AI for healthcare
Friday 09 April 2021, 8.45 to 11:00 a.m.

Artificial Intelligence and machine learning have resulted in the emergence of new solutions in healthcare, from automated computer-aided diagnosis to robotic surgery. In healthcare, AI must efficiently process dirty, heterogeneous and noisy data to learn statistical models as well as design and train complex architectures. This is crucial, for instance, to pull out patient traits and develop targeted treatments or to train AI models to forecast the spread of a disease using multi-dimensional time series data. Once the AI solution exists, an additional challenge is to ensure that medical professionals optimally use it.

This second webinar will focus on four illustrations of ongoing research at Hi! Paris related to the development of AI solutions for healthcare. Catalin Fetita (Télécom SudParis) will discuss the influence of neural network architectures and training when designing automated computer-aided diagnosis tools. In particular, he will show increased performances when using an end-to-end rather than a patch-based architecture, while also warning that it is best to avoid a naive implementation of end-to-end architecture. Next, Marc Lavielle (INRIA and Ecole Polytechnique) will describe how stochastic estimation algorithms can be used for automatic model building in mixed effect models. Model building is a time-consuming and effort-intensive process, and Marc's work suggests that it can be markedly sped up by a new algorithm. Then, in the context of assistive robotics, Adriana Tapus (ENSTA Paris) will show the importance of the user profile in the definition of robot behaviors. Focusing on the contexts of elderly people suffering from Mild Cognitive Impairment, and of people presenting sleep disorders, Adriana will show how the careful definition of the user profile can significantly bolster the robot's performance. Finally, Gaël Varoquaux (INRIA) will highlight how progress in machine learning for dirty data enables new data practices and promotes new neural network architectures when dealing with missing values, for instance.

Following their presentations, our researchers will be joined by experts in AI and machine learning to answer questions and spell out challenges related to the use of AI in healthcare. In particular, Johan Aubert (L'Oréal) and Benjamin Farcy (CapGemini) will discuss their view on the healthcare challenges their companies face. Théophile Mohr Durdez, founder of Volta Medical, brought AI right in the operating room to better guide heart surgeons, and will elaborate on how he and his co-founders managed the interaction between doctors and the AI tool. Finally, Professor Harry Sokol, a gastroenterologist and hepatologist at Hôpital Saint-Antoine, and coordinator of the FHU PaCeMM, will discuss the use of machine learning in his

research works on gut microbiota which is one of the most promising areas of medical research for the development of new treatments for many diseases.

8.45-9.55 – FOUR ILLUSTRATIONS OF RESEARCH AT HI! PARIS

8.45-8.55 a.m. — Fresh Hi! Paris news.

8.55-9.10 a.m. — Illustration 1:
Catalin Fetita— Télécom SudParis
Professor & Head of ARMEDIA team

Title: Quantification of lung lesions in fibrosing idiopathic interstitial pneumonia and extension to COVID-19

9.10-9.25 a.m. — Illustration 2:
Marc Lavielle— INRIA and Ecole Polytechnique
Director of Research at INRIA

Title: Automatic model building in mixed effect models. Applications in population pharmacology

9.25-9.40 a.m. — Illustration 3:
Adriana Tapus — ENSTA Paris
Professor & Director of the Doctoral School Institut Polytechnique of Paris

Title: Challenges in Assistive Robotics: Long-term Interaction, Personalization and Trust

9.40-9.55 a.m. — Illustration 4:
Gaël Varoquaux— INRIA
Director of Research at INRIA

Title: AI, electronic records, and health

10:00 – 11:00 – PANEL DISCUSSION

Questions from attendees will be asked to all speakers and several expert panelists:

- **Johan Aubert**
  CTO of L'Oréal Research & Innovation
- **Benjamin Farcy**
  Data Scientist at Capgemini
- **Théophile Mohr Durdez**
  CEO at Volta Medical

- **Harry Sokol**
  Professor, Gastroenterologist and hepatologist at Hôpital Saint-Antoine

ABSTRACTS OF RESEARCH ILLUSTRATIONS

Title: Quantification of lung lesions in fibrosing idiopathic interstitial pneumonia and extension to COVID-19

Presenter: Catalin Fetita (Télécom SudParis)

Fibrosing idiopathic interstitial pneumonia (fIIP) is a subclass of interstitial lung diseases, which leads to fibrosis in a continuous and irreversible process of lung function decay. Patients with fIIP require regular quantitative follow-up with computed tomography (CT) imaging in order to assess the type and amount of injured lung tissue over time. COVID-19 presents several similarities with fIIP and this condition is moreover suspected to evolve to fIIP in 10-30% of severe cases. Note also that the main difference between COVID-19 and fIIP is the presence of peripheral ground glass opacities and less or no amount of fibrosis in the lung. In this context, the development of automated computer-aided diagnosis (CAD) tools for lung lesion detection is today oriented towards machine learning approaches and exploits in particular convolutional neural networks. In this presentation, we investigate the influence of the network architecture and training strategies in the context of a limited available database, with high class imbalance and subjective and partial annotations. We show that increased performances are achieved using an end-to-end architecture versus patch-based, but also that naive implementation in the former case should be avoided. We finally discuss how the models obtained in fIIP lung lesion classification can be transferred to the follow-up of COVID-19 patients.

Title: Automatic model building in mixed effect models. Applications in population pharmacology

Presenter: Marc Lavielle (Ecole Polytechnique & INRIA)

Mixed effects models are a reference tool to describe complex biological phenomena while taking into account the variability between individuals of the same population.

These models are used to describe processes as diverse as the pharmacokinetics (PK) and pharmacodynamics (PD) of drugs, the dynamics of viruses, tumor growth, etc. Construction of a pharmacometric model is a complex process which requires confirmed expertise, advanced statistical methods, the use of sophisticated software tools, but above all time and patience. Indeed, the success of correctly identifying all the components of the model is far from straightforward: it is a question of finding the best structural model, determining the type of relationship between covariates and individual parameters, detecting possible correlations between random effects or also modelling residual errors.

I will show how an extension of the SAEM (Stochastic Approximation EM) algorithm can be used to speed up and optimize this process of model building by identifying at each step how best to improve some of the model components. The principle of this algorithm basically consists in "learning something" about the "best model", even when a "poor model" is used to fit the data. This algorithm is now implemented in Monolix, one of the most popular platforms for PKPD modeling.

Title: Challenges in Assistive Robotics: Long-term Interaction, Personalization and Trust

Presenter: Adriana Tapus (ENSTA Paris)

Recent technological advances enabled modern robots to become part of our daily life. In particular, assistive robotics emerged as an exciting research topic that can provide solutions to improve the quality of life of elderly and vulnerable individuals. All these vulnerable populations need long-term personalized assistive care and, due to serious shortage of health care professionals, socially assistive robots have been shown as viable complementary tools to address this need. We will focus on innovative perception and interaction capabilities and also discuss challenges raised by inter-individual differences and intra-individual variability over time. We will show the importance of the user profile in the definition of robot's behavior in the assistive context, and the relationship between the various parameters that define the user's profile and the task performance. This will be presented for two vulnerable populations: the elderly suffering of Mild Cognitive Impairments (MCI) and people suffering of sleep disorders.

Title: AI, electronic records, and health

Presenter: Gaël Varoquaux (INRIA)

Digital medecine is often associated with high-end biomedical sensor. Health research is typically performed on carefully curated data from closely-followed cohorts. And yet, the most complete source of data on health of populations is the mondane book-keeping of hospitals, recording prescriptions and observations. When the covid-19 pandemics stroke Paris, such data enabled us to understand in real time this emerging disease. It is "dirty", full of missing values, incorrectly-typed entries, and other biases. For these reasons, it is frowned upon as a source for medical evidence and practice, despite being available at no cost. AI, or rather machine learning, can provide powerful statistical models to extract and analysis this wealth of information.

Drawing from some lessons learned analysing electronic health records during the covid crisis, I will discuss how progress in machine learning on dirty data can enable new data practices. In particular, I will show how machine learning models can gracefully deal with missing values; and how they can model categories in the data with typos and morphological variants. The corresponding statistical models will build upon classical statistical practice, to provide well-grounded machine-learning tools, including new neural network architectures.