
Stage M2 – Intégration de données cliniques et omiques

Description

L'équipe-projet Inria MODAL est une équipe de recherche spécialisée en apprentissage statistique s'intéressant à des données complexes (hétérogènes, manquantes, fonctionnelles, ...) . Un des domaines d'application concerne les données à haut débit, par exemple issues des expériences en transcriptomique ou protéomique (données omiques). Dans le cadre d'une collaboration avec les unités de recherche METRICS et Infinite, une comparaison d'approches multi-blocs a été réalisée pour tester l'intégration de données omiques et cliniques dans un contexte de construction de score clinique. Les premiers résultats ont été présentés dans le chapitre 2 de la thèse d'H. Sarter¹.

L'objectif de ce stage est de mettre à jour les simulations en rajoutant des approches publiées dans les cinq dernières années tout en se concentrant sur l'influence de la taille d'effet et du nombre de variables par bloc dans les approches multi-blocs. Une réécriture du code R sera également nécessaire pour optimiser le temps de calcul.

Missions

- Optimiser le code R existant
- Réaliser une revue bibliographique pour suggérer d'autres analyses
- Mettre en œuvre les simulations et rédiger un article

Compétences nécessaires

- Programmation R avancée
- Apprentissage statistique dans le contexte de la grande dimension
- Connaissances de base en analyse de données omiques

Informations pratiques

Durée : 6 mois

Localisation : Inria MODAL 40 avenue Halley 59650 Villeneuve d'Ascq et ULR 2694 METRICS – Pôle recherche de la faculté de médecine – 1 place de Verdun, 59000 Lille

Contact

Envoyer CV, lettre de motivation et notes des deux dernières années à Guillemette Marot guillemette.marot@univ-lille.fr et Hélène Sarter helene.sarter@chu-lille.fr

Bibliographie

Volkman A, De Bin R, Sauerbrei W, Boulesteix AL. A plea for taking all available clinical information into account when assessing the predictive value of omics data. BMC Med Res Methodol. 2019;19(1):162.

1 H. Sarter Étude des maladies inflammatoires chroniques de l'intestin à partir d'un registre en population générale: contributions à l'épidémiologie et à l'analyse statistique de données cliniques et omiques (thèse), oct. 2024