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. 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