Perturbations in neural networks

Raza Hashmi

24-05-2021

1 Proposal

In this paper, we propose a novel way to reduce generalization error in neural networks. A standard neural network works by propagating input data through its hidden layers and improving their predictions through gradient descent on the errors. However, with limited training data and huge network size, this typically leads to over-fitting the input data. To counter this effect, we propagate a perturbation through the network along with training data. Rather than adding perturbations into the training sample itself, the perturbation is treated as a concurrent input to the network. Hence the total loss of the network becomes loss on prediction plus loss of perturbations. Similar to the popular weight regularizations, the perturbation loss is also controlled by a hyper-parameter and can be annealed as the training progresses.
There may be different kinds of perturbations. For a fully connected neural network, the perturbations might be Gaussian noise. In which case the loss function for the network becomes:

$$Total\,loss = \ loss\,on\,prediction_{\,training\,sample} + \lambda(output_{\,noise\,input})^2$$

However, for a convolutional neural network the perturbations can be extended to include augmented data for the training sample. In which case the loss of the network becomes:

$$\begin{aligned} Total\,loss = {} & loss\,on\,prediction_{\,training\,sample} + \lambda(output_{\,noise\,input})^2 \\ & + \frac{\tau}{batch\,size} \times MSE(Target, Predicted_{\,augmented\,training\,sample}) \end{aligned}$$

This loss on augmented data acts as a consistency regularization and helps the network to generalize better. Similar perturbations can also be added to different neural network architectures. The hyper parameters λ and τ both can take values in between 0 and 1. Preliminary experiments show that the hyper-parameters play a pivotal role in increasing the test accuracy.
Running a few initial experiments using MNIST dataset and a fully connected neural network with varying depths show that adding noise to the network results in significantly better test accuracy as compared to a standard neural network. Furthermore for the same number of training iterations, neural networks with perturbations on average always outperform a standard neural network. These findings are further validated by similar results on Fashion MNIST dataset.