

**ECONOMETRICS SEMINAR ON APRIL 17, 2019**

**DR. SAMET OYMAK,  
THE UCR DEPARTMENT OF  
ELECTRICAL AND COMPUTER ENGINEERING**

**“OVERPARAMETERIZATION WITHOUT OVERFITTING: FROM COMPRESSED  
SENSING TO DEEP LEARNING”**

**ABSTRACT**

Modern machine learning algorithms are often trained in the overparameterized regime where the number of parameters in the model exceeds the size of the training dataset. Despite overparameterization, these algorithms surprisingly avoid overfitting to the training data. In this talk I will present theoretical results demystifying this success by focusing on two class of problems.

In the first class of problems we avoid overfitting by using a model prior, such as sparsity, to narrow down the search space of the algorithm by using a proper regularization. For these problems, we introduce a general framework to quantify the benefit of prior knowledge in terms of problem geometry. This leads to a remarkably accurate characterization of the algorithmic behavior including estimation error, rate of convergence, and sample complexity.

The second class of problems typically arise in deep learning and require no explicit regularization during training. While neural networks have the capacity to overfit any dataset including noise, somewhat paradoxically, they continue to predict well on unseen test data. Towards explaining this phenomena, we show that, neural networks trained by the gradient descent algorithm (1) are provably robust to noise/corruption on a constant fraction of the labels and (2) provably generalize to test data despite overparametrization.