Deep Neural Networks, optimal data representation: An Information Theoretic approach

By: Naftali Tishby

http://www.cs.huji.ac.il/~tishby/

Bio

Naftali Tishby is a professor of Computer Science and the director of the Interdisciplinary Center for Neural Computation (ICNC). He is holding Ruth and Stan Flinkman Chair for Brain Research at the Edmond and Lily Safra Center for Brain Science (ELSC) at the Hebrew University of Jerusalem. He is one of the leaders of machine learning research and computational neuroscience in Israel and his numerous ex-students serve at key academic and industrial research positions all over the world. Prof. Tishby was the founding chair of the new computer-engineering program, and a director of the Leibnitz research center in computer science, at the Hebrew university.

He received his PhD in theoretical physics from the Hebrew university in 1985 and was a research staff member at MIT and Bell Labs from 1985 and 1991. Prof. Tishby was also a visiting professor at Princeton NECI, University of Pennsylvania, UCSB, and IBM research.

His current research is at the interface between computer science, statistical physics, and computational neuroscience. He pioneered various applications of statistical physics and information theory in computational learning theory. More recently, he has been working on the foundations of biological information processing and the connections between dynamics and information. He has introduced with his colleagues new theoretical frameworks for optimal adaptation and efficient information representation in biology, such as the Information Bottleneck method and the Minimum Information principle for neural coding.

Abstract

Multilayered Neural Networks and "Deep Learning" have become an incredible success for almost all real world applications of pattern recognition and machine learning. Deep neural networks are classification models that are built on many layers of linear threshold classifiers that formally mimic biological neurons and layered neural nets. They have been extensively analyzed using statistical mechanics techniques, but the theoretical understanding of the reasons they perform so well in practice and their design principles is far from satisfactory.

In this talk I will present a novel analysis of Deep Neural Networks (DNN) based on a theoretical framework for optimal data representation known as the Information Bottleneck

method. In this method we consider optimal data representations as minimal sufficient statistics, namely simplest (possibly stochastic) functions of samples that capture information on the parameters of the distribution. This we achieve through a tradeoff between compression and prediction that looks like free-energy minimization in statistical physics.

We argue that both the structure (number of layers and width of each layer) and the optimal connectivity (weights) of the layers are determined by the cascade of second order phase transitions of the Information Bottleneck tradeoff. I will explain this interesting connection and show how it can yield new design principles for DNN's and new learning algorithms.

This is a joint work with Noga Zaslavsky.