Contribution ID: **10** Type: **not specified**

Decoding the Information in Self-Supervised Learning: Pathway to Optimal Representations

Thursday 9 November 2023 11:00 (45 minutes)

Deep Neural Networks (DNNs) have excelled in many fields, largely due to their proficiency in supervised learning tasks. However, the dependence on vast labeled data becomes a constraint when such data is scarce. Self-Supervised Learning (SSL), a promising approach, harnesses unlabeled data to derive meaningful representations. Yet, how SSL filters irrelevant information without explicit labels remains unclear. In this talk, we aim to unravel the enigma of SSL using the lens of Information Theory, with a spotlight on the Information Bottleneck principle. This principle, while providing a sound understanding of the balance between compressing and preserving relevant features in supervised learning, presents a puzzle when applied to SSL due to the absence of labels during training. We will delve into the concept of 'optimal representation' in SSL, its relationship with data augmentations, optimization methods, and downstream tasks, and how SSL training learns and achieves optimal representations. Our discussion unveils our pioneering discoveries, demonstrating how SSL training naturally leads to the creation of optimal, compact representations that correlate with semantic labels. Remarkably, SSL seems to orchestrate an alignment of learned representations with semantic classes across multiple hierarchical levels, an alignment that intensifies during training and grows more defined deeper into the network. Considering these insights and their implications for class set performance, we conclude our talk by applying our analysis to devise more robust SSL-based information algorithms. These enhancements in transfer learning could lead to more efficient learning systems, particularly in data-scarce environments. Joint work with Yann LeCun, Ido Ben Shaul, and Tomer Galanti.

Presenter: Prof. SHWARTZ-ZIV, Ravid (New York University)

Session Classification: Session 5