

2018-00454 - [Campagne Post-Doctorant 2018/CRI LILLE - Efficient cross-validation for huge data sets (M/F)]

Niveau de diplôme exigé : Thèse ou équivalent

Fonction : Post-Doctorant

A propos du centre ou de la direction fonctionnelle

The Inria Lille - Nord Europe Research Centre was founded in 2008 and employs a staff of 360, including 300 scientists working in sixteen research teams. Recognised for its outstanding contribution to the socio-economic development of the Nord - Pas-de-Calais Region, the Inria Lille - Nord Europe Research Centre undertakes research in the field of computer science in collaboration with a range of academic, institutional and industrial partners.

The strategy of the Centre is to develop an internationally renowned centre of excellence with a significant impact on the City of Lille and its surrounding area. It works to achieve this by pursuing a range of ambitious research projects in such fields of computer science as the intelligence of data and adaptive software systems. Building on the synergies between research and industry, Inria is a major contributor to skills and technology transfer in the field of computer science.

Contexte et atouts du poste

Job environnements

In real-life applications, two challenging problems arise:

(i) Estimators used by practitioners always depend on unknown parameters that have to be chosen, which is called "parameter tuning". For instance in statistical learning, Lasso or SVM algorithms crucially depend on a "regularization parameter", which determines the final performance.

(ii) This tuning step is often performed by use of Cross-Validation (CV) in practice [1].

This is usually time-consuming especially when considering huge datasets, which have become ubiquitous nowadays.

From a practical point of view, a crucial question is to provide more insight in how to enhance CV to drastically reduce the traditional computational burden while achieving a reliable statistical performance. This is a particular instance of the trade-off between statistical performance and computational resources exploitation.

The goals of the post-doc are:

a- deriving a new general strategy to approximate the CV estimators in order to get fast-to-compute quantities,

b- applying this new strategy to various examples in density estimation (Gaussian mixtures, ...), in model-based clustering, variable selection with Lasso and Ridge procedures, as well as in non-parametric regression where reproducing kernels will be used,

c- providing a non-asymptotic quantification of what has been lost by the approximation,

d- identifying guidelines on how to choose the splitting parameter of the CV estimator,

e- deriving model selection results in terms of oracle-type inequalities.

Key words : machine learning

This post-doc position will be held with the **MODAL** Inria projet-team at the Inria center of Lille.

This team is already experienced in machine/statistical learning.

The post-doc will have access to any usual hardware and software resources.

Mission confiée

Assignments

The post-doc student will have to publish the results in journal and conferences papers, give talks in France and abroad. The post-doc student will have the opportunity to attend some conferences

Informations générales

- **Thème/Domaine** : Optimisation, apprentissage et méthodes statistiques Statistiques (Big data) (BAP E)
- **Ville** : Villeneuve d'Ascq
- **Centre Inria** : **CRI Lille - Nord Europe**
- **Date de prise de fonction souhaitée** : 01/11/2018
- **Durée de contrat** : 12 mois
- **Date limite pour postuler** : 29/03/2018

Contacts

- **Equipe Inria** : **MODAL**
- **Recruteur** :
Celisse Alain / alain.celisse@inria.fr

Conditions pour postuler

Instructions to apply:

Candidates will be treated firstly with a complete file: CV + letter of motivation + list of publications + 2 representative publications + one or more letters of recommendation + prospects for professional integration after the post-doc.

Sécurité défense :

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.

Attention : Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

(NIPS, COLT, ICML,...) for instance.

Principales activités

Main activities :

- 1- write papers
- 2- implement new algorithms
- 3- read journal papers and books

Compétences

The post-doc student should have a high level of English (spoken and written).

Avantages sociaux

- Subsidised catering service
- Partially-reimbursed public transport
- Social security
- Paid leave
- Sports facilities
- Flexible working hours

More information about Lille :

<http://www.lille3000.eu/portail/>

<http://www.lillemetropole.fr/mel.html>

Rémunération

Remunerating

The gross monthly salary is 2653€