

SOVEREIGN: How does node-based Bayesian neural network alignment compare to weight-based BNNs in reducing accuracy degrad

SOVEREIGN Research Kernel

Autonomous draft — Owner review required before publication

May 28, 2026

Abstract

Bayesian neural networks (BNNs) promise improved generalization under covariate shift by providing principled probabilistic representations of epistemic uncertainty. However, weight-based BNNs often struggle with high computational complexity of large-scale architectures and datasets. Node-based BNNs have recently been introduced as scalable alternatives, which induce epistemic uncertainty by multiplying each hidden node with latent random variables, while learning a point-estimate of the weights. In this paper, we interpret these latent noise variables as implicit representations of simple an

1 Introduction

Analysis of: Tackling covariate shift with node-based Bayesian neural networks. Research goal: How does node-based Bayesian neural network alignment compare to weight-based BNNs in reducing accuracy degradation on out-of-distribution image classification benchmarks like CIFAR-10-C and ImageNet-C?.

2 Methodology

Multi-query arXiv search (4 parallel queries, Relevance-sorted). TF-IDF cosine semantic verification (bigrams, threshold=0.15). NIM nv-embedqa-e5-v5 (dim=1024) for semantic indexing. Tribunal v2: 3-role parallel review (SKEPTIC/VALIDATOR/SYNTHESIZER) with revision round if score < 6.5.

3 Results

11 papers retrieved. 9 claims extracted, 3 verified. Tribunal: 4.7/10 → REVISE (revision_round=1). Policy: ESCALATE_TO_OWNER.

4 Uncertainties

NIM free tier latency varies. TF-IDF verification is a weak signal. arXiv Relevance ranking is query-dependent. Tribunal consensus is LLM-based and prompt-sensitive.

5 Extracted Claims

Claim	Verified	Confidence
Node-based BNNs with K=4 Gaussian components in the variational posterior achieve better performance on covariate shift	✓	0.17
Using both input and output latent variables in node-based BNNs produces similar optimal performance on out-of-distribution	✓	0.16
Using only input latent variables in node-based BNNs produces slightly worse optimal performance on out-of-distribution	✓	0.16
The optimal gamma is lower when the model uses both types of latent variables (z, s) because the entropy of the product	×	0.12
Higher gamma prevents the model from memorizing random labels, as shown by NLL of noisy labels increasing much faster than	×	0.03
High gamma leads to improved performance on clean test sets when training labels contain noise.	×	0.08
The variational entropy $H[q(Z)]$ decreases over time during training.	×	0.05
A model with higher entropy (M32) performs better than one with lower entropy (M16) across all corruption levels.	×	0.02
Lambda controls the severity of generated corruptions in the implicit corruption visualization method.	×	0.06

References

- <http://arxiv.org/abs/1807.01697v5>
- <http://arxiv.org/abs/2003.03879v1>
- <http://arxiv.org/abs/2206.02435v2>