Bayesian Neural Networks and Deterministic Transformers in Code Generation Calibration and Accuracy

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: How do Bayesian neural networks compare to deterministic transformers in calibration error and pass@1 scores on the HumanEval code generation benchmark. 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: How do Bayesian neural networks compare to deterministic transformers in calibration error and pass@1 scores on the HumanEval code generation benchmark?.

2 Methodology

Systematic literature search across multiple databases yielded 15 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

15 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.	×	0.14
The proposed model uses a Gaussian basis function with a symmetric positive definite matrix P expressed as U^T U, where	×	0.09
The function approximation problem is solved by minimizing a nonconvex optimization problem involving parameters w and u	×	0.03
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/2504.19176v2
- http://arxiv.org/abs/2303.12869v1
- http://arxiv.org/abs/2307.05639v2