WEIGHTED PARAMETER AVERAGING IN DEEP NEURAL NETWORK

Anonymous authors
Paper under double-blind review

ABSTRACT

We propose Weighted Parameter Averaging in Deep Neural Network, an algorithm to improve DNN accuracy and generalization. Our method is to average the neural network parameters of n best model checkpoints post training, weighted by validation loss of respective checkpoints. The resulting model generalizes better than those with a single best checkpoint model over unseen data. We present the performance of Weighted Parameter Averaging in Deep Neural Network at multiple experimental setups. For conducting experiments we use different combinations of Optimizers and Schedulers over multiple datasets and present the resulting performance scores. We demonstrate the reduction in error and good generalization performance of the resulting models over our experimental datasets.