Tool Comparison: Optuna vs Hyperopt

Audrey Bertin, Nasko Apostolov, Nelson Evbarunegbe, & Raymond Li

Introduction

In the field of machine learning, **hyperparameter tuning** is the process of selecting an optimal set of **hyperparameters** for a learning algorithm.

Hyperparameters are user-selected parameters applied to learning algorithms that affect how they are implemented–for example, constraints, weights, learning rates, number of algorithm layers/complexity, etc.

Depending on which set of hyperparameters is selected, the results and performance of a particular machine learning algorithm can vary dramatically. The possible decision space, however, is enormous. In many cases, there are essentially an infinite set of possible combinations for all of the tuning parameters in a model. It is *impossible* to try every single one of them and find the absolute best option–instead, users must select a *good* option.

Typically, in hyperparameter tuning, users will pre-define a list of combinations to try. For example, if they must provide two parameters, β and λ, they might give the following potential values for these two parameters and ask the algorithm to try every possible combination within:

- $\beta = \{3, 5, 7, 9\}$
- $\lambda = \{0.001, 0.01, 0.1, 1.0\}$

This technique is known as a **grid search**.

It can sometimes produce decent results, but has several downsides.

1. The user must themselves define which lists of values to try. Coming up with this list introduces a high amount of subjectivity.
2. It is very computationally intensive (and slow) to try every single combination, one at a time. As the number of parameters increases, the number of combinations grows exponentially.
3. There is no easy way to compare results without a lot of additional coding on top of constructing the algorithm.

This is where hyperparameter tuning tools come into play. Several tools have been developed in recent years to help simplify this process and produce better results. Some use parallel computing to test a parameters simultaneously and speed up the training process, Many have UI implementations that can help make it easier for users to compare performance across models. Additionally, modern tuning software often provides advanced techniques to help the user select the best hyperparameters, such as intelligently moving the tuning in a direction that appears to be associated with increased performance scores scores.

Optuna (released in 2020) and **Hyperopt** (released in 2011) are two industry standard parameter optimization tools designed to integrate with Python.

Optuna

Optuna was released into the hyperparameter tuning market at a time when there were already quite a few established and trusted tools available (including the second tool we will discuss–Hyperopt).

Its creators chose to introduce Optuna after they discovered limitations in many other existing tuning algorithms, including that they:

1. Had instability in some environments
2. Were not utilizing the newest technology/advancements in hyperparameter optimization
3. Did not provide a way to specify which hyperparameters should be tuned directly within the Python code (instead, requiring the user to write special code just for the optimizer)

The creators of Optuna attempted to fill those gaps, advertising the following features:

1. **Define-By-Run style API** – An API style that is beginning to become industry standard in deep learning, but had not yet been applied to hyperparameter tuning prior to Optuna. Compared to an older system Define-And-Run. Using Define-By-Run allows users to write more modular code and access more complex hyperparameter spaces.
2. **Use of learning curves to prune trials** – Optuna uses deep learning and gradient boosting algorithms to predict the end result of a training trial before it is over. This can help it quit unpromising trials before they complete, improving efficiency.
3. **Parallel distributed optimization** – Optuna supports distributed optimization, simultaneously running multiple trials. This can dramatically speed up the tuning process and allow the user to increase the scale of how many parameters they can try.
4. **Visualization dashboard** - Optuna provides a UI dashboard where users can watch the optimization process and easily compare results from optimization experiments.

The tool uses several advanced algorithms for tuning that go beyond standard grid search. These include:

- Tree-Structured Parzen Estimator (TPE) , a form of Bayesian optimization:

- Gaussian Processes (GP)
- Covariance Matrix Adaptation (CMA)
- Asynchronous Successive Halving Algorithm (ASHA)

Optuna is framework agnostic, meaning it can easily be integrated with any of Python's machine learning/deep learning frameworks: Scikit-Learn, PyTorch, Tensorflow, etc.

The code for Optuna is publicly available on GitHub:

https://github.com/optuna/optuna

Optuna also has a custom website with tutorials, examples, and links to documentation:

https://optuna.org

Hyperopt

In this project, we compare these two tools. We start with a code review of both tools, and then conduct an experimental comparison where we train and tune a machine learning model using each.

Code Review: Optuna

Code Review: Hyperopt

Experimental Comparison

Conclusion