Bandits and Hyper-parameter Optimization

Xuedong Shang[1], Emilie Kaufmann[2], Michal Valko[3]
[1] Inria Lille, SequeL team
[2] CNRS, CRIStAL
[3] Google DeepMind

Huawei Paris

Paris, Oct. 2019

What — Hyper-parameter optimization

Problem

We tackle **hyper-parameter tuning** for *supervised learning* tasks:

What — Hyper-parameter optimization

Problem

We tackle **hyper-parameter tuning** for *supervised learning* tasks:

↳ global optimisation task: $\min\{f(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in \Omega\}$;

What — Hyper-parameter optimization

Problem

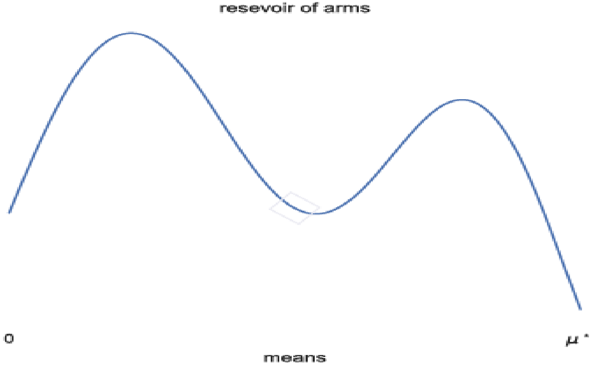
We tackle **hyper-parameter tuning** for *supervised learning* tasks:

↳ global optimisation task: $\min\{f(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in \Omega\}$;

↳ $f(\boldsymbol{\lambda}) \triangleq \mathbb{E}\left[\ell\left(\mathbf{Y}, \widehat{g}_{\boldsymbol{\lambda}}^{(n)}(\mathbf{X})\right)\right]$ measures the generalization power.

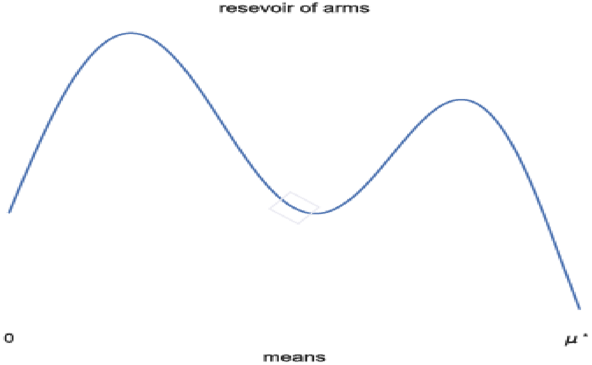
How — Best-arm identification

We see the problem as *best-arm identification* in a *stochastic infinitely-armed bandit*: arms' means are drawn from some *reservoir distribution* ν_0.



How — Best-arm identification

We see the problem as *best-arm identification* in a *stochastic infinitely-armed bandit*: arms' means are drawn from some *reservoir distribution* ν_0.



How — Best-arm identification

We see the problem as *best-arm identification* in a *stochastic infinitely-armed bandit*: arms' means are drawn from some *reservoir distribution* ν_0.

In each round

1. (optional) query a new arm from ν_0;
2. sample an arm that was previously queried.

goal: output an arm with mean close to $\mu^\star$

`D-TTTS` (Dynamic Top-Two Thompson Sampling) $\rightsquigarrow$ a dynamic algorithm built on `TTTS`.

How — D-TTTS

In this talk...

- `Beta-Bernoulli` Bayesian bandit model
- a uniform prior over the mean of new arms

Initialization: $\mu_1 \sim \nu_0$; $\mathcal{A} = \{\mu_1\}$; $m = 1$; $S_1, N_1 = 0$
while budget still available **do**
    $\forall i \in \mathcal{A}$, $\theta_i \sim \texttt{Beta}(S_i + 1, N_i - S_i + 1)$
    $I^{(1)} = \arg\max_{i=0,\ldots,m} \theta_i$ {Thompson sampling}
end while

How — D-TTTS

Initialization: $\mu_1 \sim \nu_0$; $\mathcal{A} = \{\mu_1\}$; $m = 1$; $S_1, N_1 = 0$
while budget still available **do**
    $\forall i \in \mathcal{A}$, $\theta_i \sim \mathtt{Beta}(S_i + 1, N_i - S_i + 1)$
    $I^{(1)} = \arg\max_{i=0,\ldots,m} \theta_i$ {Thompson sampling}
    if $U(\sim \mathcal{U}([0,1])) > \beta$ **then**
        while $I^{(2)} \neq I^{(1)}$ **do**
            $\forall i \in \mathcal{A}, \theta'_i \sim \mathtt{Beta}(S_i + 1, N_i - S_i + 1)$
            $I^{(2)} \leftarrow \arg\max_{i=0,\ldots,m} \theta'_i$
        end while
        $I^{(1)} \leftarrow I^{(2)}$
    end if {TTTS}
end while

How — D-TTTS

Initialization: $\mu_1 \sim \nu_0$; $\mathcal{A} = \{\mu_1\}$; $m = 1$; $S_1, N_1 = 0$
while budget still available **do**
  $\mu_{m+1} \sim \nu_0$; $\mathcal{A} \leftarrow \mathcal{A} \cup \{\mu_{m+1}\}$
  $S_{m+1}, N_{m+1} \leftarrow 0$; $m \leftarrow m + 1$
  $\forall i \in \mathcal{A}$, $\theta_i \sim \texttt{Beta}(S_i + 1, N_i - S_i + 1)$
  $I^{(1)} = \arg\max_{i=0,\dots,m} \theta_i$ {Thompson sampling}
  if $U(\sim \mathcal{U}([0,1])) > \beta$ **then**
    while $I^{(2)} \neq I^{(1)}$ **do**
      $\forall i \in \mathcal{A}, \theta'_i \sim \texttt{Beta}(S_i + 1, N_i - S_i + 1)$
      $I^{(2)} \leftarrow \arg\max_{i=0,\dots,m} \theta'_i$
    end while
    $I^{(1)} \leftarrow I^{(2)}$
  end if {TTTS}
  $Y \leftarrow \texttt{evaluate arm } I^{(1)}$; $X \sim \texttt{Ber}(Y)$
  $S_{I^{(1)}} \leftarrow S_{I^{(1)}} + X$; $N_{I^{(1)}} \leftarrow N_{I^{(1)}} + 1$
end while

How — D-TTTS c'td

D-TTTS in summary...

In each round, query a new arm endowed with a $\texttt{Beta}(1,1)$ prior, without sampling it, and run TTTS on the new set of arms.

How — D-TTTS c'td

Order statistic trick

With $\mathcal{L}_{t-1}$ the list of arms that have been effectively sampled at time t, we run TTTS on the set $\mathcal{L}_{t-1} \cup \{\mu_0\}$ where μ_0 is a pseudo-arm with posterior $\mathtt{Beta}(t - |\mathcal{L}_{t-1}|, 1)$.

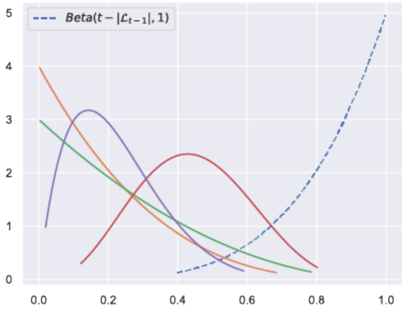


Figure: Posterior distributions of 4 arms and the pseudo-arm

Why

- TTTS is *anytime* for finitely-armed bandits;
- the flexibility of this Bayesian algorithm allows to propose a **dynamic** version for the infinite BAI;
- unlike previous approaches, D-TTTS **does not** need to fix the number of arms queried in advance, and naturally **adapts** to the difficulty of the task.

HPO as a BAI problem

BAI	HPO
query ν_0	pick a new configuration $\boldsymbol{\lambda}$˙
sample an arm˙	train the classifier $g_{\boldsymbol{\lambda}}$
reward	cross-validation loss

Experiments — Setting

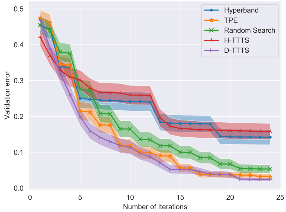
Classifier	Hyper-parameter	Type	Bounds
SVM	C	$\mathbb{R}^+$	$[10^{-5}, 10^5]$
	γ	$\mathbb{R}^+$	$[10^{-5}, 10^5]$

Table: hyper-parameters to be tuned for UCI experiments.

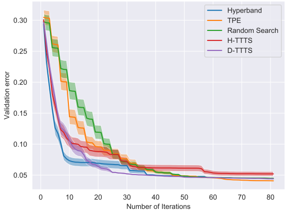
Classifier	Hyper-parameter	Type	Bounds
MLP	hidden_layer_size	Integer	$[5, 50]$
	alpha	$\mathbb{R}^+$	$[0, 0.9]$
	learning_rate_init	$\mathbb{R}^+$	$[10^{-5}, 10^{-1}]$

Table: hyper-parameters to be tuned for MNIST experiments.

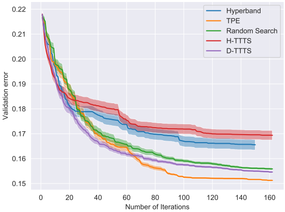
Experiments — Some results



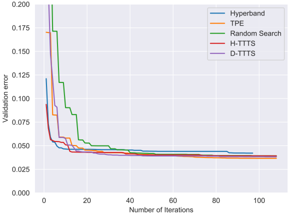
(a) wine



(b) breast cancer



(c) adult



(d) `MNIST`

Next?

- Extend to the non-stochastic setting;
- Theoretical guarantee?

Theorem (Shang, Heide, et al. 2019)

The TTTS sampling rule coupled with the Chernoff stopping rule form a δ-correct BAI strategy. Moreover, if all the arms means are distinct, it satisfies

$$\limsup_{\delta \to 0} \frac{\mathbb{E}[\tau_\delta]}{\log(1/\delta)} \leq \frac{1}{\Gamma^\star_\beta}.$$

More details

Check out [Shang, Kaufmann, et al. 2019; Shang, Heide, et al. 2019].

References

Thank you!

- Xuedong Shang, Rianne de Heide, Pierre Ménard, Emilie Kaufmann, and Michal Valko. "Fixed-confidence guarantees for Bayesian best-arm identification". 2019.
- Xuedong Shang, Emilie Kaufmann, and Michal Valko. "A simple dynamic bandit algorithm for hyper-parameter optimization". In: *6th ICML Workshop on AutoML*. 2019.