EPFL-Inria
INTERNATIONAL LAB

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Inria
INVENTEURS DU MONDE NUMÉRIQUE

Workshop EPFL-Inria
January 30 and 31, 2019, Lausanne

Vlad Nitu, EPFL

Title: « Private Machine Learning in the Edge »

Abstract:

During my presentation I will talk about two problems. The first one address the energy efficiency and performance improvements in large scale datacenters. After myriad projects, papers, products and services, we now have giant computers at our fingertips on demand that are fast, easy to use, yet still highly inefficient in their resource utilization and energy efficiency. The average compute node utilization in most cloud offerings is well below 50%. So where has this gone wrong? During my talk I will present my effort that aim to improve this historically low server utilization. Second, I will address the challenges of private machine learning in the edge. Mobile devices collectively provide a huge computing power and their number is increasing rapidly. The capabilities of these devices are also improving. Industrial big players, such as Qualcomm and Movidius, have been increasing the computing power of their mobile devices by designing energy-efficient chips, such as the Snapdragon 845 (Qualcomm). Second, mobile devices contain data that is crucial to machine learning. It is very unlikely that users will keep freely sending their data (e.g., keyboard patterns) to big ML players that make business out of it. An alternative where the users keep their data on their mobile devices, training the machine learning (ML) application locally, is clearly appealing. Thereby, in the second part of my talk I will address the challenges of training the ML models on mobile phones.