2018-00808 - Post-doctorant offer : PAC-Bayesian Analysis of artificial neural networks (M/F)

Contract type : Public service fixed-term contract  
Level of qualifications required: PhD or equivalent  
Fonction : Post-Doctoral Research Visit

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 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.

Note that the starting date of this 18-month postdoc can be postponed at the beginning of year 2019 according to the candidate availability.

Assignment

The person recruited will join a recent collaborative project initiated by two MODAL team researchers (Pascal Germain and Benjamin Guedj) with three colleagues of the Laboratoire Hubert Curien at University Jean Monnet of Saint-Etienne (Emilie Morvant, Amaury Habrard, and Rémi Emonet). 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 theoretical and practical aspects.

The person recruited will work with MODAL researchers, students and collaborators, and he/she will participate in the redaction of research articles. Travel expenses for participation to conferences and related scientific events are covered within the limits of the scale in force.

For a better knowledge of the proposed research subject:

These research papers give an overview of the PAC-Bayesian theory as practised by Pascal Germain and its collaborators:

- Pascal Germain, Alexandre Lacoste, François Laviolette, Marie Marchand, Jean-François Roy. Risk bounds for the majority vote: from a PAC-Bayesian analysis to a learning algorithm. JMLR 2015

Various PAC-Bayesian approaches are presented in the material of last NIPS Workshop on the subject:

- (Almost) 50 Shades of Bayesian Learning: PAC-Bayesian trends and insights

Finally, one can refer to these research papers that study neural networks with the help of the PAC-Bayesian theory:

- Behnam Neyshabur, Srinadh Bhojanapalli, Nathan Srebro. A PAC-Bayesian Approach to Spectrally-Normalized Margin Bounds for Neural Networks. ICLR 2018

Main activities

Scientific research:

2. Conduct empirical experiments to assess the accuracy of theory-based learning methodology and model selection criterion.
3. Collaborate with other researchers, PhD students and interns.
4. Write reports and submit research articles.
Skills
- Machine/Statistical learning
- Preparation of scientific documents (knowledge of LaTeX mandatory)
- Practical experience with neural network training (knowledge of pyTorch and/or tensorflow is a strong plus)

Benefits package
- Subsidised catering service
- Partially-reimbursed public transport

Remuneration
2653€ gross per month