LAGEPP

Laboratoire
d'automatique,
de génie des procédés,
et de génie pharmaceutique.

LABORATOIRE COMMUN DE RECHERCHE

HCL HOSPICES CIVILS DE LYON — LA RECHERCHE

UCB Lyon 1 - UMR CNRS 5007

Data scientist / Ingénieur IA (F/H)

Projet HOMEOCARE : jumeau numérique pour la prise en charge personnalisée du sepsis en réanimation

LE PROJET

Le projet HOMEOCARE est porté par le LAGEPP (Laboratoire d'Automatique, de Génie des Procédés et de Génie Pharmaceutique, UMR 5007 CNRS / Université Claude Bernard Lyon 1) et le Laboratoire commun de Recherche (LCR / Hospices Civils de Lyon) qui réunit immunologistes et réanimateurs. Il bénéficie d'un financement de prématuration dans le cadre de France 2030.

Le sepsis touche près de 49 millions de personnes chaque année dans le monde et cause environ 11 millions de décès. En réanimation, la réponse immunitaire des patients est très hétérogène et évolue au fil des jours, ce qui rend les traitements immunomodulateurs difficiles à cibler faute d'outil pour suivre cette dynamique en temps réel.

Les équipes du LAGEPP et des HCL ont montré qu'un panel restreint de quatre biomarqueurs mesurés en routine suffit à distinguer quatre trajectoires immunitaires au pronostic différent. Elles s'appuient sur deux cohortes de patients de réanimation déjà constituées, un algorithme de clustering de trajectoires validées et des modèles physico-biologiques publiés.

HOMEOCARE vise à poser les fondations d'un jumeau numérique du patient en réanimation : situer chaque patient dans sa trajectoire immunitaire à partir de ses biomarqueurs, puis simuler en temps réel l'évolution attendue pour aider le clinicien à personnaliser la prise en charge. Cette phase de prématuration construit les données consolidées, l'affectation des patients aux trajectoires et les premiers modèles d'IA qui serviront ensuite à tester virtuellement des stratégies de traitement.

Durée souhaitée : 12 mois prolongeable selon les besoins du projet, soit toute la durée du projet.

MISSIONS

Le projet sur lequel vous travaillerez vise à développer les briques logicielles et les modèles d'IA du futur jumeau numérique : consolidation des données de deux cohortes hospitalières, affectation probabiliste des patients à des clusters de trajectoires, génération de patients virtuels à partir des modèles physico-biologiques, puis entraînement et validation de modèles d'IA capables de reproduire la dynamique des biomarqueurs et d'estimer le pronostic.

UCB Lyon 1 - ESCPE - bât 308G
43 bd du 11 novembre 1918 - 69622 Villeurbanne cedex - France
Tél. : 33 (0)4 72 43 18 45 - Fax 33 (0)4 72 43 16 82
www.lagepp.univ-lyon1.fr

Lyon1 Université | cnrs

Vous travaillerez au sein du laboratoire de recherche LAGEPP (Université Claude Bernard Lyon 1, campus LyonTech-la Doua à Villeurbanne) et du laboratoire commun de Recherche (Hospices Civils de Lyon – Hôpital Edouard Herriot), sous la responsabilité de la Pr Melaz Tayakout (LAGEPP) et du Dr Samuel Le Goff (LCR-HCL), promoteurs scientifiques du projet. Vous serez en lien régulier avec les immunologistes et les réanimateurs du laboratoire commun des HCL, avec les modélisateurs du LAGEPP et avec la Direction du Système Numérique des HCL pour l'accès aux données et le déploiement de l'outil. Vous serez également encadré par un expert en data science.

En tant que Data scientist / Ingénieur IA, vous :

- Consolidez et nettoyez les données des deux cohortes (biomarqueurs mesurés dans le temps, données cliniques) pour constituer un jeu de données unique, documenté et exploitable pour la modélisation.
- Appliquez l'algorithme de clustering de trajectoires existant sur les données consolidées, le faites évoluer si nécessaire et analysez les clusters obtenus avec les cliniciens.
- Développez une méthode d'affectation probabiliste d'un patient à un cluster à partir de son état courant, et livrez le programme Python correspondant.
- Contribuez à la calibration et aux tests des modèles physico-biologiques développés par les modélisateurs du LAGEPP (estimation des paramètres par cluster, métriques de performance).
- Générez des cohortes de patients virtuels par cluster à partir de ces modèles, avec des trajectoires réalistes et des enveloppes statistiques robustes.
- Concevez et entraînez un modèle séquentiel reproduisant la dynamique des biomarqueurs, rapide à exécuter sur GPU, ainsi qu'un modèle pronostique estimant la probabilité de rétablissement à partir de la trajectoire.
- Validez ces modèles sur des patients non vus à l'entraînement, en performance comme en temps de calcul, et documentez les résultats.
- Structurez le code selon les bonnes pratiques (versionnement, tests, reproductibilité) et préparez avec la Direction du Système Numérique des HCL le transfert de l'outil dans l'environnement hospitalier.
- Participez à la rédaction des livrables, des publications et des présentations du projet.

PROFIL

Vous avez un diplôme de niveau Bac+5 (école d'ingénieurs ou master) en intelligence artificielle, mathématiques appliquées, statistiques ou data science, avec idéalement une

expérience professionnelle dans l'un de ces domaines : modélisation, machine Learning ou data science. Vous avez impérativement de l'expérience validée dans le développement de modèles d'apprentissage automatique en Python sur des données réelles.

Vous avez des connaissances approfondies en apprentissage automatique appliqué aux séries temporelles (modèles séquentiels, clustering de trajectoires, approches probabilistes) et vous pratiquez couramment un framework de deep learning tel que PyTorch, y compris pour l'entraînement sur GPU.

Vous maîtrisez la préparation et le nettoyage de données hétérogènes (pandas, SQL) ainsi que les outils de développement collaboratif (git, tests, environnements reproductibles). Une expérience en modélisation dynamique (équations différentielles, estimation de paramètres, simulation) ou en apprentissage par renforcement serait appréciée, de même qu'une première expérience sur des données de santé.

Vous avez un bon niveau en anglais, suffisant pour lire la littérature scientifique et rédiger des rapports ou des publications.

Vous êtes rigoureux.se et autonome. Vous savez présenter vos résultats à des interlocuteurs non spécialistes, en particulier des médecins et des biologistes, et vous appréciez le travail en équipe pluridisciplinaire.

POUR POSTULER

Nous vous proposons un poste en CDD pour une durée de 12 mois prolongeable selon les besoins du projet, à pourvoir dès que possible. Le salaire sera déterminé en fonction du profil et de l'expérience.

Vous serez basé.e dans les locaux du LAGEPP, Université Claude Bernard Lyon 1, bâtiment CPE, 43 boulevard du 11 novembre 1918, 69622 Villeurbanne cedex et du Laboratoire Commun de de Recherche, Hôpital Edouard Herriot 5 place d'Arsonval, Pavillon P 5ème Etage, 69003 Lyon.

CONTACTS

Melaz Tayakout : melaz.tayakout@univ-lyon1.fr
Samuel Le-Goff : samuel.le-goff@chu-lyon.fr