

Generalization error of deep learning and its learning dynamics from compression ability point of view

Taiji Suzuki

The University of Tokyo
E-mail: taiji@mist.i.u-tokyo.ac.jp

Abstract: One of biggest issues in deep learning theory is its generalization ability despite the huge model size. The classical learning theory suggests that overparameterized models cause overfitting. However, practically used large deep models avoid overfitting, which is not well explained by the classical approaches. To resolve this issue, several attempts have been made. Among them, the compression based bound is one of the promising approaches. In this talk, we give a new frame-work for compression based bounds. The bound gives even better rate than the one for the compressed network by improving the bias term. We can obtain a data dependent generalization error bound which gives a tighter evaluation than the data independent ones. Moreover, we discuss the learning dynamics of deep learning and how a compressible network is trained using the notion of neural tangent kernel.