Benchmarking and improving random search for hyperparameter optimization

Supervision:

- Thomas Moreau, MIND, Inria (thomas.moreau@inria.fr)
- Gaël Varoquaux, SODA, Inria (gael.varoquaux@inria.fr)

Context. Hyperparameter optimization is a key component of machine learning pipelines. It consists in finding the hyperparameters that gives the best generalization error for a given model and dataset. This is a challenging task as the number of hyperparameters to optimize can be large and the number of possible values for each hyperparameter can be huge. It is also a critical part for benchmarks, to show the robustness of a method to hyperparameter choice [2, 3]. In practice, the most common approach is to use a random search [1] to explore the parameter space. Random search consists in randomly sampling hyperparameters and evaluating the performance of the model, keeping the ones that give the lowest generalization error. This approach has been shown to be more efficient than classical grid search in many cases. However, this approach can be expensive as it requires to train many models. Moreover, to account for the randomness of the search, it is common to repeat the search multiple times and evaluate the variance of the results. This process makes benchmarking hyperparameter optimization even more expensive, and energy consuming. An idea to reduce the cost of repeated random search is to reuse points from previous searches to construct the next one. But this process can introduce bias in the search, and underestimate the variance of the results. In this internship, we will investigate how to make this process sound and robust, to reduce the cost of benchmarking hyperparameter optimization.

Methods. XXX.

Environment. The internship will take place at Inria Saclay (or at NeuroSpin), between the MIND team and the SODA team for maximally 6 months (April-September 2024). The MIND team is a large team focused on mathematical methods for statistical modeling of brain function using neuroimaging data (fMRI, MEG, EEG) as well as on computational imaging methods to accelerate MRI scans. Particular topics of interest include machine learning techniques, numerical and parallel optimization, human neuroimaging at ultra-high magnetic fields, applications to cognitive and clinical neuroscience, and scientific software development. The SODA team ...

Requirements. We seek candidates who are strongly motivated by challenging research topics in machine learning, with a taste for topics which have high societal impact. Applicants should have a strong mathematical background with knowledge in statistical analyis and machine learning. Basic knowledge in estimators' consistency would be a plus. In regard to software engineering, proficiency in `Python` is expected and preliminary experience in a deep-learning library is a plus.

References

[1] James Bergstra and Yoshua Bengio. Random Search for Hyper-Parameter Optimization. *Journal of Machine Learning Research (JMLR)*, 13(1):281–305, 2012.

[2] Léo Grinsztajn, Edouard Oyallon, and Gaël Varoquaux. Why do tree-based models still outperform deep learning on tabular data?, July 2022.

[3] Thomas Moreau, Mathurin Massias, Alexandre Gramfort, Pierre Ablin, Pierre-Antoine Bannier, Benjamin Charlier, Mathieu Dagréou, Tom Dupré la Tour, Ghislain Durif, Cassio F. Dantas, Quentin Klopfenstein, Johan Larsson, En Lai, Tanguy Lefort, Benoit Malézieux, Badr Moufad, Binh T. Nguyen, Alain Rakotomamonjy, Zaccharie Ramzi, Joseph Salmon, and Samuel Vaiter. Benchopt: Reproducible, efficient and collaborative optimization benchmarks. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 36, New-Orlean, LA, USA, November 2022. Curran Associates, Inc.