

Offre de stage

Sujet : Efficient learning of deep kernel architectures

Poursuite en thèse possible

Encadrement

Florence d'Alché-Buc, florence.dalche@telecom-paristech.fr

Lieu et dates du stage

Telecom ParisTech, 46 rue Barrault, 75013 Paris

Date de début du stage : à partir de février 2017

Équipe(s) d'accueil de la thèse

Equipe Statistique, Signal & Apprentissage du département TSI, département SES.

Mots clés

Deep (structured) learning, Operator-valued kernels, Random Fourier features, Massive data

Sujet détaillé

Deep learning has recently achieved impressive progress in processing millions of complex data such as images and documents. However the mathematical analysis of why deep learning approaches work so well is still missing. Recently new angles to deep learning have been explored through a few new approaches like Deep Gaussian Processes networks or Deep Kernel Learning based on a kernelization of Convnet. This internship aims at going a step further in this direction by studying how deep learning can be instantiated within vector-valued Reproducing Kernel Hilbert Spaces (vv-RKHS). Vv-RKHS provide a flexible and general framework to learn vector-valued functions and address multi-task regression, vector field learning or structured output prediction. Moreover, as in the scalar case, they benefit from a powerful approximation based on Random Fourier Features that allow to cope with 1 million of data. This internship will focus on building new deep kernel architecture based on operator-valued kernels and/or their approximations. The work will be a mix between practise and theory.

Preferably the intern will have a solid background in advanced machine learning, a strong interest for deep learning in a rigorous framework with kernels.

References :

CHAIRE MACHINE LEARNING FOR BIG DATA

C. Brouard, M. Szafranski, F. d'Alché-Buc, Input Output Kernel Regression: Supervised and Semi-Supervised Structured Output Prediction with Operator-Valued Kernels, *JMLR* 17(176):1–48, (2016).

R. Brault, M. Heinonen, F. d'Alché-Buc, Random Fourier Features for Operator-valued kernels, *ACML* 2016.

R. Brault, N. Lim, F. d'Alché-Buc, Scaling up Vector Autoregressive models with operator-valued Random Fourier features, *AALTD'16*, joint workshop to *ECML/PKDD* 2016, to appear in *LNCS*.

[Thang D. Bui](#), [Daniel Hernández-Lobato](#), [Yingzhen Li](#), [José Miguel Hernández-Lobato](#), [Richard E. Turner](#), Deep Gaussian Processes for Regression using Approximate Expectation Propagation, arxiv report (2016).

Y. Cho and Lawrence K. Saul, Kernel Methods for Deep Learning, *Advances in Neural Information Processing Systems* 22 (*NIPS* 2009), pp. 342—350, 2009.

A. Damianou and N. D. Lawrence. (2013), "Deep Gaussian Processes." *AISTATS*, 2013.

Huang, Po-Sen, Haim Avron, Tara N. Sainath, Vikas Sindhvani, and Bhuvana Ramabhadran. "Kernel methods match deep neural networks on timit." In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 205-209. IEEE, 2014.

J. Mairal, P. Koniusz, Z. Harchaoui and C. Schmid. [Convolutional Kernel Networks](#). *Adv. Neural Information Processing Systems (NIPS)*. 2014.

Zhiting Hu, Zichao Yang, Ruslan Salakhutdinov, and Eric Xing, Deep Neural Networks with Massive Learned Knowledge, *Conference on Empirical Methods in Natural Language Processing (EMNLP'16)*.

Néhémy Lim, Florence d'Alché-Buc, Cédric Auliac, George Michailidis: Operator-valued kernel-based-vector autoregressive models for network inference. *Machine Learning* 99(3): 489-513 (2015)

Maxime Sangnier, Olivier Fercoq, Florence d'Alché-Buc, Joint Quantile Regression in vector-valued RKHSs, *NIPS* 2016.

This internship will benefit from the environment of Chaire Machine Learning for Big Data.

La Chaire Machine Learning for Big Data

Le traitement statistique des masses de données convoque à la fois mathématiques appliquées et informatique, à travers une discipline en plein essor : le Machine Learning ou apprentissage statistique.

Créée en septembre 2013 avec le soutien de la Fondation Télécom et financée à hauteur de près de 2 M€ par quatre entreprises partenaires : Criteo, PSA Peugeot Citroën, Safran et BNP Paribas, la Chaire Machine Learning for Big Data est portée par le mathématicien Stéphan Cléménçon, Enseignant-Chercheur, Professeur au sein du Département du Traitement du Signal et des Images à Télécom ParisTech.



Proposant cinq axes de recherche méthodologiques, enrichis par des applications industrielles concrètes, cette Chaire a pour objectif d'animer, en interaction avec ses partenaires, une activité de recherche de pointe en Machine Learning, ainsi que de proposer des programmes de formation.

La variété des données aujourd'hui disponibles (nombres, images, textes, signaux), leur grande dimension et leur volumétrie rendent souvent inopérantes les méthodes statistiques traditionnelles reposant sur le prétraitement humain et un long travail de modélisation. Le Machine Learning vise donc à élaborer et étudier des algorithmes, à vocation prédictive le

plus souvent, permettant à des machines d'apprendre automatiquement à partir des données et à effectuer des tâches de façon performante.

Les avancées technologiques, l'omniprésence des capteurs (systèmes embarqués, objets connectés, Internet...) et l'explosion des réseaux sociaux s'accompagnent d'un véritable déluge de données, propulsant les sciences de l'information au centre du processus de valorisation des masses de données. En plus de la collecte et du stockage, l'enjeu est de pouvoir analyser ces données afin d'optimiser les décisions et mettre au point de nouvelles applications.

Au-delà du buzz médiatique dont il fait l'objet, le Big Data est donc un sujet stratégique majeur, au cœur d'enjeux économiques et sociétaux considérables. Son impact est désormais perçu dans presque tous les secteurs de l'activité humaine : de la recherche scientifique à la médecine en passant, entre autres, par la finance, le bâtiment, l'e-commerce, la défense ou les transports.

En savoir plus sur la Chaire, ses axes de recherche, ses activités, ses publications :

<http://machinelearningforbigdata.telecom-paristech.fr>

Profil du candidat

Etudiant titulaire d'un master 2 recherche

- Apprentissage statistique / reconnaissance des formes
 - Traitement de données structurées
 - Séries temporelles, traitement du langage naturel
- Bon niveau en programmation (Python, Java, C/C++)
- Bon niveau d'anglais

Candidatures

à 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

Romain Brault, Markus Heinonen, Florence d'Alché-Buc, Random Fourier Features for Operator-valued kernels, ACML 2016.

Romain Brault, Néhémy Lim, Florence d'Alché-Buc, Scaling up Vector Autoregressive models with operator-valued Random Fourier features, AALTD'16, joint workshop to ECML/PKDD 2016, to appear in LNCS.

Céline Brouard, Marie Szafranski, Florence d'Alché-Buc, Input Output Kernel Regression: Supervised and Semi-Supervised Structured Output Prediction with Operator-Valued Kernels

CHAIRE MACHINE LEARNING FOR BIG DATA

17(176):1–48, (2016).

Céline Brouard, [Huibin Shen](#), [Kai Dührkop](#), [Florence d'Alché-Buc](#), [Sebastian Böcker](#), [Juho Rousu](#): Fast metabolite identification with Input Output Kernel Regression. [Bioinformatics 32\(12\)](#): 28-36 (2016).

Maxime Sangnier, Olivier Fercoq, Florence d'Alché-Buc, Joint Quantile Regression in vector-valued RKHSs, NIPS 2016.

Néhémy Lim, Florence d'Alché-Buc, Cédric Auliac, George Michailidis: Operator-valued kernel-based-vector autoregressive models for network inference. *Machine Learning* 99(3): 489-513 (2015)