.

## References


- http://arxiv.org/abs/2412.00381v1
- http://arxiv.org/abs/2502.17119v2
- http://arxiv.org/abs/2104.11797v1