

PAC-Bayes training for neural networks: sparsity and uncertainty quantification

Wednesday, 3 August 2022 17:00 (20 minutes)

Increasing computational power and storage capacity have made high-dimensional datasets accessible to many areas of research such as medicine, natural and social sciences. While classical statistical methods are not compatible with high-dimensional data, especially due to the curse of dimensionality, machine learning methods have been successfully applied to regression problems in practice. On the theoretical level, a popular way to circumvent the curse of dimensionality is the concept of sparsity. We study the Gibbs posterior distribution from PAC-Bayes theory for sparse deep neural nets in a nonparametric regression setting. To access the posterior distribution, an efficient MCMC algorithm based on backpropagation is constructed. The training yields a Bayesian neural network with a joint distribution on the network parameters. Using a mixture over uniform priors on sparse sets of network weights, we prove an oracle inequality which shows that the method adapts to the unknown regularity and hierarchical structure of the regression function. Studying the Gibbs posterior distribution from a frequentist Bayesian perspective, we analyze the diameter and show high coverage probability of the resulting credible sets. The method is illustrated with an animation in a simulation example.

This talk is based on joint work with Mathias Trabs.

Primary author: STEFFEN, Maximilian (Karlsruhe Institute of Technology)

Co-author: Prof. TRABS, Mathias (Karlsruhe Institute of Technology)

Session Classification: Session B2 Statistics

Track Classification: Statistics