Post-doctoral fellowship

Leveraging Artificial Neural Networks to predict the cardiac action potential of molecules

- **Supervision**:
  - Pierre Barbillon (Prof. AgroParisTech, MIA Paris-Saclay)
  - Pierre Gloaguen (Associate Prof. AgroParisTech, MIA Paris-Saclay)
  - Gerhard Hessler (Head of Synthetic Molecular Design, Sanofi)
  - Friedemann Schmidt(Head Digital Toxicology, Sanofi)
- **Start**: at your earliest convenience for 12 months.
- **Gross salary**: $\sim$ 2850 euros per month
- **Labs**: MIA Paris-Saclay et Sanofi (Paris et Frankfurt).
- **Application**: Send an email with CV and motivation letter to pierre.barbillon@agroparistech.fr

Keywords: Artificial Neural Networks, Physics Informed Neural Networks, Emulation, Cardiac Action Potential, Molecular Structures.

Overview of the projet: See Figure 1.

Step 1: Prediction of the cardiac action potential as functional time series from the measures of 8 ionic currents. Sanofi has developed an ODE solver to achieve this task. However, the solver is computationally demanding. We then propose to replace it with state-of-the art emulators such as Gaussian Process emulator (GPE) (Williams and Rasmussen, 2006) or Multivariate Adaptive Regression Splines (MARS) (Friedman, 1991). In addition, we aim to investigate a physics-informed neural network (Raissi et al., 2019) which will be constrained by the differential equation system. An important challenge is to investigate how the physical constraints imposed in the loss function of the learning problem can enhance state of the art deep learning models for time series, such as recurrent neural networks (Lipton et al., 2015), or transformers (Wen et al., 2022).

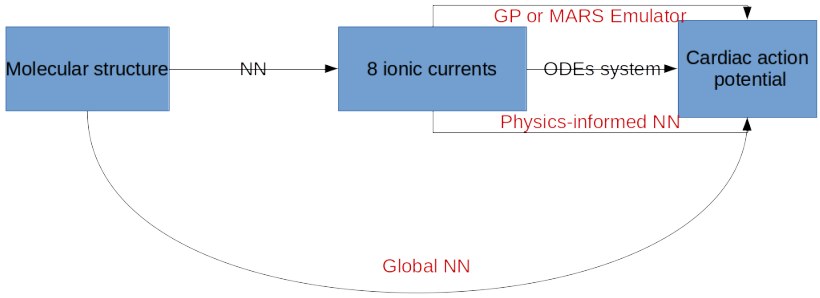

Figure 1: From the molecular structure to the cardiac action potential. Step 1 is concerned with the right hand side of the scheme i.e. predicting from ionic currents the cardiac action potential. Step 2 is concerned with the global scheme i.e. predicting the cardiac action potential from the molecular structure. Elements in black characters are already available at Sanofi, elements in red characters have to be fit during the post-doc.

Step 2: The global predictor from molecular structure to the cardiac action potential could be learned as a global neural network. A more promising solution would be to couple the NN already fitted at Sanofi linking the ionic currents with the molecular structure and the best predictor developed in Step 1.

Presentation of UMR MIA-Paris-Saclay: UMR MIA-Paris-Saclay, AgroParisTech, INRAE, Université Paris Saclay brings together statisticians and computer scientists specialised in statistical and computational modelling and learning for biology, ecology, the environment, agronomy and agri-food. Their expertise covers statistical inference methods (complex models, latent variable models, Bayesian inference, learning, model selection, etc.), and algorithmic methods (generalisation, domain transfer, knowledge representation). The unit develops original statistical and computational methods that are generic or motivated by specific problems in life sciences. Its activities are based on a good knowledge of the disciplines involved: ecology, environment, agri-food, molecular biology and systems biology. The unit is attached to the MATHNUM department of INRAE and the MMIP department of AgroParisTech.

Presentation of Sanofi: Sanofi is a global life sciences company committed to improving access to healthcare and supporting the people we serve throughout the continuum of care. From prevention to treatment, Sanofi transforms scientific innovation into healthcare solutions, in human vaccines, rare diseases, multiple sclerosis, oncology, immunology, infectious diseases, diabetes and cardiovascular solutions and consumer healthcare. More than 110,000 people in over 100 countries at Sanofi are dedicated to make a difference on patients' daily life, wherever they live and enable them to enjoy a healthier life. As a company with a global vision of drug development and a highly regarded corporate culture, Sanofi is recognized as one of the best pharmaceutical companies in the world

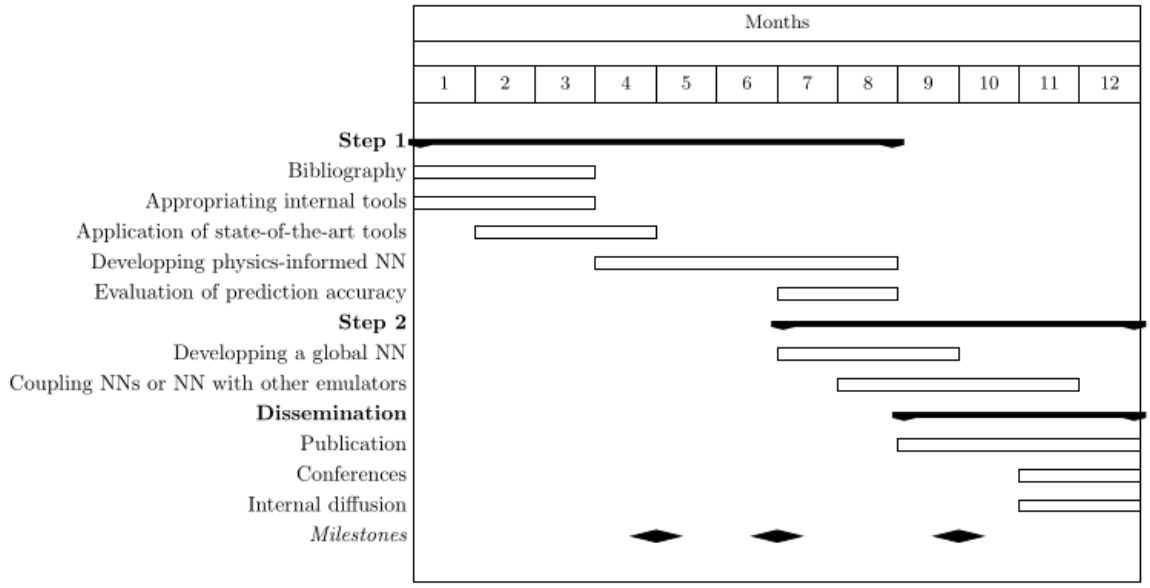

Figure 2: Gantt chart of the planned schedule for the post-doc fellow.

and is pioneering the application of Artificial Intelligence (AI) in the R& D organization including drug discovery, chemical manufacturing and control, translational research, clinical development, and regulatory document management and submission.

Organization: The post-doc fellow is expected to travel between Sanofi in Frankfurt and AgroParisTech in Palaiseau. The post-doc fellow will start in Frankfurt for 6 weeks in order to appropriate the stakes of the project and the internal tools developed by Sanofi. Then, two additional visits of a couple of weeks in Frankfurt are planned: at six months and at the end of the project. Regular interactions with the whole supervision team are already set up as the milestones (see Figure 2). Additional meetings could be organized according to the on-going of the project.

Skills : The post-doctoral fellow should have a Ph.D. in Statistical Learning / Machine Learning. She/He should also have a strong interest in applications especially in the pharmaceutical field. Programming skills in Python are mandatory. The candidate should have a good level of English, allowing for a good understanding of scientific articles, writing of articles and interactions with the full supervision team.

References

Friedman, J. H. (1991). Multivariate adaptive regression splines. *The annals of statistics*, 19(1):1–67.

Lipton, Z. C., Berkowitz, J., and Elkan, C. (2015). A critical review of recurrent neural networks for sequence learning. *arXiv preprint arXiv:1506.00019*.

Raissi, M., Perdikaris, P., and Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707.

Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., and Sun, L. (2022). Transformers in time series: A survey. *arXiv preprint arXiv:2202.07125*.

Williams, C. K. and Rasmussen, C. E. (2006). *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA.