

# Teaching Guidance in Interactive Machine Learning systems

**Domain:** Human-Computer Interaction  
**Keywords:** Intelligent Systems, Machine Learning, Human-Computer Partnerships  
**Labs:** ISIR, Sorbonne Université, Paris, France  
**Supervisor:** Baptiste Caramiaux, CNRS researcher at Sorbonne Université ([baptiste.caramiaux@sorbonne-universite.fr](mailto:baptiste.caramiaux@sorbonne-universite.fr))  
**Co-Advisor:** Téo Sanchez, PhD Student at Inria ([teo.sanchez@inria.fr](mailto:teo.sanchez@inria.fr))

## Context

Machine Learning (ML) has become ubiquitous in modern interactive technology because of the wide range of complex tasks it can handle without requiring explicit programming but providing data instead. However, most ML systems offer users little agency in the way models are trained from data: the training and testing are mostly performed offline by ML practitioners and **remain opaque to end users**, especially with deep and non-interpretable model architectures.

The growing field of Interactive Machine Learning (IML) (Fails & Olsen, 2013) field tries to understand how the development of ML can be more interactive and accessible to end users. This field focuses on building human-centered interaction to allow end-users to personalize the training and evaluate the model (Gillies et al., 2016). With novice users, we found that the way participants organize training examples both in term of order and complexity greatly affect the performances and their understanding of the system (Sanchez et al.). As a guideline for design, we suggested the use of guidance for building the training **curriculum** (Cakmak & Thomaz, 2014; Wall et al., 2019) i.e. a strategy to organize the training examples in a meaningful way, that can both improve the performances and increase the users understanding of the system.

As an example of guidance, we investigated a technique called **Active Learning** (Settles, 2009) that estimates the most informative element of unlabeled data to be given to the model. This technique relies on the uncertainty estimation of Neural Networks (Gal et al., 2017). In theory, the Active Learning framework is expected to improve the performance but is unevenly applicable depending on the interaction scenarios.

## Objectives

The goal of this internship is to explore and compare guidance for **teaching interactively** Machine Learning algorithms. We want to investigate how guidance strategy affects user understanding and model performance in a chosen interaction scenario.

The intern can choose a relevant interaction scenario that can cover various:

- Tasks, involving image (or sound) classification or generation;
- Input modalities: image, sound, or movement data.

The scenario will be decided collegially with the supervisors. Particular care will put on the fact that teaching guidance has a critical role in the chosen scenario.

## Specific Activities

During this internship, the student will be expected to:

- Study the different types of guidance from the literature (algorithmic, user-centered or other) and reflect on their properties and application scope.
- Choose a relevant Interactive Machine Learning scenario on which to apply one or several guidance.
- Design and implement a prototype application. For instance, the prototype can be developed with a framework Marcelle (<https://marcelle.netlify.app/>) dedicated to Interactive Machine Learning application and developed by Jules Françoise (LIMSI, Université Paris Saclay), Baptiste Caramiaux and Téo Sanchez.
- Evaluate the prototype and understand how different guidance affect both the performance and the user behavior and understanding.

The internship will last from **four to six months**, starting March 1<sup>st</sup>, 2021.

## Expected Results

We anticipate that this work will lead to a submission to a top-tier research conference, such as ACM CHI (deadline mid-September). It could also serve as the foundation for a Ph.D. thesis.

## Required Skills

We are looking for motivated students who are enthusiastic about Human-Computer Interaction (HCI) research and particularly about Interactive Machine Learning. Candidates may have skills in Machine Learning (ML) and interests in HCI or skills in HCI and interests in ML

Programming skills are required, preferably Javascript (or/and optionally Python).

## References

- (Cakmak & Thomaz, 2014) Cakmak, M., & Thomaz, A. L. (2014). Eliciting good teaching from humans for machine learners. *Artificial Intelligence*, 217, 198–215. <https://doi.org/10.1016/j.artint.2014.08.005>
- (Fails & Olsen, 2013) Fails, J. A., & Olsen, D. R. (2003). Interactive machine learning. *Proceedings of the 8th International Conference on Intelligent User Interfaces - IUI '03*, 39. <https://doi.org/10.1145/604045.604056>
- (Gal et al., 2017) Gal, Y., Islam, R., & Ghahramani, Z. (2017). Deep Bayesian Active Learning with Image Data. *ArXiv Preprint ArXiv:1703.02910*. Retrieved from <http://arxiv.org/abs/1703.02910>
- (Gillies et al., 2016) Gillies, M., Fiebrink, R., Tanaka, A., Garcia, J., Bevilacqua, F., Heloir, A., ... Caramiaux, B. (2016). Human-centered machine learning. *Conference on Human Factors in Computing Systems - Proceedings, 07-12-May-2016*, 3558–3565. <https://doi.org/10.1145/2851581.2856492>
- (Sanchez et al.) Sanchez, T., Françoise, J., Bevilacqua, F., Mackay, W., and Caramiaux, B. (under review). How do People Train a Machine? Strategies, Beliefs and (Mis)Understandings. *Submitted to ACM CSCW*
- (Settles, 2009) Settles, B. (2009). Active Learning Literature Survey - TR 1648. *Sciences-New York*. <https://doi.org/10.1016/j.matlet.2010.11.072>
- (Wall et al., 2019) Wall, E., Ghorashi, S., & Ramos, G. (2019). Using Expert Patterns in Assisted Interactive Machine Learning: A Study in Machine Teaching. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11748 LNCS, 578–599. [https://doi.org/10.1007/978-3-030-29387-1\\_34](https://doi.org/10.1007/978-3-030-29387-1_34)