Optimal Approximation of Classifier Functions by Deep ReLU Networks

Philipp Petersen

Department of Mathematics, Technische Universität Berlin, Germany

Email: philipp.petersen@tu-berlin.de

In this talk, we analyze function approximation by deep neural networks. In particular, we are inter-ested in the approximation of functions that assume only finitely many values, i.e. piecewise constant functions. This analysis is motivated by the recent successful applications of deep neural networks in classification tasks, where complex and high-dimensional data is mapped to only a few classifying la-bels. We restrict ourselves to the ReLU activation function, which is probably the most widely used in applications. For this activation function, we will demonstrate that deep neural networks yield optimal approximation rates in an information theoretical sense. Using a concrete construction, we show that, for n, d ∈ N arbitrary, the number of edges and the depth of a ReLU neural network approximating a piecewise Cn function on Rd with jump singularities along Cn curves, scale optimally with respect to the approximation quality. To study the optimality, we establish the asymptotic description complexity of piecewise constant functions and interpret networks as encoder-decoder pairs. Fundamental limits on the efficiency of encoding algorithms then provide lower bounds of neural networks. This is joint work with Helmut B¨olcskei, Philipp Grohs, Gitta Kutyniok, and Felix Voigtlaender.