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

UCB Lyon 1 - UMR CNRS 5007

Data scientist / AI Engineer (M/F)

HOMEOCARE project: a digital twin for personalised sepsis management in intensive care

THE PROJECT

The HOMEOCARE project is led by LAGEPP (Laboratory of Automation, Process and Pharmaceutical Engineering, UMR 5007 CNRS / Université Claude Bernard Lyon 1) and the Joint Research Laboratory (LCR / Hospices Civils de Lyon), which brings together immunologists and intensive care physicians. It is funded by a pre-maturation grant under the France 2030 programme.

Sepsis affects nearly 49 million people worldwide every year and causes around 11 million deaths. In intensive care, patients' immune responses are highly heterogeneous and change day to day, making immunomodulatory treatments difficult to target because no tool can track these dynamics in real time.

The LAGEPP and HCL teams have shown that a small panel of four routinely measured biomarkers is enough to distinguish four immune trajectories with different prognoses. Their work builds on two existing cohorts of intensive care patients, a validated trajectory clustering algorithm, and published physico-biological models.

HOMEOCARE aims to lay the foundations of a digital twin of the intensive care patient: placing each patient on their immune trajectory based on their biomarkers, then simulating the expected course in real time to help clinicians personalise care. This pre-maturation phase builds the consolidated dataset, the assignment of patients to trajectories and the first AI models, which will later be used to test treatment strategies virtually.

Desired duration: 12 months, extendable according to project needs, i.e. for the full duration of the project.

RESPONSIBILITIES

The project you will work on aims to develop the software components and AI models of the future digital twin: consolidating data from two hospital cohorts, probabilistic assignment of patients to trajectory clusters, generating virtual patients from the physico-biological models, and then training and validating AI models able to reproduce biomarker dynamics and estimate prognosis.

You will work within the LAGEPP research laboratory (Université Claude Bernard Lyon 1, LyonTech-la Doua campus, Villeurbanne) and the Joint Research Laboratory (Hospices Civils

de Lyon – Edouard Herriot Hospital), under the supervision of Prof. Melaz Tayakout (LAGEPP) and Dr Samuel Le Goff (LCR-HCL), the project's scientific leads. You will work closely with the immunologists and intensive care physicians of the HCL joint laboratory, with the LAGEPP modelling team, and with the HCL Digital Systems Department for data access and deployment of the tool. You will also be mentored by a data science expert.

As a Data Scientist / AI Engineer, you will:

- Consolidate and clean data from the two cohorts (biomarkers measured over time, clinical data) to build a single, documented dataset ready for modelling.
- Apply the existing trajectory clustering algorithm to the consolidated data, adapt it as needed, and analyse the resulting clusters with clinicians.
- Develop a method for probabilistically assigning a patient to a cluster based on their current state, and deliver the corresponding Python program.
- Contribute to the calibration and testing of the physico-biological models developed by the LAGEPP modellers (parameter estimation per cluster, performance metrics).
- Generate cohorts of virtual patients per cluster from these models, with realistic trajectories and robust statistical envelopes.
- Design and train a sequential model reproducing biomarker dynamics that runs fast on GPU, as well as a prognostic model estimating the probability of recovery from the trajectory.
- Validate these models on patients not seen during training, in terms of both performance and computation time, and document the results.
- Structure the code according to best practices (version control, testing, reproducibility) and prepare, with the HCL Digital Systems Department, the transfer of the tool to the hospital environment.
- Contribute to writing the project's deliverables, publications, and presentations.

PROFILE

You hold a Master's-level degree (engineering school or MSc) in artificial intelligence, applied mathematics, statistics or data science, ideally with professional experience in one of these fields: modelling, machine learning or data science. Proven experience in developing machine learning models in Python on real-world data is essential.

You have in-depth knowledge of machine learning applied to time series (sequential models, trajectory clustering, probabilistic approaches) and are fluent in a deep learning framework such as PyTorch, including GPU training.

You are skilled in preparing and cleaning heterogeneous data (pandas, SQL) and in collaborative development tools (git, testing, reproducible environments). Experience in dynamic modelling (differential equations, parameter estimation, simulation) or reinforcement learning would be a plus, as would prior experience with health data.

You have a good level of English, sufficient to read the scientific literature and write reports or publications.

You are rigorous and autonomous. You can present your results to non-specialists, particularly physicians and biologists, and you enjoy working in a multidisciplinary team.

HOW TO APPLY

We offer a fixed-term contract (CDD) for 12 months, extendable according to project needs, starting as soon as possible. Salary will be determined according to profile and experience.

You will be based at LAGEPP, Université Claude Bernard Lyon 1, CPE building, 43 boulevard du 11 novemebre 1918, 69622 Villeurbanne cedex, and the Joint Research Laboratory, Edouard Herriot Hospital, 5 place d'Arsonval, Pavillon P, 5t^{h} floor, 69003, Lyon.

CONTACTS

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