Edge Intelligence
Denis Trystram
February 2019

The new paradigm of edge computing is expected to improve the agility of HPC platforms and cloud service deployments by using opportunistic local computing resources. Implementing or analysing new types of data intensive applications close to the place where the data are produced is crucial for designing efficient distributed machine learning methods and as a consequence, reducing the data movements that consume most of the energy.

The challenges are to design new machine learning methods that fully exploit the distributed character of the edge and to develop algorithms and subsequent pieces of software that will allow the deployment of the edge/fog hybrid infrastructures.

The research agenda proposed in this Chair is two-fold. In the first hand, we will study new methods for distributed Machine Learning and data analytics (i). In the second hand, we will develop the models and mechanisms for the orchestration of efficient local resource management (ii). The organization reflects the good balance between academic researchers and several companies, ranging from start-ups to big companies, which will serve as actual case studies.

(i) Deploying Machine Learning (ML) on the edge : Data analytics applications based on BigData and AI technologies can greatly benefit from the Edge/Fog hybrid environments to cover their needs in low latency and scalability. Despite Deep learning, core ML tasks (such as classification, logistic regression, SVM, clustering) will be investigated. A promising direction is to consider creating coresets (sketch of data set that is small compared to the original data set) that do not need to cross the whole network and contains enough information to reconstruct good solutions. Another topic of interest is the on-line learning on data streams.

(ii) Efficient resource management for programming the edge. The motivation of the second subject of this Chair comes from the use of edge/fog computing paradigm, which requires extensions of current programming models and development of new abstractions that will allow developers to design novel applications that can benefit from such massively distributed systems. This also opens other challenges in security, privacy and trust of the distributed resources.

Following the needs of each application, the processing can be done at the hardware components of the edge or it can be spread out on different parts of the network, from local devices and private servers to gateways and mini data centers (at the Fog).
Providing efficient distributed algorithms and all underlying software operations is the only way to sustain the grows of local data production and subsequent analyses.