Scaling Synthetic Feature Dimensionality in CausalMixFT: Stability and Generalization Trade-offs in Tabular Data

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 14 peer-reviewed papers addressing the following research question: How does scaling the dimensionality of synthetic features in CausalMixFT affect the trade-off between fine-tuning stability (measured by validation loss variance) and generalization performance. 7 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.3/10. 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: How does scaling the dimensionality of synthetic features in CausalMixFT affect the trade-off between fine-tuning stability (measured by validation loss variance) and generalization performance (measured by test accuracy) on tabular benchmarks like OpenML or Yandex datasets?.

2 Methodology

Systematic literature search across multiple databases yielded 14 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.3/10.

3 Results

14 papers retrieved. 7 claims extracted; 0 independently verified. Quality review score: 3.3/10.

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.	×	0.14
The proposed model uses a Gaussian basis function with a symmetric positive definite precision matrix P expressed as P =	×	0.12
The function approximation problem is solved by minimizing a nonconvex optimization problem with respect to weights w an	×	0.03
The error function E(w, u) for the regression case is defined as the sum of squared differences between target values yn	×	0.03
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) presents feature importance scores for features x1 through x10 for a dataset with N=100.	×	0.03

References

- http://arxiv.org/abs/2105.04026v2
- http://arxiv.org/abs/2501.08306v4
- http://arxiv.org/abs/2307.05639v2