

Le Cam Data Science Colloquium

CEA Saclay, Digitéo Labs, Amphi 33 et 34

14 Juin 2017, 14h-18h

Organisateurs: A. Souloumiac, CEA LIST, E. Moulines, Ecole Polytechnique,
Y. Goude, EDF

Mark Girolami *Professor Imperial College, London, Director Lloyds Register Foundation-Turing Programme on Data Centric Engineering (The Alan Turing Institute)*

What if Euler, Runge, and Kutta had read Bayes and Lebesgue? : Probabilistic Numerical Methods

ABSTRACT : This talk will introduce our contributions to an emerging area of research defining a nexus of applied mathematics, statistical science and computer science, called "probabilistic numerics". The aim is to consider numerical problems from a statistical viewpoint, and as such provide numerical methods for which numerical error can be quantified and controlled in a probabilistic manner. This philosophy will be illustrated on problems ranging from predictive policing via crime modelling to computer vision, where probabilistic numerical methods provide a rich and essential quantification of the uncertainty associated with such models and their computation.

Éric Barat *Senior Expert (CEA LIST)*

An invitation to nonparametric Bayesian modelling through Completely Random Measures

ABSTRACT : Bayesian nonparametrics gained recently in popularity thanks to its flexibility and its ability to build refined hierarchical and nested models. The now celebrated Dirichlet Process (DP) stands as a cornerstone in the field for modelling infinite dimensional prior over probability distributions. Besides DP, other priors issued from Completely Random Measures (CRM) might be considered to cover a wide scope of statistical applications while exhibiting theoretical as well as inferential properties. We review some of these constructions.

Pierre Alquier *Professor (ENSAE)*

Estimation bounds and sharp oracle inequalities of regularized procedures with Lipschitz loss functions

ABSTRACT : We obtain estimation error rates and sharp oracle inequalities for regularization procedures of the form

$$\hat{f} = \operatorname{argmin}_{f \in \mathcal{F}} \sum_{i=1}^N L(f(X_i), Y_i) + \lambda \|f\|$$

when $\|\cdot\|$ is any norm, \mathcal{F} is a convex class of functions and L is a Lipschitz loss function satisfying a Bernstein condition over \mathcal{F} . We explore both the bounded and subgaussian stochastic frameworks for the distribution of the $f(X_i)$'s, with no assumption on the distribution of the Y_i 's. The general results rely on two main objects : a complexity function and a sparsity equation, that depend on the specific setting in hand (loss L and norm $\|\cdot\|$).

As a proof of concept, we obtain minimax rates of convergence in the following problems : 1) matrix completion with any Lipschitz loss function, including the hinge and logistic loss for the so-called 1-bit matrix completion instance of the problem, and quantile losses for the general case, which enables to estimate any quantile on the entries of the matrix ; 2) logistic LASSO and variants such as the logistic SLOPE ; 3) kernel methods, where the loss is the hinge loss, and the regularization function is the RKHS norm.

Sponsors

Supporté par la L'*Initiative Data Science* de l'Ecole Polytechnique et l'institut de convergence.

VENIR À DIGITEO SACLAY

PAR LA ROUTE

Depuis Paris

➔ Par N118 via Pont de Sèvres

Suivre la N118 direction Bordeaux/Nantes jusqu'à sortie 8 (Saclay/Gif-sur-Yvette), au rond-point prendre la 3ème sortie (RD 306) direction Gif-sur-Yvette; à 400m tourner à droite.

CEA SACLAY DIGITEO LABS

Bât. 565

91191 GIF-SUR-YVETTE CEDEX

Tél : +33 (0)1 69 08 25 01

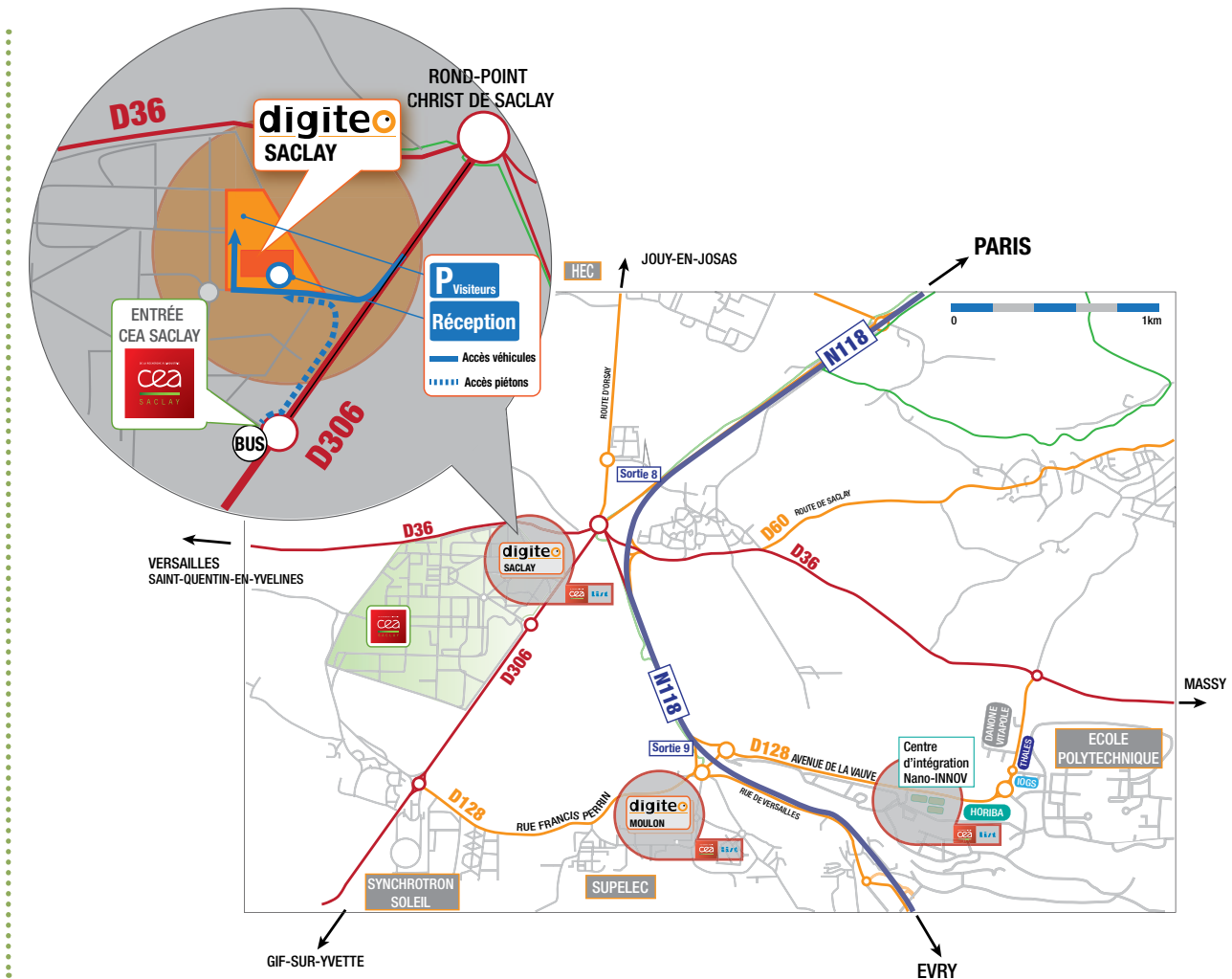
Coordonnées GPS

N 48°43,6422'

E 02°09,4097'

➔ Par A6 via Porte d'Orléans ou Porte d'Italie

Prendre l'A6a, puis E5/E50/Palaiseau/Étampes/Bordeaux/Nantes/Massy/Longjumeau; sortie à gauche direction D444/Versailles/Igny/Bièvres, rejoindre A126 puis continuer sur RD36 jusqu'au rond-point du Christ de Saclay, prendre la 4ème sortie (RD 306) direction Gif-sur-Yvette; à 400m tourner à droite.



EN TRANSPORTS EN COMMUN



➔ Depuis Roissy-CDG ou Paris : RER B (direction Saint-Rémy-les-Chevreuse), arrêt "Le Guichet" puis bus Mobicaps ligne 9 direction Gare de Jouy-en-Josas, arrêt "INSTN - CEA/Saclay RN 306".

➔ Depuis Orly : bus 91.10 (Orly ►► Massy ►► Saclay), arrêt "CEA Porte 306 (Saclay)".

➔ Depuis Massy-Palaiseau (TGV or RER) : bus 91.06 A ou B (Massy ►► Saint-Quentin-en-Yvelines), arrêt "CEA Porte 306 (Saclay)".