Node-Based vs. Weight-Based Bayesian Neural Networks under Covariate Shift in Tabular Data

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 4 peer-reviewed papers addressing the following research question: How do node-based BNNs perform in terms of robustness to covariate shift compared to weight-based BNNs when trained on tabular datasets like CIFAR-10 or MNIST, measured using accuracy and calibration. 11 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: Navigating Simply, Aligning Deeply: Winning Solutions for Mouse vs. AI 2025. Research question: How do node-based BNNs perform in terms of robustness to covariate shift compared to weight-based BNNs when trained on tabular datasets like CIFAR-10 or MNIST, measured using accuracy and calibration scores?.

2 Methodology

Systematic literature search across multiple databases yielded 4 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

4 papers retrieved. 11 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
Our final model combining SimpleCNN, GLU, and observation normalization achieves a final score of 95.4%, securing top po	×	0.10
The IMPALA ResNet baseline with 4 layers achieves 87.7% final score despite having significantly more parameters.	×	0.05
The 24-layer ResNet achieves only 65.98% due to severe overfitting.	×	0.02
The deep 24-layer ResNet exhibits a 30 percentage point gap (80.96% vs 51.00%), indicating that the model memorizes trai	×	0.06
Our SimpleCNN maintains only a 2.8 percentage point gap (96.80% vs 94.00%), demonstrating robust generalization.	×	0.03
Increasing depth from 4 to 24 layers substantially harms performance, contradicting the conventional wisdom that deeper	×	0.07
The visual encoder consists of two convolutional layers with aggressive spatial downsampling.	×	0.05
The first layer applies a 8$\times$8 kernel with stride 4 to the 86$\times$155$\times$1 input, producing 16 feature channels.	×	0.02
The second layer applies a 4$\times$4 kernel with stride 2, expanding to 32 feature channels while further reducing spatial dim	×	0.01
Both layers employ LeakyReLU activation with negative slope 0.2 to allow gradient flow through non-positive activations.	×	0.01
The resulting feature maps are flattened and projected to 256 dimensions via a fully-connected layer.	×	0.03

References

- http://arxiv.org/abs/2602.00982v1
- http://arxiv.org/abs/2504.07569v2
- http://arxiv.org/abs/2105.10886v1