arXiv:2302.00511v2 [cs.LG] 6 Feb 2023

迭代加深 Hyperband

Jasmin Brandt [*1] **Marcel Wever** [*2] **Dimitrios Iliadis** [3] **Viktor Bengs** [4] **Eyke H¨ullermeier** [2]

摘要

超参数优化（HPO）旨在自动搜索参数化机器学习算法的最适超参数配置（HPC）。Hyperband 是一种先进的 HPO 方法，但它自身也有一些会影响其性能的参数。其中一个参数——最大预算——尤其棘手：如果设得过小，就需要事后增加预算；而 Hyperband 在设计上并非增量式，因此必须重新运行整个算法。这不仅代价高昂，还会导致已经积累的宝贵知识流失。本文提出了 Hyperband 的增量式变体，以消除这些缺点，并证明这些变体满足与采用"恰当"预算的原始 Hyperband 定性相似的理论保证。此外，我们通过在基准数据集上进行实验，展示了这些变体的实用价值。

1. 引言

将机器学习（ML）算法成功应用于某个特定学习问题时，通常会面临如何选择其超参数的问题。这些参数通常具有很高的可配置性，而且往往会显著影响学习器的重要元属性，例如复杂度、学习速度或（最终）性能。因此，谨慎选择这些参数对于评估学习器是否适用于当前学习问题至关重要。然而，手动搜索良好甚至最优的超参数既耗时又费力，还可能引入人为偏差或损害可复现性。事情之所以更加复杂，是因为最优超参数选择可能因任务而异，也就是说，它可能取决于学习问题、数据集的性质以及相应的性能度量。

由于研究和实践中都迫切需要实现这一过程的自动化，超参数优化（HPO）这一研究领域应运而生（Feurer& Hutter，2019；Bischl 等人，2021）。一般而言，针对超参数优化问题的方法旨在尽可能高效地为给定的机器学习算法找到最优超参数配置（HPC），使该算法利用找到的超参数能够很好地泛化到新的、未见过的数据点。Hyperband `Hyperband`（Li 等人，2018）是一种常用的超参数优化方法，它基于多保真度概念：首先将可用计算预算用于低成本超参数配置的探索，再逐步将预算投入最有希望的超参数配置。这里的预算可以指用于训练的样本数或属性数等计算资源，也可以指迭代算法的迭代次数或神经网络的训练轮次。

然而，最大预算和探索速率之外，`Hyperband` 本身也有一些决定其在超参数优化问题中表现的参数。尤其是前者的选择颇具挑战，因为在许多情况下，一开始并不容易确定合适的取值，而取值过高则代价高昂，因为运行`Hyperband` 本身就已十分耗费资源。例如，当预算对应于机器学习算法中的某个超参数，而该参数决定算法处理数据的时长时，就会出现这种情况，比如神经网络的训练轮次数、逻辑回归的迭代次数等。恰恰在这些情况下，预算取值过小会带来诸多问题，因为所考察的高性能计算集群可能尚未发挥出其潜力，也就是说，模型尚未训练至收敛。然而，事后增加预算会导致通常代价高昂的`Hyperband` 重跑；更糟的是，已经积累的宝贵知识也会被丢弃。不用说，从生态角度来看，这同样不可取，因为计算资源以及为在较低最大预算下优化超参数所消耗的能源，基本上都被浪费了（Tornede 等人，2021）。

为克服这些问题，我们提出了 Hyperband 的增量式扩展，使其能够事后提高最大预算，并继续运行此前在较小最大预算下进行的

[*]共同贡献 [1]帕德博恩大学，帕德博恩，德-国 [2]慕尼黑机器学习中心，慕尼黑大学，慕尼黑，德国 [3]根特大-学，根特，比利时 [4]慕尼黑大学，慕尼黑，德国。通讯作者：Marcel Wever <marcel.wever@ifi.lmu.de>。

Hyperband 运行。借鉴启发式图搜索算法（如迭代加深 A* 或迭代加深搜索）的思路（Korf，1985），我们将这一扩展命名为迭代加深 Hyperband (ID-HB)，并提出了三个变体。每种变体都以不同方式考虑如何复用先前 Hyperband收集的信息，本质区别在于保守程度。我们给出了理论结果和实证结果，表明与效率提升相比，性能下降微乎其微。为此，我们针对基准数据集上的各种任务提供了理论保证和实证评估。

2. 超参数优化

在典型的监督式机器学习场景中，学习器会获得一个（训练）数据集 $\mathcal{D} = \{(\boldsymbol{x}^{(n)}, y^{(n)})\}_{n=1}^{N} \subset \mathcal{X} \times \mathcal{Y}$，其中 $\mathcal{X}$ 是某个特征空间，$\mathcal{Y}$ 是一个标签空间。$\mathcal{D}$ 中的实例-标签对被假设为来自某个未知数据生成分布 P^* 的独立同分布样本，即，每个 $z^{(n)} = (\boldsymbol{x}^{(n)}, y^{(n)})$ 都是 $Z = (X, Y) \sim P^*$ 的一次独立实现。此外，学习器还会获得一个（合适的）假设空间 $\mathcal{H}$，它是所有映射 $\mathcal{Y}^{\mathcal{X}} = \{h : \mathcal{X} \to \mathcal{Y}\}$ 从特征空间 $\mathcal{X}$ 到标签空间 $\mathcal{Y}$ 的子集。因此，每个假设 $h \in \mathcal{H}$ 都会为给定实例 $\boldsymbol{x} \in \mathcal{X}$ 分配一个预测 $h(\boldsymbol{x}) \in \mathcal{Y}$。对于给定的实例-标签对 $(\boldsymbol{x}, y)$，假设的预测质量通过 $L(h(\boldsymbol{x}), y)$ 进行评估，其中 $L : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$ 是一个鼓励正确预测的损失函数。在监督式机器学习中，目标是归纳出一个假设，使得按照 P^* 采样的实例-标签对 $(\boldsymbol{x}, y)$ 的预期损失（泛化误差）最低。形式上，学习器寻求找到

$$h^* \in \arg\min_{h \in \mathcal{H}} \mathrm{GE}(h)\,,$$

其中，$\mathrm{GE}(h)$ 表示 h 的泛化误差：

$$\mathrm{GE}(h) = \mathbb{E}_{(\boldsymbol{x},y)\sim P^*}\left[L(h(\boldsymbol{x}), y)\right].$$

为此，我们设计了一种学习器（或归纳器）$\mathcal{A}$，它针对给定的（训练）数据集返回一个被认为适用于当前学习任务（即泛化误差较低）的假设。通常，学习器由参数空间 Λ（其元素称为超参数）参数化。超参数可能是高维元组，其分量来自不同的域（连续、离散或类别型）。因此，学习器在形式上是一个映射

$$\mathcal{A} : \mathbb{D} \times \Lambda \to \mathcal{H},\ (\mathcal{D}, \lambda) \mapsto \hat{h}\,,$$

其中，$\mathbb{D} = \bigcup_{N\in\mathbb{N}} (\mathcal{X} \times \mathcal{Y})^N$ 是所有可能训练数据集的集合。值得注意的是，如果假设空间 $\mathcal{H}$ 是从特征空间 $\mathcal{X}$ 到 $\mathbb{P}(\mathcal{Y})$ 的所有映射的子集，也可以用类似的方式定义学习器，即，标签空间 $\mathcal{Y}$ 上的概率集合。唯一的区别在于，此时损失函数的签名不同，也就是说，它是一个映射 $L : \mathbb{P}(\mathcal{Y}) \times \mathcal{Y} \to \mathbb{R}$.

因此，超参数优化（HPO）的目标是为给定的学习器 $\mathcal{A}$ 和数据集 $\mathcal{D}$，找到一个最优超参数，即

$$\lambda^* \in \arg\min_{\lambda\in\Lambda} \ell(\lambda) = \arg\min_{\lambda\in\Lambda} \mathrm{GE}\big(\mathcal{A}(\mathcal{D}, \lambda)\big).$$

然而，由于 P^* 未知，泛化误差 ℓ也未知，因此，对于固定的超参数$\widehat{\ }$，人们会借助函数 $\ell : \Lambda \to \mathbb{R}$ 来估计后者，该函数通常称为验证误差。由于计算特定超参数对应的验证误差可能会消耗大量可用资源（例如墙上时钟时间、使用的数据点数量等），通常只会在某种资源分配下确定验证误差，因此其实际值取决于所用资源。$\widehat{\ }$ 鉴于此，我们用 $\ell_r(\lambda)$ 表示$\mathcal{A}$ 在使用超参数 λ 和 r 个资源单位时的验证误差。显然，r 的选择涉及权衡：使用的资源单位越多，估计就越准确，但计算成本也越高，反之亦然。

粗略地说，HPO方法旨在为 $\mathcal{A}$ 找到合适的超参数，同时尽可能少地使用资源，和/或将搜索过程中评估超参数所消耗的资源限制在最高可分配预算 R 之内。为方便起见，我们假设 R 属于 $\mathbb{N} \cup \{\infty\}$，其中 $R = \infty$ 表示资源使用不受限制。我们定义 $\ell_*(\lambda) := \lim_{r\to R} \ell_r(\lambda)$，对任意$\lambda \in \Lambda$，并定义 $\nu_* := \inf_{\lambda\in\Lambda} \ell_*(\lambda)$。因此，HPO方法的正式目标是确定一个属于 $\mathrm{argmin}_{\lambda\in\Lambda} \ell_*(\lambda) - \nu_*$. 的超参数 λ

3. Hyperband

本节简要介绍Hyperband 算法，该算法由 Li 等人 (2018) 提出，以及其子程序 SuccessiveHalving（Karnin 等人，2013）。

3.1. 连续减半

连续减半（SH）算法可在固定预算内解决非随机最佳臂识别问题，并已成功应用于超参数优化。它会迭代地将可用预算分配给一组超参数配置，评估其性能，并淘汰表现最差的一半。该过程会反复进行，直到只剩下一种超参数配置，随后算法将其作为推荐的优胜者返回。因此，我们可以为更有希望的超参数配置分配呈指数级增长的预算。

3.2. Hyperband

由 Hyperband 算法（Li 等人 （2018）将超参数优化问题视为纯探索型自适应资源分配问题来加以解决。它会遍历不同数量的超参数配置集合 n，同时保持预算 B 不变，并将连续减半算法作为子程序应用于 n 个配置，预算为 B。这样便可以考虑不同的分配策略，以权衡（i）纳入更多配置 n与（ii）为配置提供更长的训练时间 B/n。每次调用连续减半算法称为一个括号。

4. 相关工作

要达到最先进的性能，超参数优化（超参数优化）是机器学习过程中不可或缺的一步，其目标是针对给定的数据集和性能指标，为机器学习算法寻找最合适的超参数配置。将超参数优化视为黑箱优化问题，可以使用各种方法来解决。然而，超参数优化面临的一项特殊挑战是，评估超参数配置的代价高昂，因此，尽管网格搜索和随机搜索等朴素方法应用广泛，却并不实用。

超参数优化文献可分为两个分支：无模型方法和基于模型的方法。后者利用优化曲面的代理模型来抽样更有希望的候选项（Hutter 等人，2011），例如，进化算法属于前一类无模型方法。值得注意的是，网格搜索和随机搜索也属于无模型方法。虽然标准随机搜索被认为在超参数优化场景中效率过低，但在Hyperband （Li 等人，2018）中，随机搜索与多保真度候选评估流程（即连续减半（ Jamieson& Tal-walkar，2016））相结合，从而构成一种强大的超参数优化方法。此外，无模型方法Hyperband 还可以与进化算法等其他无模型方法（Awad 等人，2021）结合，或与基于模型的方法混合（Falkner 等人，2018），以提升其有效性和效率。在（ Mendes 等人，2021）中，作者提出了一种元学习方法，将评估预算进一步集中于有希望的候选项，而不是浪费在性能逊于当前最优配置的超参数配置上。

尽管这些方法旨在提高Hyperband单次运行的效能与效率 Hyperband，但本文关注的是，在提高单个超参数配置最多可分配的预算后，如何提升Hyperband的效率。一般设定已在（Li 等人，2018）中作为无限时域设置加以考虑和分析，其中Hyperband会随着最大预算的增加而反复运行 R。我们在先前已进行的 Hyperband基础上继续，而不是每次提高最大预算时都从头开始。

有关超参数优化的更详细介绍以及对相关方法更全面的概述，请读者参阅（Feurer & Hutter，2019；Bischl 等人，2021）。

5. 迭代加深 Hyperband

在 Hyperband 的无限时域设置中，一旦一次 Hyperband 运行结束，就会从头开始新一轮运行，并提高可分配预算的上限 R。然而，对于预算上限提高到 R 的 Hyperband 运行，此前评估过的超参数配置会被完全丢弃，这意味着之前投入的预算实质上被浪费了。由于多种原因，丢弃这些信息是不合理的：

- 我们忽视了已经获得的、有关有潜力的超参数配置的知识。
- 我们没有利用已收集的信息，而是评估新的候选项，以便进行更多探索；这固然是我们所希望的，但也需要更多预算。
- 从生态角度来看，先前 Hyperband 运行中投入的资源并未得到妥善利用，因为这些信息（如果确实用到的话）仅用于决定是否以更大的预算重新运行 Hyperband。

为利用已收集的信息，我们提出对Hyperband的扩展。该扩展具有状态，因此可以在更大的最大预算下继续运行 R，我们将其称为迭代加深 Hyperband（ID-HB）。我们的方法名称灵感源自迭代加深搜索（Korf，1985），例如迭代加深深度优先搜索，其中深度优先搜索会以不断增大的最大深度重复执行。该扩展的伪代码见算法 1，图5.3 展示了其核心思想。

从算法 1中可以看出，旧的最大规模 R_{t-1} 会增加 η 倍，得到 R_t。为便于理论分析，我们仅考虑最大预算按 η倍增加的情况。根据新的最大规模 R_t，变量 $s_{\max}$ 和 B 按照Li 等人(2018)中的方法计算，这将为新的最大规模 R_t 增加一个括号，并增大已有括号的池大小。因此，我们为括号补充额外的超参数配置，直到达到相对于 R_t的正确初始候选池大小。然而，由于通常新增采样的超参数配置数量超过了能够传递到某个括号下一次迭代的数量，因此，可以采用不同策略来处理先前Hyperband的状态以及新采样的超参数配置，策略包括

可以设想。

在后续章节 5.1 至 5.3中，我们将详细阐述三种可能的策略，说明如何将先前在较低最大预算下运行的`Hyperband` 延续下去 R_{t-1}，以适应更大的最大预算 R_t。我们根据这些策略对于从头开始运行 `Hyperband` 时所做决策的忠实程度，对它们进行排序。这个顺序也与效率的提升相对应，也就是说，评估超参数配置所累计消耗的预算更少。

5.1. 丢弃式迭代加深（dID-HB）

可以说，更新括号以纳入新收到的超参数配置时，最忠实但也可能效率低下的做法，是允许修正先前有关晋级到后续迭代的决策。更具体地说，当初始候选池中的超参数配置因新增配置而扩充时，我们允许丢弃先前运行中已晋级、且已在更大预算下完成评估的超参数配置。

在第一次迭代中，我们会考虑所有可用的候选项，即新采样的超参数配置、先前丢弃的配置以及先前晋级的超参数配置。前 k 名的计算方式与之前相同，选出的超参数配置会晋级至下一次迭代。因此，丢弃式迭代加深 Hyperband 能够修正先前有关晋级的决策，并丢弃在上一次运行中已晋级的超参数配置。

虽然在某次迭代中，只需评估新采样的超参数配置，以及为新的最大规模而进行的最后一次迭代中的配置 R_t，但也可能只有新的配置会晋级到后续迭代。在这种情况下，这种有状态扩展变体只能节省先前用于按括号的最小预算评估旧配置的资源。不过，我们将在第 7.2 节看到，旧候选项通常会在后续迭代中保留下来。

最终，我们得到了一种迭代加深 Hyperband 变体：在使用与原始 Hyperband 相同的初始超参数配置集合时，它有望在保持相同结果的同时显著提高效率。然而，在最坏情况下，这种变体的开销几乎与重新运行原始版本相当，也就是说，它只节省了对旧候选项进行第一次迭代时的评估开销。

5.2. 保留式迭代加深（pID-HB）

为了进一步提高效率并复用超参数配置的先前评估结果，在保留式迭代加深 Hyperband（pID-HB）中，我们从一个候选项池中晋升前 k 名超参数配置。该候选项池由已晋升的超参数配置，以及在上一次 Hyperband 运行中已晋升至此预算级别、但在后续运行中未获晋升的所有超参数配置组成。这样，我们便能保留有关这些超参数配置的信息：它们已针对该预算进行评估，但在先前的一次迭代中被舍弃。

继续考虑此类超参数配置，可以让已经被丢弃的超参数配置重新回到有前景候选池中。一方面，这样可能提高效率，因为当某个超参数配置重新进入晋级候选项集合后，除了使用新最大规模预算的最后一次迭代外，我们不需要再花费额外预算评估新的候选项。另一方面，如果此前被丢弃的旧超参数配置在更大预算下的表现优于取代它的配置，我们就可以修正当前运行中的“错误”决策。

在最坏情况下，pID-HB 的计算复杂度与 dID-HB 相同，只节省了在各个括号的起始预算上评估旧超参数配置所需的计算资源。然而，由于能够重新考虑那些已经被丢弃且在之前的 Hyperband 运行中已经完成评估的候选项，因此重复利用已评估候选项的机会有所增加。尽管 pID-HB 不一定会返回与从头重新运行 Hyperband 相同的结果，但直观而言，我们预计其性能会相近。此外，尽管 pID-HB 与 dID-HB 具有相同的最坏情况运行时间复杂度，pID-HB 至少会将所有可用信息纳入考虑；也就是说，以前评估过但被丢弃的超参数配置的评估数据不会被忽略，而是仍会被考虑。

5.3. 高效迭代加深（eID-HB）

第三种也是最后一种策略，是以最高效、因而最节省资源的方式继续之前的 Hyperband 运行。因此，我们将这一变体称为“高效迭代加深 Hyperband”（eID-HB）。然而，最高效率是以可能导致性能下降为代价的，因为它不会撤销此前做出的任何决策。

换句话说，它只会通过晋级超参数配置来填充候选池，直到达到新的层级。如果某个超参数配置已晋级到连续减半的后续迭代中，即使新采样的超参数配置集合中存在实际上可以取代它的配置，它也会保留下来。从技术上讲，在确定

连续减半下一次迭代的前 k 名时，我们会减去上一次运行中已经晋级的超参数配置数量。根据这一差值，从先前被淘汰的候选项以及新采样且已晋级至当前迭代的超参数配置中，选出新的晋级项。

原则上，与从头开始运行 Hyperband 相比，某些超参数配置可能会被错误地晋级。由于已经晋级的超参数配置会保持晋级状态，且无法修正决策，整体性能可能会下降。不过，如果性能下降可以忽略不计，eID-HB 就能让我们以仅运行一次较大最大规模 R_t 的 Hyperband 为代价，将 Hyperband 的最大规模提高到 R_t。

在后续章节中，我们将给出理论保证，并通过实验证明，所提出的扩展在保持性能相近的同时，显著提高了效率。

6. 理论结果

我们将理论结果分为两部分。首先，针对我们对 SuccessiveHalving 的扩展给出一些理论保证；其次，将这些保证扩展到迭代加深 Hyperband 算法。

6.1. 迭代加深 SuccessiveHalving

由于连续减半算法（Jamieson &Tal-walkar，2016）求解的是老虎机问题，因此在以下对连续减半迭代加深变体的分析中，我们将沿用多臂老虎机的记号。只需将超参数配置视为一个臂，我们的算法就可以轻松应用于超参数优化。若我们拉动某个臂 i 次 k，就会观察到损失 $\ell_{i,\,k}$。与 Li 等人 （2018）类似，对所提出的迭代加深-连续减半算法进行理论分析时，我们需要作出以下假设。

假设 6.1. 对于每个臂 $i \in \mathbb{N}$，极限 $\nu_i := \lim_{t\to\infty} \ell_{i,\,t}$ 存在。

此外，我们用 $\gamma\ (j) \ \geq \sup_i |\ell_{i,\,j} - \nu_i| \ \forall j \in \mathbb{N}$ 表示收敛速度，并给出以下结果，其证明见附录B.1。

定理 6.2 （迭代加深–SuccessiveHalving 的必要预算）。 固定 n 个臂，其中 $\tilde{n}$ 个臂已被晋升。令 $\nu_i = \lim_{t\to\infty} \ell_{i,\,t}$ 并假设 $\nu_1 \leq \cdots \leq \nu_n$。对于任意 $\epsilon > 0$，令

$$z_{ID-SH} = \eta\lceil \log_\eta(n)\rceil \times \max_{i=2,\ldots,n} i\Big(1 + \min\left\{R, \gamma^{-1}\left(\max\left\{\tfrac{\epsilon}{4}, \tfrac{\nu_i-\nu_1}{2}\right\}\right)\right\}\Big).$$

如果使用任意预算 *3*、算法 *4* 或 *5* 运行算法 $B \geq z_{ID-SH}$，则会返回一个满足条件的臂 $\hat{\imath}$： $\nu_{\hat{\imath}} - \nu_1 \leq \epsilon/2$。

此外，以下定理说明了增量变体相较于代价高昂的重新运行 SuccessiveHalving（SH）所带来的改进（证明见附录 B.2）。

定理 6.3 （xID-SH 拉取次数相较于 SH 的改进）。 固定 n 个臂、预算 B、最大规模 R，以及 r 和 η。假设我们已在 $\tilde{n}$ 个臂上运行过 *SuccessiveHalving*，且使用了相同的 B、 r 和 η。令 $\eta_- = \eta - 1$ 和 $s^+ = s + 1$。若我们使用上述变量，在 s 轮中运行 *SuccessiveHalving*（*SH*）、高效迭代加深 *SuccessiveHalving*（*eID-SH*），以及保留式/丢弃式迭代加深 *SuccessiveHalving*（*p/dID-SH*），则有

a) $\#\{$*eID-SH* 的总拉取次数$\}$

$$\leq \left(1 - \frac{(s^+)(\tilde{n}R + \eta^s)(\eta_-) - (\eta^{s^+} - 1)(2R + n)}{(s^+)(nR + \eta^s)(\eta_-) - (\eta^{s^+} - 1)(R + n)}\right) \times \#\{\textit{SH} \text{ 的总拉取次数}\}$$

以及 *b)*$\#\{$*p/dID-SH* 的总拉取次数$\}$

$$\leq \left(1 - \frac{(\eta_-)((s^+)\eta^s + R\tilde{n}) - (\eta^{s^+} - 1)(R + n)}{(\eta_-)(s^+)(nR + \eta^s) - (\eta^{s^+} - 1)(R + n)}\right) \times \#\{\textit{total pulls of SH}\}.$$

eID-SH 相较于 SH 的总拉取次数改进比例见图 2，其他变体的类似图表见附录B.2。

6.2. 迭代加深 Hyperband

如上所述，最优超参数配置 λ^* 可能并非总是存在。即使存在，由于超参数配置空间通常非常大，甚至可能是无限的，因此搜索它也可能不可行。因此，我们将放宽目标，转而寻求找到至少“近乎最优”的配置。与超参数优化问题相关文献类似，我们将此类近乎最优配置定义如下：对于 $\epsilon > 0$，若 λ 是一个 ϵ-最优配置 当且仅当 $\nu_\lambda - \nu_{\lambda^*} \leq \epsilon.$。为了确保搜索此类配置不会像大海捞针一样，我们需要一个假设，以保证样本集中存在 ϵ-最优配置的概率足够高。

假设 6.4。 Λ 中 ϵ最优配置的比例为 $\alpha \in (0,\ 1)$。

请注意，现在样本集中至少存在一个 ϵ 最优配置的概率至少为 $1 - \delta$，如果其大小

算法 1迭代加深 Hyperband (ID-HB)

输入： 旧最大规模 R_{t-1}、 $\eta \geq 2$，旧损失 $L_{(t-1,\cdot,\cdot)}$，被丢弃的配置 $D_{(t-1,\cdot,\cdot)}$，晋升的配置 $P_{(t-1,\cdot,\cdot)}$，模式标志 $\rho \in \{p, d, e\}$

初始化： $R_t \leftarrow \eta R_{t-1}$， $s_{\max} \leftarrow \lfloor$对数$_\eta$（$R_t$）$\rfloor$， $B \leftarrow$ （$s_{\max}+1$）R_t

对于 $s \in \{s_{\max}, s_{\max}-1, \ldots, 0\}$**，执行**

$n_t \leftarrow \lceil \frac{B}{R_t} \frac{\eta^s}{(s+1)} \rceil$

$r = R\eta^{-s}$

如果 $s > 0$ **则** ▷ 旧最大规模的配置

$n_{t-1} \leftarrow \lceil \frac{\eta^{s-1} s_{\max}}{s} \rceil$

else

$n_{t-1} \leftarrow 0$

结束如果

$\delta \leftarrow n_t - n_{t-1}$

$C \leftarrow$获取超参数配置(δ)

对于 $i \in \{0, \ldots, s\}$ **执行**

$r_i \leftarrow r\eta^i$

$L_{(t,s,i)} \leftarrow L_{(t-1,s-1,i)} \cup \{$运行，然后返回验证损失($c$， r_i)： $c \in C \setminus P_{(t-1,s-i,i-1)}\}$ ▷ 评估配置

如果 $\rho = d$**则** ▷ 丢弃 ID-HB

如果 $i = 0$**则**

$T \leftarrow D_{(t-1,s_1,i)} \cup P_{(t-1,s-1,i)} \cup C$

else

$T \leftarrow C$

结束条件

$k \leftarrow \lfloor n_t / \eta^{i+1} \rfloor$

否则如果 $\rho = p$ **则** ▷ 保留 ID-HB

$T \leftarrow D_{(t-1,s_1,i)} \cup P_{(t-1,s-1,i)} \cup C$

$k \leftarrow \lfloor n_t / \eta^{i+1} \rfloor$

否则如果 $\rho = e$ **则** ▷ 高效 ID-HB

$T \leftarrow D_{(t-1,s-1,i)} \cup C$

$k \leftarrow \lfloor n_t / \eta^{i+1} \rfloor - \lfloor n_{t-1} / \eta^{i+1} \rfloor$

结束条件判断

$C \leftarrow$顶端$_k$(T， $L_{(t,s,i)}$， k)

$D_{(t,s,i)} \leftarrow T \setminus (C \cup P_{(t-1,s-1,i)})$ ▷ 更新被丢弃的配置

$P_{(t,s,i)} \leftarrow P_{(t-1,s-1,i)} \cup C$ ▷ 更新晋升的配置

结束循环

结束循环

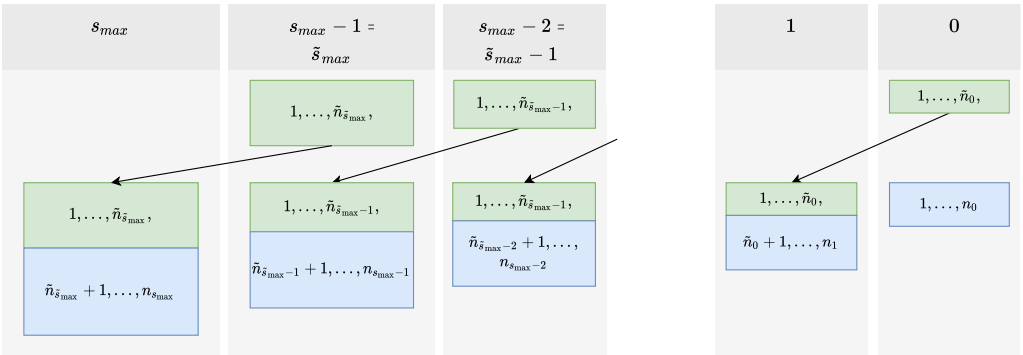

图*1.* 接管先前采样的配置示意图。

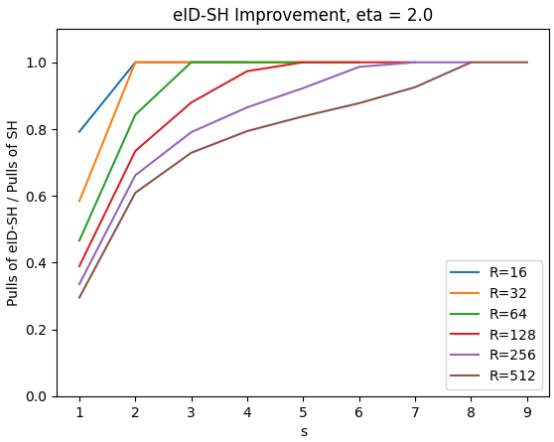

图 *2*。不同 SH 轮数 s 和每轮最大预算 R下，eID-SH 和逐步减半的拉取次数占比。

样本集为 $\lceil \log_{1-\alpha}(\delta) \rceil$，对于固定的失败概率$\delta \in (0,1)$。由此，我们可以陈述以下定理，其证明见附录B.3。

定理 6.5。令 η、R、α 和 δ 为固定值，使得

$$R \geq \max \Bigg\{ \lceil \log_{1-\alpha}(\delta) \rceil (\eta - 1) + 1, \\ \eta \Big(\log_\eta(\log_\eta(R)) + 4 + \frac{\lfloor \log_\eta(R) \rfloor}{2} \\ - \log_\eta \left((\lfloor \log_\eta(R) \rfloor + 1)! \right) / (\lfloor \log_\eta(R) \rfloor + 1) \Big) \bar{\gamma}^{-1} \Bigg\}$$

$$for \quad \bar{\gamma}^{-1} := \max_{s=0,\ldots,\lfloor \log_\eta(R) \rfloor} \max_{i=2,\ldots,n_s} i \Big(1 \\ + \min \Big\{ R, \gamma^{-1} \Big(\max \Big\{ \frac{\epsilon}{4}, \frac{\nu_i - \nu_1}{2} \Big\} \Big) \Big\} \Big).$$

那么，*ID-HB* 以至少 $1-\delta$ 的概率找到一个 ϵ-最优配置。

7. 实证评估

本节对所提出的`Hyperband` 扩展进行实证评估，并将第 5 节提出的三种策略与原始的`Hyperband` 应用方式进行比较，即增加最大规模 R，如无限时域设置中所做的那样。具体而言，我们关注以下两个研究问题：

RQ1 ID-HB 的各个变体 dID-HB、pID-HB 和 eID-HB 能否保持所返回超参数配置的质量？**RQ2** dID-HB 和 pID-HB 能在多大程度上减少计算开销？

为回答**RQ1** 和**RQ2**，我们进行了一系列全面的实验，其设置详见第7.1节。

表 *1*。所考虑的 YAHPO-Gym 基准测试列表、学习器类型、所考虑的数据集数量、目标函数，以及可用作保真度参数的预算类型。

基准测试	模型	# 个实例	目标	保真度
lcbench	神经网络	34	验证准确率	训练轮次
nb301	神经网络	1	验证准确率	训练轮次
rbv2 支持向量机	SVM	106	acc	比例
rbv2 ranger	随机森林	119	acc	比例
rbv2 xgboost	XGBoost	119	acc	比例

随后，我们在第7.2节中呈现实验结果并进行讨论。

7.1. 实验设置

在实验评估中，为回答研究问题**RQ1** 和**RQ2**，我们将所提出的 ID-HB 方法的三种变体分别与作为基线的原始 Hyperband 进行比较。为此，我们开展了大量实验，涵盖各类超参数优化任务，包括神经网络、支持向量机、XGBoost、随机森林的超参数优化，以及神经架构搜索。

我们使用 YAHPO Gym（Pfisterer 等人，2022）作为基准测试库。该库提供可快速评估的代理基准测试，用于超参数优化，并特别支持多保真优化，因此非常适合本研究。从 YAHPO Gym 中，我们选取了表 1所列的基准测试。除了 `nb301` 仅包含 CIFAR10 这一个数据集外，其他所有基准测试均包含多个数据集，从而能够广泛比较 ID-HB 与原始 `Hyperband`（下文简称 IH-HB）。我们针对 $\eta = 2$ 和 $\eta = 3$评估所有基准测试，但受篇幅限制，此处仅展示 $\eta = 2$的结果。关于 $\eta = 3$ 的结果，以及各个数据集的详细结果，见附录C。

此外，我们将初始最大规模设置为 $R_{t-1} = 16$，并在首次运行后将其扩大 η 倍，达到 $R_t = R_{t-1}\eta$。对于将训练数据集的一部分作为保真度参数的基准测试，我们将预算 r 通过 r/R_t 转换为 0 到 1 之间的比例。此外，我们对算法、参数化和基准实例的每种组合分别使用 30 个随机种子进行重复，从而总计进行 $30 \times 4 \times 2 \times 379 = 90$，960 次超参数优化运行。所有实验均在一台配备 2xIntel Xeon Gold 5122 和 256GB 随机存取存储器的工作站上完成。

代码和数据可通过 GitHub[1] 公开获取。

[1]https://github.com/mwever/iterative-deepening-hyperband

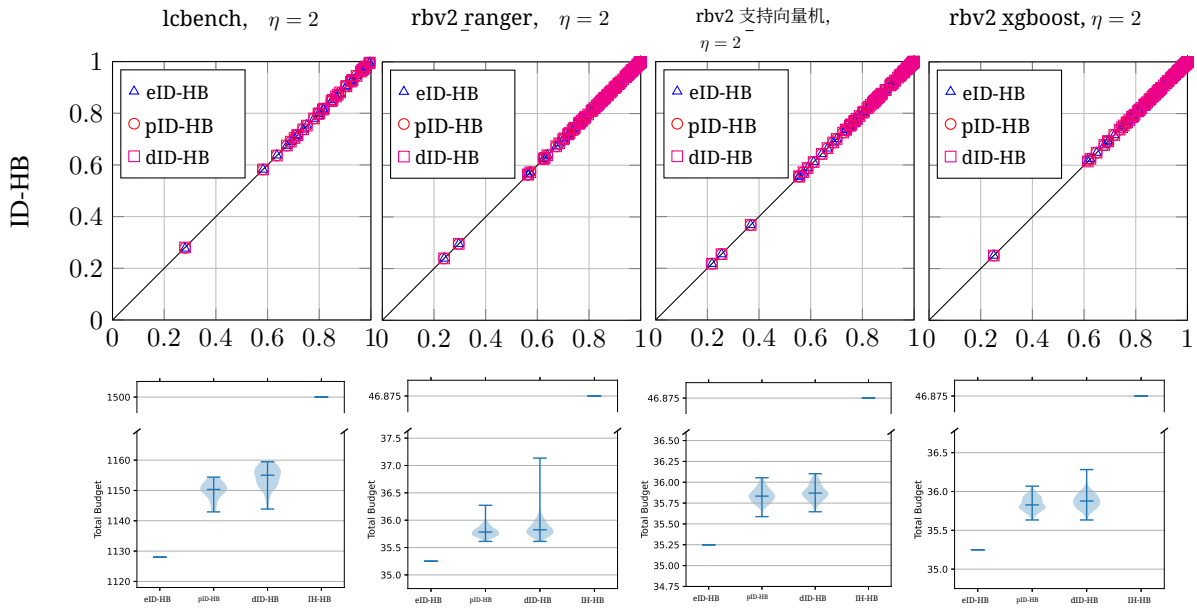

图*3*。ID-HB 与原始 Hyperband 版本的比较。上图：散点图展示了采用 ID-HB 策略与 IH-HB 获得的最终当前最优配置的准确率。下图：小提琴图展示了单次运行消耗的平均总预算。

7.2. 实证结果

在图 3中，我们展示了包含多个数据集的基准测试结果：`lcbench`、rbv2 支持向量机、`rbv2 ranger`和`rbv2 xgboost`。在上排中，我们展示了散点图：横轴为 IH-HB 获得的最终候选模型性能，纵轴为 ID-HB 策略获得的性能。对角线表示两者性能相当。从这些图中可以明显看出，任何 ID-HB 策略的表现都与原始 `Hyperband` 变体相近。对于 `nb301` 基准测试，也能观察到类似的情况：所有方法的准确率均为0.9061 （见附录C）。

关于 **RQ1**，对于此处考虑的基准测试，我们由此可以通过实证确认，所提出的 ID-HB 扩展确实能够保持返回的超参数配置的质量。尤其值得注意的是 eID-HB 的性能，因为它不会修订任何决策，直觉上似乎很可能导致性能下降；然而，在此处考虑的超参数优化中，这种情况从未发生。所有 Hyperband 变体在返回解的质量方面无一例外地表现同样出色。此外，这些结果对于 $\eta = 3$也具有代表性。

不过，若考虑单次运行消耗的平均预算，ID-HB 仍能带来显著改进。如图 3底行所示，该行展示了单次运行消耗的总预算分布，eID-HB 是效率最高的策略。与直觉预期一致，pID-HB 比 dID-HB 更有效率，不过差异通常相当有限。或许更令人意外的是，在所有基准测试中，pID-HB 和 dID-HB 的效率都显著高于 IH-HB。即使是最差的运行，总预算消耗仍能减少 20%，从而回答了**RQ2**。尽管 pID-HB 和 dID-HB 的理论最坏情况分析仅显示出微小、几乎可以忽略的改进，但在实际中，这些策略似乎相当高效，且只需重新进行少量评估。

8. 结论与未来工作

本文提出了一种对广为人知的超参数优化方法 `Hyperband` 的扩展，称为迭代加深 Hyperband（ID-HB），旨在提升当事后需要增大 `Hyperband` 的最大规模超参数时的效率。我们推导出了三种策略，它们在从头开始运行 `Hyperband` 时，复用同一组超参数配置样本的程度各不相同。对于这三种策略，我们都给出了最终选择质量以及继续执行先前 `Hyperband` 运行时所节省预算的理论保证。在实证研究中，我们还发现，这三种策略的结果与成本高得多的 `Hyperband`基线变体相近。事实上，在效率最高的策略中，我们的方法只需为增大最大规模后的那次运行分配预算。未来，我们计划将更高效的 `Hyperband` 扩展与更先进的超参数配置采样方法相结合，例如（Awad 等人，2021）或（Falkner

等人，2018）以及 HyperJump，以进一步提升 ID-HB 的效能和效率。

软件与数据

软件和实验数据可通过 GitHub 公开获取：
https://github.com/mwever/iterative-deepening-hyperband。

致谢

本研究得到了研究培训小组 Dataninja（可信赖人工智能，助力无缝问题求解：新一代智能与稳健数据分析相结合）的支持。该项目由德国北莱茵-威斯特法伦州资助。

参考文献

Awad, N. H., Mallik, N., and Hutter, F. DEHB：面向可扩展、稳健且高效超参数优化的进化式 Hyperband。载于第三十届国际人工智能联合会议论文集，第 2147–2153 页，2021。Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A.-L., 等。超参数优化：基础、算法、最佳实践与未解挑战。*arXiv* 预印本 *arXiv:2107.05847*，2021。Falkner, S., Klein, A., and Hutter, F. BOHB：大规模稳健且高效的超参数优化。载于第 *35* 届机器学习国际会议论文集，机器学习研究论文集第 80 卷，第 1436–1445 页。PMLR，2018。Feurer, M. and Hutter, F. 超参数优化。载于自动化机器学习，第 3–33 页。施普林格，2019。Hutter, F., Hoos, H. H., and Leyton-Brown, K. 用于通用算法配置的序贯模型优化。载于学习与智能优化——第五届国际会议，计算机科学讲义第 6683 卷，第 507–523 页。施普林格，2011。Jamieson, K. and Talwalkar, A. 非随机最佳臂识别与超参数优化。载于第 *19* 届人工智能与统计国际会议论文集，机器学习研究论文集第 51 卷，第 240–248 页。PMLR，2016。Karnin, Z., Koren, T., and Somekh, O. 多臂老虎机中的近乎最优探索。载于机器学习国际会议，第 1238–1246 页。PMLR，2013。Korf, R. E. 深度优先迭代加深：一种最优的可采纳树搜索。人工智能，27(1):97– 109，1985。Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., 和 Talwalkar, A. Hyperband：一种基于新型赌博机的超参数优化方法。机器学习研究期刊，18(185):1–52，2018。Mendes, P., Casimiro, M., 和 Romano, P. Hyperjump：通过风险建模加速 Hyperband。*arXiv* 预印本 *arXiv:2108.02479*，2021。Pfisterer, F., Schneider, L., Moosbauer, J., Binder, M., 和 Bischl, B. YAHPO gym—— 一种用于超参数优化的高效多目标多保真基准。在国际自动化机器学习会议，*AutoML 2022*，机器学习研究论文集第188卷，第3/1–39页。PMLR，2022。Tornede, T., Tornede, A., Hanselle, J., Wever, M., Mohr, F., 和 H¨ ullermeier, E. 迈向绿色自动化机器学习：现状与未来方向。*arXiv* 预印本 *arXiv:2111.05850*，2021。

A. 伪代码

在理论分析中，我们考虑以下算法，其中 Hyperband 部分与连续减半部分严格分离。所有不同于（Jamieson 和 Talwalkar，2016）提出的原始连续减半算法以及（Hyperband 算法（由Li 等人，2018）的部分，均以蓝色文字标示。

算法 2迭代加深 Hyperband (ID-HB)

输入： 最大规模 R，η $\geq$ 2，旧 最大尺寸 $\tilde{R}$ $\in$ $\{0, R/\eta\}$，旧的配置样本序列 $((C_{s,k})_{k\in\{0,\dots,s\}})_{s\in\{0,\dots,\log_\eta(\tilde{R})\}}$ 和损失值 $((L_{s,k})_{k\in\{0,\dots,s\}})_{s\in\{0,\dots,\log_\eta(\tilde{R})\}}$

初始化： $s_{max} = \lfloor \log_\eta(R) \rfloor$, $B = (s_{max}+1)R$

if $\tilde{R} > 0$**然后**

$\tilde{s}_{max} = \lfloor \log_\eta(\tilde{R}) \rfloor = s_{max} - 1$, $\tilde{B} = (\tilde{s}_{max}+1)$ $\tilde{R}$

结束 if

对于 $s \in \{s_{max}, s_{max}-1, \dots, 0\}$ **执行**

$n_s = \lceil \frac{B}{R} \frac{\eta^s}{(s+1)} \rceil$, $r_s = R/\eta^s$

if $\tilde{R} > 0$和 $s > 0$ **然后**

$\tilde{s} = s - 1$, $\tilde{n}_s = \lceil \frac{\tilde{B}}{\tilde{R}} \frac{\eta^{\tilde{s}}}{(\tilde{s}+1)} \rceil$, $\tilde{r}_s = \tilde{R}/\eta^{\tilde{s}} = r_s$

else

$\tilde{n}_s = 0$

结束如果

$S \leftarrow$ 样本 $n_s - \tilde{n}_s$ 配置

xID-SuccessiveHalving(S, B, r_s, R, η, $(C_{\tilde{s},\ k})_{k\in\{0,\dots,\tilde{s}\}}$, $(L_{\tilde{s},\ k})_{k\in\{0,\dots,\tilde{s}\}}$)

结束循环

输出： 中间损失最小的配置

算法 3高效迭代加深逐次减半（eID-SH）

输入： S 臂集合，预算 B, r，最大规模 R, η，$(C_k)_k$ 旧配置序列，$(L_k)_k$ 旧损失序列 初始化： $S_0 \leftarrow S$, $\tilde{n} = |C_0|$, $n = |S_0| + |C_0|$对于 $k \in \{0, 1, \dots, s\}$，执行$n_k = \lfloor n/\eta^k \rfloor - \lfloor \tilde{n}/\eta^k \rfloor$, $r_k = r\eta^k$在 S_k 中对每个臂进行拉取 r_k 次 如果 $k \leq s-1$，则$S_{k+1} \leftarrow$ 保留 $S_k \cup C_k \backslash C_{k+1}$ 中最好的 $\lfloor n/\eta^{k+1} \rfloor - \lfloor \tilde{n}/\eta^{k+1} \rfloor$ 个臂；否则$S_{k+1} \leftarrow$ 保留 $S_k \cup C_k$中最好的 $\lfloor n/\eta^{k+1} \rfloor$ 个臂 结束条件判断 结束循环 输出：剩余配置

算法 4 丢弃式迭代加深逐次减半（dID-SH）输入： S 臂集合、预算 B、 r，最大大小 R、 η，$(C_k)_k$ 旧配置序列，$(L_k)_k$ 旧损失序列初始化： $S_0 \leftarrow S \cup C_0$， $n = |S_0|$对于 $k \in \{0, 1, \dots, s\}$， 执行$n_k = \lfloor n/\eta^k \rfloor$， $r_k = r\eta^k$对 $S_k \backslash C_k$ 中的每个臂进行 r_k 次拉取$S_{k+1} \leftarrow$ 保留 $\lfloor n/\eta^{k+1} \rfloor$ 个最佳臂（来自 S_k）结束循环输出：剩余配置

算法 5 保持型迭代加深逐次减半（pID-SH）输入： S 臂集合、预算 B、 r，最大大小 R、 η，$(C_k)_k$ 旧配置序列，$(L_k)_k$ 旧损失序列初始化： $S_0 \leftarrow S \cup C_0$， $n = |S_0|$对于 $k \in \{0, 1, \ldots, s\}$，执行$n_k = \lfloor n/\eta^k \rfloor$， $r_k = r\eta^k$对 $S_k \backslash C_k$ 中的每个臂进行 r_k 次拉取$S_{k+1} \leftarrow$ 保留 $\lfloor n/\eta^{k+1} \rfloor$ 个最佳臂（来自 $S_k \cup C_k$）结束循环输出：剩余配置

B. 证明

B.1. 定理证明6.2

定理证明 *6.2*。 首先，我们将重点讨论算法 3 中给出的高效迭代加深逐次减半算法。

步骤 1：算法 thm 3 从不超出预算 B，如下所示。所用预算的上界为 y

$$\begin{aligned}\sum_{k=0}^{s} n_k r_k &= \sum_{k=0}^{s} \left(\lfloor n/\eta^k \rfloor - \lfloor \tilde{n}/\eta^k \rfloor\right) \frac{R\eta^k}{\eta^s} \\ &\leq \sum_{k=0}^{s} \left(\lfloor n/\eta^k \rfloor\right) \frac{R\eta^k}{\eta^s} \\ &\leq \sum_{k=0}^{s} \frac{(s_{\max}+1)\eta^s}{(s+1)} \frac{R}{\eta^s} \\ &\leq (s_{\max}+1)R = B.\end{aligned}$$

步骤 2：令 $n_k = |S_k| + |C_k|$ 和 $\tilde{n}_k = |C_k|$ 满足 $n_0 = n$ 和 $\tilde{n}_o = \tilde{n}_\circ$ 不失一般性，我们假设损失的极限值按如下顺序排列：$\nu_1 \leq \nu_2 \leq \cdots \leq \nu_{n\circ}$ 请注意，由于上述条件， S_k 中的臂的极限值以及 C_k 中的臂的极限值也按顺序排列，例如，对于 ν_i， $\nu_j \in S_k$，其中 $i < j$，我们有 $\nu_i \leq \nu_{j\circ}$ 下文令$i'_k = \min\{\lfloor n_k/\eta \rfloor + 1, \lfloor \tilde{n}_k/\eta \rfloor + \lfloor n_k/\eta \rfloor + 1\}_\circ$ 假设 $B \geq z_{eID-SH}$，则对于每一轮次 k

$$\begin{aligned}r_k &\geq \frac{B}{(n_k - \tilde{n}_k)\lceil \log_\eta(n) \rceil} - 1 \\ &\geq \frac{\eta}{n_k - \tilde{n}_k} \max_{i=2,\ldots,n} i\left(1 + \min\left\{R, \gamma^{-1}\left(\max\left\{\frac{\epsilon}{4}, \frac{\nu_i - \nu_1}{2}\right\}\right)\right\}\right) - 1 \\ &\geq \frac{\eta}{n_k - \tilde{n}_k} i'_k\left(1 + \min\left\{R, \gamma^{-1}\left(\max\left\{\frac{\epsilon}{4}, \frac{\nu_{i'_k} - \nu_1}{2}\right\}\right)\right\}\right) - 1 \\ &\overset{(*)}{\geq} \frac{\eta}{n_k - \tilde{n}_k} (n_k - \tilde{n}_k)/\eta\left(1 + \min\left\{R, \gamma^{-1}\left(\max\left\{\frac{\epsilon}{4}, \frac{\nu_{i'_k} - \nu_1}{2}\right\}\right)\right\}\right) - 1 \\ &= \left(1 + \min\left\{R, \gamma^{-1}\left(\max\left\{\frac{\epsilon}{4}, \frac{\nu_{i'_k} - \nu_1}{2}\right\}\right)\right\}\right) - 1 \\ &= \min\left\{R, \gamma^{-1}\left(\max\left\{\frac{\epsilon}{4}, \frac{\nu_{i'_k} - \nu_1}{2}\right\}\right)\right\},\end{aligned}$$

其中第四行（$*$）由以下内容得出：

- **情况 1：** $i'_k = \lfloor \tilde{n}_k/\eta \rfloor + 1_\circ$ 我们有 $i'_k \geq n_k/\eta \geq (n_k - \tilde{n}_k)/\eta_\circ$

- **情况 2：** $i'_k = \lfloor \tilde{n}_k/\eta \rfloor + \lfloor n_{k-1}/\eta \rfloor + 1$。如果 $\tilde{n}_k = 0$，我们有

$$i'_k = \left\lfloor \frac{n_{k-1}}{\eta} \right\rfloor + 1 \geq \frac{n_{k-1}}{\eta} \geq \frac{n_k}{\eta} \geq \frac{n_k - \tilde{n}_k}{\eta}.$$

如果 $\tilde{n}_k \geq 1$，我们有

$$\begin{aligned} i'_k &\geq \frac{\tilde{n}_k}{\eta} - 1 + \frac{n_{k-1}}{\eta} - 1 + 1 \\ &= \frac{\tilde{n}_k - n_{k-1} - \eta}{\eta} \\ &\geq \frac{n_k + (\eta - 1)n_k + \tilde{n}_k - \eta}{\eta} \\ &\geq \frac{n_k}{\eta} \geq \frac{n_k - \tilde{n}_k}{\eta}, \end{aligned}$$

其中，第 3 行由 $n_k = \lfloor n_{k-1}/\eta \rfloor$ 推出，第 4 行由 $n_k \geq \tilde{n}_k \geq 1$ 和 $\eta \geq 2$ 推出，因此我们有 $\eta - 1 \geq 1$，于是可以估计 $n_k \geq 1$。

接下来，我们证明对于所有 $t \geq \tau_i := \gamma^{-1}\left(\frac{\nu_i - \nu_1}{2}\right)$，都有 $\ell_{i,t} - \ell_{1,t} > 0$。根据 γ 的定义，对于所有 $i \in [n]$，我们有 $|\ell_{i,t} - \nu_i| \leq \gamma(t) \leq \frac{\nu_i - \nu_1}{2}$，其中最后一个不等式对 $t \geq \tau_i$ 成立。因此，对于 $t \geq \tau_i$，我们有

$$\begin{aligned} \ell_{i,t} - \ell_{1,t} &= \ell_{i,t} - \nu_i + \nu_i - \nu_1 + \nu_1 - \ell_{1,t} \\ &= \ell_{i,t} - \nu_i - (\ell_{1,t} - \nu_1) + \nu_i - \nu_1 \\ &\geq -2\gamma(t) + \nu_i - \nu_1 \\ &\geq -2\frac{\nu_i - \nu_1}{2} + \nu_i - \nu_1 \\ &= 0. \end{aligned}$$

在这种情况下，我们会先淘汰臂 i，再淘汰臂 1，因为每一轮都会按经验损失对各臂排序，并舍弃排名靠前的一半。注意，根据假设，ν_i 的界限随 i 非递减，因此 τ_i 的值随 i 非递增。

固定一个轮次 k，并假设 $1 \in S_k \cup C_k$（注意，$1 \in S_0 \cup C_0$）。上述计算表明

$$t \geq \tau_i \implies \ell_{i,t} \geq \ell_{1,t}. \tag{1}$$

接下来我们考虑两种不同的情形。

- **情况 1：** $k \leq s - 1$。在这种情况下，我们从集合 $S_k \cup C_k \backslash C_{k+1}$ 中保留最优的 $\lfloor n_k/\eta \rfloor - \lfloor \tilde{n}_k/\eta \rfloor$ 个臂，并且已经晋级了最优的

$\lfloor \tilde{n}_k/\eta \rfloor$来自 C_k。

$$
\begin{aligned}
&\{1 \in S_k \cup C_k,\ 1 \notin S_{k+1} \cup C_{k+1}\} \\
&\quad \Longleftrightarrow \left\{ \sum_{i \in S_k \cup C_k \backslash C_{k+1}} \mathbb{1}\{\ell_{i,r_k} < \ell_{1,r_k}\} \geq \lfloor n_k/\eta \rfloor - \lfloor \tilde{n}_k/\eta \rfloor, \sum_{i \in C_k} \mathbb{1}\{\ell_{i,r_k} < \ell_{1,r_k}\} \geq \lfloor \tilde{n}_k/\eta \rfloor \right\} \\
&\quad \Longrightarrow \left\{ \sum_{i \in S_k \cup C_k \backslash C_{k+1}} \mathbb{1}\{r_k < \tau_i\} \geq \lfloor n_k/\eta \rfloor - \lfloor \tilde{n}_k/\eta \rfloor, \sum_{i \in C_k} \mathbb{1}\{r_k < \tau_i\} \geq \lfloor \tilde{n}_k/\eta \rfloor \right\} \\
&\quad \Longrightarrow \left\{ \sum_{i=2}^{\lfloor n_k/\eta \rfloor - \lfloor \tilde{n}_k/\eta \rfloor + \lfloor \tilde{n}_k/\eta \rfloor + 1} \mathbb{1}\{r_k < \tau_i \wedge i \in S_k \cup C_k \backslash C_{k+1}\} \geq \lfloor n_k/\eta \rfloor - \lfloor \tilde{n}_k/\eta \rfloor, \right. \\
&\qquad \left. \sum_{i=2}^{\lfloor \tilde{n}_k/\eta \rfloor + \lfloor n_{k-1}/\eta \rfloor + 1} \mathbb{1}\{r_k < \tau_i \wedge i \in C_k\} \geq \lfloor \tilde{n}_k/\eta \rfloor \right\} \\
&\quad \Longrightarrow \left\{ r_k < \min\left\{ \tau_{\lfloor n_k/\eta \rfloor + 1}, \tau_{\lfloor \tilde{n}_k/\eta \rfloor + \lfloor n_{k-1}/\eta \rfloor + 1} \right\} \right\} \\
&\quad \Longleftrightarrow \left\{ r_k < \tau_{\max\{\lfloor n_k/\eta \rfloor + 1, \lfloor \tilde{n}_k/\eta \rfloor + \lfloor n_{k-1}/\eta \rfloor + 1\}} \right\},
\end{aligned}
$$

第一行根据算法的定义得出，第二行则由公式1得出。第三行中，我们假设处于最坏情形：$\lfloor n_k/\eta \rfloor - \lfloor \tilde{n}_k/\eta \rfloor$ 中最好的臂 $S_k \cup C_k \backslash C_{k+1}$ 都劣于$\lfloor \tilde{n}_k/\eta \rfloor$ 中的最佳臂 C_k（这些臂会保留在集合 C_{k+1}中）；反过来，$\lfloor \tilde{n}_k/\eta \rfloor$ 中的最佳臂 C_k 也劣于 S_k中的所有臂。第四行是因为 τ_i 单调不增（对所有 $i < j$，都有 $\tau_i \geq \tau_j$，因此$\mathbb{1}\{r_k < \tau_i\} \geq \mathbb{1}\{r_k < \tau_j\}$，所以求和中的第一个指标（不包括1）会先于其他任何 i在$S_k \subset [n]$中零散出现的 $[n]$顶开启。

- **情况 2：** $k = s$。

  在这种情况下，我们 保留最优的 $\lfloor n_k/\eta \rfloor$ 臂，并从 $S_k \cup C_k$ 中取出，且有 $C_{k+1} = \emptyset$，因此可以像上文一样得到 ve

$$
\begin{aligned}
\{1 \in S_k \cup C_k,\ 1 \notin S_{k+1}\} &\Longleftrightarrow \left\{ \sum_{i \in S_k \cup C_k} \mathbb{1}\{\ell_{i,r_k} < \ell_{1,r_k}\} \geq \lfloor n_k/\eta \rfloor \right\} \\
&\Longrightarrow \left\{ \sum_{i \in S_k \cup C_k} \mathbb{1}\{r_k < \tau_i\} \geq \lfloor n_k/\eta \rfloor \right\} \\
&\Longrightarrow \left\{ \sum_{i=2}^{\lfloor n_k/\eta \rfloor + 1} \mathbb{1}\{r_k < \tau_i\} \geq \lfloor n_k/\eta \rfloor \right\} \\
&\Longleftrightarrow \left\{ r_k < \tau_{\lfloor n_k/\eta \rfloor + 1} \right\}.
\end{aligned}
$$

总之，我们可以得出结论：$1 \in S_k \cup C_k$ 且 $1 \notin S_{k+1} \cup C_{k+1}$ 如果
$r_k < \tau_{\max\{\lfloor n_k/\eta \rfloor + 1, \lfloor \tilde{n}_k/\eta \rfloor + \lfloor n_k/\eta \rfloor + 1\}}$。这意味着

$$
\{1 \in S_k \cup C_k,\ r_k \geq \tau_{i'_k}\} \implies \{1 \in S_{k+1} \cup C_{k+1}\}. \tag{2}
$$

回想一下，$r_k \geq \gamma^{-1}\big($最大值 $\{\frac{\epsilon}{4}, \frac{\nu_{i'_k} - \nu_1}{2}\}\big)$ 且
$\tau_{i'_k} = \gamma^{-1}\big(\frac{\nu_{i'_k} - \nu_1}{2}\big)$，我们考察以下三个穷尽情形：

- **Case 1:** $\frac{\nu_{i'_k} - \nu_1}{2} \geq \frac{\epsilon}{4}$ and $1 \in S_k \cup C_k$

  In this case, $r_k \geq \gamma^{-1}\big(\frac{\nu_{i'_k} - \nu_1}{2}\big) = \tau_{i'_k}$. By Equation 2 we have that $1 \in S_{k+1} \cup C_{k+1}$ since $1 \in S_k \cup C_k$.

- **Case 2:** $\frac{\nu_{i'_k} - \nu_1}{2} < \frac{\epsilon}{4}$ and $1 \in S_k \cup C_k$

在此情形下， $r_k \geq \gamma^{-1}\left(\frac{\epsilon}{4}\right)$ 但 $\gamma^{-1}\left(\frac{\epsilon}{4}\right) < \tau_{i'_k}$。方程2表明， $1 \in S_k \cup C_k$ 但 $1 \notin S_{k+1} \cup C_{k+1}$。在 $1 \in S_{k+1} \cup C_{k+1}$ 这一良好事件发生时，算法会继续执行，而在下一轮次中，情况 1 或情况 2 都可能成立。因此，假设 $1 \notin S_{k+1} \cup C_{k+1}$。这里我们证明 $\{1 \in S_k \cup C_k,\ 1 \notin S_{k+1} \cup C_{k+1}\} \overset{\frown}{\Longrightarrow}$ 最大值 $_{i \in S_{k+1} \cup C_{k+1}} \nu_i \leq \nu_1 + \epsilon/2$。由于 $1 \in S_0 \cup C_0$，这保证算法3 要么以臂 $i = 1$ 退出，要么存在某个臂 $\widehat{i}$ 满足 $\nu_{\widehat{i}} \leq \nu_1 + \epsilon/2$。

令 $p = \min\{i \in [n] : \frac{\nu_i - \nu_1}{2} \geq \frac{\epsilon}{4}\}$。注意，根据该情况的判据， $p > i'_k$

$$r_k \geq \gamma^{-1}\left(\frac{\epsilon}{4}\right) \geq \gamma^{-1}\left(\frac{\nu_i - \nu_1}{2}\right) = \tau_i, \quad \forall i \geq p.$$

因此，由方程 1($t \geq \tau_i \implies \ell_{i,t} \geq \ell_{1,t}$) 可知，臂 $i \geq p$ 始终会满足 $\ell_{i,r_k} \geq \ell_{1,r_k}$，并且在臂 1 被淘汰之前或同时被淘汰，前提是 $1 \in S_k \cup C_k$。总之，如果臂 1 被淘汰，使得 $1 \in S_k \cup C_k$ 但 $1 \notin S_{k+1} \cup C_{k+1}$，那么根据 p 的定义，$\max_{i \in S_{k+1} \cup C_{k+1}} \nu_i \leq \max_{i<p} \nu_i < \nu_1 + \epsilon/2$。

- **情况 3：** $1 \notin S_k \cup C_k$

  由于 $1 \in S_0 \cup C_0$，存在某个 $r < k$，使得 $1 \in S_r \cup C_r$ 且 $1 \notin S_{r+1} \cup C_{r+1}$。对于这个 r，只能是情况 2，因为情况 1 会使 $1 \in S_{r+1} \cup C_{r+1}$ 激增。然而，在情况 2 下，如果 $1 \notin S_{r+1} \cup C_{r+1}$，那么 $\max_{i \in S_{r+1} \cup C_{r+1}} \nu_i \leq \nu_1 + \epsilon/2$。

由于 $1 \in S_0 \cup C_0$，我们要么有：对于所有 k，直到 $\widehat{\ }$，臂 1 始终留在 $S_k \cup C_k$ 中（可能在情况 1 和情况 2 之间交替），此后算法以最优臂 1 退出；要么存在某个 k，使得情况 3 成立，且算法以某个臂 i 退出，并且 $\nu_{\widehat{i}} \leq \nu_1 + \epsilon/2$。

接下来，我们将证明，对于4号算法和5号算法中给出的丢弃型和保留型迭代加深-连续减半算法，同样的保证也成立。

因此，我们分两步进行：首先，将 dID-SH 算法化简为连续减半算法，以沿用其理论保证。其次，我们将说明需要如何修改连续减半算法的证明，才能为我们的 pID-SH 算法获得相同的理论保证。

步骤 1：接下来，我们将区分两种不同情况，以便将丢弃型迭代加深-连续减半算法4 化简为（Jamieson 和 Talwalkar，2016）（或（Karnin 等人，2013）提出的原始连续减半算法。

- **情况 1：** $(C_k)_k = \emptyset$。

  如果我们有 $(C_k)_k = \emptyset$，那么我们得到的就是（Jamieson 和 Talwalkar，2016）提出的连续减半算法，因此可以沿用其理论保证。

- **情况 2：** $(C_k)_k \neq \emptyset$.

  因此，接下来我们要考虑的有趣情况是情况 $(C_k)_k \neq \emptyset$。假设算法 4 被算法 2 作为子程序调用。由于 $(C_k)_k \neq \emptyset$，之前已经以臂数 $\tilde{n}$ 和预算 $\tilde{r}_s = \tilde{R}/\eta^{\tilde{s}} = \frac{R}{\eta}/\eta^{s-1} = R/\eta^s = r_k$ 为 $s \in \{0, \ldots, \lfloor \log_\eta(R) \rfloor\}$ 调用过算法 4。因此，$(C_k)_k$ 中的臂已经被拉取了 r_k 次，并且观测到了它们的损失值 $(L_k)_k$。将这些损失值与算法 4 中每次迭代 k 时对 $S_k \backslash C_k$ 中各臂进行 r_k 次拉取所观测到的损失值结合起来，我们可以根据对 $S_k \backslash C_k$ 最近拉取所观测到的损失，以及对 C_k 此前观测到的损失，从 S_k 中保留最好的 $\lfloor n/\eta^{k+1} \rfloor$ 个臂。因此，我们得到的 S_{k+1} 中的臂，与从头开始运行算法 5、使用 $(C_k)_k = \emptyset$ 和 $S = S \cup C_0$ 所得到的臂相同，因而可以应用情况 1。

总结这两种情况，我们可以保留由（Li 等人，2018）针对有限时域设置下的原始版本连续减半证明的理论结果（$R < \infty$）。

步骤 2：为使保持型迭代加深逐次减半算法也能获得相同的保证，我们可以采用与（Li 等人，2018）证明连续减半时类似的方法。对于固定的轮次 k 和 $1 \in S_k \cup C_k$，由于 $1 \in S_0 \cup C_0$，

我们有

$$\begin{aligned}\{1 \in S_k \cup C_k, 1 \notin S_{k+1}\} &\Leftrightarrow \left\{ \sum_{i \in S_k \cup C_k} \mathbb{1}\{\ell_{i,r_k} < \ell_{1,r_k}\} \geq \lfloor n_k/\eta \rfloor \right\} \\ &\Rightarrow \left\{ \sum_{i \in S_k \cup C_k} \mathbb{1}\{r_k < \tau_i\} \geq \lfloor n_k/\eta \rfloor \right\} \\ &\Rightarrow \left\{ \sum_{i=2}^{n_k + |C_k \setminus (S_k \cap C_k)| + 1} \mathbb{1}\{r_k < \tau_i\} \geq \lfloor n_k/\eta \rfloor \right\} \\ &\Leftrightarrow \{r_k < \tau_{\lfloor n_k/\eta \rfloor + 1}\}.\end{aligned}$$

证明的其余部分与（Li 等人，2018）中关于连续减半的证明相同。 □

B.2. 逐步减半与 ID-SH 的比较

定理 *6.3* 的证明。 首先，比较运行 SuccessiveHalving(n, r) 时的总拉取次数与运行 xID-SuccessiveHalving(n, r) 时的总拉取次数，其中假设我们已经运行过 SuccessiveHalving($\tilde{n}, \tilde{r}$)。以下我们重点讨论 SuccessiveHalving(n, r) 的拉取次数下界。

$$\begin{aligned}\sum_{k=0}^{s} n_k r_k &= \sum_{k=0}^{s} \left\lfloor \frac{n}{\eta^k} \right\rfloor \cdot \left\lfloor \frac{R\eta^k}{\eta^s} \right\rfloor \\ &\geq \sum_{k=0}^{s} \left(\frac{n}{\eta^k} - 1 \right) \left(\frac{R\eta^k}{\eta^s} - 1 \right) \\ &= \sum_{k=0}^{s} \frac{nR}{\eta^s} - \frac{R\eta^k}{\eta^s} - \frac{n}{\eta^k} + 1 \\ &= \frac{(s+1)(nR+\eta^s)}{\eta^s} - \frac{R}{\eta^s} \sum_{k=0}^{s} \eta^k - n \sum_{k=0}^{s} \left(\frac{1}{\eta} \right)^k \\ &= \frac{(s+1)(nR+\eta^s)}{\eta^s} - \frac{R(\eta^{s+1}-1)}{\eta^s(\eta-1)} - \underbrace{\frac{n(1-(1/\eta)^{s+1})}{1-1/\eta}}_{= \frac{n\left(\frac{\eta^{s+1}-1}{\eta^{s+1}}\right)}{\frac{\eta-1}{\eta}} = \frac{n(\eta^{s+1}-1)}{\eta^s(\eta-1)}} \\ &= \frac{(s+1)(nR+\eta^s)}{\eta^s} - \frac{(\eta^{s+1}-1)(R+n)}{\eta^s(\eta-1)} \\ &= \frac{(s+1)(nR+\eta^s)(\eta-1) - (\eta^{s+1}-1)(R+n)}{\eta^s(\eta-1)},\end{aligned}$$

其中第五行使用了几何级数的闭式表达，而其他各行则进行了简单的变换。

a) eID-SH 的拉取次数与逐步减半的比较。

为了进行比较　此外，我们还必须计算 eID-SuccessiveHalving(n, r) 的总拉取次数上界：

$$\begin{aligned}\sum_{k=0}^{s} n_k r_k &= \sum_{k=0}^{s} \left(\lfloor n/\eta^k \rfloor - \lfloor \tilde{n}/\eta^k \rfloor\right) \left\lfloor \frac{R\eta^k}{\eta^s} \right\rfloor \\ &\leq \sum_{k=0}^{s} \left(\frac{n-\tilde{n}}{\eta^k} + 1\right) \cdot \frac{R\eta^k}{\eta^s} \\ &= \frac{(s+1)(n-\tilde{n})R}{\eta^s} + \frac{R}{\eta^s}\sum_{k=0}^{s} \eta^k \\ &= \frac{(s+1)(n-\tilde{n})R}{\eta^s} + \frac{R(\eta^{s+1}-1)}{\eta^s(\eta-1)} \\ &= \frac{(s+1)(n-\tilde{n})R(\eta-1) + R(\eta^{s+1}-1)}{\eta^s(\eta-1)}.\end{aligned}$$

为了比较两者，我们构造商

$$\begin{aligned}\frac{\#\{\text{total pulls of eID-SH}(n,r)\}}{\#\{\text{total pulls of SH}(n,r)\}} &\leq \frac{(s+1)(n-\tilde{n})R(\eta-1) + R(\eta^{s+1}-1)}{(s+1)(nR+\eta^s)(\eta-1) - (\eta^{s+1}-1)(R+n)} \\ &= 1 - \frac{(s+1)(\tilde{n}R+\eta^s)(\eta-1) - (\eta^{s+1}-1)(2R+n)}{(s+1)(nR+\eta^s)(\eta-1) - (\eta^{s+1}-1)(R+n)}.\end{aligned}$$

b) dID-SH 的拉取次数与 SH 的比较。

与 a) 类似，我们首先需要得到 dID-SuccessiveHalving(n, r) 一次运行中总拉取次数的上界。虽然我们只在 dID-SH 的第一轮采样 $n-\tilde{n}$ 个新臂，但最优的 n/η 个臂可能全部来自新采样的臂，因此在 SH($\tilde{n}$, $\tilde{r}$) 的运行中，被保留到 dID-SH 下一轮的臂都可能尚未以更高预算拉取过。在这种最坏情况下，我们可以估算

$$\begin{aligned}\sum_{k=0}^{s} n_k r_k &= (n-\tilde{n}) \left\lfloor \frac{R}{\eta^s} \right\rfloor + \sum_{k=1}^{s} \left\lfloor \frac{n}{\eta^k} \right\rfloor \left\lfloor \frac{R\eta^k}{\eta^s} \right\rfloor \\ &\leq \frac{(n-\tilde{n})R}{\eta^s} + \sum_{k=1}^{s} \frac{nR}{\eta^s} \\ &= \frac{(n-\tilde{n})R}{\eta^s} + \frac{snR}{\eta^s} \\ &= \frac{R((s+1)n-\tilde{n})}{\eta^s}.\end{aligned}$$

同样，我们现在可以如下计算拉取次数的商。

$$\begin{aligned}\frac{\#\{\text{total pulls of dID-SH}(n,r)\}}{\#\{\text{total pulls of SH}(n,r)\}} &\leq \frac{(\eta-1)R((s+1)n-\tilde{n})}{(\eta-1)(s+1)(nR+\eta^s) - (\eta^{s+1}-1)(R+n)} \\ &= 1 - \frac{(\eta-1)((s+1)\eta^s + R\tilde{n}) - (\eta^{s+1}-1)(R+n)}{(\eta-1)(s+1)(nR+\eta^s) - (\eta^{s+1}-1)(R+n)}.\end{aligned}$$

请注意，对于 pID-SH 的拉取次数，我们也可以进行同样的处理，因为这里存在相同的最坏情形：在 pID-SH 的下一轮中，我们只保留新采样的配置，而不保留任何先前晋升的配置。 □

为了直观了解总拉取次数的改进情况，我们在图 4 和图 5 中展示了不同轮次数 s、每轮最大预算 R 以及淘汰比例 η下的上述各项。请注意，上述结果假设了 pID-SH 和 dID-SH 算法的最坏情形，即所有先前晋升的配置都劣于所有新采样的配置。在大多数问题场景中，pID-SH 和 dID-SH 的总拉取次数的平均改进幅度将位于图 4 所示最坏情形曲线与最佳情形之间，而最佳情形与 eID-SH 一致，并展示于图 5 和2中。由于我们提出的方法所需的总拉取次数绝不会多于逐步减半，因此我们绘制了 1 与定理 6.3中推导出的分数的最小值。

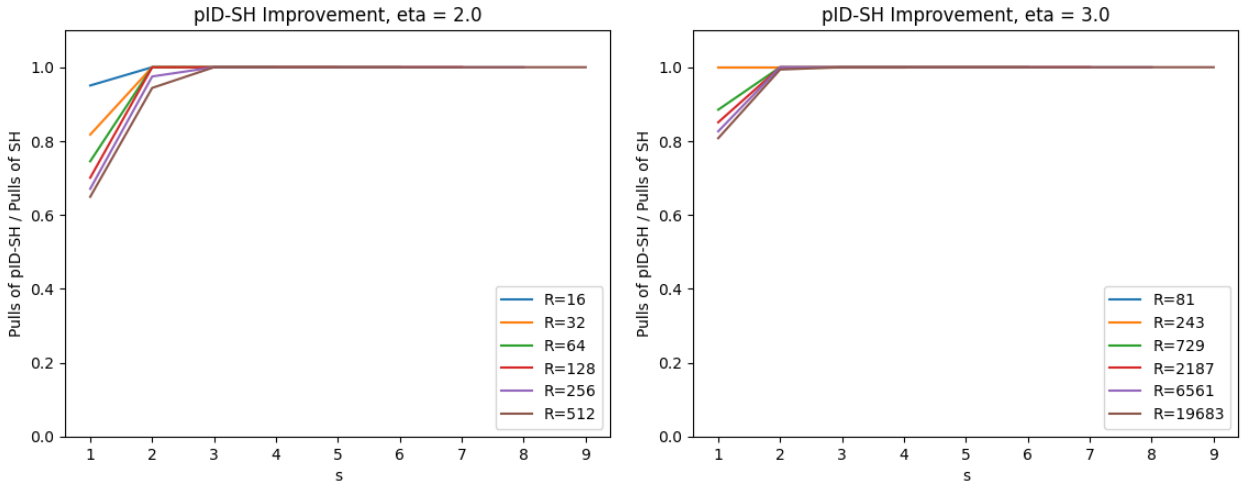

图*4*。 在不同的 SH 轮数 s、每轮最大预算 R 和丢弃比例 η下，pID-SH（分别为 dID-SH）与逐步减半的总拉取次数之比。

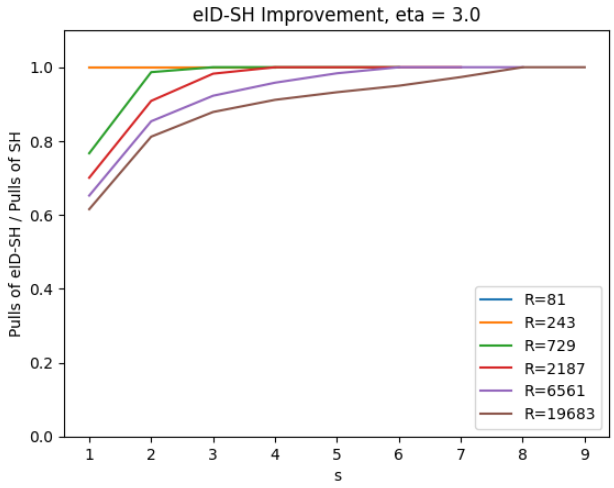

图*5*。 在不同的 SH 轮数 s 和每轮最大预算 R下，eID-SH 的总拉取次数之比。

B.3. 迭代加深 Hyperband

定理 6.5 的证明。 要推导算法 2中 ID-HB 所需的预算，只需将每次调用 xID-SH 所需的预算相加即可。幸运的是，eID-SH、dID-SH 和 pID-SH 所需的预算并无差异，因此，无论调用哪一种连续减半算法的变体，ID-HB 的一次运行总共使用的预算为

$$
\begin{aligned}
&\sum_{s=0}^{\lfloor \log_\eta(R) \rfloor} \text{Budget_xID-SH}(n_s, r_s) \\
&\quad = \sum_{s=0}^{\lfloor \log_\eta(R) \rfloor} \eta \left\lceil \log_\eta \left(\left\lceil (\lfloor \log_\eta(R) \rfloor + 1) \cdot \frac{\eta^s}{s+1} \right\rceil \right) \right\rceil \cdot \max_{i=2,\dots,n_s} i \left(1 + \min \left\{ R, \gamma^{-1} \left(\max \left\{ \frac{\epsilon}{4}, \frac{\nu_i - \nu_1}{2} \right\} \right) \right\} \right) \\
&\quad = (*).
\end{aligned}
$$

通过简单的估算和变换，我们得到

$$
\begin{aligned}
\eta \left\lceil \log_\eta \left(\left\lceil (\lfloor \log_\eta(R) \rfloor + 1) \cdot \frac{\eta^s}{s+1} \right\rceil \right) \right\rceil &\leq \eta \left\lceil \log_\eta \left(\left\lceil (\log_\eta(R) + 1) \right\rceil \cdot \left\lceil \frac{\eta^s}{s+1} \right\rceil \right) \right\rceil \\
&= \eta \left\lceil \log_\eta \left(\left\lceil \log_\eta(R) + 1 \right\rceil \right) + \log_\eta \left(\left\lceil \frac{\eta^s}{s+1} \right\rceil \right) \right\rceil \\
&\leq \eta \left\lceil \log_\eta \left(\log_\eta(R) + 2 \right) + \log_\eta \left(\frac{\eta^s}{s+1} + 1 \right) \right\rceil \\
&\leq \eta \left\lceil \log_\eta \left(\log_\eta(R) \right) + 2 + \log_\eta \left(\frac{\eta^s}{s+1} \right) + 1 \right\rceil \\
&= \eta \left\lceil \log_\eta \left(\log_\eta(R) \right) + \log_\eta \left(\eta^s \right) - \log_\eta \left(s+1 \right) + 3 \right\rceil \\
&\leq \eta \left(\log_\eta \left(\log_\eta(R) \right) + s - \log_\eta \left(s+1 \right) + 4 \right).
\end{aligned}
$$

请注意，第四行由以下内容得出

$$
\begin{aligned}
&\log_\eta(x+1) \leq \log_\eta(x) + 1 \\
&\Leftrightarrow x + 1 \leq \eta \cdot x \\
&\Leftrightarrow x \geq \frac{1}{\eta - 1}.
\end{aligned}
$$

在我们的设定中，我们有 $\eta \geq 2$，因此 $\log_\eta(x+1) \geq \log_\eta(x) + 1$ 当且仅当 $x \geq 1$。对于 $s \geq 0$，我们有 $\frac{\eta^s}{s+1} \geq 1$，并且不失一般性地，$\log_\eta(R) \geq 2$；否则，$s_{max}$ 的值以及 Hyperband 的运行都会变得毫无意义。我们可以继续

$$
\begin{aligned}
(*) \leq &\sum_{s=0}^{\lfloor \log_\eta(R) \rfloor} \eta \left(\log_\eta \left(\log_\eta(R) \right) + s - \log_\eta \left(s+1 \right) + 4 \right) \\
&\cdot \underbrace{\max_{s=0,\dots,\lfloor \log_\eta(R) \rfloor} \max_{i=2,\dots,n_s} i \left(1 + \min \left\{ R, \gamma^{-1} \left(\max \left\{ \frac{\epsilon}{4}, \frac{\nu_i - \nu_1}{2} \right\} \right) \right\} \right)}_{=:\bar{\gamma}^{-1}} \\
= \eta &\left(\left(\lfloor \log_\eta(R) \rfloor + 1 \right) \left(\log_\eta(\log_\eta(R)) + 4 \right) + \sum_{s=0}^{\lfloor \log_\eta(R) \rfloor} s - \sum_{s=0}^{\lfloor \log_\eta(R) \rfloor} \log_\eta(s+1) \right) \bar{\gamma}^{-1} \\
= \eta &\left(\left(\lfloor \log_\eta(R) \rfloor + 1 \right) \left(\log_\eta(\log_\eta(R)) + 4 \right) + \frac{\lfloor \log_\eta(R) \rfloor \left(\lfloor \log_\eta(R) \rfloor + 1 \right)}{2} - \log_\eta \left(\prod_{s=0}^{\lfloor \log_\eta(R) \rfloor} (s+1) \right) \right) \bar{\gamma}^{-1} \\
= \eta &\left(\left(\lfloor \log_\eta(R) \rfloor + 1 \right) \left(\log_\eta(\log_\eta(R)) + 4 \right) + \frac{\lfloor \log_\eta(R) \rfloor \left(\lfloor \log_\eta(R) \rfloor + 1 \right)}{2} - \log_\eta \left(\left(\lfloor \log_\eta(R) \rfloor + 1 \right)! \right) \right) \bar{\gamma}^{-1}.
\end{aligned}
$$

由于我们在 ID-HB 算法中选择预算 B 为 $B = (s_{\max}+1)R = \left(\lfloor \log_\eta(R) \rfloor + 1\right) R$，我们可以将两边同时除以 $\left(\lfloor \log_\eta(R) \rfloor + 1\right)$，得到

$$R \geq \eta \left(\log_\eta(\log_\eta(R)) + 4 + \frac{\lfloor \log_\eta(R) \rfloor}{2} - \frac{\log_\eta\left(\left(\lfloor \log_\eta(R) \rfloor + 1\right)!\right)}{\lfloor \log_\eta(R) \rfloor + 1} \right) \bar{\gamma}^{-1}.$$

回想一下，在 ID-HB 的第 s 轮中，每次调用 xID-SH 时，我们会比较 $n_s = (\lfloor \log_\eta(R) \rfloor + 1)$ 个超参数配置，因此样本总数为

$$\begin{aligned}
&\left(\lfloor \log_\eta(R) \rfloor + 1\right) \sum_{s=0}^{\lfloor \log_\eta(R) \rfloor} \frac{\eta^s}{s+1} \\
&\geq \sum_{s=0}^{\lfloor \log_\eta(R) \rfloor} \eta^s \\
&\overset{\text{Geometric Sum}}{=} \frac{\eta^{\lfloor \log_\eta(R) \rfloor + 1} - 1}{\eta - 1} \\
&\geq \frac{R-1}{\eta-1}.
\end{aligned}$$

根据假设 6.4 我们在样本集中以至少 ϵ的概率拥有一个满足 - 最优的超参数配置，当且仅当 $1-\delta$

$$\begin{aligned}
&\frac{R-1}{\eta-1} \geq \left\lceil \log_{1-\alpha}(\delta) \right\rceil \\
\Leftrightarrow \quad &R \geq \left\lceil \log_{1-\alpha}(\delta) \right\rceil (\eta - 1) + 1.
\end{aligned}$$

□

C. 详细实证结果

本节将更详细地介绍实证研究的结果。为此，我们从不同方面考察神经网络的配置：针对分类问题调整学习算法的超参数、针对多目标预测问题调整学习算法的超参数，以及搜索神经网络架构。

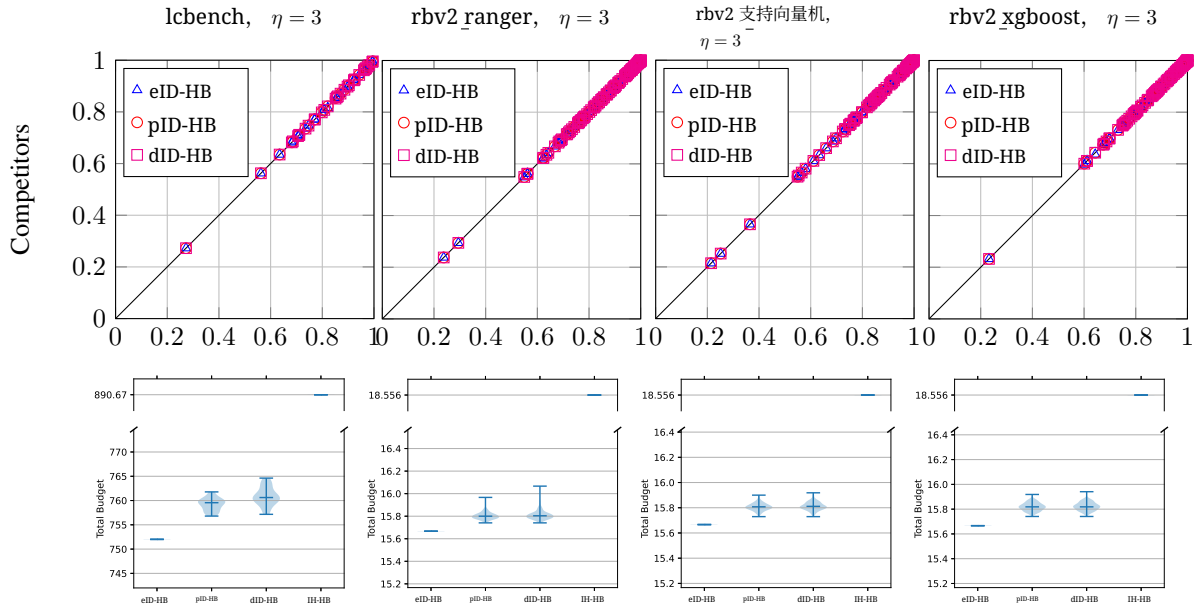

图 *6*。ID-HB 与原始 Hyperband 版本的比较。上：散点图展示通过 ID-HB 策略获得的最终当前最优解的准确率与 IH-HB 的对比。下：小提琴图展示单次运行所消耗的平均总预算。

nb301, $\eta = 2$				
OpenML ID	IH-HB	dID-HB	pID-HB	eID-HB
CIFAR10	.906±.00	.906±.00	.906±.00	.906±.00

nb301, $\eta = 3$				
OpenML ID	IH-HB	dID-HB	pID-HB	eID-HB
CIFAR10	.939±.00	.939±.00	.939±.00	.939±.00

rbv2_ranger, $\eta = 2$				
OpenML ID	IH-HB	dID-HB	pID-HB	eID-HB
3	.995±.00	.995±.00	.995±.00	.995±.00
6	.950±.01	.950±.01	.950±.01	.950±.01
11	.899±.01	.899±.01	.899±.01	.899±.01
12	.972±.00	.972±.00	.972±.00	.972±.00
14	.836±.01	.836±.01	.836±.01	.836±.01
15	.978±.01	.978±.01	.978±.01	.978±.01
16	.951±.00	.951±.00	.951±.00	.951±.00
18	.721±.02	.721±.02	.721±.02	.721±.02
22	.785±.01	.785±.01	.785±.01	.785±.01
23	.565±.02	.565±.02	.565±.02	.565±.02
24	.999±.00	.999±.00	.999±.00	.999±.00

28	.977±.00	.977±.00	.977±.00	.978±.00
29	.881±.03	.881±.03	.881±.03	.879±.03
31	.763±.01	.763±.01	.763±.01	.763±.01
32	.993±.00	.993±.00	.993±.00	.993±.00
37	.780±.03	.780±.03	.780±.03	.780±.03
38	.993±.00	.993±.00	.993±.00	.993±.00
42	.968±.01	.968±.01	.968±.01	.968±.01
44	.949±.00	.949±.00	.949±.00	.949±.00
46	.965±.00	.965±.00	.965±.00	.965±.00
50	.994±.00	.994±.00	.994±.00	.994±.00
54	.761±.02	.761±.02	.761±.02	.761±.02
60	.860±.01	.860±.01	.860±.01	.860±.01
151	.906±.01	.906±.01	.906±.01	.905±.01
181	.628±.03	.628±.03	.631±.03	.629±.03
182	.919±.01	.919±.01	.919±.01	.919±.01
188	.719±.01	.719±.01	.719±.01	.718±.01
300	.946±.01	.946±.01	.946±.01	.946±.01
307	.943±.01	.943±.01	.943±.01	.943±.01
312	.985±.00	.985±.00	.986±.00	.986±.00
334	.815±.11	.815±.11	.815±.11	.815±.11
375	.968±.00	.968±.00	.968±.00	.968±.00
377	.988±.00	.988±.00	.988±.00	.988±.00
458	.986±.00	.986±.00	.986±.00	.986±.00
469	.239±.03	.239±.03	.239±.03	.239±.03
470	.684±.01	.685±.01	.685±.01	.685±.01
554	.999±.00	.999±.00	.999±.00	.999±.00
1040	.998±.00	.998±.00	.998±.00	.998±.00
1049	.914±.01	.914±.01	.914±.01	.914±.01
1050	.903±.00	.903±.00	.903±.00	.903±.00
1053	.822±.01	.822±.01	.822±.01	.822±.01
1056	.994±.00	.994±.00	.994±.00	.994±.00
1063	.850±.03	.850±.03	.850±.03	.850±.03
1067	.870±.01	.870±.01	.870±.01	.870±.01
1068	.941±.00	.941±.00	.941±.00	.940±.00
1111	.000±.00	.000±.00	.000±.00	.000±.00
1220	.840±.00	.840±.00	.840±.00	.840±.00
1457	.773±.03	.773±.03	.772±.03	.773±.03
1461	.918±.01	.918±.01	.918±.01	.918±.01
1462	.997±.00	.997±.00	.997±.00	.997±.00
1464	.786±.01	.786±.01	.785±.01	.786±.01
1468	.952±.01	.952±.01	.952±.01	.952±.01
1475	.639±.01	.639±.01	.639±.01	.639±.01
1476	.996±.00	.996±.00	.996±.00	.996±.00
1478	.979±.00	.979±.00	.979±.00	.979±.00
1479	.573±.03	.573±.03	.573±.03	.573±.03
1480	.725±.02	.725±.02	.725±.02	.725±.02
1485	.873±.01	.873±.01	.873±.01	.873±.01
1486	.967±.00	.967±.00	.967±.00	.967±.00
1487	.943±.00	.943±.00	.943±.00	.943±.00
1489	.902±.01	.902±.01	.902±.01	.902±.01
1493	.831±.02	.831±.02	.831±.02	.830±.02
1494	.873±.01	.873±.01	.873±.01	.873±.01
1497	.997±.00	.997±.00	.997±.00	.997±.00
1501	.932±.01	.932±.01	.932±.01	.932±.01

1510	.967±.01	.967±.01	.967±.01	.967±.01
1515	.898±.02	.898±.02	.898±.02	.898±.02
1590	.889±.01	.889±.01	.889±.01	.889±.01
4134	.830±.02	.830±.02	.830±.02	.830±.02
4135	.963±.01	.963±.01	.963±.01	.962±.01
4154	.998±.00	.998±.00	.998±.00	.998±.00
4534	.961±.00	.961±.00	.961±.00	.961±.00
4538	.673±.02	.673±.02	.673±.02	.673±.02
4541	.863±.12	.863±.12	.863±.12	.863±.12
6332	.810±.05	.810±.05	.810±.05	.809±.05
23381	.624±.02	.624±.02	.624±.02	.625±.02
23512	.810±.05	.810±.05	.810±.05	.810±.05
23517	.564±.02	.564±.02	.564±.02	.564±.02
40496	.745±.03	.745±.03	.745±.03	.745±.03
40498	.700±.02	.700±.02	.700±.02	.700±.02
40499	.979±.00	.979±.00	.979±.00	.980±.00
40536	.913±.01	.913±.01	.913±.01	.913±.01
40668	.846±.02	.846±.02	.846±.02	.846±.02
40670	.959±.00	.959±.00	.959±.00	.959±.00
40685	.997±.00	.997±.00	.997±.00	.997±.00
40701	.967±.00	.967±.00	.967±.00	.966±.00
40900	.995±.00	.995±.00	.995±.00	.995±.00
40923	.000±.00	.000±.00	.000±.00	.000±.00
40927	.000±.00	.000±.00	.000±.00	.000±.00
40966	.980±.01	.980±.01	.980±.01	.980±.01
40975	.977±.01	.977±.01	.977±.01	.977±.01
40978	.979±.00	.979±.00	.979±.00	.979±.00
40979	.979±.00	.979±.00	.979±.00	.979±.00
40981	.885±.02	.885±.02	.884±.02	.884±.02
40982	.796±.01	.796±.01	.796±.01	.796±.01
40983	.990±.00	.990±.00	.990±.00	.990±.00
40984	.976±.00	.976±.00	.976±.00	.976±.00
40994	.925±.01	.925±.01	.925±.01	.925±.01
40996	.999±.00	.999±.00	.999±.00	.999±.00
41027	.856±.01	.856±.01	.857±.01	.857±.01
41138	.998±.00	.998±.00	.998±.00	.998±.00
41142	.785±.02	.785±.02	.785±.02	.785±.02
41143	.821±.00	.821±.00	.822±.01	.822±.01
41146	.943±.00	.943±.00	.943±.00	.943±.00
41150	.998±.00	.998±.00	.998±.00	.998±.00
41156	.858±.01	.858±.01	.858±.01	.858±.01
41157	.772±.14	.772±.14	.773±.14	.773±.14
41159	.000±.00	.000±.00	.000±.00	.000±.00
41161	.000±.00	.000±.00	.000±.00	.000±.00
41162	.954±.02	.954±.02	.954±.02	.954±.02
41163	.955±.01	.955±.01	.955±.01	.955±.01
41164	.742±.02	.742±.02	.742±.02	.742±.02
41165	.761±.11	.761±.11	.761±.11	.761±.11
41166	.704±.04	.704±.04	.704±.04	.705±.04
41168	.736±.02	.736±.02	.736±.02	.736±.02
41169	.742±.18	.742±.18	.742±.18	.742±.18
41212	.851±.01	.851±.01	.851±.01	.851±.01
41216	.296±.00	.296±.00	.296±.00	.296±.00

41278	.766±.01	.766±.01	.767±.01	.767±.01

rbv2_ranger, $\eta = 3$

OpenML ID	IH-HB	dID-HB	pID-HB	eID-HB
3	.994±.00	.994±.00	.994±.00	.994±.00
6	.946±.01	.946±.01	.946±.01	.946±.01
11	.897±.01	.897±.01	.897±.01	.897±.01
12	.971±.00	.971±.00	.971±.00	.971±.00
14	.833±.01	.833±.01	.833±.01	.833±.01
15	.977±.01	.977±.01	.977±.01	.977±.01
16	.951±.00	.951±.00	.951±.00	.951±.00
18	.716±.02	.716±.02	.716±.02	.716±.02
22	.784±.01	.784±.01	.784±.01	.784±.01
23	.561±.02	.561±.02	.561±.02	.561±.02
24	.999±.00	.999±.00	.999±.00	.999±.00
28	.976±.00	.976±.00	.976±.00	.976±.00
29	.877±.04	.877±.04	.877±.04	.877±.04
31	.762±.01	.762±.01	.762±.01	.762±.01
32	.993±.00	.993±.00	.993±.00	.993±.00
37	.777±.03	.777±.03	.777±.03	.777±.03
38	.993±.00	.993±.00	.993±.00	.993±.00
42	.964±.01	.964±.01	.964±.01	.964±.01
44	.947±.00	.947±.00	.947±.00	.947±.00
46	.964±.00	.964±.00	.963±.00	.963±.00
50	.990±.01	.990±.01	.990±.01	.990±.01
54	.757±.02	.757±.02	.757±.02	.757±.02
60	.858±.01	.858±.01	.858±.01	.858±.01
151	.899±.01	.899±.01	.899±.01	.899±.01
181	.624±.03	.624±.03	.624±.03	.624±.03
182	.916±.01	.916±.01	.916±.01	.916±.01
188	.714±.01	.714±.01	.714±.01	.714±.01
300	.944±.01	.944±.01	.944±.01	.944±.01
307	.939±.01	.939±.01	.939±.01	.939±.01
312	.983±.01	.983±.01	.983±.01	.983±.01
334	.759±.11	.759±.11	.759±.11	.759±.11
375	.966±.00	.966±.00	.966±.00	.966±.00
377	.988±.00	.988±.00	.988±.00	.988±.00
458	.985±.00	.985±.00	.985±.00	.985±.00
469	.238±.03	.238±.03	.238±.03	.238±.03
470	.682±.01	.682±.01	.682±.01	.682±.01
554	.999±.00	.999±.00	.999±.00	.999±.00
1040	.997±.00	.997±.00	.997±.00	.997±.00
1049	.913±.01	.913±.01	.913±.01	.913±.01
1050	.901±.00	.901±.00	.901±.00	.901±.00
1053	.819±.00	.819±.00	.819±.00	.819±.00
1056	.994±.00	.994±.00	.994±.00	.994±.00
1063	.848±.03	.848±.03	.848±.03	.848±.03
1067	.868±.01	.868±.01	.868±.01	.868±.01
1068	.939±.00	.939±.00	.939±.00	.939±.00
1111	.999±.00	.999±.00	.999±.00	.999±.00
1220	.839±.01	.839±.01	.839±.01	.839±.01
1457	.762±.02	.762±.02	.762±.02	.762±.02

1461	.914±.00	.914±.00	.914±.00	.914±.00
1462	.995±.00	.995±.00	.995±.00	.995±.00
1464	.780±.01	.780±.01	.780±.01	.780±.01
1468	.947±.01	.947±.01	.947±.01	.947±.01
1475	.635±.01	.635±.01	.635±.01	.635±.01
1476	.996±.00	.996±.00	.996±.00	.996±.00
1478	.978±.00	.978±.00	.978±.00	.978±.00
1479	.562±.03	.562±.03	.562±.03	.562±.03
1480	.722±.02	.722±.02	.722±.02	.722±.02
1485	.868±.01	.868±.01	.868±.01	.868±.01
1486	.965±.00	.965±.00	.965±.00	.965±.00
1487	.941±.00	.941±.00	.941±.00	.941±.00
1489	.899±.01	.899±.01	.899±.01	.899±.01
1493	.826±.02	.826±.02	.826±.02	.826±.02
1494	.871±.01	.871±.01	.871±.01	.871±.01
1497	.996±.00	.996±.00	.996±.00	.996±.00
1501	.929±.01	.929±.01	.929±.01	.929±.01
1510	.965±.01	.965±.01	.965±.01	.966±.01
1515	.893±.02	.893±.02	.893±.02	.893±.02
1590	.885±.01	.885±.01	.885±.01	.885±.01
4134	.822±.02	.822±.02	.822±.02	.822±.02
4135	.959±.01	.959±.01	.959±.01	.959±.01
4154	.998±.00	.998±.00	.998±.00	.998±.00
4534	.959±.00	.959±.00	.959±.00	.959±.00
4538	.668±.02	.668±.02	.668±.02	.668±.02
4541	.801±.13	.801±.13	.801±.13	.801±.13
6332	.808±.05	.808±.05	.808±.05	.808±.05
23381	.622±.02	.622±.02	.622±.02	.621±.02
23512	.782±.04	.782±.04	.782±.04	.781±.04
23517	.549±.02	.549±.02	.549±.02	.548±.02
40496	.743±.03	.743±.03	.743±.03	.743±.03
40498	.691±.02	.691±.02	.692±.02	.692±.02
40499	.978±.00	.978±.00	.978±.00	.978±.00
40536	.904±.02	.904±.02	.904±.02	.904±.02
40668	.838±.02	.838±.02	.838±.02	.838±.02
40670	.957±.00	.957±.00	.957±.00	.957±.00
40685	.997±.00	.997±.00	.997±.00	.997±.00
40701	.965±.00	.965±.00	.965±.00	.965±.00
40900	.995±.00	.995±.00	.995±.00	.995±.00
40923	.000±.00	.000±.00	.000±.00	.000±.00
40927	.000±.00	.000±.00	.000±.00	.000±.00
40966	.977±.01	.977±.01	.977±.01	.977±.01
40975	.975±.01	.975±.01	.975±.01	.975±.01
40978	.976±.00	.976±.00	.976±.00	.976±.00
40979	.978±.00	.978±.00	.978±.00	.978±.00
40981	.883±.01	.883±.01	.883±.01	.883±.01
40982	.792±.01	.792±.01	.792±.01	.792±.01
40983	.989±.00	.989±.00	.989±.00	.989±.00
40984	.974±.00	.974±.00	.974±.00	.974±.00
40994	.921±.01	.921±.01	.921±.01	.921±.01
40996	.997±.00	.997±.00	.997±.00	.997±.00
41027	.852±.01	.852±.01	.852±.01	.851±.01
41138	.995±.00	.995±.00	.996±.00	.996±.00
41142	.770±.02	.770±.02	.771±.02	.771±.02

41143	.818±.01	.818±.01	.818±.01	.818±.01
41146	.941±.00	.941±.00	.941±.00	.941±.00
41150	.995±.01	.995±.01	.995±.01	.995±.01
41156	.857±.01	.857±.01	.857±.01	.857±.01
41157	.767±.14	.767±.14	.767±.14	.767±.14
41159	.000±.00	.000±.00	.000±.00	.000±.00
41161	.000±.00	.000±.00	.000±.00	.000±.00
41162	.944±.02	.944±.02	.944±.02	.944±.02
41163	.952±.01	.952±.01	.952±.01	.952±.01
41164	.734±.02	.734±.02	.734±.02	.734±.02
41165	.695±.09	.695±.09	.695±.09	.695±.09
41166	.685±.04	.687±.04	.687±.04	.687±.04
41168	.727±.02	.727±.02	.727±.02	.727±.02
41169	.643±.20	.643±.20	.643±.20	.643±.20
41212	.847±.01	.847±.01	.847±.01	.847±.01
41216	.294±.00	.294±.00	.294±.00	.294±.00
41278	.760±.01	.760±.01	.760±.01	.760±.01

rbv2 支持向量机，$\eta = 2^-$

OpenML ID	IH-HB	dID-HB	pID-HB	eID-HB
3	.994±.00	.994±.00	.994±.00	.994±.00
6	.971±.01	.971±.01	.971±.01	.971±.01
11	.969±.02	.969±.02	.969±.02	.969±.02
12	.981±.00	.981±.00	.981±.00	.981±.00
14	.845±.01	.846±.01	.846±.01	.846±.01
15	.973±.01	.973±.01	.973±.01	.973±.01
16	.978±.00	.978±.00	.978±.00	.978±.00
18	.770±.01	.770±.01	.770±.01	.770±.01
22	.889±.01	.889±.01	.889±.01	.889±.01
23	.556±.03	.556±.03	.556±.03	.556±.03
24	.000±.00	.000±.00	.000±.00	.000±.00
28	.993±.00	.993±.00	.993±.00	.993±.00
29	.872±.03	.872±.03	.872±.03	.872±.03
31	.756±.01	.756±.01	.756±.01	.756±.01
32	.997±.00	.997±.00	.997±.00	.997±.00
37	.778±.04	.778±.04	.778±.04	.778±.04
38	.975±.00	.975±.00	.975±.00	.975±.00
42	.943±.02	.943±.02	.943±.02	.943±.02
44	.938±.00	.938±.00	.938±.00	.938±.00
46	.959±.01	.959±.01	.959±.01	.959±.01
50	.999±.00	.999±.00	.999±.00	.999±.00
54	.842±.02	.842±.02	.842±.02	.842±.02
60	.873±.00	.873±.00	.873±.00	.873±.00
151	.840±.01	.840±.01	.840±.01	.840±.01
181	.611±.03	.611±.03	.611±.03	.612±.03
182	.917±.01	.917±.01	.917±.01	.917±.01
188	.687±.03	.687±.03	.687±.03	.687±.03
300	.977±.00	.977±.00	.977±.00	.977±.00
307	.991±.00	.991±.00	.991±.00	.991±.00
312	.976±.01	.976±.01	.976±.01	.976±.01
334	.910±.06	.910±.06	.909±.06	.909±.06
375	.994±.00	.994±.00	.994±.00	.994±.00

377	.991±.00	.990±.00	.991±.00	.991±.00
458	.996±.00	.996±.00	.996±.00	.996±.00
469	.218±.02	.218±.02	.218±.02	.218±.02
470	.738±.02	.738±.02	.738±.02	.738±.02
1040	.995±.00	.995±.00	.995±.00	.995±.00
1049	.911±.01	.911±.01	.911±.01	.911±.01
1050	.913±.01	.913±.01	.913±.01	.913±.01
1053	.819±.00	.819±.00	.819±.00	.819±.00
1056	.997±.00	.997±.00	.997±.00	.997±.00
1063	.849±.03	.849±.03	.849±.03	.849±.03
1067	.858±.00	.858±.00	.857±.00	.857±.00
1068	.946±.00	.946±.00	.946±.00	.946±.00
1111	.987±.00	.987±.00	.987±.00	.987±.00
1220	.847±.00	.847±.00	.847±.00	.847±.00
1457	.825±.01	.825±.01	.825±.01	.825±.01
1461	.908±.01	.908±.01	.908±.01	.907±.01
1462	.999±.00	.999±.00	.999±.00	.999±.00
1464	.778±.01	.778±.01	.779±.01	.779±.01
1468	.947±.01	.947±.01	.947±.01	.947±.01
1475	.572±.01	.572±.01	.572±.01	.572±.01
1476	.995±.00	.995±.00	.995±.00	.995±.00
1478	.989±.00	.989±.00	.989±.00	.989±.00
1479	.768±.02	.768±.02	.768±.02	.768±.02
1480	.725±.01	.725±.01	.725±.01	.725±.01
1485	.588±.03	.588±.03	.588±.03	.588±.03
1486	.962±.00	.962±.00	.962±.00	.962±.00
1487	.946±.00	.946±.00	.946±.00	.946±.00
1489	.891±.01	.891±.01	.891±.01	.891±.01
1493	.863±.01	.863±.01	.863±.01	.863±.01
1494	.876±.01	.876±.01	.876±.01	.876±.01
1497	.927±.01	.927±.01	.927±.01	.927±.01
1501	.962±.01	.962±.01	.962±.01	.962±.01
1510	.976±.01	.976±.01	.976±.01	.976±.01
1515	.806±.03	.806±.03	.806±.03	.805±.04
1590	.861±.00	.861±.00	.861±.00	.861±.00
4134	.771±.00	.771±.00	.771±.00	.771±.00
4135	.951±.00	.951±.00	.951±.00	.951±.00
4154	.999±.00	.999±.00	.999±.00	.999±.00
4534	.971±.00	.971±.00	.971±.00	.971±.00
4538	.555±.01	.555±.01	.555±.01	.555±.01
6332	.833±.05	.833±.05	.833±.05	.834±.05
23381	.643±.04	.643±.04	.643±.04	.643±.04
40496	.752±.04	.752±.04	.752±.04	.752±.04
40498	.665±.02	.665±.02	.664±.02	.664±.02
40499	.997±.00	.997±.00	.997±.00	.997±.00
40536	.866±.00	.866±.00	.866±.00	.866±.00
40668	.860±.03	.860±.03	.860±.03	.859±.03
40670	.962±.00	.962±.00	.962±.00	.962±.00
40685	.998±.00	.998±.00	.998±.00	.998±.00
40701	.918±.00	.918±.00	.918±.00	.918±.00
40900	.996±.00	.996±.00	.996±.00	.995±.00
40966	.997±.00	.997±.00	.997±.00	.997±.00
40975	.995±.01	.995±.01	.995±.01	.995±.01
40978	.975±.01	.975±.01	.975±.01	.975±.01

40979	.984±.00	.984±.00	.984±.00	.984±.00
40981	.866±.02	.866±.02	.866±.02	.866±.02
40982	.754±.02	.754±.02	.754±.02	.754±.02
40983	.985±.00	.985±.00	.985±.00	.985±.00
40984	.960±.01	.960±.01	.960±.01	.959±.01
40994	.957±.01	.957±.01	.957±.01	.957±.01
41027	.848±.02	.848±.02	.848±.02	.848±.02
41138	.991±.00	.991±.00	.991±.00	.991±.00
41142	.736±.01	.736±.01	.736±.01	.735±.01
41143	.802±.02	.802±.02	.802±.02	.802±.02
41146	.925±.01	.925±.01	.925±.01	.924±.01
41156	.853±.01	.853±.01	.853±.01	.853±.01
41157	.841±.11	.841±.11	.841±.11	.841±.11
41162	.917±.01	.917±.01	.917±.01	.917±.01
41163	.985±.00	.985±.00	.985±.00	.985±.00
41164	.700±.01	.700±.01	.700±.01	.700±.01
41169	.368±.01	.368±.01	.368±.01	.368±.01
41212	.789±.02	.789±.02	.789±.02	.789±.02
41216	.255±.00	.255±.00	.255±.00	.255±.00
41278	.757±.00	.757±.00	.757±.00	.756±.00

rbv2 支持向量机，$\eta = 3$

OpenML ID	IH-HB	dID-HB	pID-HB	eID-HB
3	.993±.00	.993±.00	.993±.00	.993±.00
6	.968±.01	.968±.01	.968±.01	.968±.01
11	.957±.03	.957±.03	.957±.03	.957±.03
12	.980±.00	.980±.00	.980±.00	.980±.00
14	.844±.01	.844±.01	.844±.01	.844±.01
15	.972±.01	.972±.01	.972±.01	.972±.01
16	.975±.00	.975±.00	.975±.00	.975±.00
18	.766±.02	.766±.02	.766±.02	.765±.01
22	.885±.01	.885±.01	.885±.01	.885±.01
23	.554±.03	.554±.03	.554±.03	.554±.03
24	.000±.00	.000±.00	.000±.00	.000±.00
28	.992±.00	.992±.00	.992±.00	.992±.00
29	.867±.04	.867±.04	.867±.04	.867±.04
31	.754±.01	.754±.01	.754±.01	.754±.01
32	.997±.00	.997±.00	.997±.00	.997±.00
37	.777±.04	.777±.04	.777±.04	.777±.04
38	.974±.00	.974±.00	.974±.00	.974±.00
42	.941±.02	.941±.02	.941±.02	.941±.02
44	.936±.01	.936±.01	.936±.01	.936±.01
46	.958±.01	.958±.01	.958±.01	.958±.01
50	.995±.01	.995±.01	.995±.01	.995±.01
54	.833±.02	.833±.02	.833±.02	.833±.02
60	.872±.00	.872±.00	.872±.00	.872±.00
151	.832±.02	.832±.02	.832±.02	.832±.02
181	.610±.03	.610±.03	.610±.03	.610±.03
182	.914±.01	.914±.01	.914±.01	.914±.01
188	.686±.03	.686±.03	.686±.03	.686±.03
300	.977±.00	.977±.00	.977±.00	.977±.00
307	.984±.02	.984±.02	.984±.02	.984±.02

312	.974±.01	.974±.01	.974±.01	.974±.01
334	.893±.06	.893±.06	.889±.05	.889±.05
375	.993±.00	.993±.00	.993±.00	.993±.00
377	.988±.01	.988±.01	.988±.01	.988±.01
458	.996±.00	.996±.00	.996±.00	.996±.00
469	.214±.02	.214±.02	.214±.02	.214±.02
470	.734±.02	.734±.02	.734±.02	.734±.02
1040	.995±.00	.995±.00	.995±.00	.995±.00
1049	.910±.01	.910±.01	.910±.01	.910±.01
1050	.911±.01	.911±.01	.911±.01	.911±.01
1053	.818±.00	.818±.00	.818±.00	.818±.00
1056	.997±.00	.997±.00	.997±.00	.997±.00
1063	.849±.03	.849±.03	.848±.03	.848±.03
1067	.857±.00	.857±.00	.857±.00	.857±.00
1068	.944±.00	.944±.00	.944±.00	.944±.00
1111	.986±.00	.986±.00	.986±.00	.986±.00
1220	.845±.00	.845±.00	.845±.00	.845±.00
1457	.824±.01	.824±.01	.824±.01	.824±.01
1461	.906±.01	.906±.01	.906±.01	.906±.01
1462	.999±.00	.999±.00	.999±.00	.999±.00
1464	.776±.01	.776±.01	.776±.01	.776±.01
1468	.945±.01	.945±.01	.945±.01	.945±.01
1475	.565±.02	.565±.02	.565±.02	.565±.02
1476	.994±.00	.994±.00	.994±.00	.994±.00
1478	.988±.00	.988±.00	.988±.00	.988±.00
1479	.767±.02	.767±.02	.767±.02	.767±.02
1480	.724±.01	.724±.01	.724±.01	.724±.01
1485	.580±.03	.580±.03	.580±.03	.580±.03
1486	.958±.01	.958±.01	.958±.01	.958±.01
1487	.945±.00	.945±.00	.945±.00	.945±.00
1489	.886±.02	.886±.02	.886±.02	.886±.02
1493	.861±.01	.861±.01	.861±.01	.861±.01
1494	.874±.01	.874±.01	.873±.01	.873±.01
1497	.921±.01	.921±.01	.921±.01	.921±.01
1501	.959±.01	.959±.01	.959±.01	.959±.01
1510	.976±.01	.976±.01	.976±.01	.976±.01
1515	.799±.04	.799±.04	.799±.04	.799±.04
1590	.860±.00	.860±.00	.860±.00	.860±.00
4134	.770±.01	.770±.01	.770±.01	.770±.01
4135	.949±.00	.949±.00	.949±.00	.949±.00
4154	.999±.00	.999±.00	.999±.00	.999±.00
4534	.969±.00	.969±.00	.969±.00	.969±.00
4538	.549±.02	.549±.02	.549±.02	.549±.02
6332	.827±.05	.827±.05	.827±.05	.827±.05
23381	.633±.04	.633±.04	.633±.04	.633±.04
40496	.748±.05	.748±.05	.749±.05	.749±.05
40498	.660±.02	.660±.02	.659±.02	.660±.02
40499	.997±.00	.997±.00	.997±.00	.997±.00
40536	.864±.00	.864±.00	.864±.00	.864±.00
40668	.845±.03	.845±.03	.845±.03	.845±.03
40670	.961±.00	.961±.00	.961±.00	.961±.00
40685	.997±.00	.997±.00	.997±.00	.997±.00
40701	.916±.01	.916±.01	.916±.01	.916±.01
40900	.995±.00	.995±.00	.995±.00	.995±.00

40966	.996±.00	.996±.00	.996±.00	.996±.00
40975	.988±.02	.988±.02	.988±.02	.988±.02
40978	.975±.01	.975±.01	.975±.01	.975±.01
40979	.983±.00	.983±.00	.983±.00	.983±.00
40981	.863±.02	.863±.02	.863±.02	.863±.02
40982	.749±.02	.749±.02	.749±.02	.749±.02
40983	.984±.00	.984±.00	.984±.00	.984±.00
40984	.958±.01	.958±.01	.958±.01	.958±.01
40994	.956±.01	.956±.01	.956±.01	.956±.01
41027	.836±.02	.836±.02	.836±.02	.836±.02
41138	.990±.00	.990±.00	.990±.00	.990±.00
41142	.732±.01	.732±.01	.732±.01	.732±.01
41143	.798±.02	.798±.02	.798±.02	.798±.02
41146	.923±.01	.923±.01	.923±.01	.922±.01
41156	.852±.01	.852±.01	.852±.01	.852±.01
41157	.834±.12	.834±.12	.834±.12	.834±.12
41162	.913±.01	.913±.01	.913±.01	.913±.01
41163	.980±.01	.980±.01	.981±.01	.981±.01
41164	.698±.01	.698±.01	.698±.01	.698±.01
41169	.365±.01	.365±.01	.365±.01	.365±.01
41212	.779±.02	.779±.02	.779±.02	.779±.02
41216	.252±.01	.252±.01	.252±.01	.251±.01
41278	.755±.01	.755±.01	.755±.01	.755±.01

rbv2_xgboost, $\eta = 2$

OpenML ID	IH-HB	dID-HB	pID-HB	eID-HB
3	.985±.01	.985±.01	.985±.01	.985±.01
6	.000±.00	.000±.00	.000±.00	.000±.00
11	.857±.02	.857±.02	.857±.02	.857±.02
12	.983±.00	.983±.00	.983±.00	.983±.00
14	.852±.02	.852±.02	.852±.02	.852±.02
15	.984±.01	.984±.01	.984±.01	.984±.01
16	.966±.01	.966±.01	.966±.01	.966±.01
18	.775±.05	.775±.05	.775±.05	.775±.05
22	.857±.03	.857±.03	.857±.03	.857±.03
23	.614±.04	.614±.04	.617±.04	.617±.04
24	.999±.00	.999±.00	.999±.00	.999±.00
28	.993±.01	.993±.01	.993±.01	.993±.01
29	.891±.01	.891±.01	.891±.01	.891±.01
31	.754±.02	.754±.02	.754±.02	.753±.02
32	.999±.00	.999±.00	.999±.00	.999±.00
37	.786±.02	.786±.02	.786±.02	.786±.02
38	.989±.00	.989±.00	.989±.00	.989±.00
42	.892±.09	.892±.09	.892±.09	.892±.09
44	.964±.01	.964±.01	.964±.01	.964±.01
46	.981±.00	.981±.00	.981±.00	.981±.00
50	.959±.04	.959±.04	.959±.04	.959±.04
54	.810±.03	.810±.03	.810±.03	.810±.03
60	.917±.02	.917±.02	.917±.02	.917±.02
151	.938±.03	.938±.03	.938±.03	.938±.03
181	.649±.07	.649±.07	.649±.07	.649±.07
182	.953±.04	.953±.04	.953±.04	.952±.04

188	.694±.04	.694±.04	.694±.04	.694±.04
300	.000±.00	.000±.00	.000±.00	.000±.00
307	.844±.08	.844±.08	.844±.08	.844±.08
312	.961±.01	.961±.01	.961±.01	.961±.01
334	.678±.01	.678±.01	.678±.01	.678±.01
375	.992±.01	.992±.01	.992±.01	.992±.01
377	.981±.01	.981±.01	.981±.01	.981±.01
458	.997±.00	.997±.00	.997±.00	.997±.00
469	.251±.06	.251±.06	.251±.06	.251±.06
470	.692±.01	.692±.01	.692±.01	.692±.01
554	.000±.00	.000±.00	.000±.00	.000±.00
1040	.997±.00	.997±.00	.997±.00	.997±.00
1049	.925±.01	.925±.01	.925±.01	.925±.01
1050	.909±.00	.909±.00	.909±.00	.909±.00
1053	.839±.01	.839±.01	.839±.01	.839±.01
1056	.993±.00	.993±.00	.993±.00	.992±.00
1063	.885±.02	.885±.02	.885±.02	.885±.02
1067	.876±.01	.876±.01	.876±.01	.876±.01
1068	.938±.00	.938±.00	.938±.00	.938±.00
1111	.990±.01	.990±.01	.990±.01	.990±.01
1220	.927±.06	.927±.06	.927±.06	.927±.06
1457	.000±.00	.000±.00	.000±.00	.000±.00
1461	.977±.03	.977±.03	.977±.03	.977±.03
1462	.992±.00	.992±.00	.992±.00	.992±.00
1464	.795±.01	.795±.01	.795±.01	.795±.01
1468	.892±.03	.892±.03	.892±.03	.892±.03
1475	.717±.16	.717±.16	.717±.16	.717±.16
1476	.999±.00	.999±.00	.999±.00	.999±.00
1478	.997±.00	.997±.00	.997±.00	.997±.00
1479	.747±.04	.747±.04	.747±.04	.747±.04
1480	.738±.01	.738±.01	.738±.01	.737±.01
1485	.822±.03	.822±.03	.822±.03	.823±.03
1486	.979±.01	.979±.01	.979±.01	.979±.01
1487	.948±.01	.948±.01	.949±.01	.949±.01
1489	.897±.01	.897±.01	.898±.01	.898±.01
1493	.000±.00	.000±.00	.000±.00	.000±.00
1494	.872±.01	.872±.01	.873±.01	.873±.01
1497	.998±.00	.998±.00	.998±.00	.998±.00
1501	.920±.02	.920±.02	.920±.02	.920±.02
1510	.970±.01	.970±.01	.970±.01	.970±.01
1515	.790±.12	.790±.12	.793±.12	.793±.12
1590	.891±.05	.891±.05	.891±.05	.891±.05
4134	.816±.02	.816±.02	.816±.02	.815±.02
4135	.935±.02	.935±.02	.936±.02	.936±.02
4154	.998±.00	.998±.00	.998±.00	.998±.00
4534	.968±.01	.968±.01	.968±.01	.967±.01
4538	.868±.14	.868±.14	.868±.14	.868±.14
4541	.000±.00	.000±.00	.000±.00	.000±.00
6332	.813±.04	.813±.04	.813±.04	.813±.04
23381	.625±.02	.625±.02	.625±.02	.624±.02
23512	.943±.08	.943±.08	.943±.08	.943±.08
23517	.865±.15	.865±.15	.865±.15	.865±.15
40496	.764±.04	.764±.04	.764±.04	.764±.04
40498	.786±.17	.786±.17	.786±.17	.787±.17

40499	.993±.01	.993±.01	.993±.01	.993±.01
40536	.886±.01	.886±.01	.887±.01	.887±.01
40668	.990±.04	.990±.04	.990±.04	.990±.04
40670	.982±.01	.982±.01	.982±.01	.982±.01
40685	.999±.00	.999±.00	.999±.00	.999±.00
40701	.941±.01	.941±.01	.941±.01	.941±.01
40900	.994±.00	.994±.00	.994±.00	.994±.00
40923	.000±.00	.000±.00	.000±.00	.000±.00
40927	.000±.00	.000±.00	.000±.00	.000±.00
40966	.963±.02	.963±.02	.963±.02	.963±.02
40975	.925±.05	.925±.05	.926±.05	.926±.05
40978	.948±.02	.948±.02	.948±.02	.948±.02
40979	.980±.00	.980±.00	.980±.00	.980±.00
40981	.876±.01	.876±.01	.876±.01	.876±.01
40982	.834±.02	.834±.02	.834±.02	.834±.02
40983	.983±.00	.983±.00	.983±.00	.983±.01
40984	.984±.01	.984±.01	.984±.01	.984±.01
40994	.900±.01	.900±.01	.900±.01	.900±.01
40996	.000±.00	.000±.00	.000±.00	.000±.00
41027	.990±.03	.990±.03	.990±.03	.990±.03
41138	.975±.01	.975±.01	.975±.01	.975±.01
41142	.767±.02	.767±.02	.767±.02	.767±.02
41143	.809±.01	.809±.01	.809±.01	.808±.02
41146	.952±.01	.952±.01	.952±.01	.952±.01
41150	.999±.01	.999±.01	.999±.01	.999±.01
41156	.883±.01	.883±.01	.883±.01	.883±.01
41157	.834±.12	.834±.12	.834±.12	.834±.12
41159	.898±.04	.898±.04	.898±.04	.898±.04
41161	.998±.00	.998±.00	.998±.00	.998±.00
41162	.986±.03	.986±.03	.986±.03	.986±.03
41163	.993±.01	.993±.01	.993±.01	.993±.01
41164	.897±.14	.897±.14	.897±.14	.897±.14
41165	.000±.00	.000±.00	.000±.00	.000±.00
41166	.000±.00	.000±.00	.000±.00	.000±.00
41168	.000±.00	.000±.00	.000±.00	.000±.00
41169	.000±.00	.000±.00	.000±.00	.000±.00
41212	.834±.03	.834±.03	.834±.03	.833±.03
41216	.964±.11	.964±.11	.964±.11	.964±.11
41278	.984±.06	.984±.06	.984±.06	.984±.06

rbv2_xgboost, $\eta = 3$

OpenML ID	IH-HB	dID-HB	pID-HB	eID-HB
3	.979±.01	.979±.01	.979±.01	.979±.01
6	.999±.01	.999±.01	.999±.01	.999±.01
11	.845±.02	.845±.02	.845±.02	.845±.02
12	.982±.00	.982±.00	.982±.00	.982±.00
14	.848±.02	.848±.02	.848±.02	.848±.02
15	.982±.01	.982±.01	.982±.01	.982±.01
16	.965±.01	.965±.01	.965±.01	.965±.01
18	.753±.05	.753±.05	.753±.05	.753±.05
22	.847±.03	.847±.03	.847±.03	.847±.03
23	.601±.04	.601±.04	.601±.04	.597±.04

24	.999±.00	.999±.00	.999±.00	.999±.00
28	.987±.01	.987±.01	.987±.01	.987±.01
29	.887±.01	.887±.01	.887±.01	.887±.01
31	.748±.02	.748±.02	.748±.02	.748±.02
32	.997±.00	.997±.00	.997±.00	.997±.00
37	.784±.02	.784±.02	.784±.02	.784±.02
38	.987±.01	.987±.01	.987±.01	.987±.01
42	.861±.13	.861±.13	.861±.13	.861±.13
44	.961±.01	.961±.01	.961±.01	.961±.01
46	.980±.01	.980±.01	.980±.01	.980±.01
50	.924±.06	.924±.06	.924±.06	.924±.06
54	.797±.03	.797±.03	.797±.03	.797±.03
60	.912±.02	.912±.02	.912±.02	.911±.02
151	.922±.04	.922±.04	.922±.04	.921±.04
181	.643±.06	.643±.06	.643±.06	.637±.06
182	.936±.03	.936±.03	.936±.03	.936±.03
188	.678±.04	.678±.04	.682±.04	.681±.04
300	.999±.01	.999±.01	.999±.01	.999±.01
307	.806±.10	.806±.10	.806±.10	.802±.10
312	.957±.01	.957±.01	.957±.01	.956±.01
334	.675±.01	.675±.01	.675±.01	.675±.01
375	.983±.01	.983±.01	.983±.01	.983±.01
377	.979±.01	.979±.01	.979±.01	.979±.01
458	.997±.00	.997±.00	.997±.00	.997±.00
469	.232±.05	.232±.05	.232±.05	.232±.05
470	.687±.01	.687±.01	.688±.01	.688±.01
554	.000±.00	.000±.00	.000±.00	.000±.00
1040	.996±.00	.996±.00	.996±.00	.996±.00
1049	.922±.01	.922±.01	.922±.01	.922±.01
1050	.907±.00	.907±.00	.907±.00	.907±.00
1053	.836±.01	.836±.01	.836±.01	.836±.01
1056	.992±.00	.992±.00	.992±.00	.992±.00
1063	.882±.02	.882±.02	.882±.02	.882±.02
1067	.875±.01	.875±.01	.875±.01	.874±.01
1068	.937±.01	.937±.01	.937±.01	.937±.01
1111	.986±.01	.986±.01	.986±.01	.986±.01
1220	.899±.05	.899±.05	.899±.05	.899±.05
1457	.000±.00	.000±.00	.000±.00	.000±.00
1461	.956±.03	.956±.03	.956±.03	.956±.03
1462	.990±.00	.990±.00	.990±.00	.990±.00
1464	.791±.01	.791±.01	.791±.01	.791±.01
1468	.885±.03	.885±.03	.886±.03	.886±.03
1475	.611±.09	.611±.09	.611±.09	.611±.09
1476	.996±.01	.996±.01	.996±.01	.996±.01
1478	.994±.01	.994±.01	.994±.01	.994±.01
1479	.699±.08	.699±.08	.699±.08	.697±.08
1480	.731±.02	.731±.02	.731±.02	.731±.02
1485	.793±.05	.793±.05	.795±.05	.795±.05
1486	.977±.01	.977±.01	.977±.01	.977±.01
1487	.947±.01	.947±.01	.947±.01	.947±.01
1489	.890±.01	.890±.01	.890±.01	.890±.01
1493	.989±.06	.989±.06	.989±.06	.989±.06
1494	.865±.02	.865±.02	.865±.02	.865±.02
1497	.998±.00	.998±.00	.998±.00	.998±.00

1501	.919±.02	.919±.02	.919±.02	.918±.02
1510	.966±.01	.966±.01	.966±.01	.966±.01
1515	.674±.20	.674±.20	.674±.20	.674±.20
1590	.870±.04	.870±.04	.870±.04	.870±.04
4134	.807±.02	.807±.02	.807±.02	.807±.02
4135	.926±.01	.926±.01	.926±.01	.926±.01
4154	.998±.00	.998±.00	.998±.00	.998±.00
4534	.964±.01	.964±.01	.964±.01	.964±.01
4538	.757±.14	.757±.14	.757±.14	.757±.14
4541	.989±.06	.989±.06	.989±.06	.989±.06
6332	.796±.05	.796±.05	.796±.05	.796±.05
23381	.612±.03	.612±.03	.612±.03	.612±.03
23512	.875±.11	.875±.11	.875±.11	.875±.11
23517	.767±.16	.767±.16	.767±.16	.767±.16
40496	.750±.04	.750±.04	.750±.04	.750±.04
40498	.698±.15	.698±.15	.698±.15	.698±.15
40499	.987±.01	.987±.01	.987±.01	.985±.01
40536	.883±.01	.883±.01	.883±.01	.883±.01
40668	.981±.05	.981±.05	.981±.05	.981±.05
40670	.980±.01	.980±.01	.980±.01	.980±.01
40685	.999±.00	.999±.00	.999±.00	.999±.00
40701	.931±.02	.931±.02	.931±.02	.931±.02
40900	.993±.00	.993±.00	.993±.00	.993±.00
40923	.000±.00	.000±.00	.000±.00	.000±.00
40927	.000±.00	.000±.00	.000±.00	.000±.00
40966	.954±.03	.954±.03	.954±.03	.954±.03
40975	.900±.05	.900±.05	.900±.05	.900±.05
40978	.937±.03	.937±.03	.937±.03	.937±.03
40979	.980±.01	.980±.01	.980±.01	.980±.01
40981	.866±.02	.866±.02	.866±.02	.866±.02
40982	.820±.03	.820±.03	.820±.03	.820±.03
40983	.982±.00	.982±.00	.982±.00	.982±.00
40984	.980±.01	.980±.01	.980±.01	.980±.01
40994	.898±.01	.898±.01	.898±.01	.898±.01
40996	.000±.00	.000±.00	.000±.00	.000±.00
41027	.973±.04	.973±.04	.973±.04	.973±.04
41138	.972±.01	.972±.01	.972±.01	.972±.01
41142	.761±.01	.761±.01	.761±.01	.761±.01
41143	.801±.02	.801±.02	.801±.02	.801±.02
41146	.949±.01	.949±.01	.949±.01	.949±.01
41150	.997±.01	.997±.01	.997±.01	.996±.01
41156	.879±.01	.879±.01	.879±.01	.878±.01
41157	.826±.12	.826±.12	.826±.12	.826±.12
41159	.880±.04	.880±.04	.880±.04	.880±.04
41161	.997±.00	.997±.00	.997±.00	.997±.00
41162	.960±.04	.960±.04	.960±.04	.960±.04
41163	.987±.01	.987±.01	.987±.01	.987±.01
41164	.775±.17	.775±.17	.775±.17	.775±.17
41165	.000±.00	.000±.00	.000±.00	.000±.00
41166	.000±.00	.000±.00	.000±.00	.000±.00
41168	.997±.02	.997±.02	.997±.02	.997±.02
41169	.000±.00	.000±.00	.000±.00	.000±.00
41212	.821±.04	.821±.04	.821±.04	.821±.04
41216	.941±.15	.941±.15	.941±.15	.941±.15

41278	.963±.08	.963±.08	.963±.08	.963±.08

lcbench, $\eta = 2$				
OpenML ID	IH-HB	dID-HB	pID-HB	eID-HB
3945	.993±.00	.993±.00	.993±.00	.993±.00
7593	.716±.03	.716±.03	.716±.03	.716±.03
34539	.981±.01	.981±.01	.981±.01	.981±.01
126025	.881±.02	.881±.02	.879±.02	.877±.02
126026	.966±.01	.966±.01	.966±.01	.965±.01
126029	.926±.01	.926±.01	.926±.01	.926±.01
146212	.975±.01	.975±.01	.976±.01	.975±.01
167104	.922±.01	.922±.01	.922±.01	.922±.01
167149	.971±.00	.971±.00	.971±.00	.970±.00
167152	.967±.00	.967±.00	.967±.00	.967±.00
167161	.818±.03	.818±.03	.818±.03	.816±.03
167168	.780±.02	.780±.02	.780±.02	.779±.02
167181	.904±.01	.904±.01	.904±.01	.903±.01
167184	.800±.01	.800±.01	.800±.01	.800±.02
167185	.961±.00	.961±.00	.961±.00	.961±.00
167190	.847±.01	.847±.01	.847±.01	.845±.01
167200	.715±.02	.715±.02	.715±.02	.713±.02
167201	.738±.04	.738±.04	.738±.04	.738±.04
168329	.281±.03	.281±.03	.279±.03	.277±.03
168330	.636±.02	.636±.02	.637±.02	.637±.02
168331	.582±.04	.582±.04	.582±.04	.582±.04
168335	.929±.01	.929±.01	.929±.01	.929±.01
168868	.996±.00	.996±.00	.996±.00	.996±.00
168908	.814±.01	.814±.01	.815±.01	.815±.01
168910	.675±.02	.675±.02	.675±.02	.675±.02
189354	.704±.03	.704±.03	.704±.03	.704±.03
189862	.852±.02	.852±.02	.853±.02	.852±.02
189865	.944±.01	.944±.01	.944±.01	.944±.01
189866	.690±.03	.690±.03	.690±.03	.690±.03
189873	.797±.03	.797±.03	.797±.03	.797±.03
189905	.905±.02	.905±.02	.905±.02	.905±.02
189906	.865±.02	.865±.02	.865±.02	.865±.02
189908	.872±.01	.872±.01	.872±.01	.872±.01
189909	.752±.02	.752±.02	.753±.02	.753±.02

lcbench, $\eta = 3$				
OpenML ID	IH-HB	dID-HB	pID-HB	eID-HB
3945	.993±.00	.993±.00	.993±.00	.993±.00
7593	.711±.04	.711±.04	.711±.04	.712±.04
34539	.982±.01	.982±.01	.982±.01	.982±.01
126025	.886±.02	.886±.02	.887±.02	.888±.02
126026	.965±.01	.965±.01	.965±.01	.965±.01
126029	.923±.02	.923±.02	.924±.02	.924±.02
146212	.975±.01	.975±.01	.975±.01	.975±.01
167104	.922±.01	.922±.01	.922±.01	.922±.01
167149	.969±.01	.969±.01	.969±.01	.969±.01
167152	.967±.00	.967±.00	.967±.00	.967±.00
167161	.821±.03	.821±.03	.822±.03	.816±.03
167168	.770±.02	.770±.02	.770±.02	.770±.02
167181	.903±.02	.903±.02	.903±.02	.903±.02
167184	.798±.03	.798±.03	.799±.03	.796±.03
167185	.961±.00	.961±.00	.961±.00	.961±.00
167190	.851±.02	.851±.02	.853±.01	.848±.02
167200	.710±.02	.710±.02	.710±.02	.711±.02
167201	.735±.05	.735±.05	.736±.05	.736±.05
168329	.273±.04	.273±.04	.273±.04	.273±.04
168330	.635±.03	.635±.03	.635±.03	.635±.03
168331	.562±.05	.562±.05	.562±.05	.562±.05
168335	.926±.02	.926±.02	.926±.02	.926±.02
168868	.996±.00	.996±.00	.996±.00	.996±.00
168908	.808±.01	.808±.01	.808±.02	.808±.02
168910	.682±.03	.682±.03	.682±.03	.682±.03
189354	.706±.03	.706±.03	.706±.03	.706±.03
189862	.856±.02	.856±.02	.856±.02	.854±.02
189865	.943±.01	.943±.01	.943±.01	.943±.01
189866	.686±.03	.686±.03	.686±.03	.686±.03
189873	.771±.08	.771±.08	.771±.08	.771±.08
189905	.898±.02	.898±.02	.898±.02	.898±.02
189906	.860±.02	.860±.02	.860±.02	.860±.02
189908	.871±.02	.871±.02	.871±.02	.871±.02
189909	.747±.03	.747±.03	.748±.03	.748±.03