A Model-Agnostic Randomized Learning Framework based on Random Hypothesis Subspace Sampling (RHSS)

– Yiting Cao, Chao Lan
University of Oklahoma

Background

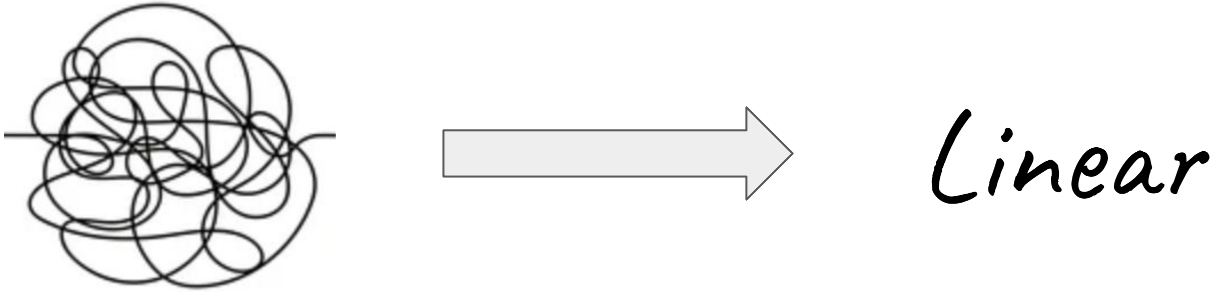
Randomized Machine Learning (RML) studies how to randomize regular learning processes and has achieved outstanding learning efficiency.

However, most RML techniques are model-specific.

- Kernel: Random Fourier Features (RFF) [2008] approximates k(x,y) with the inner product of x and y's Random Fourier Features.
- Tree: Extra tree [2006] splits tree nodes based on randomly selected features instead of optimally selected ones.
- MLP: Random Vector Functional Link (RVFL) [1994] randomly generates some network weights of instead of training all weights.

We propose a model-agnostic RML framework named RHSS.

Given any hypothesis class, RHSS randomly samples k hypotheses and learns a optimal model from their linear span by simply solving a linear least square problem in $O(nk^2)$ time, where n is the number of training instances.



We propose a model-agnostic randomized learning framework.

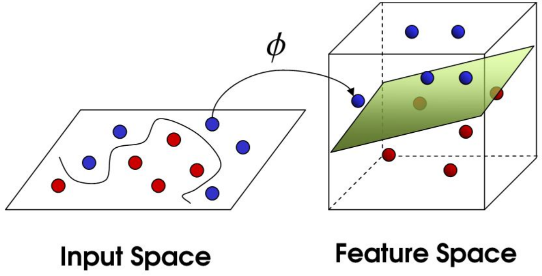
How is this changing the current paradigm?

- Model-agnostic
- **Removes the "weakly learned" assumption**

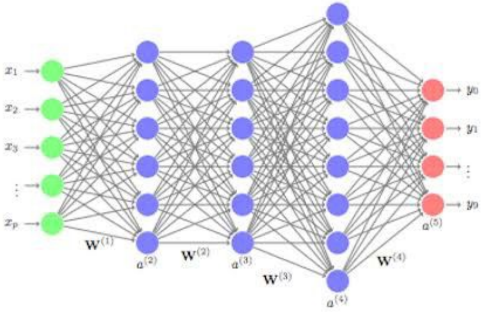
Traditional ensemble learning methods hinges on the assumption that base models are weakly learned (better than random guessing), we remove such assumption: base models can just be randomly generated, without any learning.

We show applications of RHSS with three hypothesis classes.

Kernel Regression



Multi-Layer Perceptron

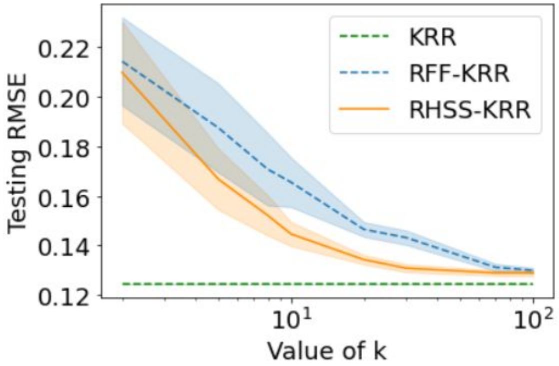


Decision Tree

Performance of RHSS and its model-specific counterpart.

RHSS-KRR converges more efficiently than RFF-KRR as k increases.

Kernel Regression



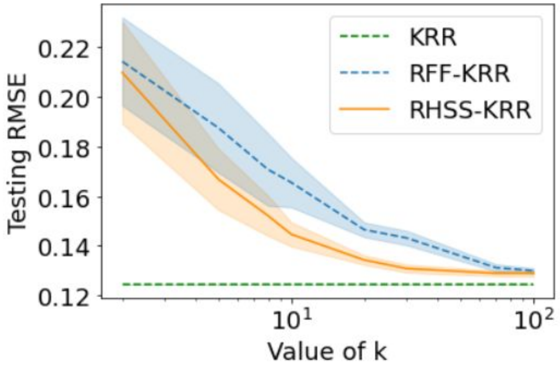
Multi-Layer Perceptron

Decision Tree

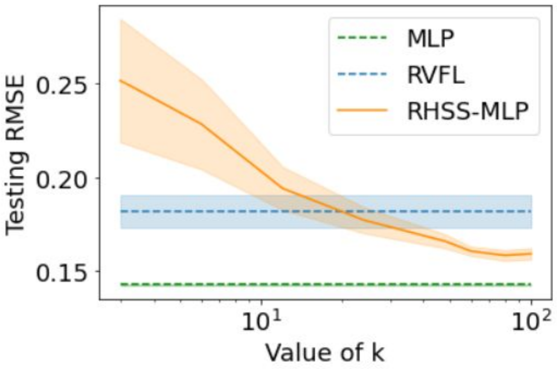
Performance of RHSS and its model-specific counterpart.

RHSS-MLP outperforms RVFL with number of hypothesis > 12.

Kernel Regression



Multi-Layer Perceptron

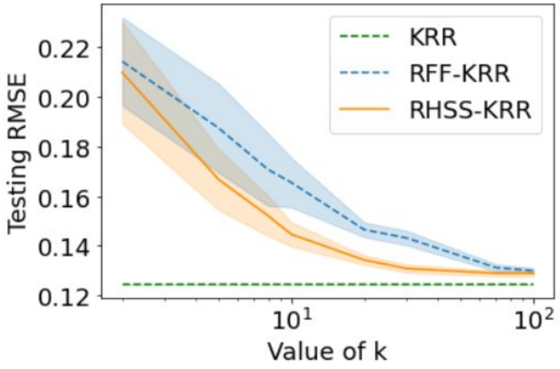


Decision Tree

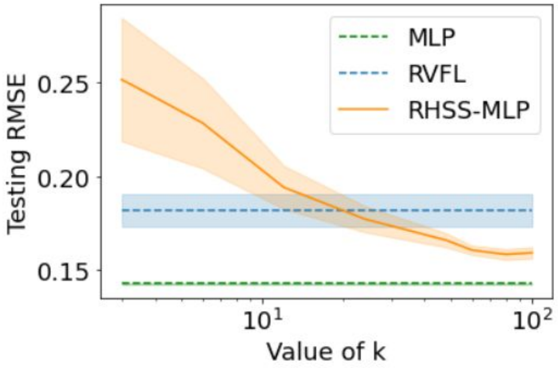
Performance of RHSS and its model-specific counterpart.

RHSS-Tree is similar to extra tree and comparable to random forest.

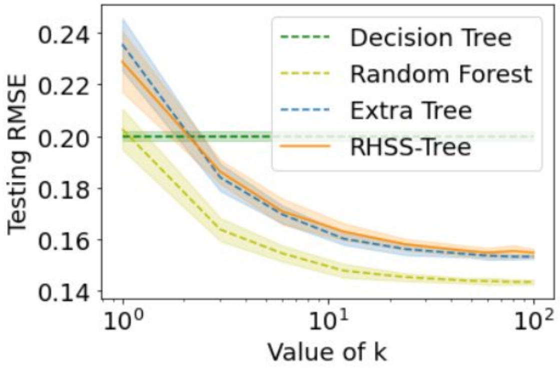
Kernel Regression



Multi-Layer Perceptron



Decision Tree



We derive theoretical guarantees for RHSS.

Under proper conditions, we show w.h.p. the generalization error of RHSS has

$$er(f) \leq er_n(f) + 8TMk\,\mathcal{R}_n(H) + T^2\sqrt{\frac{8\log\frac{1}{\delta}}{n}}$$

, where $er_n(f) = O(\frac{k^{-1/n}}{n})$

Taking the optimal k, the error bound becomes

$$O\left(\frac{1}{n}\right) + O\left(\frac{1}{n^2}\right) + O\left(\frac{1}{\sqrt{n}}\right)$$

Thanks!