# TREEES

**Eli Ben-Michael** EBENMICHAEL@BERKELEY.EDU
*Department of Statistics*
*University of California*
*Berkeley, CA 94720-1776, USA*

**Editor:**

## Abstract

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

**Keywords:** TREEES

## 1. Introduction

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

# 2. Bayesian Optimization and Bandit-Based Algorithms

## 2.1 Bayesian Optimization

## 2.2 SuccessiveHalving and Hyperband

# 3. A Trio of Tree Algorithms

## 3.1 SequentialHalving with Decision Tree inputs

## 3.2 SequentialTree

### 3.2.1 Algorithm

**Algorithm 1** SequentialTree

```
procedure SequentialTree(Box constrained space 𝒳 ⊂ ℝ^d, budget T, η, m)
    Assign 𝒫 = {𝒳}
    Assign n_nodes = m
    for r = 1 ... log_η(m) do
        The number of pulls per element of 𝒫  n_r = ⌊ T / (|𝒫|⌈log_η(m)⌉) ⌋
        𝒜_r = ∅
        B̂_r = ∅
        for p ∈ 𝒫 do
            Remove p from 𝒫
            Sample {x_p1, ..., x_pn_r} independently and uniformly over p
            Let {y_p1, ..., y_pn_r} be the result of querying the function at these points
            Train a decision tree with n_nodes nodes, t_p = DecisionTree(y_p ~ X_p)
            Add the n_nodes partitions defined by t_p to 𝒜_r
            Add the prediction from t_p for each partition to B̂_r.
        end for
        if |𝒫| > η^2 then
            Set 𝒫 to be the elements of 𝒜_r with predictions in the top |𝒫|/η^2 of B̂_r
        else
            Let p̂ be the element of 𝒫 with the best prediction
            return MidPoint(p̂)
        end if
        n_nodes = η
    end for
end procedure
```

### 3.3 PartitionTree

#### 3.3.1 Algorithm

**Algorithm 2** PartitionTree

**procedure** PartitionTree(Box constrained space $\mathcal{X} \subset \mathbb{R}^d$, budget $T$, $\eta$, rounds $R$)
    $m = \eta^2$
    $\mathcal{P} = \{\mathcal{X}\}$
    **for** $r = 1, \ldots, R$ **do** The number of pulls per element of $\mathcal{P}$ $n_r = \left\lfloor \frac{T}{|\mathcal{P}|R} \right\rfloor$
        $\mathcal{A}_r = \emptyset$
        $\hat{B}_r = \emptyset$
        **for** $p \in \mathcal{P}$ **do**
            Remove $p$ from $\mathcal{P}$
            Sample $\{x_{p1}, \ldots, x_{pn_r}\}$ independently and uniformly over $p$
            Let $\{y_{p1}, \ldots, y_{pn_r}\}$ be the result of querying the function at these points
            Train a decision tree with $m$ nodes, $t_p = \text{DecisionTree}(y_p \sim X_p)$
            Add the $m$ partitions defined by $t_p$ to $\mathcal{A}_r$
            Add the prediction from $t_p$ for each partition to $\hat{B}_r$.
        **end for**
        Set $\mathcal{P}$ to be the elements of $\mathcal{A}_r$ with predictions in the top $\frac{|\mathcal{P}|}{\eta}$ of $\hat{B}_r$
        Let $\hat{p}$ be the element of $\mathcal{P}$ with the best prediction
        $m = \eta$
    **end for**
    **return** MidPoint($\hat{p}$)
**end procedure**

## 4. Empirical Evaluation

## 5. Conclusion

### Acknowledgments

We would like to acknowledge

## Appendix A.