Federated Learning and its Applications in Communication

Abstract:

Deep neural networks have recently demonstrated their incredible ability to solve complex tasks. Today's models are trained on millions of examples and are able to reliably annotate images, translate text, understand spoken language or play strategic games such as chess or Go. At the same time the number of intelligent devices on wireless networks (e.g. smartphones, IoT devices) has rapidly grown. These devices are equipped with sensors that allow them to collect and process data at unprecedented scales. However, due to limited resources (e.g., bandwidth and power), latency constraints, and data privacy concerns, centralized training schemes, which require all the data to reside at a central location and were the basis of all above-mentioned successes, are no longer available in the wireless network setting. Thus, these training schemes are increasingly substituted by Federated Learning (FL), which allows multiple parties to jointly train a model on their combined data, without any of the participants having to reveal their local data to other parties, or to a centralized server. This new form of collaborative training concentrates learning in locations where the models are actually used (i.e., on the network edge), and thus minimizes latency and resource consumption.

This talk will discuss recently developments in Federated Learning and will present compression techniques, which can reduce the communication overhead by orders of magnitude, making these new training schemes practical for communications applications. Furthermore, the talk will present a recently proposed extension to FL called "Clustered Federated Learning", which for the first time allows model-agnostic distributed multi-task training under privacy constraints. Finally, the talk will touch upon compression methods for neural networks and will summarize the activities of the MPEG AHG on Compression of Neural Networks for Multimedia Content Description and Analysis.