



Offre n°2024-07548

Post-Doctoral Research Visit F/M Robust and effective stochastic optimization methods for training deep learning models

Le descriptif de l'offre ci-dessous est en Anglais

Type de contrat : CDD

Niveau de diplôme exigé : Thèse ou équivalent

Fonction : Post-Doctorant

Contexte et atouts du poste

The project will be funded by the PRAIRIE 3IA Institute -- ANR-19-P3IA-0001.

Travel expenses are covered within the limits of the scale in force.

Mission confiée

The main objective of this project is the understanding and development of robust and effective stochastic optimization methods for training deep learning models.

State-of-the-art optimization techniques require careful tuning of several hyperparameters, such as learning rate (schedule), momentum, and weight decay. One aspect of this research project is to investigate and develop adaptive techniques for selecting hyperparameters, in order to reduce the tuning effort. For instance, the stochastic Polyak step size and its recent variants have shown promising results for adaptively setting the learning rate.

One challenge in this regard is that theoretically optimal hyperparameter values depend on quantities that are unknown before training. A core part of the project consists of studying online estimation of unknown quantities for hyperparameter selection and exploring its applicability for modern deep learning problem instances.

The project also aims to develop a theoretical framework for robust optimization methods that are resilient to outliers and heavy-tailed noise in gradient distributions. This is particularly relevant for transformer models, which have dominated the field in recent years but often suffer from training instabilities. We aim to investigate the interactions between model architectures and data domains in deep learning, focusing on disentangling the effects of transformer models and their input data distribution on gradient outliers.

The recruited person will be co-advised with Adrien Taylor and Francis Bach.

Principales activités

Main activities :

Conduct theoretical research
Conduct experiments for empirical verification
Write scientific articles
Disseminate the scientific work in appropriate venues.

Compétences

Technical skills and level required :

Languages : High-level of professional/academic English

Coding skills : Good level of coding in Python and related deep learning libraries

Avantages

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking and flexible organization of working hours (after 12 months of

- employment)
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Informations générales

- **Thème/Domaine** : Optimisation, apprentissage et méthodes statistiques Statistiques (Big data) (BAP E)
- **Ville** : Paris
- **Centre Inria** : [Centre Inria de Paris](#)
- **Date de prise de fonction souhaitée** : 2024-09-01
- **Durée de contrat** : 2 ans
- **Date limite pour postuler** : 2024-05-18

Contacts

- **Équipe Inria** : [SIERRA](#)
- **Recruteur** :
Simsekli Umut / umut.simsekli@inria.fr

A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

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.

Consignes pour postuler

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.