MSIAM – 2nd year research internship

Bayesian nonparametric models for hidden Markov random fields on count variables and application to disease mapping

Supervisors: Julyan Arbel, Jean-Baptiste Durand and Florence Forbes
E-mail: [Julyan.Arbel|Jean-Baptiste.Durand|Florence.Forbes]@inria.fr

Location of internship: Inria Mistis and Laboratoire Jean Kuntzmann[1]
(1) Inria, Equipe Projet Mistis and Laboratoire Jean Kuntzmann, Saint Martin d'Hères

Allowance : about 550 euros per month
Duration: 5-6 months (starting in February)

Context:

Hidden Markov random fields (HMRFs) have been widely used in image segmentation and more generally, for clustering of data indexed by graphs. Dependent hidden variables (states) represent the cluster identities and determine their interpretations. Dependencies between state variables are induced by notions of neighbourhood in the graph. A difficult and crucial problem in HMRFs is identification of the number of possible state values K. Recently selection methods based on Dirichlet processes have been developed. These do not assume that K is bounded a priori, thus allowing its adaptive selection with respect to the quantity of available data and avoiding explicit estimation of models with different fixed values for K.
Our team's work has focused on Bayesian nonparametric models for HMRFs for continuous, Gaussian observations. However, extensions have to be developed for other types of applications characterized by discrete observed data (typically issued from counts).

Tasks:

The aim of this internship is to define Bayesian nonparametric models for HMRFs with Poisson observation distributions, to implement their estimation procedures and to apply them to disease mapping in epidemiology. Depending on the obtained results, various extensions may be considered, to take into account possible over- or under-dispersion of data with respect to the Poissonian assumption, or zero-inflated cases.

Prerequisites:

Multivariate statistical analysis, Bayesian statistics, programming.

Related Master programs and tracks: MSIAM Data Science (research)

This work may be continued as a PhD thesis.

References:

[1] Albughdadi, M., Chaari, L., Tourneret, J.-Y., Forbes, F. and Ciuciu, P (2017). A Bayesian non-parametric hidden Markov random model for hemodynamic brain parcellation. *Signal Processing* **135**, 132-146.
[2] Forbes, F., Charras-Garrido, M., Azizi, L., Doyle, S. and Abrial, D (2013). Spatial risk mapping for rare disease with hidden Markov fields and variational EM. *The Annals of Applied Statistics* **7**(2), 1192-1216.
[3] Lü, H., Arbel, J., Forbes, F. (2018). Bayesian Nonparametric Priors for Hidden Markov Random Fields. 50e Journées de la Statistique de la SFdS.
https://toltex.u-ga.fr/users/RCqls/Workshop/jds2018/resumesLongs/subm339.pdf