

Stage 2019-2020 Master 2 ou 3ème année d'école d'ingénieur

Dirigé par Mohamed Amine Benadjaoud et Ian Morilla.

Co-dirigé par Fabien Millat.

Service de recherche en radiobiologie et en médecine régénérative (SERAMED) de l'IRSN.

Lieu :	IRS[N] Institut de Radioprotection et de Sûreté Nucléaire BP 17 92262 Fontenay-aux-Roses CEDEX
Unité :	PSE-SANTE/SERAMED/ Laboratoire de Radiobiologie des expositions Médicales (LRMed)
Durée/Rémunération:	6 à 9 mois / 1200 euros par mois
Date de disponibilité :	A partir de Mars/Avril 2020

L'IRSN, Etablissement Public à caractère Industriel et Commercial (EPIC) – dont les missions sont désormais définies par la Loi n° 2015-992 du 17 août 2015 relative à la transition énergétique pour la croissance verte (TECV) – est l'expert public national des risques nucléaires et radiologiques. L'IRSN concourt aux politiques publiques en matière de sûreté nucléaire et de protection de la santé et de l'environnement au regard des rayonnements ionisants. Organisme de recherche et d'expertise, il agit en concertation avec tous les acteurs concernés par ces politiques, tout en veillant à son indépendance de jugement.

Titre : Traitement statistique des données transcriptomiques sur cellules uniques « Single Cell » pour l'étude de la réponse endothéliale aux radiations.

Contexte :

La radiobiologie est l'étude des effets biologiques des rayonnements sur le vivant. Une des missions de l'IRSN est de développer des programmes de recherche afin de comprendre les effets biologiques des rayonnements ionisants (RI) dans le cadre de leur utilisation dans le domaine médical et notamment la radiothérapie.

Dans le But de mieux comprendre et de modéliser la réponse moléculaire des cellules endothéliales (cellules qui tapissent la surface des vaisseaux sanguins) à l'irradiation, le service de recherche en radiobiologie et en médecine régénérative (SERAMED) de l'IRSN a récemment mis en place une technique de séquençage sur cellule unique «single cell RNA seq ». L'intérêt fondamental de cette approche est de connaître l'hétérogénéité de la réponse cellulaire aux RI et d'identifier les signatures moléculaires des sous-populations cellulaires.

Le volume de données en présence est gigantesque (de l'ordre 10^6 données par expérience) nécessitant alors une forte réduction dimensionnelle pour visualiser l'information biologique.

Missions :

Le stage proposé vise à mobiliser un ensemble d'outils de réduction dimensionnelle et d'apprentissage adapté à la très grande dimensionnalité des données en présence. Les différentes techniques d'analyses seront conduites sur une base de données récemment acquise regroupant plusieurs conditions expérimentales. L'étudiant se familiarisera tout d'abord avec les techniques de normalisation et de réduction dimensionnelle classiquement utilisées dans ce contexte : analyse en composantes principales parcimonieuse (ACP), t-distributed stochastic neighbor embedding (t-SNE) et Uniform Manifold Approximation and Projection (UMAP). Par la suite, d'autres méthodes permettant une plus grande prise en compte de la nature des données en présence (sur-dispersion et inflation de zéro) seront explorées, en particulier : l'ACP probabiliste reposant sur des distributions de comptage, des réductions dimensionnelles non linéaires basés sur un apprentissage profond (deep learning) dits « auto-encodeurs » ainsi que des approches d'analyse de données fonctionnelles. Ces différentes méthodes seront testées afin de dégager les points forts et les limites de chacune. En particulier, une analyse de sensibilité sur l'impact des différentes méthodes de normalisation sera conduite.

Intérêt du stage :

Ce stage permettra à l'étudiant d'acquérir de solides connaissances en statistique et en apprentissage profond (deep learning) dans un contexte d'analyse de données à grande dimension (big data) à fort développement méthodologique actuellement.

Il est également prévu que ce travail se poursuive par une thèse qui débutera à l'automne 2020.

Personne à contacter : Benadjaoud Mohamed Amine (mohamedamine.benadjaoud@irsn.fr).