

On the implicit bias of Deep Learning algorithms

Olivier Bousquet

One of the most intriguing observations of Deep Learning in the last decade is that one can effectively train very large models that generalize well despite their extreme non-convexity and huge number of parameters. While this should largely be credited to significant advances in algorithms, software and hardware, the principles that make this possible remain poorly understood. The theoretical understanding of optimization-based learning algorithms has historically focused on the criterion they optimize, yet it appears that the generalization capabilities of Deep Learning algorithms stems primarily from the implicit regularization that they perform. In this talk, we will present an intuitive explanation of this phenomenon of implicit regularization, expose what is known about it and discuss the various questions that remain to be explored.