

CRUNCH Seminars at Brown, Division of Applied Mathematics

Friday – February 22, 2019

Paper Review: Elimination of All Bad Local Minima in Deep Learning by Kenji Kawaguchi and Leslie Pack Kaelbling
Yeonjong Shin

In this paper, we theoretically prove that we can eliminate all suboptimal local minima by adding one neuron per output unit to any deep neural network, for multi-class classification, binary classification, and regression with an arbitrary loss function. At every local minimum of any deep neural network with added neurons, the set of parameters of the original neural network (without added neurons) is guaranteed to be a global minimum of the original neural network. The effects of the added neurons are proven to automatically vanish at every local minimum. Unlike many related results in the literature, our theoretical results are directly applicable to common deep learning tasks because the results only rely on the assumptions that automatically hold in the common tasks. Moreover, we discuss several limitations in eliminating the suboptimal local minima in this manner by providing additional theoretical results and several examples.