Internship Proposal
Master 2 or engineering degree (Computer Science or Bioinformatics)

Title: Multi-objective optimization and deep learning for healthcare applications
Keywords: Deep learning, Multi-objective optimization, Transformers, EHR.
Supervisors: Farida Zehraoui, AROB@S Team, IBISC laboratory, Paris-Saclay University, Univ. Evry
Eric Angel, AROB@S Team, IBISC laboratory, Paris-Saclay University, Univ. Evry
Massinissa Hamidi, AROB@S Team, IBISC laboratory, Paris-Saclay University, Univ. Evry

Contact: farida.zehraoui@univ-evry.fr
Duration of internship: 6 months
Location: IBISC Laboratory, IBGBI, University of Evry, 23 Boulevard de France, 91000 Evry.

Description:
Stochastic gradient descent methods are widely used in the field of machine learning and more particularly, to optimize the parameters of deep neural networks (deep learning).
Most gradient descent methods aim to optimize a single function representing a single objective or a linear combination of multiple objectives. Recently, variants of gradient descent for multiobjective optimization have been proposed in the literature [1]. These methods have been used for multi-task learning [2].
We developed, in the AROB@S team, a multi-objective approach that aims to optimize the architecture of a neural network by removing neurons and connections between neurons during the learning phase to improve efficiency and model interpretability. The goal of this internship is to adapt this approach to multi-task learning using real data from electronic health records of patients admitted to intensive care units (MIMIC-IV) [4]. Hypernetworks [5] or multi-objective gradient descent [2] will be used for multi-objective optimization.

Steps of the project:

- Study of the multi-objective approach proposed in our team as well as state-of-the-art deep learning architectures based on attention mechanisms like *Transformers* [3].
- Adapt and implement the proposed approach to process time series and generate a Pareto front.
- Apply the implementation to real-world time series data of patients while considering multiple tasks, including early predictions of sepsis.

Bibliography

[1] S. Liu, L.N. Vicente, The stochastic multi-gradient algorithm for multi-objective optimization and its application to supervised machine learning, arXiv:1907.04472, 2021.
[2] Sener, O., & Koltun, V. Multi-task learning as multi-objective optimization. *NeurIPS,* 2018.
[3] Rasmy, L., Xiang, Y., Xie, Z., Tao, C., & Zhi, D. (2021). Med-BERT: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. *NPJ digital medicine*, *4*(1), 1-13.
[4] Johnson, A.E.W., Bulgarelli, L., Shen, L. et al. MIMIC-IV, a freely accessible electronic health record dataset. Sci Data 10, 1 (2023). https://doi.org/10.1038/s41597-022-01899-x
[5] Navon, Aviv & Shamsian, Aviv & Chechik, Gal & Fetaya, Ethan. (2020). Learning the Pareto Front with Hypernetworks.