2018-00731 - Bayesian deep learning for model selection and approximate inference (Campagne Doctorant Grenoble Rhône-Alpes PhD Campaign)

**Level of qualifications required**: Graduate degree or equivalent

**Function**: PhD Position

**Level of experience**: Recently graduated

**About the research centre or Inria department**

Grenoble Rhône-Alpes Research Center groups together a few less than 800 people in 35 research teams and 9 research support departments.

Staff is localized on 5 campuses in Grenoble and Lyon, in close collaboration with labs, research and higher education institutions in Grenoble and Lyon, but also with the economic players in these areas.

Present in the fields of software, high-performance computing, Internet of things, image and data, but also simulation in oceanography and biology, it participates at the best level of international scientific achievements and collaborations in both Europe and the rest of the world.

**Context**

The topic of this PhD position is on the interface of two of the most important machine learning paradigms: Bayesian inference and deep learning. The main objective is to investigate how these paradigms can mutually benefit each other. On the one hand Bayesian learning provides a theoretically sound framework to formalize the estimation of the architecture and parameters of deep neural network models. On the other hand, deep neural networks offer new tools in Bayesian modeling, e.g., to learn flexible (non-parametric) priors or computationally efficient posterior distribution approximations.

The position is in the **Mistis** research team at Inria Grenoble Rhône-Alpes headed by Florence Forbes. The research will be done in collaboration with the **Thoth** research team at Inria Grenoble Rhône-Alpes. Mistis is a joint team of Inria and Laboratoire Jean Kuntzmann, a joint research unit of Centre National de Recherche Scientifique (CNRS), Institut National Polytechnique de Grenoble (G-INP) and Université Grenoble-Alpes (UGA). The context of our work is the analysis of structured stochastic models with statistical tools. The idea underlying the concept of structure is that stochastic systems that exhibit great complexity can be accounted for by combining simple local assumptions in a coherent way. This provides a key to modelling, computation, inference and interpretation. Our goal is to contribute to statistical modelling by offering theoretical concepts and computational tools to handle properly some of these issues that are frequent in modern data. So doing, we aim at developing innovative techniques for high scientific, societal, economic impact applications and in particular via image processing and spatial data analysis in environment, biology and medicine. For more information see [https://team.inria.fr/mistis/](https://team.inria.fr/mistis/)

Thoth is a joint team of Inria and Laboratoire Jean Kuntzmann, it is motivated by today’s context in which the quantity of digital images and videos available on-line continues to grow at a phenomenal speed. The main objectives of the team are: (i) designing and learning structured models capable of representing this visual information; (ii) learning visual models from minimal supervision or unstructured meta-data; and (iii) large-scale learning and optimisation. For more information see [http://thoth.inrialpes.fr/](http://thoth.inrialpes.fr/)
Both teams are internationally recognized in their expertise fields, statistical modeling and Bayesian statistics for Mistis, deep learning and visual recognition for Thoth. They provide an excellent research environment, organising seminars at Inria, including the biweekly Deep Learning Reading Group (https://project.inria.fr/deeplearning/) and the monthly Bayes in Grenoble Seminar (https://sites.google.com/view/bigseminar/), thus regularly attracting international collaborators. The prospective student will be embedded in a structure with about 40 people working in areas related to the proposed topic. This environment provides access to compute cluster with about 60 GPUs, hosted by the Thoth team, as required to practically evaluate the developed learning models.

Supervision:

- Julyan Arbel (Mistis team), http://www.julyanarbel.com/ (firstname.lastname@inria.fr)
- Jakob Verbeek (Thoth team), http://lear.inrialpes.fr/people/verbeek/

Assignment

Context and objective

The field of machine learning has recently been drastically impacted by deep learning. Deep neural networks are now at the basis of the state-of-the-art in computer vision, speech recognition, natural language processing, and many other areas. While very effective, these models are computationally costly and require large quantities of data to accurately estimate their many parameters. Bayesian statistics offer a theoretically well founded framework to reason about uncertainty, and it is one of the cornerstones of modern machine learning. Although seemingly quite remote from deep neural networks which are theoretically poorly understood, there is a great potential of cross-fertilization between deep learning and Bayesian statistics. Yet, the interaction between these two learning paradigms is relatively underexplored so far. The goal of this thesis is to contribute new theory and practical techniques that lie at the interface of these two techniques.

Approach

We want to explore the Bayesian learning framework to address two problems in learning deep neural networks. The first is the design of energy and memory efficient neural network models that are suitable for deployment in devices such as mobile phones or drones. Current solutions, such as [1], only considered using sparsity inducing priors to suppress redundant network parameters in a given architecture. See also [2] for a posterior concentration approach to sparse deep networks. To make progress, and also learn the overall architecture of convolutional neural networks, we want to explore Bayesian learning of the convolutional neural fabric meta-architecture [3]. The second problem we want to address is the reliance of deep learning on large quantities of training data, which are typically time consuming and expensive to acquire. In the context of visual object recognition, our goal is to learn to recognize a new object class from as few examples as possible [4], in a “small data” context. Within the Bayesian learning framework we can harness the experience gathered from learning to recognize N previous object classes in a posterior distribution over the network parameters, and use this as a prior distribution over the parameters for the new object category.

Probabilistic graphical models, such as Bayesian networks and Markov random fields, provide a framework to represent the conditional independence structure of complex multivariate distributions. One of the most important problems in such models is to characterize the posterior distribution over the latent random variables given observed data. For many models of interest, the posterior is however intractable to compute exactly. Existing approximate inference techniques, such as loopy belief propagation or mean-field inference, are limited in the sense that they are (i) generic and not adapted to a specific graphical model of interest, and (ii) computationally costly due to their iterative nature. The second topic of this thesis is the exploration of deep neural networks as trainable approximate inference models: given observed data the neural network outputs a distribution over the latent variables. The inference networks can be trained using a variational maximum likelihood criterion [5] or from approximate posteriors computed using conventional iterative techniques.

Applications

To validate the practical effectiveness of the developed Bayesian learning techniques we will consider deep convolutional networks for visual recognition tasks. To assess the effectiveness of the proposed approximate inference techniques we will evaluate them for Markov random field models for image
processing and latent aspect models, such as Latent Dirichlet Allocation, for textual data.

**Keywords** *(keywords mentioned with ** pertain to Inria Scientific Strategic Plan):*

** Machine learning
** Deep learning
** Neural networks
** Big data & small data
** Limited energy and storage for small objects

* Bayesian statistics
* Approximate inference
* Model selection
* Non-parametric statistics

**Main activities**

**Skills**

We are looking for outstanding candidates with the following profile:

- Master degree, preferably in Computer Science or Applied Mathematics
- Solid mathematical knowledge, especially (Bayesian) statistics, linear algebra, and optimization
- Solid programming skills; the project involves programming in python and PyTorch, R, or similar deep learning framework
- Highly creative and motivated
- Fluent in English, both written and spoken
- Prior knowledge of / experience in deep learning

**Benefits package**

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

**Remuneration**


Monthly salary after taxes: around 1596,05€ for 1st and 2nd year. 1678,99€ for 3rd year. (medical insurance included).

**General Information**

- **Theme/Domain**: Optimization, machine learning and statistical methods
  Statistics (Big data) (BAP E)
- **Town/city**: Montbonnot
- **Inria Center**: CRI Grenoble - Rhône-Alpes
- **Starting date**: 2018-10-01
- **Duration of contract**: 3 years
Deadline to apply: 2018-05-01

Contacts

- **Inria Team**: MISTIS
- **Recruiter**: Arbel Julyan / julyan.arbel@inria.fr

Conditions for application

**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). Authorisation 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.