GREGORY NAISAT
Department of Statistics
The University of Chicago

Topology of Deep Learning

FRIDAY, June 9, 2017, at 9:30 AM
Jones 304, 5747 S. Ellis Avenue

ABSTRACT

Deep neural networks give rise to state-of-the-art solutions to many hard problems in machine learning. Nevertheless, there has been relatively little understanding of what makes it such a successful scheme. We provide evidence and attempt an explanation for two features that we think are critical in deep neural networks: (i) a nonsmooth activation functions (such as $\max(0,x)$) as opposed to a smooth one such as $\tanh(x)$; (ii) a large number of layers despite the well-known fact that any function can be approximated arbitrarily well by a two-layer neural network. We will argue that deep neural networks operate by changing the topology of the input space—transforming a topologically complicated data set into a topologically simple one. Additionally, adopting Euler characteristic as a proxy for a topological measure of complexity, we express expected changes in the Euler characteristics of the decision regions as those pass through a general layer in a neural network.