A Simple Dynamic Bandit Algorithm for Hyper-parameter Tuning

Xuedong Shang[1,2], Emilie Kaufmann[2,3] and Michal Valko[4]

[1]Inria SequeL [2]Univ. Lille [3]CNRS [4]DeepMind Paris

Problem and Objectives

We treat the **hyper-parameter tuning** problem for *supervised learning* tasks.

- global optimisation task: $\min\{f(\boldsymbol{\lambda}) : \boldsymbol{\lambda} \in \Omega\}$;
- $f(\boldsymbol{\lambda}) \triangleq \mathbb{E}\left[\ell\left(\mathbf{Y}, g_{\boldsymbol{\lambda}}^{(n)}(\mathbf{X})\right)\right]$ measures the generalization power;
- goal: a simple, robust, (almost) parameter-free bandit algorithm.

How and Why

How?

- We see the problem as a *stochastic infinitely many-armed bandit*
- Beta-Bernoulli bandit model
- A Beta resevoir ν_0 over the means of the arms
- A uniform prior Π_0 over the arms $\to$ posterior:

$$\Pi_t = \bigotimes_{i=1}^{k_t} \mathtt{Beta}(1 + S_{t,i}, N_{t,i} - S_{t,i} + 1)$$

- At each round, D-TTTS either **samples a new arm** or **re-samples a previous one**, and runs TTTS on the increasing set of arms

Why?

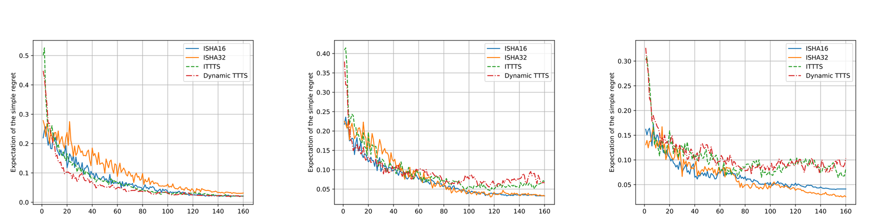
$\to$ TTTS is *anytime* for finitely-armed bandits

$\to$ The number of arms added by D-TTTS depends on the difficulty of the task (the resevoir)

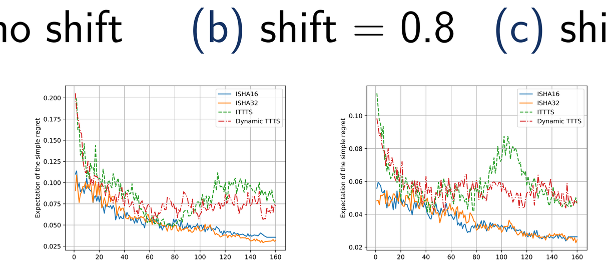
$\to$ e.g. if the resevoir is difficult (like Beta(5,1)), the pseudo-arm μ_0 will be sampled more often

$\to$ D-TTTS **does not** need to fix the number of arms sampled in advance, and naturally **adapts** to the difficulty of the task

When it fails? $\leftarrow$ If $\mu^\star \neq 1$?



(a) no shift (b) shift = 0.8 (c) shift = 0.6



(d) shift = 0.4 (e) shift = 0.2

Notation and Glossary

- Ω is the hyper-parameter space
- $\boldsymbol{\lambda}$ is a hyper-parameter configuration
- $g_{\boldsymbol{\lambda}}$ is a classifier
- ℓ is the cross-validation error in this work
- $\mu_1, \mu_2, \cdots$ denote the true means
- Θ is the parameter space
- $\Theta_i \triangleq \{\boldsymbol{\theta} \in \Theta \,|\, \theta_i > \max_{j \neq i} \theta_j\}$ is the subset of arm i being optimal

Recommendation Rule

We recommend the arm with the largest *posterior probability of being optimal*:

$$\bar{I}_n \triangleq \arg\max_{i \in \mathcal{A}} \Pi_n(\Theta_i).$$

Sampling Rule

Input: β
Initialization: $\mu_1 \sim \nu_0$; $\mathcal{A} = \{\mu_1\}$; $m = 1$; $S_1, N_1 = 0$
while budget still available **do**
  $\mu_{m+1} \sim \nu_0$; $\mathcal{A} \leftarrow \mathcal{A} \cup \{\mu_{m+1}\}$
  $S_{m+1}, N_{m+1} \leftarrow 0$; $m \leftarrow m + 1$
  $\forall i \in \mathcal{A}$, $\theta_i \sim \mathtt{Beta}(S_i + 1, N_i - S_i + 1)$
  $I^{(1)} = \arg\max_{i=0,\ldots,m} \theta_i$
  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
  $Y \leftarrow$ `evaluate arm` $I^{(1)}$; $X \sim \mathtt{Ber}(Y)$
  $S_{I^{(1)}} \leftarrow S_{I^{(1)}} + X$; $N_{I^{(1)}} \leftarrow N_{I^{(1)}} + 1$
end while

Some Tricks

- Binarization trick: When a reward $Y_{t,i} \in [0,1]$ is observed, the algorithm is updated with a fake reward $Y'_{t,i} \sim \mathtt{Ber}(Y_{t,i}) \in \{0,1\}$.
- Order statistic trick: At time t, let $\mathcal{L}_{t-1}$ be the list of arms that have been efficiently sampled, we run TTTS on the set $\mathcal{L}_{t-1} \cup \{\mu_0\}$ where μ_0 is a pseudo-arm with posterior distribution $\mathtt{Beta}(t - |\mathcal{L}_{t-1}|, 1)$.

Experimental Setting

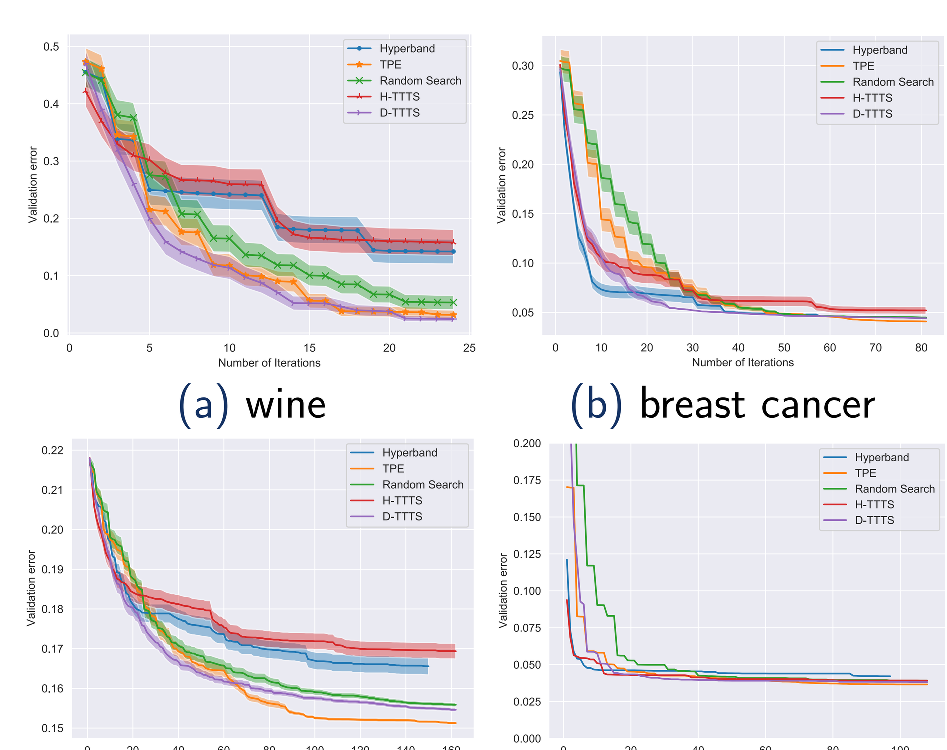
Classifier	Hyper-parameter	Type	Bounds
SVM	C	$\mathbb{R}^+$	$[10^{-5}, 10^5]$
	γ	$\mathbb{R}^+$	$[10^{-5}, 10^5]$

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`	$\mathbb{R}^+$	$[10^{-5}, 10^{-1}]$

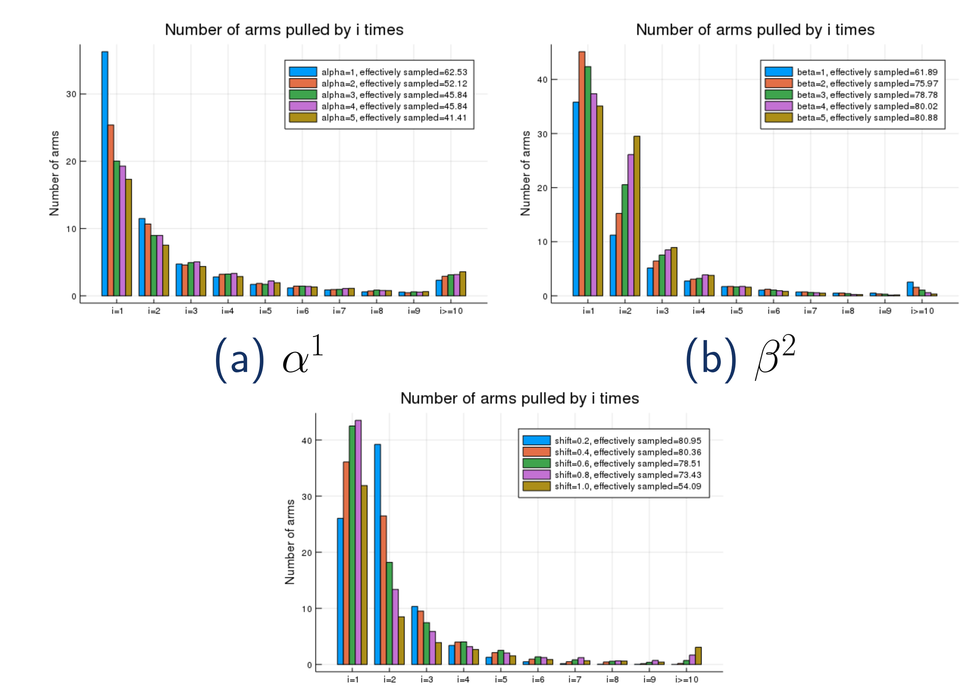
Table: hyper-parameters to be tuned for MNIST experiments.

Results for HPO



(a) wine (b) breast cancer (c) adult (d) MNIST

Illustrations of Efficiently Sampled Arms



(a) α^1 (b) β^2 (c) shift[3]

1. efficiently sampled arms for $\mathtt{Beta}(\alpha, 1)$ resevoirs
2. efficiently sampled arms for $\mathtt{Beta}(1, \beta)$ resevoirs
3. efficiently sampled arms for shifted Beta resevoirs

References

[1] Daniel Russo.
Simple Bayesian algorithms for best arm identification.
In *Proceedings of the 29th Conference on Learning Theory (CoLT)*, 2016.

[2] Lisha Li, Kevin Jamieson, Giulia DeSalvo, Ameet Talwalkar, and Afshin Rostamizadeh.
Hyperband: Bandit-based configuration evaluation for hyperparameter optimization.
In *Proceedings of the 5th International Conference on Learning Representations (ICLR)*, 2017.

Acknowledgements

The research presented was supported by European CHIST-ERA project DELTA, French Ministry of Higher Education and Research, Nord-Pas-de-Calais Regional Council, Inria and Otto-von-Guericke-UniversitÃdt Magdeburg associated-team north-European project Allocate, and French National Research Agency projects BADASS (n.ANR-16-CE40-0002), and BoB (n.ANR-16-CE23-0003).