



Contribution ID: 28

Type: **not specified**

## **Mikhail Belkin- Neural networks, wide and deep, singular kernels and Bayes optimality**

*Monday, October 3, 2022 4:30 PM (1 hour)*

Wide and deep neural networks are used in many important practical settings.

In this talk I will discuss some aspects of width and depth related to optimization and generalization.

I will first discuss what happens when neural networks become infinitely wide, giving a general result for the transition to linearity (i.e., showing that neural networks become linear functions of parameters) for a broad class of wide neural networks corresponding to directed graphs.

I will then proceed to the question of depth, showing equivalence between infinitely wide and deep fully connected networks trained with gradient descent and Nadaraya-Watson predictors based on certain singular kernels.

Using this connection we show that for certain activation functions these wide and deep networks are (asymptotically) optimal for classification but, interestingly, never for regression.

Based on joint work with Chaoyue Liu, Adit Radhakrishnan, Caroline Uhler and Libin Zhu.