A Study of Bayesian Neural Network: Literature Review

Shaoyu Feng (sf865)
Yunjia Zeng (yz682)

March 19, 2020

In advancement of computer hardware and computational power, there has been a tremendous research trend on employing deep learning techniques, represented by neural networks, to solve challenging problems like computer vision, natural language understanding etc. And we have observed unprecedented performance of those techniques in different functional areas. However, critiques have been demonstrated that traditional neural network architectures are prone to overfitting (N.Srivastava et al. 1), and thus various regularization methods has been developed such as early stopping, weight decay and dropout (N.Srivastava. et al. 1; Blundell et al., 1). Moreover, traditional neural networks do not address the uncertainties in model weights, and provides no useful measures in the form of confidence in predictions or results (Blundell et al., 1).

In light of those shortcomings for traditional neural networks, research works have been carried out to apply Bayesian paradigm into the neural network architecture, which make it possible to assign a probability distribution towards model weights of the network. (Blundell et al., 2) Those types of neural networks with uncertainties over the weights is normally known as Bayesian Neural Network (BNN). However, there are also a few challenges in constructing an effective BNN, including setting up the reasonable prior, selecting the variational inference method and thus approximating the loss function and performing back-propagation on the network to perform gradient descent. (Neal, 22).

In Bayesian inference in BNN, the very starting point is to choose the prior distribution over the model parameters, which in BNN are the connections weights between different layers and the bias. Most past works on Bayesian inference for neural network have used independent Gaussian distribution, which make it possible to make the training process to converge to a Gaussian process.(Neal, 34; MacKay, D, 448). Recent researchers have also focused on more intricate effect at unit level, state that the induced prior distribution becomes increasingly heavy-tailed with the depth of the neural network layers. (Mariia et al ,1). However, the selection of priors remains ad-hoc. (Neal, 17), mixed priors could lead to a significant improvement in model performance (Blundell et al., 5).

BNN normally use standard normal as prior and calculate the posterior distribution of weights and biases using the training data. The entire posterior distribution of weights and biases are used to predict on test data, by taking the expectations of the weights, which is equivalent to take an ensemble of infinite networks. (Blundell et al., 3) Despite the theoretical advantage, this adds extra difficulty to train the model in practice. Bayesian inference on the weights of a neural network is intractable as the number of parameters is very large and the functional form of a neural network does not allow easy

integration for weights update. (Miller et al, 1) Therefore, Bayesian variational inference method has to be applied in approximating posterior distribution on the weights. In particular, the usage of variational approximation attempts to find the distribution over model weights, that minimizes the Kullback-Leibler (KL) Divergence. (Miller et al, 3; Blundell et al.,3).

However, KL divergence optimization is intractable as it directly depends on posterior distribution. (Miller et al, 3) Another quantity called Evidence Lower Bound (ELBO) has to be added which could deduce a lower bound for KL divergence, and we could use reparameterization gradient estimators (RGEs) to calculate the gradient of ELBO and enable gradient descent optimization during network training. (Miller et al, 3-6). The resulting cost function is variously known as the variational free energy or the expected lower bound (Blundell et al.,3).

BNNs win advantages over traditional neural network in following ways: regularization over the weights to reduce overfitting (Blundell et al., 8), posterior inference on the weights for prediction, and useful measures of uncertainty attributed to unseen predictor space or heteroskedasticity (Neal, 47). BNNs have been recently adopted in various areas in deep learning research. The uncertainty in BNNs can be used to drive exploration in contextual bandit problems using Thompson sampling in reinforcement learning (Blundell et al.,5). Bayesian GAN has been shown to outperform many GAN architectures like DCGAN, DCGAN ensembles on many benchmark datasets (Sattchi, 1).

1 Reference

1. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. "Dropout: a simple way to prevent neural networks from overfitting." The Journal of Machine Learning Research, 15(1):1929–1958, 2014.
2. C. Blundell, J. Cornebise, K. Kavukcuoglu, and D. Wierstra. "Weight uncertainty in neural network." Proceedings of the 32nd ICML volume 37 of Proceedings of Machine Learning Research, pages 1613–1622. PMLR, 2015.
3. Alex Graves. "Practical variational inference for neural net- works". In Advances in Neural Information Processing Systems (NIPS), pages 2348–2356, 2011.
4. Tran, D., Kucukelbir, A., Dieng, A. B., Rudolph, M., Liang, D., and Blei, D. M.. "Edward: A library for probabilistic modeling, inference, and criticism." arXiv:1610.09787.
5. Neal, R. M. "Bayesian learning for neural networks." PhD thesis, University of Toronto. 1995
6. MacKay, D. "A practical Bayesian framework for backpropagation networks". Neural Computation, 4(3):448–472, 1992.
7. Mariia Vladimirova, Jakob Verbeek, Pablo Mesejo, and Julyan Arbel. "Understanding priors in Bayesian neural networks at the unit level". volume 97 of Proceedings of Machine Learning Research, pages 6458–6467, 2019.
8. Andrew C Miller, Nicholas J Foti, Alexander D'Amour, and Ryan P Adams. "Reducing reparameterization gradient variance. " arXiv preprint arXiv:1705.07880, 2015.
9. Saatchi, Y. and Wilson, A. G. "Bayesian gan". In Advances in Neural Information Processing Systems, pp. 3625–3634, 2017.