Bayesian View on Deep Learning Training: Variational Inference

SHUHEI, ISHIWATARI[1]., Dr. HUANG XIAOWEI & Dr. XINGYU ZHAO.

[1]*Department of Computer Science, University of Liverpool, Liverpool, UK*

A traditional neural network only produces point estimates of parameters; hence it fails to introduce the notion of uncertainty, making it unsuitable for certain applications. Point estimates are also associated with overfitting, as an extensive amount of training observations are required to calibrate what is often millions of parameters of a network. Bayesian neural network (BNN), a combination of deep neural networks and probabilistic models, provides the solution to the above limitations by representing the parameters as the posterior distributions. However, the computation of the true posterior is often intractable, hence variational inference is applied to approximate the posterior with the variational distribution. This project implemented Bayes by Backprop, a variational inference method to identify the optimal parameters of the approximate distribution in backpropagation. The project further explored how the choice of prior distributions affects the BNN's performance. Hyperparameters of a prior were also investigated in the project. The findings suggest that Bayes by Backprop can be successfully applied to learn the weight distributions in the MNIST dataset. The choice of a prior and its hyperparameter is imperative to the BNN's performance. Finally, the BNN model successfully introduced the notion of uncertainty on the previously unknown CIFAR10 dataset.