[s25] BA-Seminar: Probabilistic ML (Bothmann)

Outline: “Symmetry & Mode Connectivity” (Topic 11) – Thomas Witzani

1. Introduction & Motivation
   Motivation: What are Bayesian neural networks (BNNs) and why do posterior modes matter?
   Problem: Symmetric modes in the posterior make uncertainty quantification hard.
   Objective: Build a DEI-MCMC pipeline with an optional symmetric-mode-removal step before MCMC initialization, enabling efficient and reliable posterior exploration and uncertainty quantification.

2. Related Work
   Existing BNN posterior sampling methods (e.g., HMC, SG-MCMC).
   Deep-ensemble initialization (DEI-MCMC) and benefits of ensemble-based MCMC starts.
   Symmetry detection and elimination in BNNs (Wiese et al. 2023).
   Mode connectivity and feasibility of sample-based inference in BNNs (Sommer et al. 2024).

3. Methods (Description & Comparison)
   DEI-MCMC: Use a small deep ensemble to initialize MCMC chains.
   Non-centered parameterization and reduce_sum likelihood in Stan for parallel sampling.
   Symmetric mode removal: Canonicalize each weight configuration by sorting neurons and fixing sign conventions. Apply to the deep ensemble by collapsing permutation- and sign-flip-equivalent members (symmetry group members) into a single representative before MCMC.

4. Simulation Study
   Generate synthetic datasets (linear, sinusoidal, quadratic & sharp piecewise).
   Train small deep ensembles on each synthetic set.
   Setup the Stan (BNN) model.
   Run DEI-MCMC (with/without symmetry removal).
   Measure runtime, convergence diagnostics, and posterior-predictive performance.
   Check whether posterior-predictive curves recover the original generating functions.
   Goal: Tune the pipeline for later application to real data.

5. Application to a Real Dataset
   Select a UCI dataset.
   Adapt architecture and parameters to real data.
   Run DEI-MCMC (with/without symmetry removal).
   Evaluate: Predictive accuracy (RMSE, log PPD) and calibration (credibility interval coverage).

6. Discussion
   Summarize findings from the simulations and the real-data case.
   Assess DEI-MCMC without symmetry removal.
   Assess DEI-MCMC with symmetry removal (uncertainty estimates, runtime overhead).

7. Conclusion
   Key takeaway: DEI-MCMC enables robust BNN sampling; symmetry removal can further improve uncertainty estimates.

8. Outlook
   Fully integrate symmetry removal and validate on larger real datasets. Investigate GPU acceleration and alternative parallelization.