TELECOM ParisTech

Une école de l'IMT

CHAIRE MACHINE LEARNING FOR BIG DATA

PROGRAM

International Workshop

Machine Learning & Artificial Intelligence

The stimulating challenges

September 17th & 18th, 2018

BNP PARIBAS La banque d'un monde qui change

criteo

PSA GROUPE

SAFRAN AEROSPACE · DEFENCE · SECURITY

Valeo SMART TECHNOLOGY FOR SMARTER CARS

Aim of the Workshop

The joint availability of computational power and huge datasets has considerably changed the landscape of Artificial Intelligence. In many fields, applications (self-driving cars, cybersecurity, e-health...) that seemed out of reach in the past are now closer to becoming a reality. Recent advances in Machine Learning, the key component of AI, show the growing maturity of algorithms that are now able to handle an increasing number of new tasks. However, simple adversarial attacks can still easily defeat a learning algorithm and the potentially massive deployment of AI tools in various environments raises many new concerns. Additionally to scalability and versatility of algorithms, awareness of drifting or fake data, privacy, interpretability, accountability are now all features that a learning and decision system should take into account. This workshop aims at providing a non-exhaustive overview of recent approaches developed to cope with these stimulating challenges of AI. This first instance of the International Workshop on Machine Learning and AI will serve as a forum for academics and practitioners working on both theoretical and the practical aspects of learning systems for AI.

About us

Technological advances, the ubiquity of sensors and the boom of social networks come with a real data deluge, putting information sciences and technologies at the center of the big data valorisation process. The statistical processing of this huge amount of data brings together applied maths and computer science through a **quickly expanding discipline: Machine Learning**. The volume and variety of available data make traditional statistical methods ineffective. It is the purpose of machine learning to elaborate and study algorithms that enable machines to learn automatically from data and perform tasks in an efficient way.

It is the goal of the "**Machine Learning for Big Data**" Chair to produce methodological research tackling the challenges of the statistical analysis of big data and to liven up the higher education program in that field at Télécom ParisTech. Created in September 2013 with the support of the Fondation Mines-Télécom, the Chair is funded by five companies: **Safran, PSA Group, Criteo, BNP Paribas, and Valeo** who joined the Chair in June 2017. The Chair is supported by the mathematician Stephan Clémençon, Professor in the **Image, Data, Signal** department of Télécom ParisTech.

Monday September 17, 2018

08h30 - *Welcome Coffee & Croissants*

09h00 - Introduction by Yves Poilane, director of Télécom ParisTech

Session 1 *(Amphi Thévenin)*

09h15 - **The Splendors and Miseries of AI: Overview and Challenges**
by Olivier Teytaud, Facebook AI Research

10h15 - **Natural Language Processing for social computing : from opinion mining to human-agent interaction**
by Chloé Clavel, LTCI, Télécom ParisTech

11h00 - *Coffee Break - Poster Session*

11h30 - **Machine Learning: What's next?**
by Patrice Simard, Microsoft Research AI

12h30 - *Lunch Break – Poster Session (Hall Barrault)*

Session 2 *(Amphi Thévenin)*

14h30 - **Invariance and Stability to Deformations of Deep Convolutional Representations**
by Julien Mairal, Inria

15h30 - **What can a statistician expect from GANs?**
by Maxime Sangnier, Sorbonne Université

16h15 - *Coffee Break - Poster Session*

16h45 - **Criteo AI Lab: from applied to fundamental AI**
Jérémie Mary, Criteo

17h45 - *Poster Session*

Tuesday September 18, 2018

08h45 - *Coffee & Croissants*

Session 3 *(Amphi Thévenin)*

09h15 - **Regularization for Optimal Transport and Dynamic Time Warping Distances**
by Marco Cuturi, CREST/ENSAE

10h15 - **Do we have to revisit electric power consumption data analytics in the era of deep learning?**
by George Hebrail, EDF

11h00 - *Coffee Break*

11h30 - **Causal challenges in Artificial Intelligence**
by David Lopez-Paz, Facebook AI Research

12h30 - *Lunch Break*

Session 4 *(Amphi Thévenin)*

14h30 - **Machine Learning Models: Explaining their Decisions, and Validating the Explanations**
by Gregoire Montavon, TU Berlin

15h30 - **Generating Natural Language Explanations for Visual Decisions**
by Lisa Anne Hendricks, UC Berkeley

16h15 - **Panel Session with the Keynote Speakers**
Stephan Clémençon, Florence d'Alché-Buc, LTCI, Télécom ParisTech

17h00 - *Poster Session - Closing Cocktail (E200)*

The Splendors and Miseries of AI: Overview and Challenges

by Olivier Teytaud, Facebook AI Research

We give an overview of AI; we will in particular discuss

- the breakthrough of recent combinations of deep learning & Monte Carlo Tree Search (AlphaZero);
- evolutionary algorithms for AI;
- transfer learning and training with moderate data (after all humans can learn to recognize a platypus from just one single image and play decently at Pong after a few seconds of training...);
- common sense (we can play Angry Birds correctly with almost no training because we understand the semantics);
- reinforcement learning without simulator (i.e. cases in which Monte Carlo Tree Search approaches can not be applied);
- reality gap (if learning on a simulator was enough for learning to drive!);
- structured, complex & huge action spaces;
- privacy, fairness & verification, which are critical issues for widely applying artificial intelligence;
- adversarial examples.

Natural Language Processing for social computing : from opinion mining to human-agent interaction

by Chloé Clavel, LTCI, Télécom ParisTech

The Social Computing topic aims to gather research around computational models for the analysis of social interactions whether for web analysis or social robotics. The peculiarity of this theme is its multidisciplinary approach: computational models are established in close collaboration with research fields such as psychology, sociology, and linguistics. They are based on methods from various fields in signal processing (eg speech signal processing for the recognition of emotions), in machine learning (e.g. structured output learning for the detection of opinions in texts), in computer science (ex: the automatic processing of the natural language for the detection of opinions, the integration of the socio-emotional component in the human-machine interactions). This presentation will describe examples of studies conducted around Social Computing topic.

In particular, we will examine the role of natural language processing in human-agent interaction by presenting our progress on the different research topics we are currently working on, such as the analysis of the likes and dislikes of the user during her interactions with a virtual agent using symbolic methods (Langlet & Clavel, 2016) and machine learning methods (Barriere et al., 2018). Opinion mining methods and their challenges in terms of machine learning will also be tackled (Garcia et al., 2018).

Machine Learning: What's next?

by Patrice Simard, Microsoft Research AI

For many Machine Learning (ML) problems, labeled data is readily available. When this is the case, algorithms and training time are the performance bottleneck. This is the ML researcher's paradise! Vision and Speech are good examples of such problems because they have a stable distribution and additional human labels can be collected each year. Problems that extract their labels from history, such as click prediction, data analytics, and forecasting are also blessed with large numbers of labels. Unfortunately, there are only a few problems for which we can rely on such an endless supply of free labels. They receive a disproportionally large amount of attention from the media.

We are interested in tackling the much larger class of ML problems where labeled data is sparse. For example, consider a dialog system for a specific app to recognize specific commands such as: "lights on first floor off", "increase spacing between 2nd and 3rd paragraph", "make doctor appointment after Hawaii vacation". Anyone who has attempted building such a system has soon discovered that generalizing to new instances from a small custom set of labeled instances is far more difficult than they originally thought. Each domain has its own generalization challenges, data exploration and discovery, custom features, and decomposition structure. Creating labeled data to communicate custom knowledge is inefficient. It also leads to embarrassing errors resulting from over-training on small sets. ML algorithms and processing power are not a bottleneck when labeled data is scarce. The bottleneck is the teacher and the teaching language.

To address this problem, we change our focus from the learning algorithm to teachers. We define "Machine Teaching" as improving the human productivity given a learning algorithm. If ML is the science and engineering of extracting knowledge from data, Machine Teaching is the science and engineering of extracting knowledge from teachers. A similar shift of focus has happened in computer science. While computing is revolutionizing our lives, systems sciences (e.g., programming languages, operating systems, networking) have shifted their foci to human productivity. We expect a similar trend will shift science from Machine Learning to Machine Teaching.

The aim of this talk is to convince the audience that we are asking the right questions. We provide some answers and some spectacular results. The most exciting part, however, is the research opportunities that come with the emergence of a new field.

Invariance and Stability to Deformations of Deep Convolutional Representations

by Julien Mairal, Inria

The success of deep convolutional architectures is often attributed in part to their ability to learn multiscale and invariant representations of natural signals. However, a precise study of these properties and how they affect learning guarantees is still missing. In this work, we consider deep convolutional representations of signals; we study their invariance to translations and to more general groups of transformations, their stability to the action of diffeomorphisms, and their ability to preserve signal information. This analysis is carried by introducing a multilayer kernel based on convolutional kernel networks and by studying the geometry induced by the kernel mapping. We then characterize the corresponding reproducing kernel Hilbert space (RKHS), showing that it contains a large class of convolutional neural networks with homogeneous activation functions. This analysis allows us to separate data representation from learning, and to provide a canonical measure of model complexity, the RKHS norm, which controls both stability and generalization of any learned model. In addition to models in the constructed RKHS, our stability analysis also applies to convolutional networks with generic activations such as rectified linear units, and we discuss its relationship with recent generalization bounds based on spectral norms. This is a joint work with Alberto Bietti.

What can a statistician expect from GANs?

by Maxime Sangnier, Sorbonne Université

Generative Adversarial Networks (GANs) are a class of generative algorithms that have been shown to produce state-of-the art samples, especially in the domain of image creation. The fundamental principle of GANs is to approximate the unknown distribution of a given data set by optimizing an objective function through an adversarial game between a family of generators and a family of discriminators. In this talk, we illustrate some statistical properties of GANs, focusing on the deep connection between the adversarial principle underlying GANs and the Jensen-Shannon divergence, together with some optimality characteristics of the problem. We also analyze the role of the discriminator family and study the large sample properties of the estimated distribution.

Criteo AI Lab : from applied to fundamental AI

by Jérémie Mary, Criteo

This talk will present two recent work done at Criteo AI Lab. The first is about the Film layers in order to modulate deep neural networks trained for dialog systems. The second one is about how to leverage the auction theory and the machine learning from a buyer point of view in order to reduce the impact of personalized reserve prices. Then I'll conclude the talk by a quick overview of the topic of interest of our lab.

Regularization for Optimal Transport and Dynamic Time Warping Distances

by Marco Cuturi, CREST/ENSAE

Machine learning deals with mathematical objects that have structure. Two common structures arising in applications are point clouds / histograms, as well as time series. Early progress in optimization (linear and dynamic programming) have provided powerful families of distances between these structures, namely Wasserstein distances and dynamic time warping scores. Because they rely both on the minimization of a linear functional over a (discrete) space of alignments and a continuous set of couplings respectively, both result in non-differentiable quantities. We show how two distinct smoothing strategies result in quantities that are better behaved and more suitable for machine learning applications, with applications to the computation of Fréchet means.

Do we have to revisit electric power consumption data analytics in the era of deep learning?

by George Hebrail, EDF

Electric power consumption data are more and more available at different aggregation levels (from individual customer to national level) and at different sampling rates (from one measure every second to every year). In this talk, we will present the main available datasets at EDF and the different applications of machine learning methods, both supervised and unsupervised. This includes for instance decomposition of consumption by usage and customer segmentation by clustering their consumption data. The improvement of deep learning methods for these applications will be discussed.

Causal challenges in Artificial Intelligence

by David Lopez-Paz, Facebook AI Research

The route from machine learning to artificial intelligence remains uncharted. The goal of this talk is to investigate how much progress is possible by framing machine learning beyond learning correlations: that is, by uncovering and leveraging causal relations. To this end, we will first identify multiple failure cases in modern machine learning pipelines and try to understand such failures as instances of mistaking correlation by causation. If convinced, we will explore three different ways to reveal causation from data, with some preliminary results. I hope to motivate further research by relating how advances in understanding causation from data would allow machines to ignore confounding effects and spurious correlations, generalize across distributions, leverage structure to reason, design efficient interventions, benefit from compositionality, and build causal models of the world in an unsupervised way.

Machine Learning Models: Explaining their Decisions, and Validating the Explanations

by Gregoire Montavon, TU Berlin

Machine learning models such as deep neural networks have been successful at solving complex tasks in image recognition, text understanding, or physics. There is also a high demand to use these models for assisting humans in taking decisions, e.g. medical diagnosis, or autonomous driving. For this, one needs to be able to trust the learned model, and it is therefore necessary to thoroughly validate it. In particular, we should ensure that its decisions are based on the correct input features.
In this talk, the deep Taylor decomposition framework for explaining decisions in terms of input features will be presented. The framework is applicable to a wide range of neural network architectures, including highly complex ones such as GoogleNet. It works by propagating the model's decision backwards in the network until the input variables are reached. The propagation mechanism at each layer is based on a Taylor expansion principle.
Explanation techniques can be used to validate a trained model. But we also need to validate the explanation technique itself. Ground-truth explanations are usually not available. However one can still test the explanation technique for a number properties considered as desirable. We will show how free parameters of the Taylor expansion allow to induce these desirable properties.

Generating Natural Language Explanations for Visual Decisions

by Lisa Anne Hendricks, UC Berkeley

Clearly explaining a rationale for a classification decision to an end-user can be as important as the decision itself. For example, providing a textual explanation like « This is a cardinal because it is red with a black cheek patch » can help human users better trust and interact with an AI agent. In this talk, I will first present a model to generate explanations for a fine-grained classification task. In particular, I will consider how grounding visual evidence can lead to better explanations, including explanations which accurately discuss visual evidence which is not present in an image (e.g., « This is not a cardinal because it is not a red bird »). Finally, I will outline a new dataset collected specifically for the task of generating explanations and discuss how deictic explanations (pointing to important regions of an image) are complementary to textual explanations.

TELECOM ParisTech

Une école de l'IMT

CHAIRE MACHINE LEARNING FOR BIG DATA

workshopmlai.wp.imt.fr

Réseau EDUROAM
login : telecom809162-4312
mot de passe : fCFx9Qhv

#MLAI

September 2018 - International Workshop Machine Learning & Artificial Intelligence - Conception graphique : Pôle Communication Scientifique - Direction de la recherche - Contact : Samntha Bonnet