

Generalization error analysis of deep learning: avoiding curse of dimensionality and practical application

Taiji Suzuki

Department of Mathematical Informatics, University of Tokyo

Abstract:

In this talk, we discuss the generalization error of deep learning by analyzing the complexity of deep neural network models in some settings. Especially, we consider how to avoid the curse of dimensionality from theoretical and practical point of view. First, we define a data dependent intrinsic dimensionality of deep neural network models and show how it affects generalization performance. To analyze that, we define a reproducing kernel Hilbert space for each internal layer and borrow theoretical techniques from kernel methods. It gives a fast learning rate which has mild dependence on the network size unlike the previous analysis. We also develop a simple compression algorithm which is applicable to wide range of network models and show that it gives state-of-the-art compression performance. Second, we will talk about a wavelet approach to analyze the generalization error of deep learning. Through the analysis, we discuss how the curse of dimensionality is avoided by deep learning under suitable smoothness assumptions.