MA 578: Bayesian Statistics

Final Project Proposal

Bayesian Temperature Scaling for Neural Network Calibration

November 11, 2025

Problem Description and Significance. Neural networks often have poor calibration. Their predicted probabilities do not match actual correctness rates. A model might assign 95% confidence to a prediction that is only correct 70% of the time. This is a problem in safety-critical applications like healthcare, autonomous vehicles, and finance where accurate uncertainty is needed. This project develops a Bayesian temperature scaling method that improves calibration and quantifies uncertainty in the calibration parameter.

Dataset and Methodology. I will use the CIFAR-10 dataset, which contains 60,000 32×32 color images across 10 classes. A small CNN will be trained to produce logits. The temperature parameter T will be treated as a random variable with a prior distribution (log-normal or gamma) and its posterior will be estimated using MCMC in PyMC. This extends standard temperature scaling by providing a full posterior distribution instead of just a point estimate, which is useful when validation data is limited.

Expected Results. The project will report posterior distributions and credible intervals for the temperature parameter. Results will be evaluated using Expected Calibration Error (ECE), Brier score, and reliability diagrams, compared against standard temperature scaling. The analysis will include sensitivity analysis for different priors, posterior predictive checks, and visualizations of calibration improvements and uncertainty. This will show how Bayesian methods can improve neural network calibration and uncertainty quantification.