

## LABORATOIRE NATIONAL DE METROLOGIE ET D'ESSAIS

### STAGE

Direction de la Métrologie Scientifique et Industrielle  
Département Sciences des données et incertitudes

Réf : STA/INCEBD/DMSI

Lieu : Bassin de St Quentin-en-Yvelines / Trappes (78)

Durée : 6 mois à compter de MARS/AVRIL 2025

### Estimation de la distribution des données pour les modèles de Deep Learning utilisés pour la segmentation d'instances

Le LNE : [www.lne.fr](http://www.lne.fr)

*Leader dans l'univers de la mesure et des références, jouissant d'une forte notoriété en France et à l'international, le LNE soutient l'innovation industrielle et se positionne comme un acteur important pour une économie plus compétitive et une société plus sûre.*

*Au carrefour de la science et de l'industrie depuis sa création en 1901, le LNE offre son expertise à l'ensemble des acteurs économiques impliqués dans la qualité et la sécurité des produits.*

*Pilote de la métrologie française, notre recherche est au cœur de notre mission de service public et constitue un facteur fondamental au soutien de la compétitivité des entreprises.*

*Nous avons à cœur de répondre aux exigences des industriels et du monde académique, pour des mesures toujours plus justes, effectuées dans des conditions de plus en plus extrêmes ou sur des sujets innovants tels que les véhicules autonomes, les nanotechnologies ou la fabrication additive.*

#### Contexte du stage :

Des travaux de recherche sont menés depuis plusieurs années au sein du département Science des Données et Incertitudes sur l'évaluation des incertitudes associées aux prédictions des algorithmes d'apprentissage. Une thèse portant sur deux sources d'incertitude (incertitude prédictive induite par l'incertitude des entrées [1] et incertitude dite de répétabilité correspondant à l'incertitude liée au réentraînement du même réseau) a été soutenue fin 2023. A présent, nos efforts se portent sur l'incertitude du jeu de données d'entraînement et plus spécifiquement l'incertitude d'échantillonnage. Pour se faire, le (ou la) stagiaire s'attachera à estimer précisément la distribution des données d'entraînement par différentes méthodes : VAEs [2], GANs [3], Normalizing Flows [4], .... Cette étape devra lui permettre de simuler de nouveaux échantillons issus de cette distribution afin de (1) tester la robustesse des modèles d'ores-et-déjà disponibles et (2) statuer sur leurs plages de fonctionnement. L'application industrielle attachée à ces développements est la segmentation d'instances de particules à partir de mesures en microscopie électronique à balayage.

#### Missions :

Intégré(e) au sein du département Science des Données et Incertitudes, les développements à réaliser au cours de ce stage s'articulent de la manière suivante :

- Etude bibliographique des méthodes à l'état de l'art pour l'estimation de la distribution des données d'entraînement adaptée aux tâches de segmentation d'instances
- Sélection et implémentation des différentes méthodes
- Simulation de nouveaux échantillons issus des distributions estimées à partir des différentes méthodes
- Evaluation de la robustesse des modèles existants de segmentation d'instances à partir des échantillons simulés
- Rédiger un rapport scientifique synthétisant vos résultats.
- Fournir les codes Python développés

### **Profil :**

Étudiant(e) en M2 ou en dernière année d'école d'ingénieur, spécialisé(e) en mathématiques ou statistiques appliquées. Doté(e) d'une forte curiosité scientifique et d'un goût pour le travail en équipe, vous souhaitez mettre en œuvre et enrichir vos compétences en apprentissage automatique. La maîtrise du langage de programmation Python est essentielle et la connaissance de PyTorch et des librairies couramment utilisées dans le domaine du Deep Learning est souhaitée.

### **Gratification :**

1255 € brut/mois pour une formation Bac +5.

### **Pour candidater :**

**Envoyez votre candidature à : [loic.coquelin@lne.fr](mailto:loic.coquelin@lne.fr) en rappelant en objet du mail la référence de l'offre indiquée en 1<sup>ère</sup> page (STA/INCEBD/DMSI)**

### **Références :**

[1] Monchot, P., Coquelin, L., Petit, S. J., Marmin, S., Le Pennec, E., & Fischer, N. (2023). Input uncertainty propagation through trained neural networks. In International Conference on Machine Learning 2023.

[2] Kingma, D.P., Welling, M., 2014. Auto-encoding variational bayes. International Conference on Learning Representations (ICLR) .

[3] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. Advances in neural information processing systems, 27.

[4] Papamakarios, G., Nalisnick, E., Rezende, D. J., Mohamed, S., & Lakshminarayanan, B. (2021). Normalizing flows for probabilistic modeling and inference. Journal of Machine Learning Research, 22(57), 1-64.