

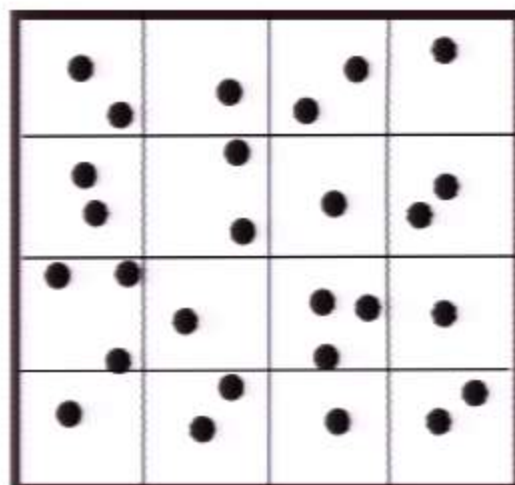
Regression-tree Tuning in a Streaming Setting

Samory Kpotufe^{1,2} and Francesco Orabona¹

¹*Toyota Technological Institute at Chicago*

²*Max Planck Institute for Intelligent Systems*

Tree-based regression



As data size increases ...

- Refine the partition of $\{x_1, x_2, \dots, \dots, x_t\}$
- Recompute $\hat{f}_t \equiv$ average Y in each cell.
Naive way takes time $\Omega(t)$.

Our result

Time $O(\log t)$ with nearly-optimal regression rate!!

Our result

Time $O(\log t)$ with nearly-optimal regression rate!!

Other Algorithm highlights:

- Handles general metric space \mathcal{X} of unknown complexity.
- Adaptation to unknown dimension(\mathcal{X}).
- Robustness to adversarial data (theory + simulations).

For details come to our poster **Sat40**