

Objet / Ingénieur de recherche *Méthodes d'apprentissage statistique rapides, sobres et asymptotiquement efficaces / Octobre 2024 (1 an)*

L'utilisation des données massives (big-data) pour les collectivités et les industries est un enjeu majeur aujourd'hui pour améliorer leurs services et leurs performances dans un monde en transition avec ses nouvelles contraintes environnementales. Notamment, l'empreinte carbone du numérique et de l'intelligence artificielle que nous sommes amenés à utiliser plus largement est un enjeu majeur pour nos organisations. Les progrès technologiques permettent de réduire l'impact CO₂, mais les modèles grossissent également en parallèle. Il faut donc pouvoir proposer des méthodologies d'apprentissage statistique sobres pour atteindre l'objectif de neutralité carbone.

Le projet « Méthodes d'apprentissage statistique rapides, sobres et asymptotiquement efficaces » aborde le concept d'efficacité asymptotique, le développement d'outils rapides et sobres d'aide à la décision basés sur ces recherches et leur transfert vers le monde socio-économique.

Grâce à la notion d'efficacité asymptotique dans la théorie de la *normalité locale asymptotique*, il est possible de définir, pour un problème impliquant beaucoup de données, ce qu'est la « meilleure » procédure d'estimation des paramètres d'un modèle ou la « meilleure » décision statistique (tests d'hypothèses). Pour certaines expériences statistiques, les procédures efficaces ne sont pas connues. Pour d'autres expériences, ces procédures efficaces sont connues mais elles sont généralement complexes et très longues à calculer numériquement pour des données massives ou des données acquises à haute-fréquence. Des alternatives doivent donc être trouvées pour gérer des ensembles de données volumineux ou acquis à haute fréquence disponibles dans les différents domaines d'application. Le projet ANR EFFI 2022-2025 (www.ffi-stats.fr) vise à améliorer les connaissances sur l'efficacité asymptotique de plusieurs expériences statistiques et à fournir de nouveaux estimateurs et de nouvelles procédures de test efficaces et innovants pour des applications réelles.

L'ingénieur de recherche participera aux développements des méthodes d'apprentissage statistique rapides, sobres et efficaces au laboratoire de mathématiques de l'Université du Mans. Notamment,

1. Développement des méthodes de machine learning et deep learning basées sur la méthode one-step dans la continuité des articles [1,2] qui serviront à l'analyse des données du programme de recherche « Efficience et Sobriété Numérique » avec la société EREN Groupe (<https://www.eren-groupe.com/fr>) qui développe une offre de services autour de la décarbonation (production d'énergie renouvelable, solution de stockage d'énergie, production d'hydrogène, ...) au travers de ses différentes sociétés.
2. Développement du package OneStep pour le logiciel R [3,4].
3. Développement des méthodologies rapides et efficaces pour les séries temporelles [5] qui serviront à l'analyse des données énergétiques de Le Mans Métropole.



Laboratoire Manceau de Mathématiques
<http://lmm.univ-lemans.fr>
contact-lmm@univ-lemans.fr

Le Mans Université
Avenue Olivier Messiaen
72085 - LE MANS cedex 9
univ-lemans.fr

Des compétences en statistique et en programmation en R et/ou en tensorflow/PyTorch (ou équivalent) seront demandées.

Laboratoire Manceau de Mathématiques Le Laboratoire Manceau de Mathématiques (LMM) développe à Le Mans Université une recherche en mathématiques théorique et appliquée, principalement à l'assurance et aux énergies. Cette recherche s'articule historiquement autour de deux axes thématiques : probabilités, finance et risques d'une part et statistique des processus et applications d'autre part. Plus récemment, des activités de recherche transverses en actuariat, risque et assurance se sont développées avec la fondation de l'Institut du Risque et de l'Assurance (IRA). Le LMM est membre de la Fédération de recherche CNRS Henri Lebesgue (<https://www.fpl.math.cnrs.fr>) et du LabEx Centre Henri Lebesgue (<https://www.lebesgue.fr/fr>).

Laboratoire d'Innovation Territorial Le Laboratoire d'Innovation Territorial est une démarche collective de réflexion, innovation, expérimentation et diffusion à destination des acteurs publics, économiques et de la recherche initiée par Le Mans Métropole. Il vise à croiser les savoirs scientifiques, expertises opérationnelles et expériences du territoire pour construire des savoirs et expérimenter des solutions en faveur de la neutralité carbone. Dans ce cadre, Le Mans Métropole soutient la recherche à travers son appel à projets Allocations doctorales Neutralité Carbone, et a rejoint avec Le Mans Université le programme national de recherche-action POPSU Transitions pour la période 2023-2026.

Contact : Alexandre.Brouste@univ-lemans.fr

[1] A. Brouste, N. Dugué and W. Meskini *Speeding up the training of neural networks with the one-step procedure*, Neural Processing Letters, 56(178).

[2] A. Brouste and Y. Esstafa. *Asymptotically efficient one-step corrected stochastic gradient descent for statistical estimation*, hal-04122876.

[3] A. Brouste, C. Dutang and D. Noutsu Mieniedou (2021) *OneStep - Le Cam's one-step estimation procedure*, R Journal, 13(1), 366-377.

[4] A. Brouste, Y. Esstafa and C. Malique. *Fast and asymptotically efficient estimation for t and log t*, hal-04612255

[5] S. Ben-Hariz, A. Brouste, Y. Esstafa and M. Soltane (2023) *Fast calibration of weak FARIMA models*, ESAIM PS, 27, 156-173.