

Deep Learning Lab Assignment 4

Settings

Optimize the hyperparameters of a CNN using the following algorithms:

Bayesian optimization
Hyperband
BOHB

Hyperparameters to optimize and their ranges:

learning rate	$[10^{-6}, 10^{-1}]$
batch size	$[32, 512]$
layer 1 filter numbers	$[2^4, 2^{10}]$
layer 2 filter numbers	$[2^4, 2^{10}]$
layer 3 filter numbers	$[2^4, 2^{10}]$

Noted that the validation error of the CNN is approximated by a surrogate to speed up the entire process.

Results

The validation error of the best configurations found by each algorithm are shown as follows. We can see that BOHB outperformed the other two.

Bayesian optimization	0.176
Hyperband	0.187
BOHB	0.154