08/06/2020

Novel Approaches For Tuning Models' Hyper Parameters

Description:

One of the most importance step in machine learning, after gathering data, is finding the hyper parameters. Whatever models which need parameters will surely perform badly if these latter are not found via the tuning process. Moreover, as we are heading to bigger models with deep learning trends, finding those might be very time consuming. It emphasizes the importance of finding quickly parameters which might be among the optimal space.

Some of famous existing methods such that grid search or random search [1] are commonly used for the hyper parameter optimization (tuning). While the first method explores all the possible combination of parameters, the number of experiments grows exponentially. Random search uses randomness to sample a set of parameters among the parameter space and should find a descent combination in a reasonable amount of time. However, sometimes a model can take few hours/days to realize a full training and thus, the tuning would need a large amount of time to be accomplished.

A new kind of optimization methods has been proposed since a couple of years: Bayesian optimization and population-based optimization. The idea is rather to try different parameters via grid search or random search, select a set of parameters in an adaptive manner. These methods aim to identify good configurations more quickly.

In this project, we would like to take few machine learning problems and try different tuning methods using the library [Ray Tune](). We would like to empirically quantify and qualify the speed for each method as well as the efficiency of the parameters found.

Proposed plan:

1. Choose few machine learning problems deep-learning.
2. Implement a benchmark tool to measure various optimization algorithms.
3. Analyze the results and determine if bandit-based, Bayesian optimization, or any other approaches perform better in average than random search.

Prerequisites: Knowledge about Machine Learning.

Supervisor: Diego Antognini (diego.antognini@epfl.ch)

References:

[1] http://www.jmlr.org/papers/volume13/bergstra12a/bergstra12a.pdf