HP Tuning

Star

$\alpha = 10$ Train-ratio
$m = 10$ explorers

10 à N
↳ best 10 N

10 à 10N
HP from best
↳

Timesteps

$10 \cdot N + 10 \cdot 10^1 N + 10 \cdot 10^2 N + 1 \, 10^3 N$
10   100   1000   1000 N

Overhead $\frac{2110}{1000} \approx 2,1 \Rightarrow 110\%$

$t_{ove} = m \cdot N + m \alpha N + m \alpha^2 N + \dots + m \alpha^{T-1} N$

$t_{base} = 1 \cdot \alpha^T N$

Overhead $\frac{t_{ove}}{t_{base}} = \frac{m N \sum_{i=0}^{T-1} \alpha^i}{\alpha^T N} = m \cdot \sum_{i=0}^{T-1} \alpha^{i-T} \approx \alpha^{-1}$ für $\alpha \gg 1$

Does Pruning already realize these gains?
As unsuccessful trials are early aborted
(thinking: Better overfall + fast long-train)

When does Pruning kick in? $m \gg$ Pruning min parameter
seems ≤ 5

Stepwise

\+
⇒ Schrittweite muss nicht getestet werden
⇒ kontinuierliches Lernen

\- nicht auf 1 Schlag reproduzierbar
\+ modelle speichern/laden

\+ HP können sich nach Fortschritt verändern/anpassen
z.B. Learning rate

m=4 Anzahl Äste
N_e
N_t

$\frac{N_t}{N_e} = \alpha$

Star with Optuna:

- Study optimize
- Train long

Loop

evtl. nicht notwendig, da nächste Study den Train long beinhaltet
Letzte Study ist dann einfach nur 1mal trainieren