Learning from examples and domain models in the internet of things

In general, the human brain only needs a few examples to learn concepts. What about machine learning processes?

- *E.g.* when we meet someone, we do not usually pay attention to all the details related to this person. Rather, we rely on only some relevant information;
- Similarly, the human brain adapts easily and quickly to circumstances it has never experienced before, based only on a few examples.

This ability is particularly made possible by prior knowledge that the brain has accumulated, and which allow for building analogies, making simplifications, establishing parallels, etc. In other words, examples are just an element of the puzzle.

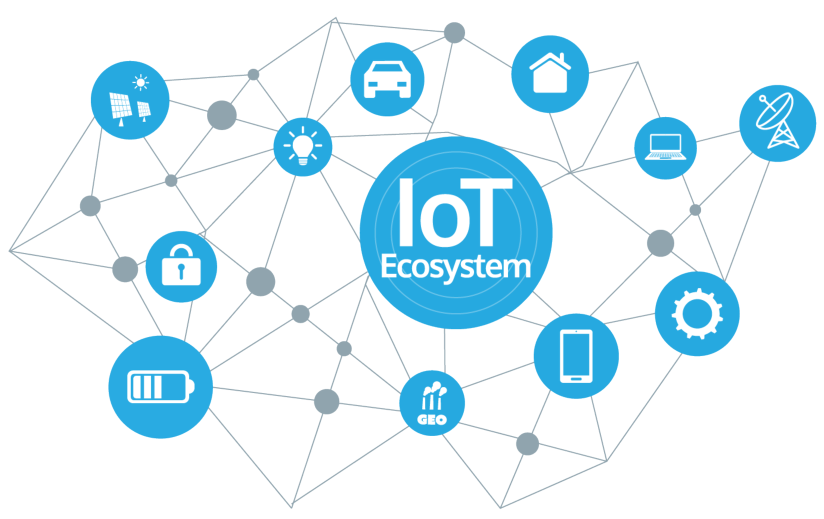
But, is it possible to replicate these capabilities to machine learning processes? Can we make them learn new concepts and adapt to new situations from a few examples?

In the Big Data era . . .

- Data is needed in abundance, *eg.*:
  - Monitoring of civil engineering structures: over **1.2 trillion** observations per year;
  - Autonomous vehicle: **40 terabytes** of data every 8 hours of driving.
- Put simply, data has a predominant and central role in machine learning processes.
- What if we incorporate prior knowledge, the same way as data, in these processes.

. . . and the internet of things

- In the IoT ecosystem, we rather speak about the concept of data source deployments. These are characterized by:
  - **variety of modalities**: temperature, pressure, sound, *etc.*;
  - **diversity of sensors**: accuracy, response time, operating conditions, *etc.*;
  - **multitude of topologies**: sensors placed at different positions of the space, around the concepts of interest;
  - **nature of deployments** that are dynamic (non-fixed) in terms of components and topologies;

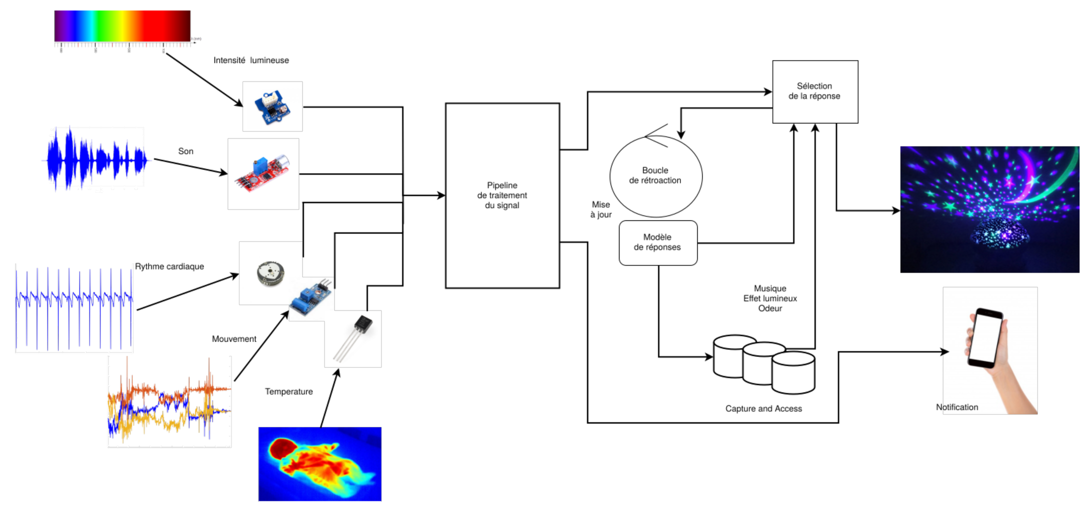


- Often, data sources and the information they generate are flattened and attributed the same status regarding the concept to learn;
- Clearly, data are the end-result of a generation process involving the aforementioned characteristics, each of which having an impact on the concept to learn. So, what if we model these data generation processes and incorporate their subtleties into the learning processes.

learning from examples . . . and domain models

We propose to shape learning processes as multi-layered architectures where we leverage modeling of the domain knowledge, including the data generation processes, in order to serve as a support for the basic learning setting from examples.

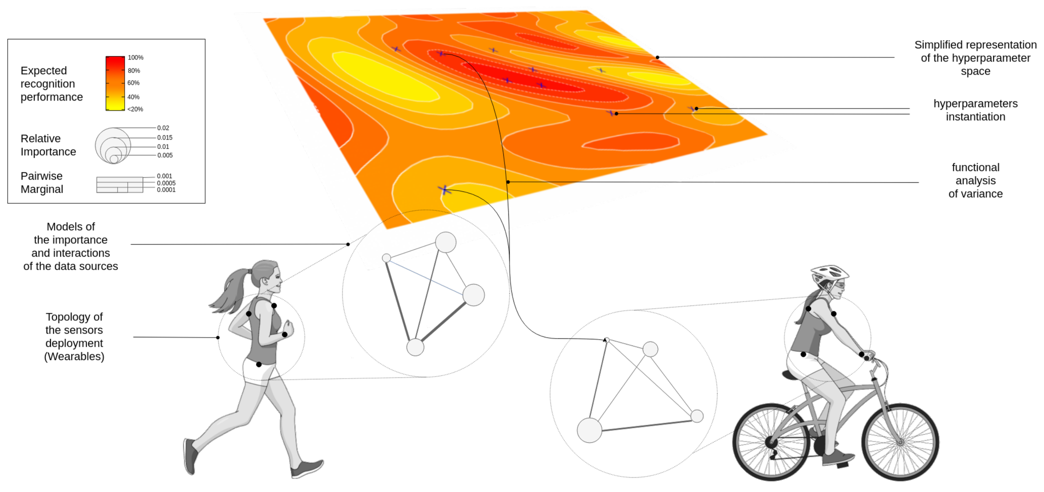
Recognition of infant discomfort [OHC17b, OHC17a]



In this first application, we explore how various domain knowledge can be composed together and serve to sooth infants:

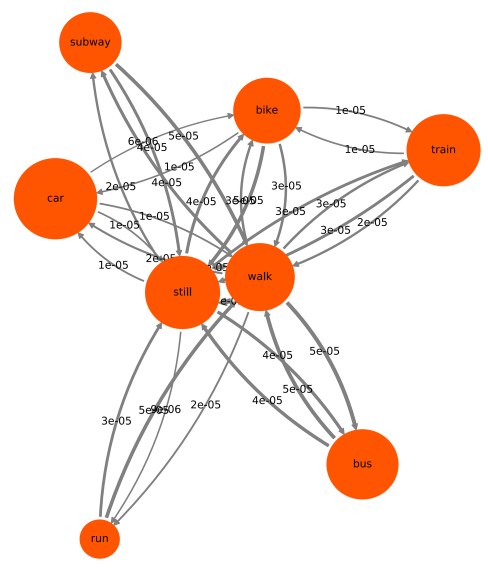
- Model of a domain expertise (pre-cries easily distinguishable by pediatricians);
- Infant cries recognition model (sound);
- Model of biophysiological signals (heart rate, temperature, *etc.*);
- Model of the system's soothing responses.

Human activity recognition [HO19, OH19, HO20]



Results related to the modeling of the topology of the on-body sensors deployment:

- Model of importance of data sources;
- Model of interactions between data sources;
- Model of transitions between activities;



These models serve as a support for learning human activities by directing the sampling from data sources adequately. This has, noticeably, led to a substantial reduction in the amounts of data used to learn.

Monitoring of industrial equipments [OHB19]

Investigation regarding the modeling of vibrations in turbocompressors.

- Model of the natural evolution of the vibration phenomena through time;
- Model of physical re-calibration;
- Model controlling and validating learning examples according to an ISO standard;

Future works

Achieve a high degree of integration between the stages of the learning pipeline through a layer that coordinates the contribution of the different support models.

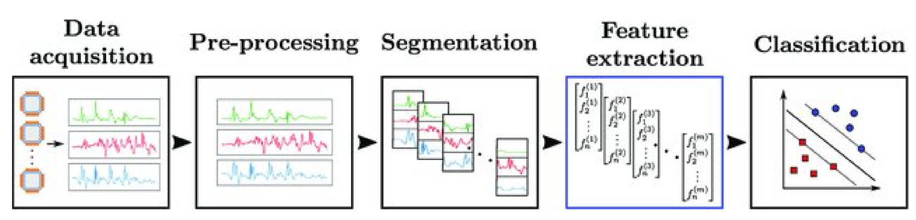


Figure 1: Traditional learning pipeline.

References

[HO19] Massinissa Hamidi and Aomar Osmani. Improving human activity recognition with data sources integration. *Deep learning for human activity recognition, DLforHAR@IJCAI*, 2019.

[HO20] Massinissa Hamidi and Aomar Osmani. Learning human activities from sensor-rich environments. *to appear in Neurocomputing*, 2020.

[OH19] Aomar Osmani and Massinissa Hamidi. Bayesian optimization of neural architectures for human activity recognition. In Nishio N. Roggen D. Inoue S. Pirttikangas S. Van Laerhoven K. Kawaguchi, N., editor, *Human Activity Sensing Corpus and Applications*. Springer Nature, 2019.

[OHB19] Aomar Osmani, Massinissa Hamidi, and Salah Bouhouche. Monitoring of a dynamical system based on autoencoders. In *proceedings of the 28th International Joint Conference on Artificial Intelligence, IJCAI*, 2019.

[OHC17a] Aomar Osmani, Massinissa Hamidi, and Abdelghani Chibani. Machine learning approach for infant cry interpretation. In *Tools with Artificial Intelligence (ICTAI), 2017 IEEE 29th International Conference on*, pages 182–186. IEEE, 2017.

[OHC17b] Aomar Osmani, Massinissa Hamidi, and Abdelghani Chibani. Platform for assessment and monitoring of infant comfort. In *2017 AAAI Fall Symposia, Arlington, Virginia, USA, November 9-11, 2017*, pages 36–44, 2017.

Massinissa Hamidi and Aomar Osmani

hamidi@lipn.univ-paris13.fr

(équipe A3)