

Final Project Proposal

Final Project: Proposal

By James Camacho and Joseph Camacho

Option Five: Paper Implementation

We will be reimplementing the paper below, which uses covariant dropout to increase accuracy and convergence speed. After testing we will add it to Tensorflow and/or Pytorch.

Paper

Abstract Dropout has been witnessed with great success in training deep neural networks by independently zeroing out the outputs of neurons at random. It has also received a surge of interest for shallow learning, e.g., logistic regression. However, the independent sampling for dropout could be suboptimal for the sake of convergence. In this paper, we propose to use multinomial sampling for dropout, i.e., sampling features or neurons according to a multinomial distribution with different probabilities for different features/neurons. To exhibit the optimal dropout probabilities, we analyze the shallow learning with multinomial dropout and establish the risk bound for stochastic optimization. By minimizing a sampling dependent factor in the risk bound, we obtain a distribution-dependent dropout with sampling probabilities dependent on the second order statistics of the data distribution. To tackle the issue of evolving distribution of neurons in deep learning, we propose an efficient adaptive dropout (named `\textbf{evolutional dropout}`) that computes the sampling probabilities on-the-fly from a mini-batch of examples. Empirical studies on several benchmark datasets demonstrate that the proposed dropouts achieve not only much faster convergence and but also a smaller testing error than the standard dropout. For example, on the CIFAR-100 data, the evolutional dropout achieves relative improvements over 10% on the prediction performance and over 50% on the convergence speed compared to the standard dropout.