Artificial Intelligence and Sustainability Research Group
Prof. Dr. Ralf Herbrich

MASTER'S PROJECT WINTER TERM 2023

BAYESIAN DEEP LEARNING ALGORITHMS

DESCRIPTION

Over the past 10 years, deep learning algorithms have become the de-facto standard learning algorithms when large amounts of data ($\approx 10^9$ data) are available to train from. Neural networks are not only able to learn a mapping of a data item x to the target variable y, but also (latent) representations $r(x)$ of the data by sequentially mapping the data x to vectorial representations $z_1 = h(W_1 x)$ and $z_{l+1} = h(W_{l+1} z_l)$, where h are the component-wise activation functions. The weights $W_1, W_2, \ldots, W_L$ are most commonly learned via (stochastic) gradient-based optimization, resulting in a single set of weights.

One disadvantage of such learning algorithms is that the weights are estimated without any uncertainty, even though a finite dataset provides only finite information about the most probable weight values. As a consequence, predictions of the target value y are overly confident, which is also referred to as overfitting. In this project, we will explore Bayesian approximate inference methods that attempt to remedy this problem, with a focus on algorithms that take into account the model structure.

The most promising approach to overcome this problem is to introduce **two** parameters (of a Gaussian belief distribution) for each weight $w_{l,i,j}$, namely $\mu_{l,i,j}$ and $\sigma^2_{l,i,j}$ assuming that $w_{l,i,j} \sim N\left(\mu_{l,i,j}, \sigma^2_{l,i,j}\right)$. Then, based on large amounts of training data, instead of learning a single value $w_{l,i,j}$ for each weight of the deep neural network, we attempt to find the "best fitting" values $\mu_{l,i,j}$ and $\sigma^2_{l,i,j}$ for each weight.

In this project, we will explore two algorithmic ideas with a special focus on large-scale: (1) Gibbs sampling where we pick a subset ("block") S of the weights $w_{l,i,j}$ and iteratively compute new samples of $w_{l,i,j}$ with $(l, i, j) \in S$ assuming all other weights are kept constant. As one can show, if such a "block" consists of all the weights pointing into one "inner node" $z_{l,i}$, then this results in a closed form linear-time sampling algorithm; (2) Variational inference and (iterative) Laplace approximation with block coordinate descent for a block S of the weights $w_{l,i,j}$, again assuming samples from all other weights are kept constant. Furthermore, we optimize local objectives for target representations, which are projections from the target observations. To the best of our knowledge, neither of these two ideas have been explored in depth, in particularly not with a focus on scalability.

In this project, we will:

- Develop an evaluation framework for Bayesian deep learning algorithms in Python/Julia
- Derive and develop the update algorithm for both block-Gibbs sampling and block-variational inference

- Evaluate the data-efficiency, predictive accuracy and calibration of existing (e.g. Laplace approximation and variational inference) and novel Bayesian deep learning algorithms on large-scale datasets

LEARNING GOALS

Through successful completion of this project, you will:

- Pass the entire cycle of a research project, from formulating initial research questions, to prototyping solutions and conducting experimental results, to deriving novel insights and formulating them in a scientific text
- Improve your programming and teamwork skills
- Learn to familiarize yourself with and work on an existing large software project with a focus on deep learning systems and Bayesian inference
- Deepen your probabilistic inference and deep learning knowledge
- Improve your research methodology and academic writing

At the end of the project, we will condense our insights and results into a research submission to be submitted to a top AI/ML conference.

PREREQUISITES

- Prior understanding of the fundamentals of deep learning (e.g., linear algebra, systems)
- Prior understanding of the fundamentals of probabilistic learning (e.g., probability theory)
- Familiarity with Python, C++ and/or Julia

INITIAL RELATED WORK

[1] Osawa et al. Practical Deep Learning with Bayesian Principles. NeurIPS 2019.

[2] Daxberger et al. Laplace Redux -- Effortless Bayesian Deep Learning. NeurIPS 2021.

[3] Piccioli et al. Gibbs Sampling the Posterior of Neural Networks. Arxiv preprint 2023.

[4] Kurle et al. On the detrimental effect of invariances in the likelihood for variational inference. NeurIPS 2022.

CONTACT

You are welcome to contact **Prof. Dr. Ralf Herbrich** (ralf.herbrich@hpi.de) or **Richard Kurle** (Richard.kurle@hpi.de) or to visit us in the F building, ground floor (on Campus II).