Combining adaptive-computation-time and learning-to-learn approaches for optimizing loss functions of base-learners

J. Sander

Abstract

Learning to learn approaches can be used to train a recurrent neural network (RNN) that learns an optimization algorithm. Adaptive computation time for RNNs allows a network to adjust its depth at each time step to the input received so far. In this work we developed two adaptive meta-learners that combine the learning to learn and adaptive computation time approaches in order to optimize loss functions of base-learners. Integrating both approaches is motivated by the idea that a learned, iterative optimization algorithm benefits from being able to adjust the number of optimization steps to the input, by learning to weight the time step losses. Our newly proposed models could be trained with less computational effort on convex and non-convex optimization tasks compared to the baseline optimizer introduced by Andrychowicz et al. (2016). Moreover the adaptive optimizers developed their own training regime that trades off computational effort against accuracy based on a prior injected preference.