Llama-2 Model Size and Robustness in Adversarial Code Generation on MBPP Pro

Assignee Research

June 1, 2026

Abstract

This report synthesises findings from 10 peer-reviewed papers addressing the following research question: What is the impact of model size on the robustness of self-invoking code generation in Llama-2 models when evaluated against adversarial perturbations in MBPP Pro benchmarks. Generating high-fidelity and biologically plausible synthetic single-cell RNA sequencing (scRNA-seq) data, especially with conditional control, is challenging due to its high dimensionality, sparsity, and complex biological variations. Existing generative models often struggle. 7 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: LapDDPM: A Conditional Graph Diffusion Model for scRNA-seq Generation with Spectral Adversarial Perturbations. Research question: What is the impact of model size on the robustness of self-invoking code generation in Llama-2 models when evaluated against adversarial perturbations in MBPP Pro benchmarks?.

2 Methodology

Systematic literature search across multiple databases yielded 10 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.5/10.

3 Results

10 papers retrieved. 7 claims extracted; 0 independently verified. Quality review score: 3.5/10.

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
VAEs and GANs have been adapted for generating synthetic cellular profiles from scRNA-seq data.	×	0.14
Flow-based models have been explored for scRNA-seq data generation, offering exact likelihood estimation and invertible	×	0.13
GNNs have been applied to various tasks in single-cell biology, including cell type annotation, trajectory inference, an	×	0.06
Diffusion Probabilistic Models (DPMs) have demonstrated state-of-the-art performance in image synthesis and audio genera	×	0.07
The LapDDPM model combines diffusion, reconstruction, and KL divergence losses, with the encoder trained on graphs pertu	×	0.12
The LapDDPM model uses a k-NN graph constructed on cells, with PCA applied to log-transformed gene expression data to re	×	0.06
The adjacency matrix A is constructed using Euclidean distance in the PCA-reduced space.	×	0.01

References

- `http://arxiv.org/abs/2507.22398v3`
- `http://arxiv.org/abs/2412.21199v2`

- `http://arxiv.org/abs/2506.13344v1`