Node-Based Bayesian Neural Networks and Gradient Boosting Machines under Domain Shift on OpenML Datasets

Assignee Research

June 7, 2026

Abstract

This report synthesises findings from 13 peer-reviewed papers addressing the following research question: How do node-based Bayesian neural networks perform relative to gradient boosting machines in terms of accuracy degradation when evaluated on OpenML datasets with simulated domain shift. 15 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 2.4/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Tackling covariate shift with node-based Bayesian neural networks. Research question: How do node-based Bayesian neural networks perform relative to gradient boosting machines in terms of accuracy degradation when evaluated on OpenML datasets with simulated domain shift?.

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 2.4/10.

3 Results

13 papers retrieved. 15 claims extracted; 0 independently verified. Quality review score: 2.4/10.

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
The study uses CIFAR and TINYIMAGENET datasets with corrupted test set versions provided by Hendrycks & Dietterich (2019	×	0.03
The architectures used in the experiments are VGG16, RESNET18, and PREACTRESNET18.	×	0.02
Three latent variable structures were tested: 'in' (input only), 'out' (output only), and 'both' (input and output).	×	0.04
The number of Gaussian components (K) in the variational posterior was set to 1, 2, or 4.	×	0.04
On OOD data, the optimal performance of using both input and output latent variables is similar to using only output lat	×	0.08
Using only input latent variables produces slightly worse optimal performance on OOD data compared to using output or bo	×	0.08
The optimal gamma (γ) is lower when the model uses both types of latent variables (z, s) compared to other configuration	×	0.05
In Figure 8 experiments, RESNET18 was tested on CIFAR10 subsets with 20%, 40%, and 80% corrupted training labels using K	×	0.05
As gamma (γ) increases, the Negative Log-Likelihood (NLL) of noisy labels increases much faster than that of clean label	×	0.02
Higher gamma (γ) prevents the model from memorizing random labels.	×	0.03
High gamma (γ) leads to improved performance on clean test sets because it prevents learning from noisy labels.	×	0.05
Figure 3 results are reported as the mean and standard deviation over 25 runs.	×	0.06
The model with higher entropy (M32) performs better than the model with lower entropy (M16) across all corruption levels	×	0.03
The parameter lambda (λ) controls the severity of the generated implicit corruptions.	×	0.06
Explicit image corruptions g(x) are approximated by finding xc that minimizes the loss function L(xc) defined in Equatio	×	0.03

References

- http://arxiv.org/abs/1512.00242v1
- http://arxiv.org/abs/2206.02435v2
- http://arxiv.org/abs/2007.09855v5