Learning by Overfitting: A Statistical Learning view of Interpolation Learning

Abstract:

The classic view of statistical learning tells us that we should balance model fit with model complexity instead of insisting on training error that’s much lower than what we can expect to generalize to, or even lower than the noise level or Bayes error. And that this balance, and control on model complexity ensures good generalization. But in recent years we have seen that in many situations we can learn and generalize without such a balance, and despite (over?) fitting the training set well below the noise level. This has caused us to rethink the basic principles underlying statistical learning theory. In this talk I will discuss how much of our theory we can salvage and how much of it needs to be revised, focusing on the role of uniform convergence in understanding interpolation learning. Based on joint work with Lijia Zhou, Fred Koehler, and Danica Sutherland.