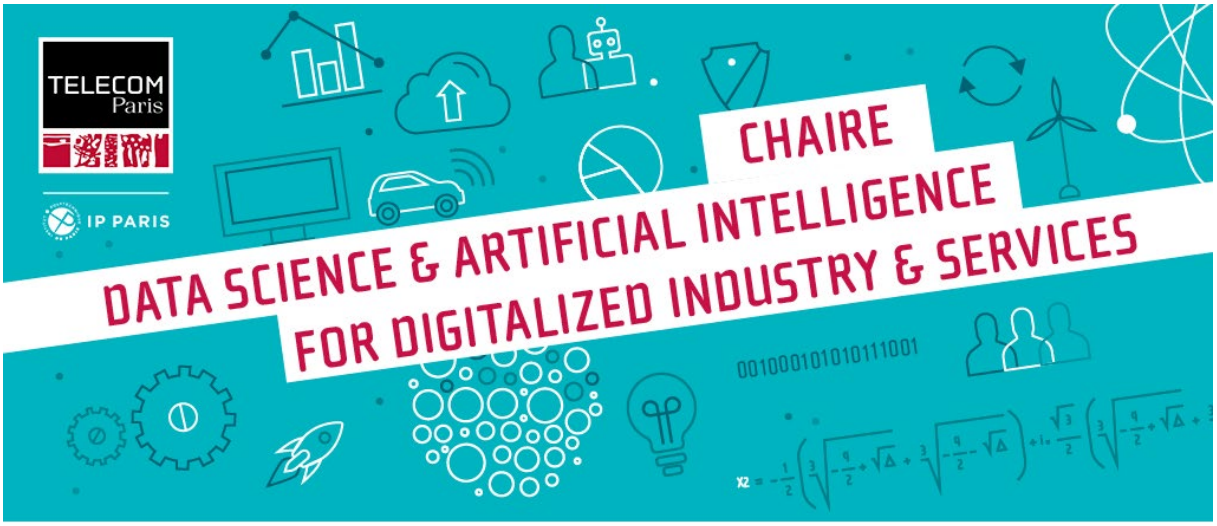

Offre de stage

The best of the two worlds: building deep kernel machines

Poursuite en thèse souhaitable

La Chaire Data Science and Artificial Intelligence for Digitized Industry and Services

Portée par Florence d'Alché-Buc, enseignante-chercheur dans le département Image, Données, Signal de Télécom ParisTech, la chaire DSAI réunit cinq partenaires industriels : Airbus Defence & Space, Engie, Idemia, Safran et Valeo Finance. Son objectif général est de développer, en liaison étroite entre les Parties, une formation et une recherche de niveau international.

Ses quatre principaux axes de recherche sont :

1. Analyse et prévision de séries temporelles (Predictive Analytics on Time Series) ;

2. Exploitation de données hétérogènes, massives et partiellement étiquetées (Exploiting Large Scale and Heterogeneous, Partially Labelled Data) ;

3. Apprentissage pour une prise de décision robuste et fiable (Learning for Trusted and Robust Decision) ;

4. Apprentissage dans un environnement dynamique (Learning through Interactions with a Changing Environment).

Plus d'information sur la chaire : **www.telecom-paris.fr/dsaidis**

Description du stage

Encadrement

Florence d'Alché-Buc

Lieu et dates du stage

Telecom ParisTech, 46 rue Barrault, 75013 Paris

Date de début du stage : Avril/Mai 2019

Équipe d'accueil de la thèse

Département TSI, équipe S2A

Mots clés

Deep learning, kernel methods, vector-valued RKHS, output kernel regression, structured output learning

Sujet détaillé

Deep neural learning has recently achieved impressive progress in many recognition tasks by exploiting millions of training complex data such as images and documents. One of the best property of neural networks is their ability to learn representation of inputs without *prior knowledge*. On the other side, kernel methods are first of all well known for their capacity to deal with structured data but they require to define the kernel. A second property of kernel methods less known deals with the use of the kernel trick in the output space. Using a kernel in the output space allows to tackle in an elegant way complex regression problems such as structured output prediction [1,2,3,4,5], function-valued function learning [6,7,8] and infinite task learning [9]. The common ingredient of all these approaches consists in learning a model with values in a Hilbert space, specifically in a Reproducing Kernel Hilbert space associated to a kernel between outputs.

We recently began to develop these approaches in the deep learning perspective [9,10] and wish to upgrade these tools to larger scale applications and to a wide rande of input data.

The goal of this internship is to investigate a new family of learning task, we call Infinite Classifier Learning that allows to cope with information retrieval tasks as well as multilabel classification in an innovative way. A new family of hybrid deep architectures based on first layers composed of neural layers and a last layer composed of a vector-valued function defined within a vector-valued Reproducing Kernel Hilbert Space will be studied.

The intern will also contribute to an on-going effort on a PyTorch Library about Deep Kernel Machines and will benefit from our first results on the task. He will join a team of PhD students and postdoc who are working on the topic.

This intern will contribute to the efforts of the group to explore this new area. Interest for a PhD thesis on the continuation of the internship is highly wishable.

Profil du candidat

Etudiant d'un master 2 recherche en Apprentissage statistique / IA / Statistiques / reconnaissance des formes,

- Très Bon niveau en apprentissage statistique
- Très Bon niveau en programmation (Python),
- Très Bon niveau d'anglais.

Candidatures

A envoyer à florence.dalche@telecom-paristech.fr

- Curriculum Vitae,

- Lettre de motivation personnalisée expliquant l'intérêt du candidat sur le sujet (directement dans le corps du mail),

- Relevés de notes des années précédentes,

- Contact d'une personne de référence.

Les candidatures incomplètes ne seront pas examinées.

Références

Maruan Al-Shedivat, Andrew Gordon Wilson, Yunus Saatchi, Zhiting Hu, Eric P. Xing:
Learning Scalable Deep Kernels with Recurrent Structure. Journal of Machine Learning Research 18: 82:1-82:37 (2017).

Mikhail Belkin, Siyuan Ma, Soumik Mandal: To Understand Deep Learning We Need to Understand Kernel Learning. ICML 2018: 540-548.

Alberto Bietti, Grégoire Mialon, Dexiong Chen, Julien Mairal:
A Kernel Perspective for Regularizing Deep Neural Networks. ICML 2019: 664-674.

Céline Brouard, Marie Szafranski, Florence d'Alché-Buc:
Input Output Kernel Regression: Supervised and Semi-Supervised Structured Output Prediction with Operator-Valued Kernels. Journal of Machine Learning Research 17: 176:1-176:48 (2016)

Romain Brault, Alex Lambert, Zoltán Szabó, Maxime Sangnier, Florence d'Alché-Buc: Infinite-Task Learning with Vector-Valued RKHSs. CoRR abs/1805.08809 , AISTAT 2019.

Pierre Laforgue, Stéphan Clémençon, Florence d'Alché-Buc: Autoencoding any Data through Kernel Autoencoders. CoRR abs/1805.11028 , AISTAT 2019.

Pierre Laforgue, Alex Lambert, Luc Motte, Florence d'Alché-Buc,On the Dualization of Operator-Valued Kernel Machines, CoRR-abs/1910.04621,2019.

Wenliang Li, Dougal J. Sutherland, Heiko Strathmann, Arthur Gretton:
Learning deep kernels for exponential family densities. ICML 2019: 6737-6746.

Junhong Lin, Lorenzo Rosasco: Generalization properties of doubly stochastic learning algorithms. J. Complexity 47: 42-61 (2018).

Alessandro Rudi, Luigi Carratino, Lorenzo Rosasco: FALKON: An Optimal Large Scale Kernel Method. NIPS 2017: 3891-3901.

Matthias Seeger, Asmus Hetzel, Zhenwen Dai, Eric Meissner, Neil D. Lawrence