

# Generalization Analysis for Mechanism of Deep Learning via Nonparametric Statistics

Masaaki Imaizumi

*The Institute of Statistical Mathematics*

*E-mail: imaizumi@ism.ac.jp*

**Abstract:** We theoretically investigate an advantage of deep neural networks (DNNs) which empirically perform better than other standard methods. While DNNs have empirically shown higher performance than other methods, understanding its mechanism is still a challenging problem. From an aspect of the nonparametric statistics, it is known many standard methods attain the optimal rate of errors for standard settings such as smooth functions, and thus it has not been straightforward to find theoretical advantages of DNNs. Our study fills this gap by extending a class for data generating processes. We mainly consider the following two points; non-smoothness of functions and intrinsic structures of data distributions. We derive the generalization error of estimators by DNNs with a ReLU activation, and show that convergence rates of the generalization error can describe an advantage of DNNs over some of the other methods. In addition, our theoretical result provides guidelines for selecting an appropriate number of layers and edges of DNNs. We provide numerical experiments to support the theoretical results.