Fully decentralized learning under fairness and heterogeneity constraints

Machine Learning is becoming ubiquitous in our everyday lives. It is now used in digital assistants, for medical diagnosis, for autonomous vehicles, etc. Its success can be explained by the good performances of learned models, sometimes reaching human-level capabilities. However, simply being accurate is not sufficient if these models are to be largely deployed. Hence, the notion of trustworthiness has to be considered as soon as human are involved in the loop. Among all the existing trustworthiness notions, fairness is especially important. For example, a model used for medical diagnosis should not be biased against sub-groups of the population. Similarly, a model designed to predict whether someone should receive a loan should give the same opportunity of receiving one to every creditworthy person regardless of their gender, ethnicity, or any other sensitive attribute.

Fairness has been extensively studied in the centralized case (Caton and Haas, 2020), that is when all the data is available in a single place. However, it has received far less attention in the decentralized setting (Lian et al., 2017; Kairouz et al., 2021), that is when the data is owned by multiple entities that would like to collaborate to learn efficient models but do not wish to share their data. In this context, data can be heterogeneous across agents and fairness can be defined at two different levels. On the one hand, locally at the data owners level and on the other hand, globally at the level of all data. Because the data can be non identically distributed across agents, these two fairness notions can differ much, making the fairness to be satisfied at the global level difficult. In this context, some questions that the intern will attempt to answer are as follows. Is it possible to design a globally fair algorithm in the fully decentralized setting? Can the proposed algorithm reduce the gap between local and global fairness? A possible way to answer the first question is to propose an extension of the FairGrad algorithm (Maheshwari and Perrot, 2022) to the decentralized setting. Regarding the latter question, an interesting starting point could be to investigate the use of the communication topology proposed in Le Bars et al. (2022).

Objectives: The goal of this 6 months internship is to study fairness at the data level in the context of Decentralized Learning. The main objectives are (i) to review some of the existing literature in the field, (ii) to design new algorithms to learn fair models in a fully decentralized and heterogeneous context and (iii) to derive theoretical guarantees on the fairness and utility levels of the obtained models.

Requirements: Successful candidates should have a solid background in Machine Learning with an interest in Fairness and Decentralized Optimization. Some knowledge in Statistical Learning Theory would be a plus but is not mandatory. A good understanding of the Python programming language is expected. Finally, proficiency in English is required as it will be one of the main working language.

Keywords: Machine Learning, Fairness, Decentralized Learning, Learning Theory, Optimization

Contact: The recruited student will be based in the INRIA Lille - Nord Europe research center and will be supervised by Michaël Perrot and Batiste Le Bars. The interested students should send an e-mail (in english) with a CV and a Motivation Letter to *michael.perrot@inria.fr* and *batiste.le-bars@inria.fr*.

References

Caton, S. and Haas, C. (2020). Fairness in machine learning: A survey. *arXiv preprint arXiv:2010.04053*.

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R., et al. (2021). Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning*, 14(1–2):1–210.

Le Bars, B., Bellet, A., Tommasi, M., Lavoie, E., and Kermarrec, A. (2022). Refined convergence and topology learning for decentralized optimization with heterogeneous data.

Lian, X., Zhang, C., Zhang, H., Hsieh, C.-J., Zhang, W., and Liu, J. (2017). Can decentralized algorithms outperform centralized algorithms? a case study for decentralized parallel stochastic gradient descent. *Advances in Neural Information Processing Systems*, 30.

Maheshwari, G. and Perrot, M. (2022). Fairgrad: Fairness aware gradient descent. *arXiv e-prints*, pages arXiv–2206.