

Techniques d'apprentissage automatique de données cliniques pour le traitement personnalisé du sepsis

Stage de formation de niveau Master 2

I Encadrement

Responsables du stage

- Agnès Grimaud, Laboratoire de Mathématiques de Versailles (LMV, Université Paris-Saclay, UVSQ).
- Charles Tillier, Laboratoire de Mathématiques de Versailles (LMV, Université Paris-Saclay, UVSQ).

Dates, localisation, rémunération du stage et contact

- Si possible, le stage débutera début mars pour une durée de 6 mois (35 heures / semaine).
- Il se déroulera au laboratoire LMV de l'Université Paris-Saclay, UVSQ à Versailles (<https://lmv.math.cnrs.fr/>).
- La gratification attribuée au/à la stagiaire est basée sur un taux horaire de 4,05 euros (soit autour de 550 euros/ mois selon la date de début de contrat).
- Pour candidater, envoyer votre CV détaillé et vos relevés de notes de Master à Agnès Grimaud (agnes.grimaud@uvsq.fr) et Charles Tillier (charles.tillier@uvsq.fr).

II Description du stage

Contexte

Ce stage s'inscrit dans le projet ANR *RHU RECORDS* (Rapid Recognition of Corticosteroid Resistant or Sensitive Sepsis)¹ qui vise à identifier et à valider des biomarqueurs de prédiction de la réponse thérapeutique aux corticoïdes dans le cadre du sepsis. Le sepsis se définit comme un état aigu de dysrégulation de la réponse de l'organisme à une infection (souvent bactérienne ou virale) pouvant entraîner la perte de fonction des organes et un risque vital pour le patient. En particulier, un des objectifs du projet est de développer des méthodes de détection rapide des patients sensibles ou résistants aux corticostéroïdes afin de pouvoir mettre en place une meilleure prise en charge des patients ayant un sepsis en situation pandémique ou non. Dans ce contexte, les laboratoires d'informatique de l'UVSQ (laboratoire DAVID) et de mathématiques (laboratoire LMV) sont en charge de la partie analyse de données du projet et disposent pour ce faire de deux bases de données qui contiennent un ensemble de caractéristiques de patients qui ont fait ou non l'objet d'une corticothérapie.

Objectifs du stage

Depuis 2020 - date à laquelle le projet RECORDS a débuté - l'équipe ADAM du laboratoire DAVID a développé une approche basée sur l'intelligence artificielle et l'apprentissage automatique pour évaluer la réactivité à la corticothérapie (CS) chez les patients atteints de sepsis, le plus tôt possible. Dans la continuité de ces travaux, le stage effectué au LMV sera constitué de plusieurs phases, comportant des aspects théoriques et appliqués avec une forte orientation vers l'analyse des données et l'algorithme :

1. Compréhension de la base de données par le stagiaire.
2. Analyse générale de ces données (sélection de variables, étude de corrélation, étude temporelle).
3. Application des modèles d'apprentissage non supervisé afin d'effectuer des tâches de clustering et établir des caractéristiques communes/spécifiques à chacun des sous-groupes (groupe traité par CS, groupe non traité par CS).
4. A la suite de la 3ème phase, faire le lien avec les modèles supervisés déjà établis par le laboratoire DAVID et confronter les différents résultats obtenus.
5. Mettre en place des techniques de type *transfer learning* en utilisant de l'information auxiliaire (temporelle par exemple) pour essayer d'améliorer la modélisation et la performance de prédiction.

Compétences requises

Master en statistiques théoriques et appliquées, maîtrise du langage R et/ou Python et appétence pour les applications relatives à la médecine/santé.

1. Site du projet