Synthetic Data Dimensionality Impacts on Multimodal Model Convergence and Loss

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 13 peer-reviewed papers addressing the following research question: What is the effect of synthetic data dimensionality on the convergence speed and validation loss of multimodal models fine-tuned on sparse structured datasets. 7 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Learning Active Subspaces and Discovering Important Features with Gaussian Radial Basis Functions Neural Networks. Research question: What is the effect of synthetic data dimensionality on the convergence speed and validation loss of multimodal models fine-tuned on sparse structured 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. 7 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
For Gaussian Radial Basis Functions, the matrix Φ is not singular if all data points are distinct and N > 2.	×	0.13
The proposed model uses a Gaussian basis function with a symmetric positive definite precision matrix P expressed as U^T	×	0.11
The function approximation problem is solved by minimizing a nonconvex optimization problem involving weights w and the	×	0.02
The error function E(w, u) for the regression case is defined as half the sum of squared differences between the target	×	0.02
The total number of parameters to optimize in the proposed model is calculated as P = M + D + D$\times$(D-1).	×	0.04
In the specific case described, the number of basis functions M is equal to the number of data points N.	×	0.04
Table (p19) displays feature importance scores for features x1 through x10 for a dataset with N=100.	×	0.04

References

- http://arxiv.org/abs/2411.15497v3
- http://arxiv.org/abs/2512.03307v1
- http://arxiv.org/abs/2307.05639v2