2018-01050 - Research intern - PAC-Bayesian theory, neural networks, deep learning, representation learning (M/F)

Contract type : Internship agreement
Level of qualifications required : Master's or equivalent
Function : Internship Research

About the research centre or Inria department

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.

Context

The MODAL team is active in both statistics and machine learning communities. One main research topic is to provide theoretical justifications on learning procedures. To this end, a part of the team (brought together by the researchers Pascal Germain and Benjamin Guedj) is fond of the PAC-Bayesian theory, which allows to express statistical guarantees on the quality of machine learning algorithms. These guarantees are typically stated as probabilistic upper bounds on the generalization loss, known as generalization bounds. There are two appealing characteristics of these bounds: (i) the bounds are computable from the training loss; that is, without relying on a testing set (ii) one can derive efficient algorithms to optimize them.

Assignment

The person recruited will join a recent collaborative ANR-funded project initiated by two MODAL team researchers (Pascal Germain and Benjamin Guedj), called APRIORI. The primary aim is to contribute to the theoretical understanding of representation learning successes—and in particular (deep) neural networks—with a heavy use of the PAC-Bayesian statistical learning theory. Among others, the conducted research will imply deriving generalization bounds for neural networks, empirically validating the bounds accuracy as model selection criteria (ideally they would allow us to get rid of the validation set), and potentially proposing new training methodologies based on the developed theory. Thus, the conducted research will cover both a theoretical and practical aspects.

The person recruited will work with Modal researchers, students and collaborators, and she/he would contribute to the writing of research articles.

For a better knowledge of the proposed research subject: Pascal Germain and Benjamin Guedj have organized a NIPS workshop on PAC-Bayesian learning, called “(Almost) 50 Shades of Bayesian Learning: PAC-Bayesian trends and insights”. The material (slides, videos) may be found here: https://bguedj.github.io/nips2017/50shadesbayesian.html

In particular, the tutorial by François Laviolette and the talk by Dan Roy are relevant to the internship.

Main activities

Scientific research:
1. Contribute to the statistical analysis of deep neural networks and/or other representation learning algorithms.
2. Write code to conduct empirical experiments to assess the accuracy of theory-based learning methodology and model selection criterion.
3. Write scientific reports and provide frequent feedback to the team.

Skills

Technical skills and level required : machine learning, statistics. Proficiency with Python and LaTeX would be appreciated.

Languages: full English proficiency. French is not required.

Benefits package

- Subsidised catering service
- Partially-reimbursed public transport
- Paid leave

Remuneration

Average monthly stipend : 550€

General Information

- Theme/Domain : Optimization, machine learning and statistical methods
- Statistics (Big data) (BAP E)
- Town/City : Villeneuve d'Ascq
- Inria Center : CRI Lille - Nord Europe
- Starting date : 2019-03-01
- Duration of contract : 6 months
- Deadline to apply : 2019-03-03

Contacts

- Inria Team : MODAL
- Recruiter : Guedj Benjamin / benjamin.guedj@inria.fr

About Inria

Inria, the French national research institute for the digital sciences, promotes scientific excellence and technology transfer to maximise its impact. It employs 2,400 people. Its 200 agile project teams, generally with academic partners, involve more than 3,000 scientists in meeting the challenges of computer science and mathematics, often at the interface of other disciplines. Inria works with many companies and has assisted in the creation of over 160 startups. It strives to meet the challenges of the digital transformation of science, society and the economy.

Conditions for application

Applications must include an up-to-date CV, a cover letter, and grade transcripts.

Defence Security :

This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST). Authorization to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

Recruitment Policy : As part of its diversity policy, all Inria positions are accessible to people with disabilities.

Warning : you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.