Neurocomputing, vol.471, pp.260-274, January, 2022.

S. Rodríguez Santana; D. Hernández Lobato

Abstract-

machine learning problems. They are often trained via back-propagation to find a value of the weights that correctly predicts the observed data. Back-propagation has shown good performance in many applications, however, it cannot easily output an estimate of the uncertainty in the predictions made. Estimating this uncertainty is a critical aspect with important applications. One method to obtain this information consists in following a posterior distribution of the model parameters. This posterior distribution summarizes which parameter values are compatible with the observed data. However, the posterior is often intractable and has to be approximated. Several methods have been devised for this task. Here, we propose a general method for approximate Bayesian inference that is based on minimizing α-divergences, and that allows for flexible approximate distributions. We call this method adversarial α-divergence minimization (AADM). We have evaluated AADM in the context of Bayesian neural networks. Extensive experiments show that it may lead to better results in terms of the test log-likelihood, and sometimes in terms of the squared error, in

Index Terms- Bayesian neural networks; Approximate inference; Alpha divergences; Adversarial variational Bayes

Citation: