

CRUNCH Seminars at Brown, Division of Applied Mathematics

Monday – April 22, 2019

An analysis of training and generalization errors in shallow and deep networks

Hrushikesh Mhaskar

An open problem around deep networks is the apparent absence of over-fitting despite large over-parametrization which allows perfect fitting of the training data. In this paper, we analyze this phenomenon in the case of regression problems when each unit evaluates a trigonometric polynomial. It is well understood that a trigonometric monomial can be synthesized with a good degree of approximation by neural networks with fixed weights and thresholds. Approximation by trigonometric polynomials serves as a “role model” for every other approximation process, including that by neural and RBF networks. In this paper, we argue that the maximum loss functional is necessary to measure the generalization error. We give estimates on exactly how many parameters ensure both zero training error as well as a good generalization error. We prove that a solution of a regularization problem is guaranteed to yield a good training error as well as a good generalization error and estimate how much error to expect at which test data.