AnoMed

S-GBDT: Frugal Differentially Private Gradient Boosting Decision Trees

Moritz Kirschte†, **Thorsten Peinemann**†, Joshua Stock⋆, Carlos Cotrini⋄, Esfandiar Mohammadi†.

The first two authors contributed equally to this work.

† Universität zu Lübeck ⋆ Universität Hamburg ⋄ ETH Zürich

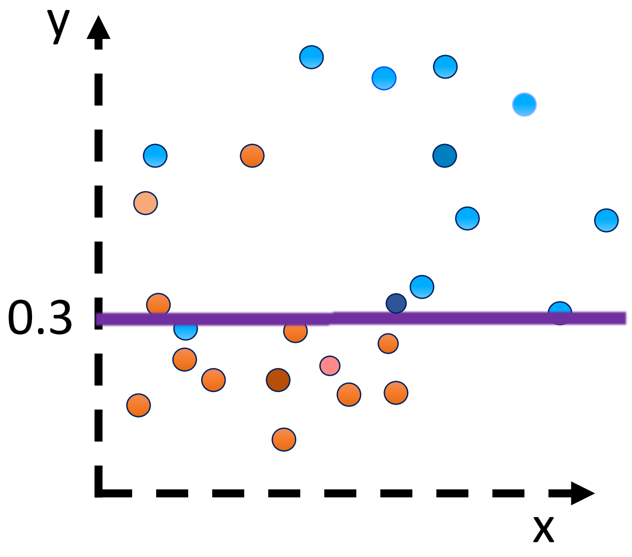
UNIVERSITÄT ZU LÜBECK
INSTITUT FOR IT SECURITY
PRIVACY & SECURITY GROUP

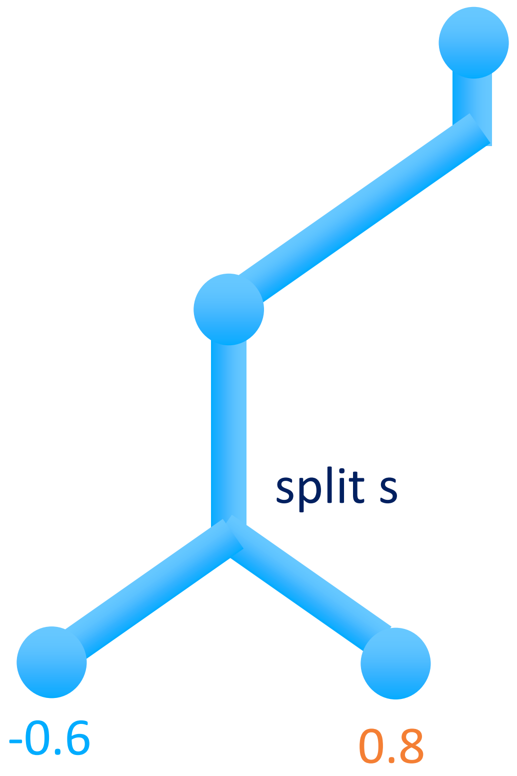
Finanziert von der Europäischen Union
NextGenerationEU

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

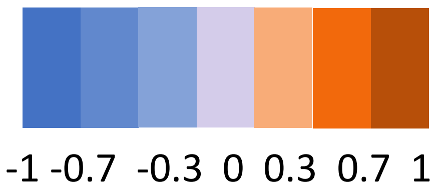
Regression using decision trees

- Task: Regress data points (x, y : features)
- Step 1: Split



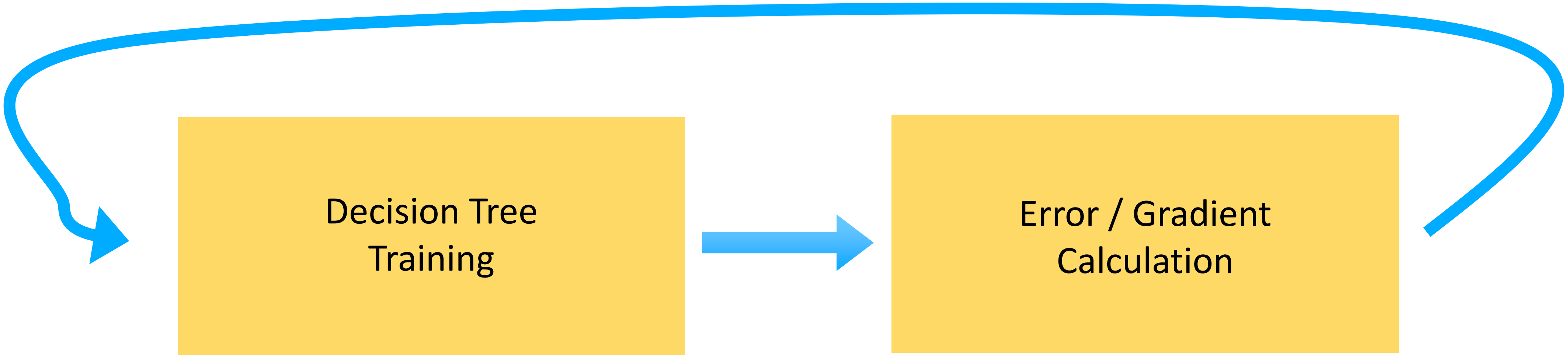


- Step 2: Predict
  - If feature y >= 0.3, then blueish, i.e. -0.6
  - If feature y < 0.3, then orangeish, i.e. 0.8



Gradient Boosting Decision Trees Ensemble

Training data:
Inputs (x,y); regression values / gradients

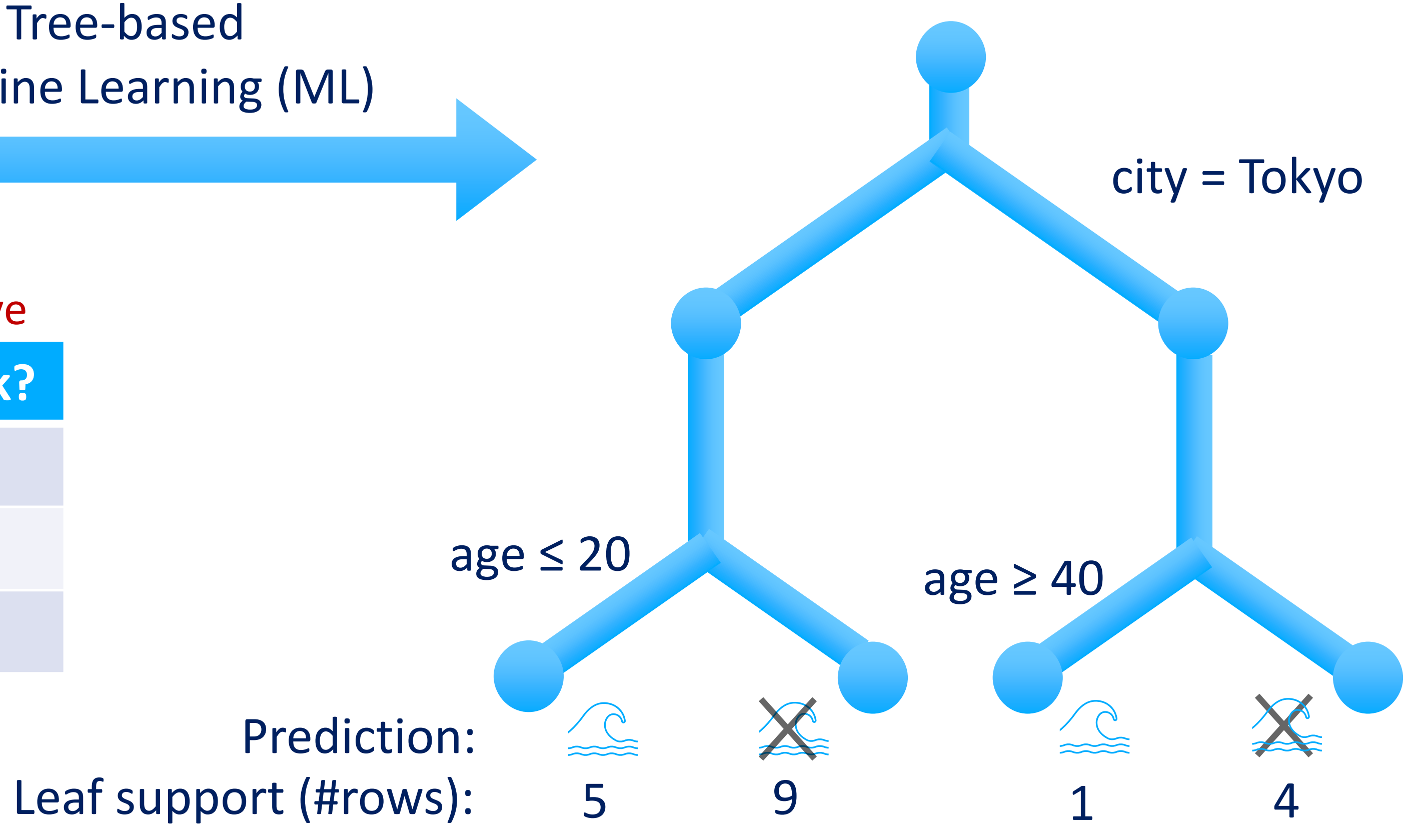


Tree-based ML applications can pose a privacy risk [1]

Training data
e. g. sensitive medical
survey data

Tree-based
Machine Learning (ML)

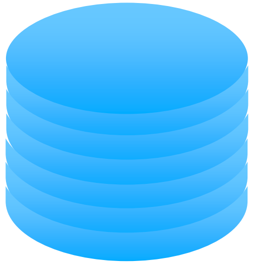
Public	Public	Public	Sensitive
Name	**City**	**Age**	**Seasick?**
Manu	Tokyo	42	
Nyasha	Hamburg	21	
…	…	…	…



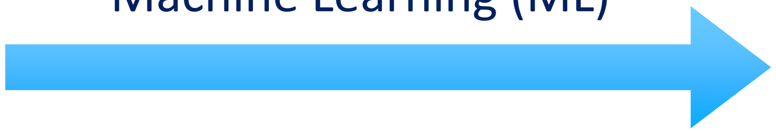
[1] Frederikson, Jha, and Ristenpart, *Model inversion attacks that exploit confidence information and basic countermeasures.* In: ACM CCS, 2015.

Tree-based ML applications can pose a privacy risk [1]

Training data
e. g. sensitive medical
survey data

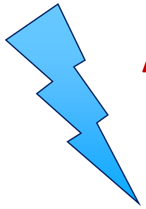


Tree-based
Machine Learning (ML)

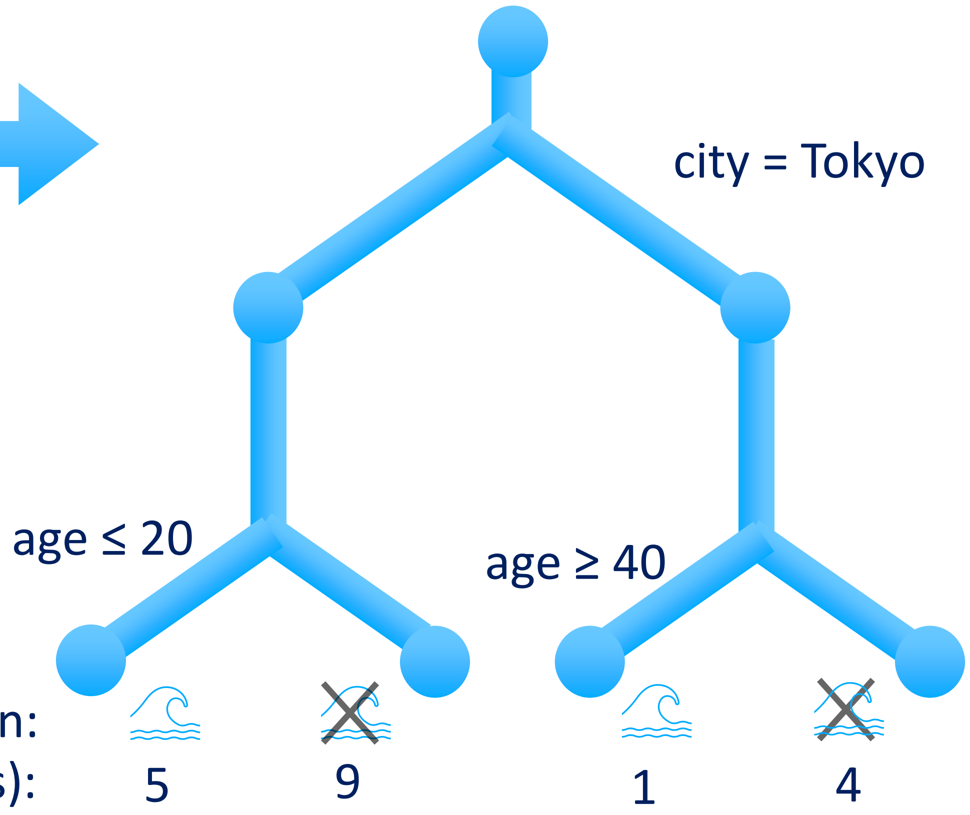


Public	Public	Public	Sensitive
Name	**City**	**Age**	**Seasick?**
Manu	Tokyo	42	
Nyasha	Hamburg	21	
…	…	…	…

Privacy Attack
on tree



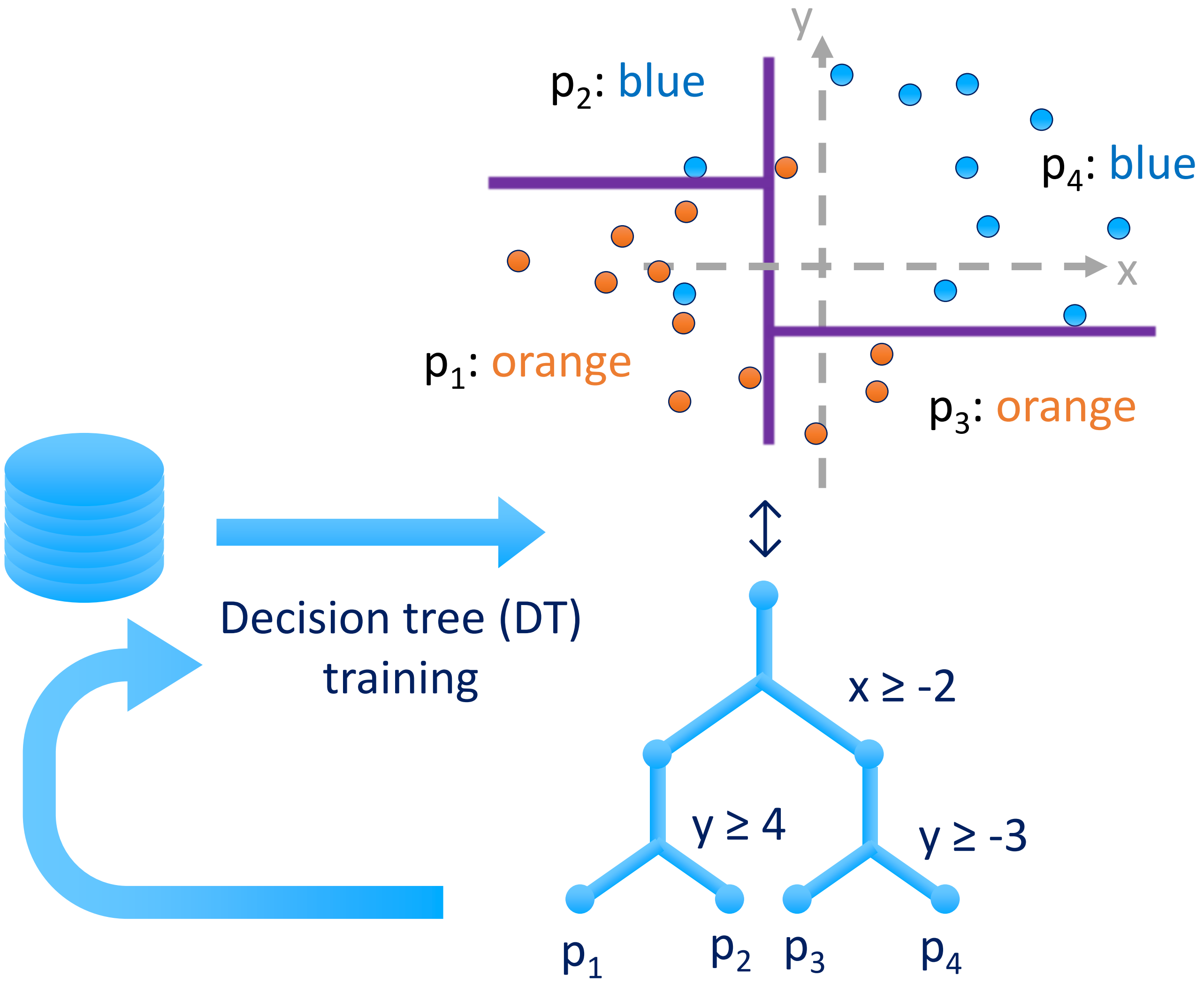
Attacker learns:
Manu must be
seasick.



[1] Frederikson, Jha, and Ristenpart, *Model inversion attacks that exploit confidence information and basic countermeasures.* In: ACM CCS, 2015.

Prior work [2] counters tree-based privacy attacks as follows and sacrifices utility

Repeat *m*-times for DT ensemble with *Gradient Boosting*



: 2 classes

[2] Q. Li, Z. Wu, Z. Wen, and B. He, *Privacy-preserving gradient boosting decision trees*. AAAI, vol. 34, no. 01, pp. 784–791, 2020.

Prior work [2] counters tree-based privacy attacks as follows and sacrifices utility

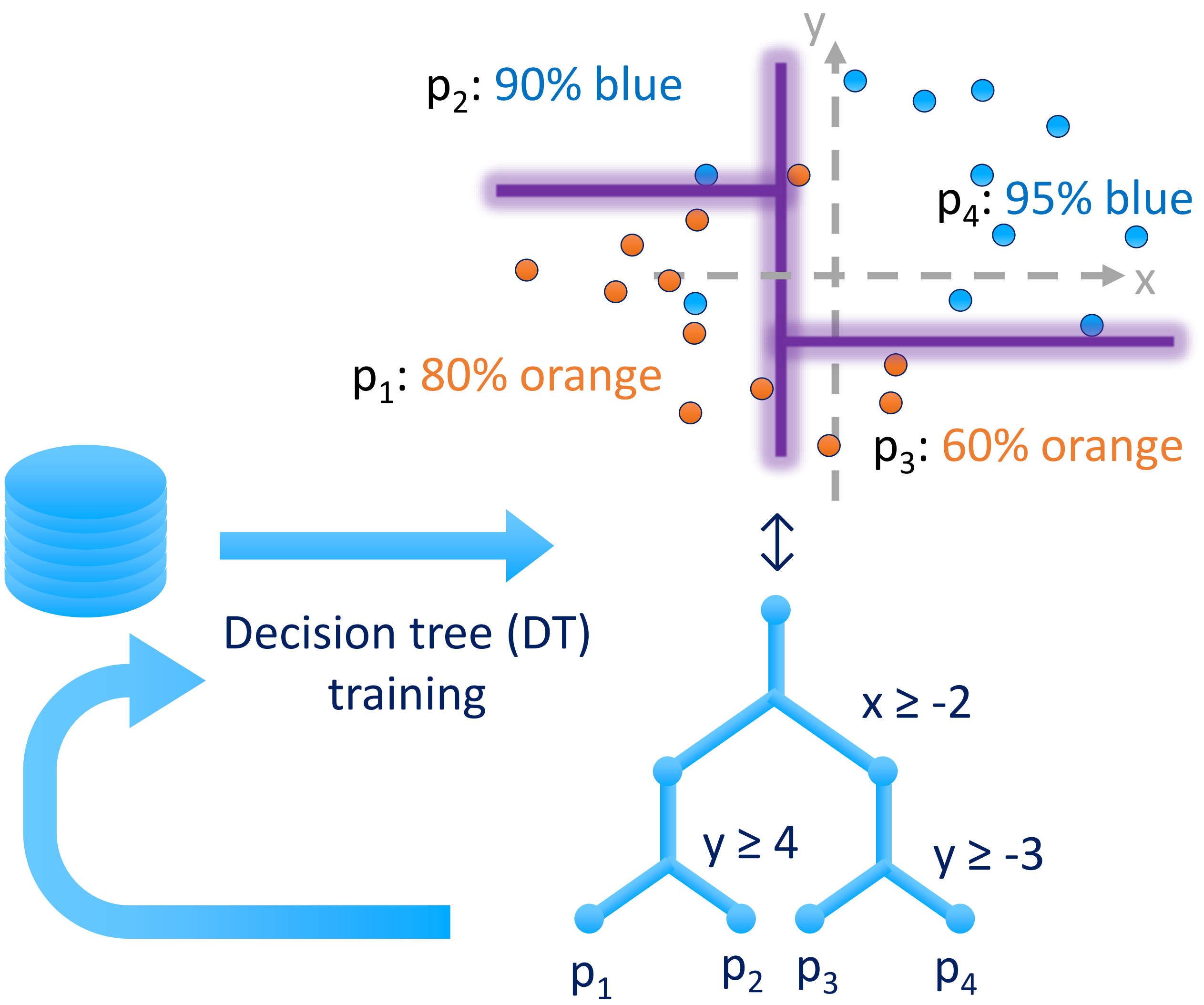
1. Noise optimal split

→ sample of PDF

2. Noise leaf predictions

$p_{\{1,2,3,4\}} + \mathcal{N}(0, \sigma)$

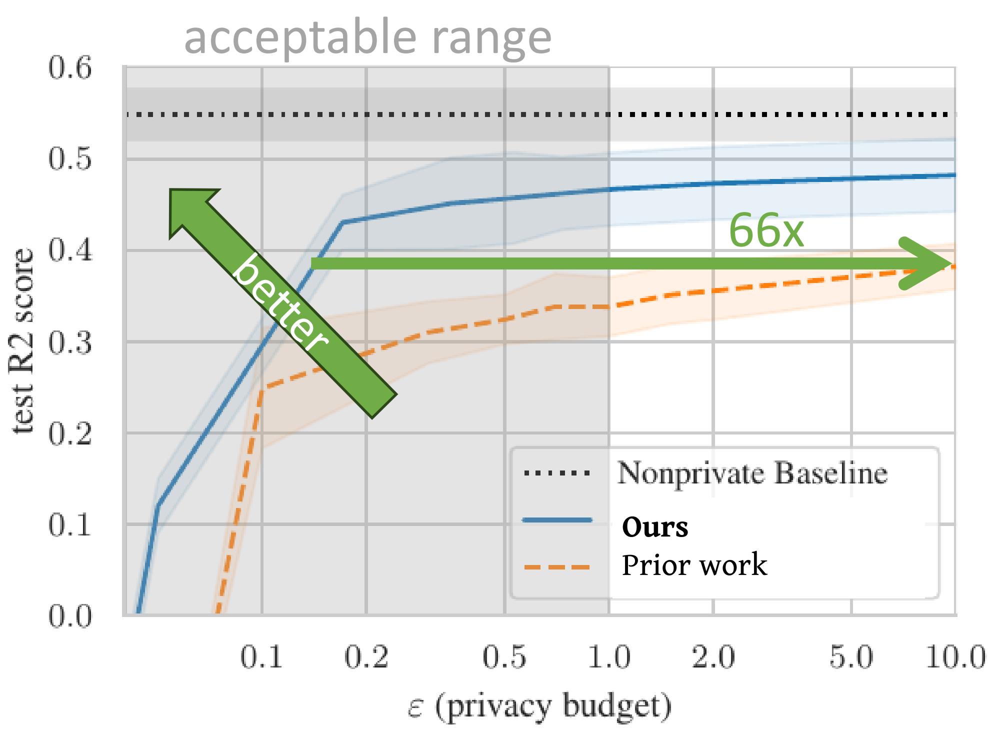
Repeat *m*-times for DT ensemble with *Gradient Boosting*



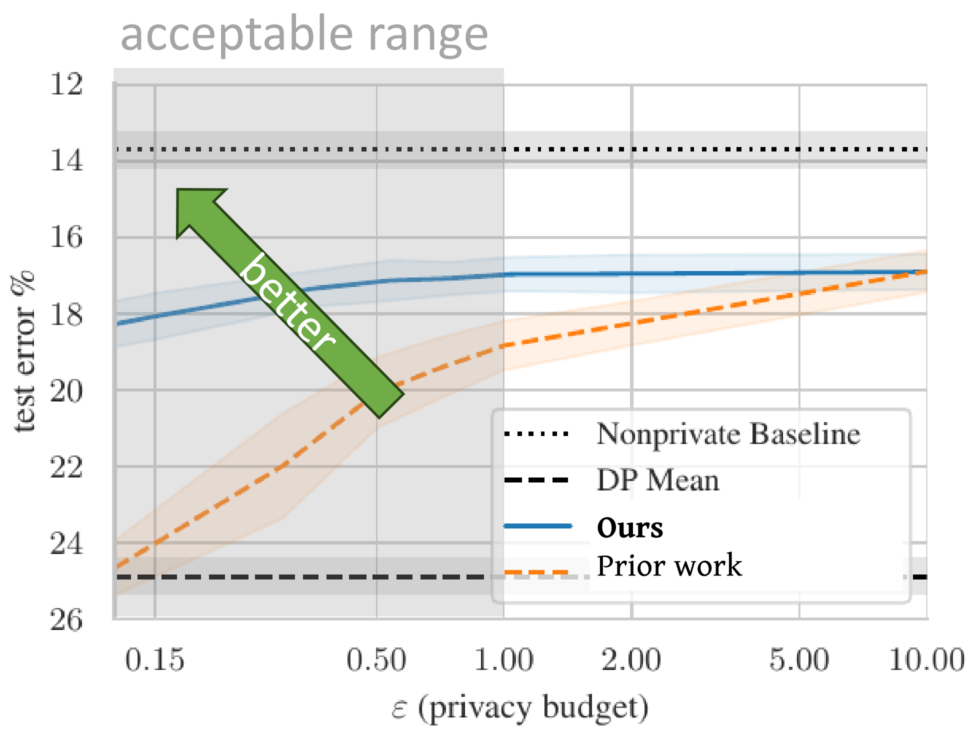
: 2 classes

[2] Q. Li, Z. Wu, Z. Wen, and B. He, *Privacy-preserving gradient boosting decision trees*. AAAI, vol. 34, no. 01, pp. 784–791, 2020.

Experimental Results



(a) Abalone (Regression, < 4k data points)



(b) Adult (Binary classification, 48k data points)

(ε, δ)-Differential Privacy

- State of the art notion for provable privacy protection
- Deflects many privacy attacks
- Requires that the impact of single data points is limited and deniable
- We consider unbounded DP (add/remove relationship)

$$\Pr[M(D) = o] \leq e^{\varepsilon} \Pr[M(D \cup \{x\}) = o] + \delta$$

randomized algorithm

worst-case dataset and challenge element

$(\alpha, \rho(\alpha))$-Rényi Differential Privacy

- Rényi divergence of order α for any two probability distributions P, Q

density of P at x

$$D_\alpha(P||Q) = \frac{1}{\alpha - 1} \log \int_{-\infty}^{\infty} \frac{P(x)^\alpha}{Q(x)^{\alpha-1}} dx$$

density of Q at x

- $(\alpha, \rho(\alpha))$ -Rényi DP

$$D_\alpha(M(D)||M(D \cup \{x\})) \leq \rho(\alpha)$$

randomized algorithm

worst-case dataset and challenge element

Our Improvements

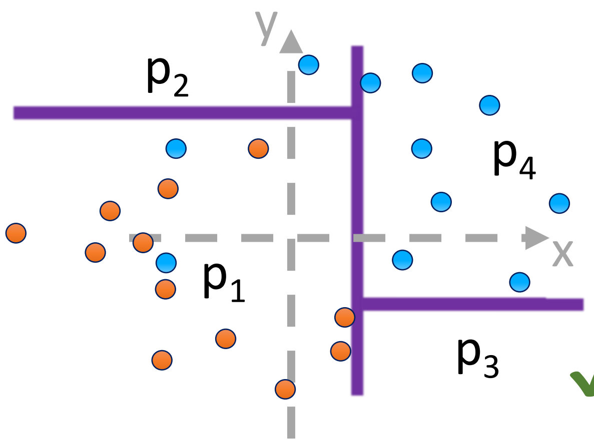
Finally usable DP-GBDT!

Our Improvements

- **Rényi DP bound for Gaussian mechanism with non-spherical noise**
- **Dynamic leaf noise scaling**
- **Learning on streams of non-i.i.d. data via an individual Rényi filter**

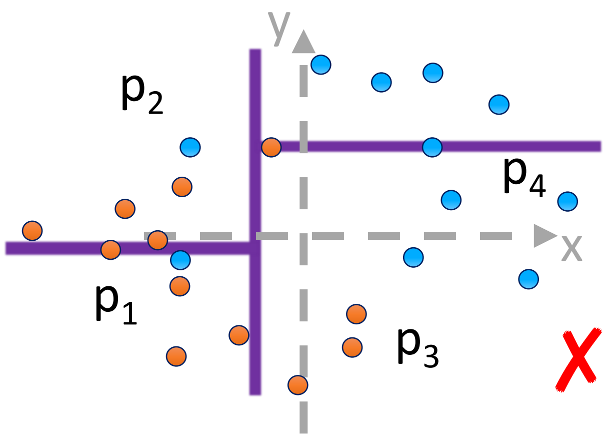
- **Incorporate random decision trees**
- Subsampled tree learning with individual Rényi filter tailored to S-GBDT
- Extension for distributed learning
- Fixing secondary privacy leakages

Random splits enjoy little utility loss while being privacy protective



✓ Good tree by chance

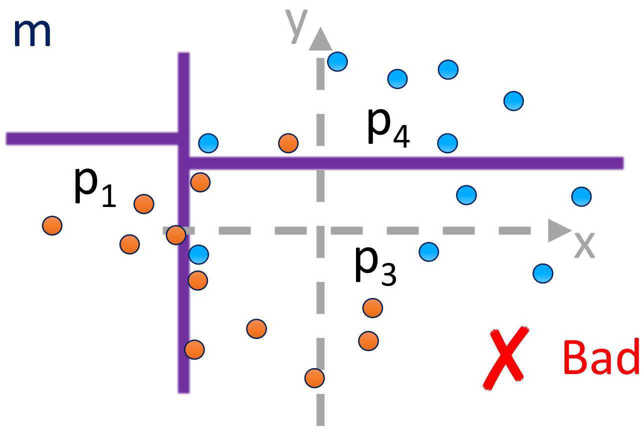
Tree 1 saves the ensemble.



✗ Bad tree

…

Tree m



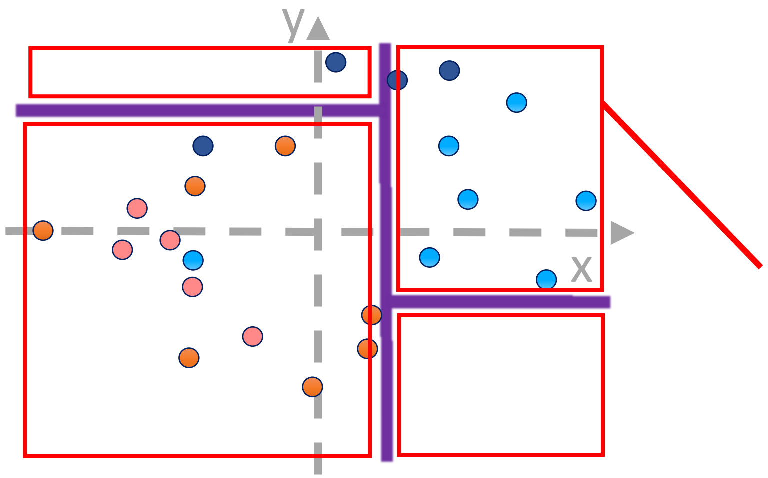
✗ Bad tree

In general: It suffices[3] that only few trees of the ensemble are useful.

[3] M. Bojarski, A. Choromanska, K. Choromanski, and Y. LeCun, *Differentially-and non-differentially-private random decision trees*. In: arXiv preprint, arXiv:1410.6973, 2014.

Differentially private leaf computation

: data points, shade of color indicates gradient



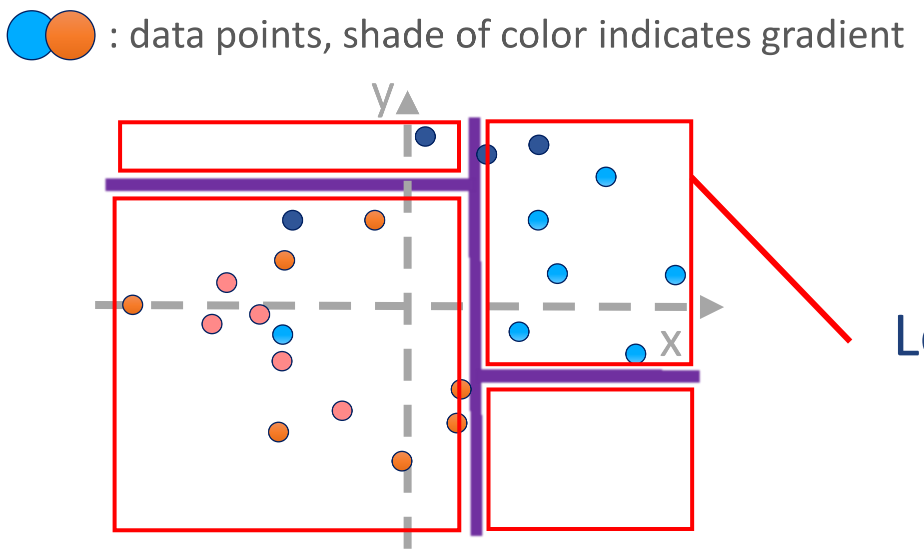
Leaf value: $\dfrac{\sum_{i=1}^{|\mathrm{leaf}|} \mathrm{gradient_i}}{\max(\lambda, |\mathrm{leaf}|)}$

gradient sum

Prior work DP leaf value: $\dfrac{\sum_{i=1}^{|\mathrm{leaf}|} \mathrm{Clip}(\mathrm{gradient_i}, g^*)}{\max(\lambda, |\mathrm{leaf}|)} + \mathcal{N}\left(0, O(\frac{g^*}{\varepsilon})\right)$

leaf support

Differentially private leaf computation



Leaf value: $\frac{\sum_{i=1}^{|\text{leaf}|} \text{gradient}_\text{i}}{\max(\lambda, |\text{leaf}|)}$

sensitivity $= g^*$

Prior work DP leaf value: $\frac{\sum_{i=1}^{|\text{leaf}|} \text{Clip}(\text{gradient}_\text{i}, g^*)}{\max(\lambda, |\text{leaf}|)} + \mathcal{N}\left(0, O(\frac{g^*}{\varepsilon})\right)$

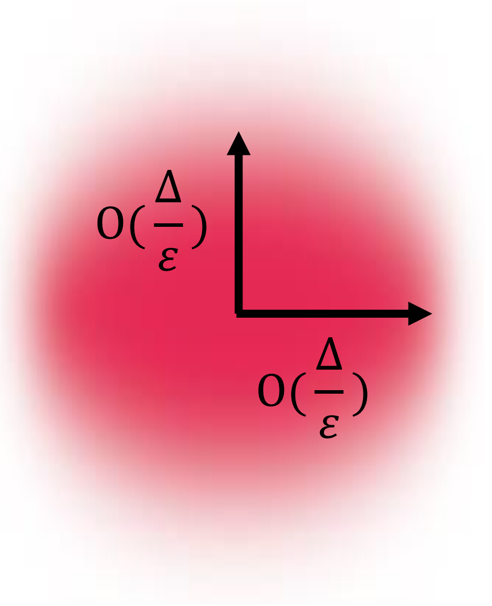
sensitivity $= 1$

- Noise does not scale with number of data points $\boldsymbol{n}$
- Privacy budget can not be shifted between gradient sum and leaf support

Dynamic leaf noise scaling

- Release both gradient sum and leaf support instead of averaged sum

DP Leaf value: $$\frac{\sum_{i=1}^{|\mathrm{leaf}|} \mathrm{Clip}(\mathrm{gradient}_{\mathrm{i}},\ g^*) + \mathcal{N}\left(0, \mathrm{O}(\frac{\Delta}{\varepsilon})\right)}{\max(\lambda, |\mathrm{leaf}| + \mathcal{N}\left(0, \mathrm{O}(\frac{\Delta}{\varepsilon})\right))}$$



$$\Delta = \sqrt{1 + (g^*)^2}$$

Dynamic leaf noise scaling

Gradient sum divided by (noisy) leaf support

Gradient sum noise divided by (noisy) leaf support

DP Leaf value: $\frac{\sum_{i=1}^{|\text{leaf}|} \text{Clip}(\text{gradient}_\text{i},\ g^*)}{\widetilde{leaf}} + \frac{\mathcal{N}\left(0, \mathrm{O}(\frac{\Delta}{\varepsilon})\right)}{\widetilde{leaf}}$

$\Delta = \sqrt{1 + (g^*)^2}$

$\widetilde{leaf} = \max(\lambda, |\text{leaf}| + \mathcal{N}\left(0, \mathrm{O}(\frac{\Delta}{\varepsilon})\right))$

- ✓ Noise scales in $O(\frac{1}{n})$ for number of data points $\boldsymbol{n}$
- ❖ Privacy budget can not be shifted between gradient sum and leaf support

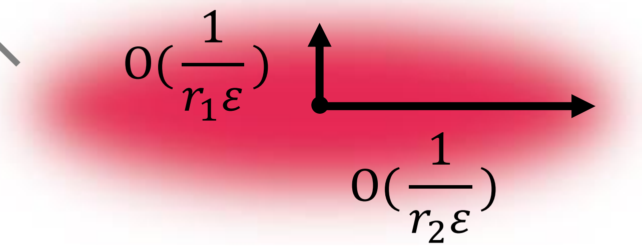
Non-spherical noise

sensitivity $= g^*$

Clipped leaf value: $$\frac{\sum_{i=1}^{|\mathrm{leaf}|} \mathrm{Clip}(\mathrm{gradient_i},\ g^*)}{\max(\lambda, |\mathrm{leaf}|)}$$

sensitivity $= 1$

$$\mathcal{N}\left(\mathbf{0}, \begin{pmatrix} O(\frac{1}{r_1 \varepsilon}) & 0 \\ 0 & O(\frac{1}{r_2 \varepsilon}) \end{pmatrix}\right),\ s.t.\ r_1 + r_2 = 1$$



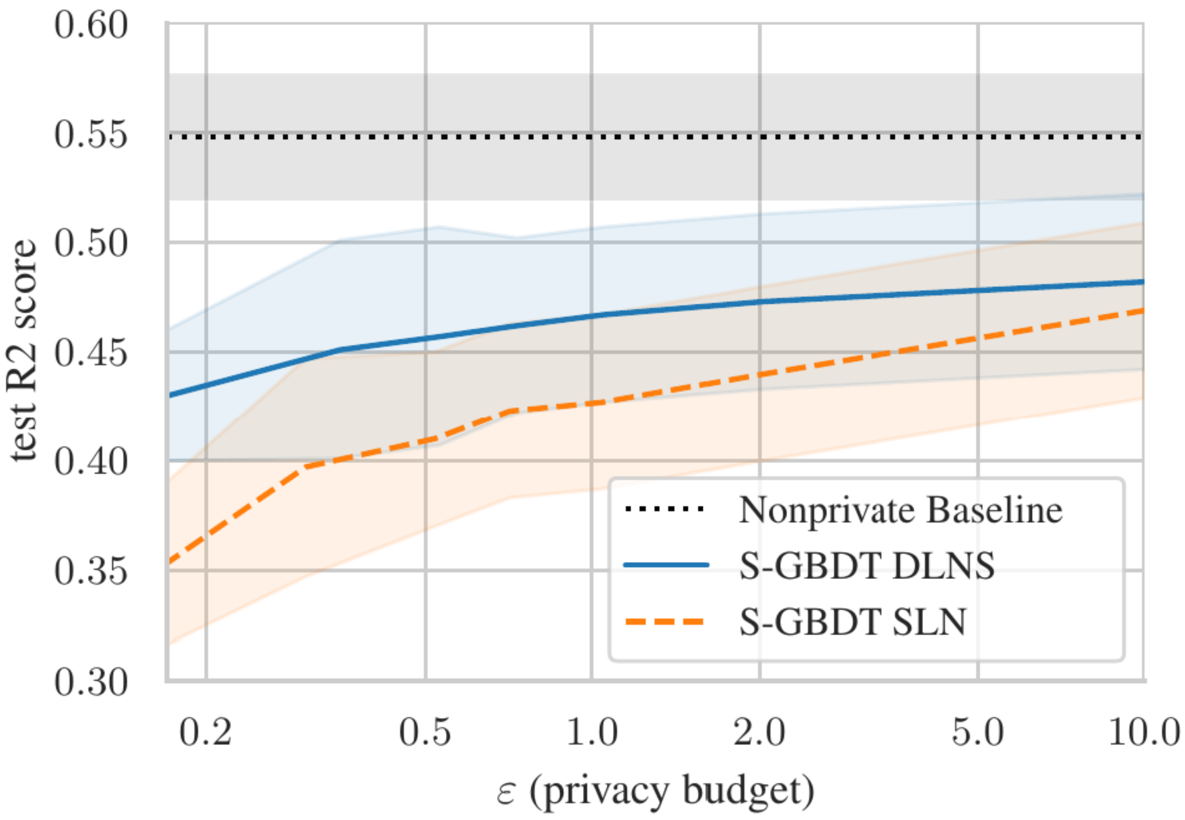
DP Leaf value: $$\frac{\sum_{i=1}^{|\mathrm{leaf}|} \mathrm{Clip}(\mathrm{gradient_i},\ g^*) + \mathcal{N}\left(0, O(\frac{1}{r_1 \varepsilon})\right)}{\max(\lambda, |\mathrm{leaf}| + \mathcal{N}\left(0, O(\frac{1}{r_2 \varepsilon})\right))}$$

RDP bound: $$\rho(\alpha) = \alpha \frac{r_1 + r_2 * (g^*)^2}{\sigma^2}$$

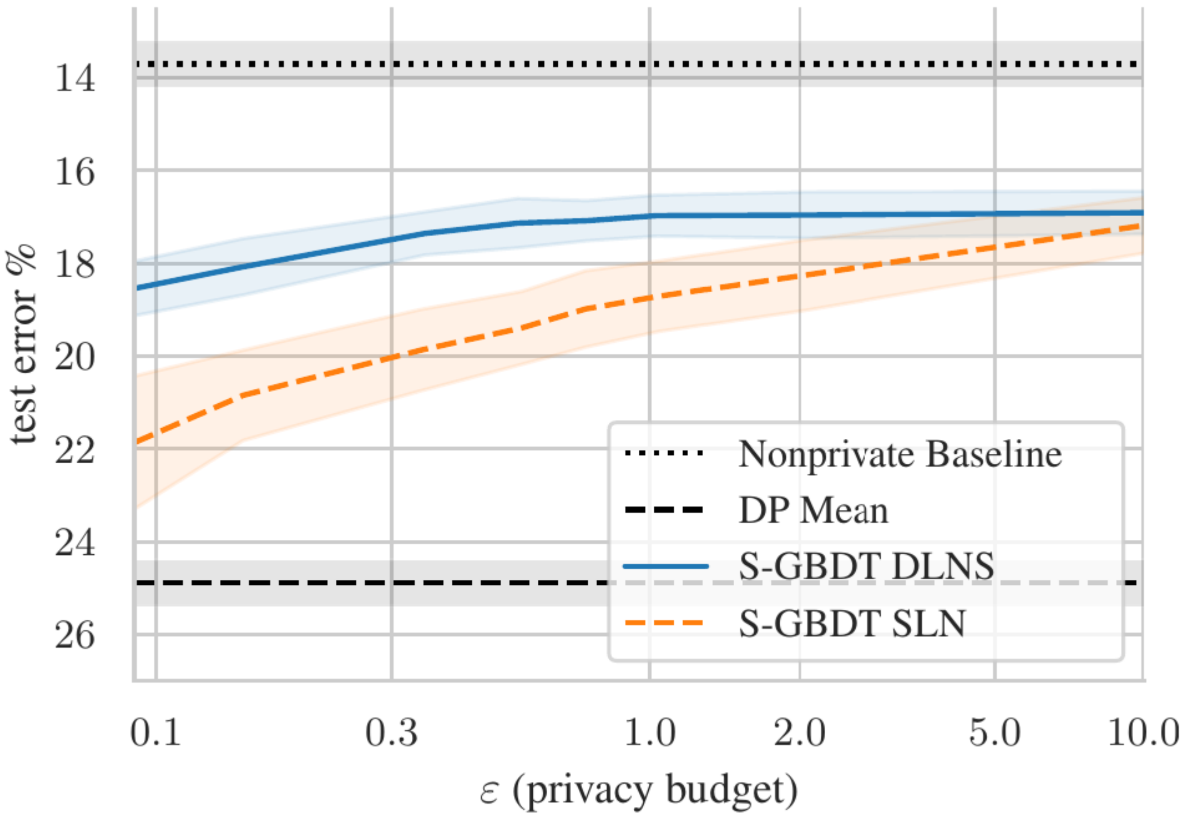
see paper for proof
or ask me

✓ Privacy budget can be shifted between gradient sum and leaf support

Experimental Results



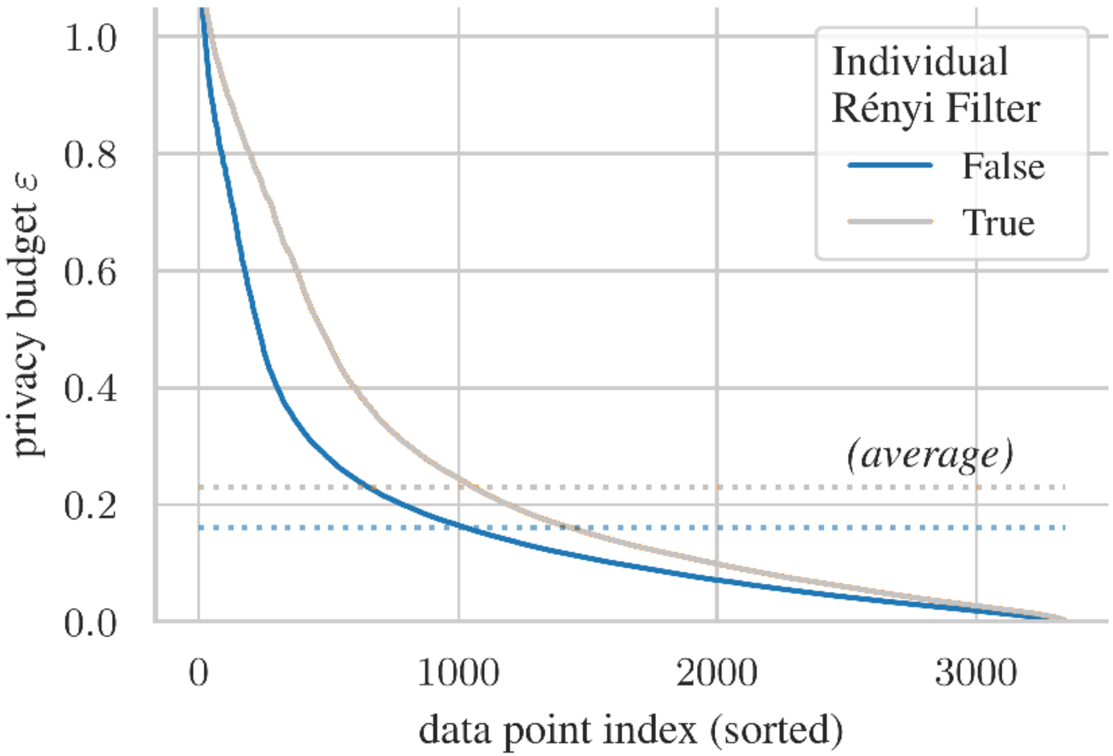
(a) **Abalone**



(b) **Adult**

Conventional Privacy Accounting

- Conventional approach for privacy accounting: same worst case analysis applied to all data points
- Can result in overly conservative estimation of privacy loss for many data points
- Data points in S-GBDT often do not fully utilize sensitivity (i.e. $g_i < g^*$)



Individual Rényi DP [4]

Individual $(\alpha, \rho(\alpha))$-Rényi DP for data point $\boldsymbol{x_i}$: $D_\alpha(M(D)||M(D \cup \{\boldsymbol{x_i}\})) \leq \rho^{(i)}(\alpha)$

worst case dataset

Individual RDP bound for releasing leaf value: $\rho^{(i)}(\alpha) = \alpha \frac{r_1 + r_2 * (\boldsymbol{g_i})^2}{\sigma^2}$

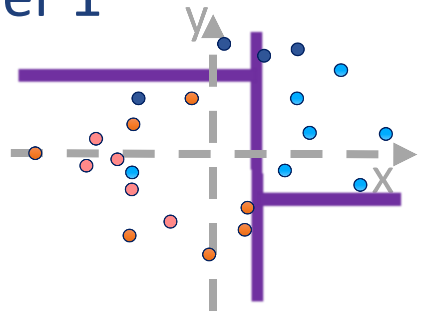
value has leakage when released,
but can be utilized by individual Rényi filter

[4] V. Feldman and T. Zrnic, *Individual Privacy Accounting via a Renyi Filter.* In: NeurIPS, pp. 28080–28091, 2021.

Tailoring individual Rényi filters [4] to S-GBDT

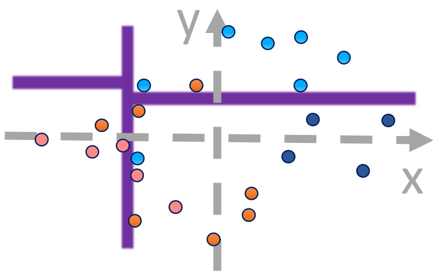
Regular rounds

Classifier 1



...

Classifier m



Datapoint	Budget
$(x_1, y_1, \mathrm{label}_1)$	0%
$(x_2, y_2, \mathrm{label}_2)$	0%
...	...

Datapoint	Budget
$(x_1, y_1, \mathrm{label}_1)$	50%
$(x_2, y_2, \mathrm{label}_2)$	90%
...	...

individual Rényi divergence budget $\rho^{(i)}(\alpha)$

Upper bound on RDP privacy loss for m trees: $m * \rho(\alpha)$

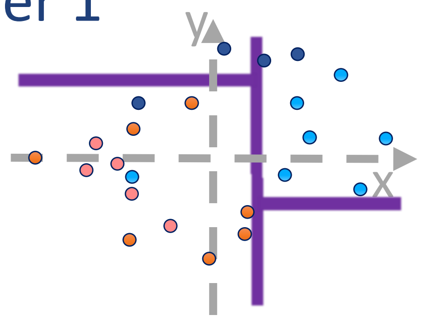
[4] V. Feldman and T. Zrnic, *Individual Privacy Accounting via a Renyi Filter.* In: NeurIPS, pp. 28080–28091, 2021.

Tailoring individual Rényi filters [4] to S-GBDT

Regular rounds

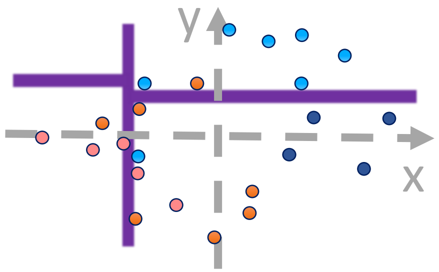
Extra rounds due to Rényi Filter

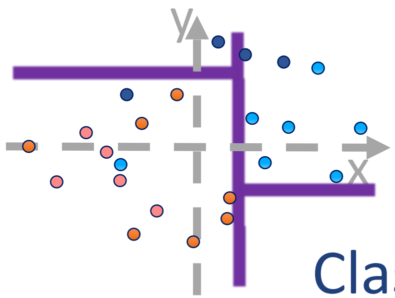
Classifier 1



...

Classifier m





Classifier m+5

Datapoint	Budget
$(x_1, y_1, \text{label}_1)$	0%
$(x_2, y_2, \text{label}_2)$	0%
...	...

Datapoint	Budget
$(x_1, y_1, \text{label}_1)$	50%
$(x_2, y_2, \text{label}_2)$	90%
...	...

Datapoint	Budget
$(x_1, y_1, \text{label}_1)$	75%
$\cancel{(x_2, y_2, \text{label}_2)}$	100%
...	...

Excluded from training

individual Rényi divergence budget $\rho^{(i)}(\alpha)$

Upper bound on RDP privacy loss for m trees: $m * \rho(\alpha)$

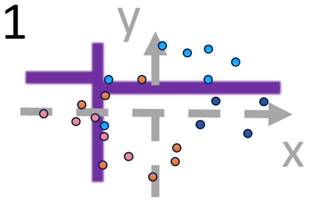
[4] V. Feldman and T. Zrnic, *Individual Privacy Accounting via a Renyi Filter.* In: NeurIPS, pp. 28080–28091, 2021.

Insight: Individual Rényi filters [4] are effective for streams of non-i.i.d. Data

Regular rounds

Extra rounds due to Rényi Filter

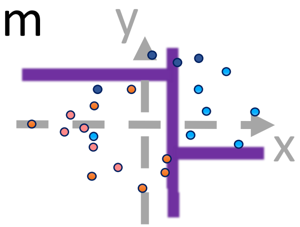
Classifier 1



Datapoint	Budget
$(x_1, y_1, \text{label}_1)$	0%
$(x_2, y_2, \text{label}_2)$	0%
…	…

…

Classifier m



Datapoint	Budget
$(x_1, y_1, \text{label}_1)$	50%
$(x_2, y_2, \text{label}_2)$	30%
…	…

new data arrives

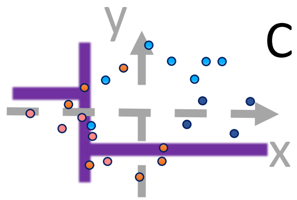
$(x_a, y_a, \text{label}_a)$	0%
$(x_b, y_b, \text{label}_b)$	0%

new data can have a different distribution

i.e. different value range for attributes and labels

…

Classifier m+5



…

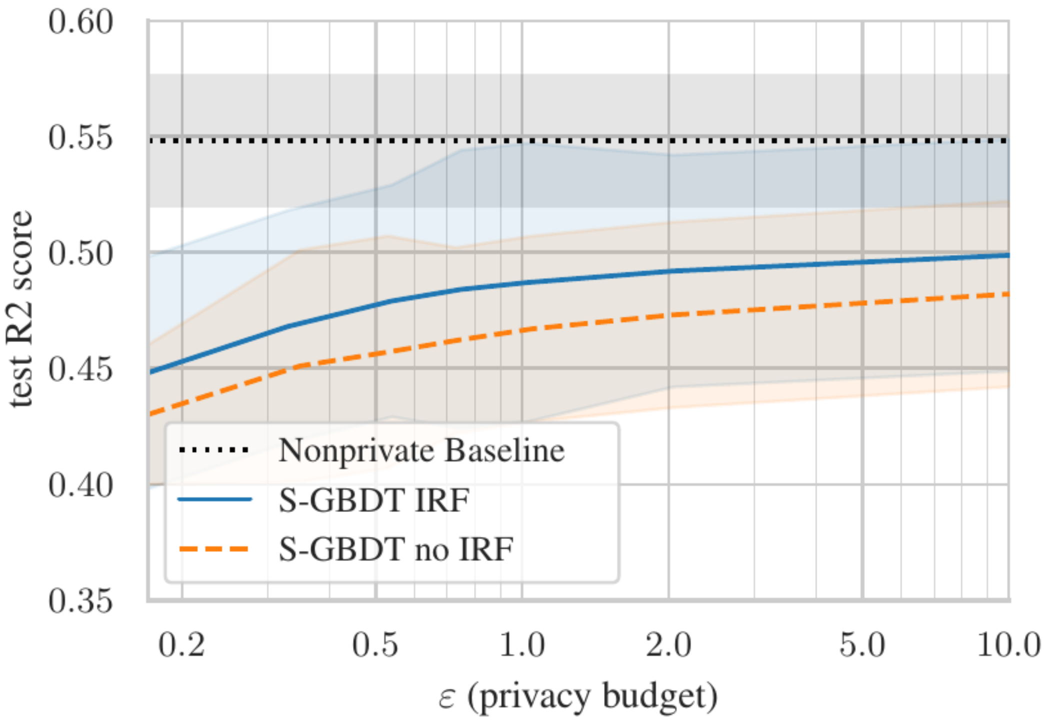
Datapoint	Budget
$(x_1, y_1, \text{label}_1)$	85%
$(x_2, y_2, \text{label}_2)$	45%
…	…
$(x_a, y_a, \text{label}_a)$	30%
$(x_b, y_b, \text{label}_b)$	40%

with Rényi filter: incorporate new data in training

without Rényi filter: retrain (expensive) and potentially double privacy budget

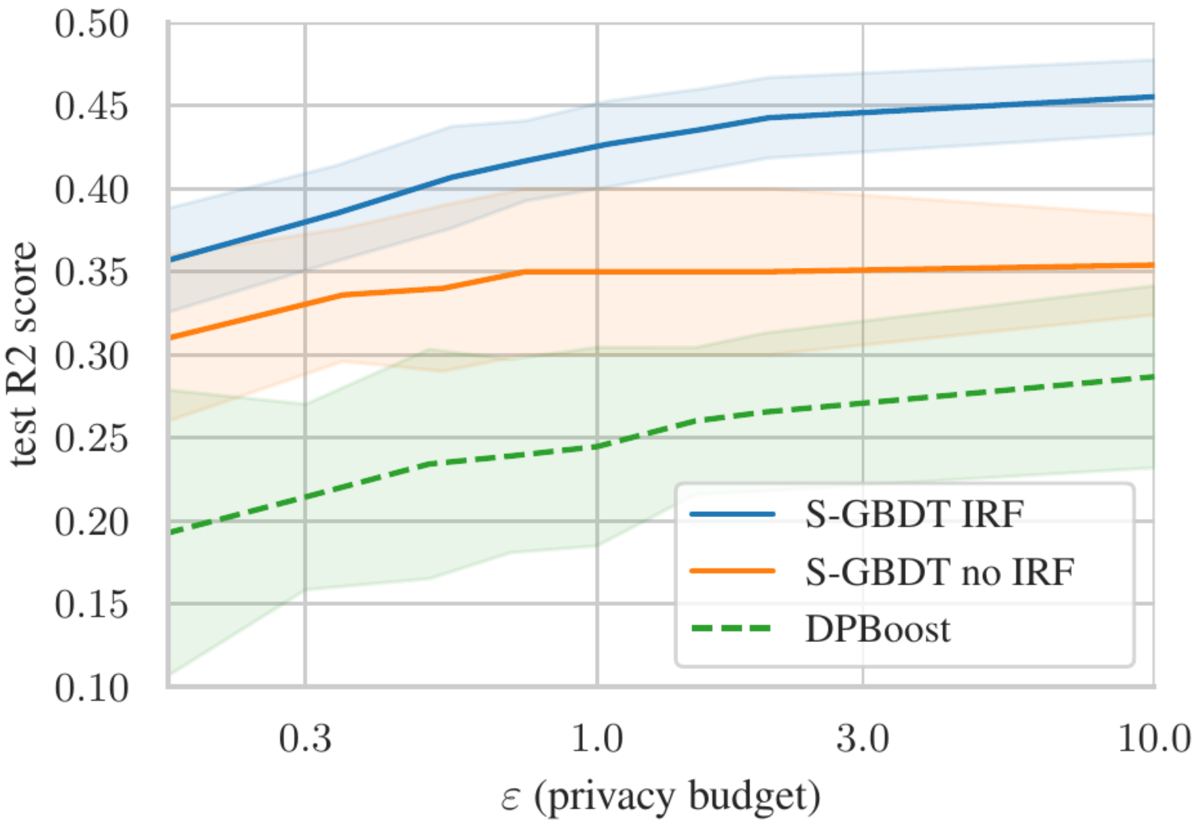
[4] V. Feldman and T. Zrnic, *Individual Privacy Accounting via a Renyi Filter.* In: NeurIPS, pp. 28080–28091, 2021.

Experimental Results



(a) **Abalone**

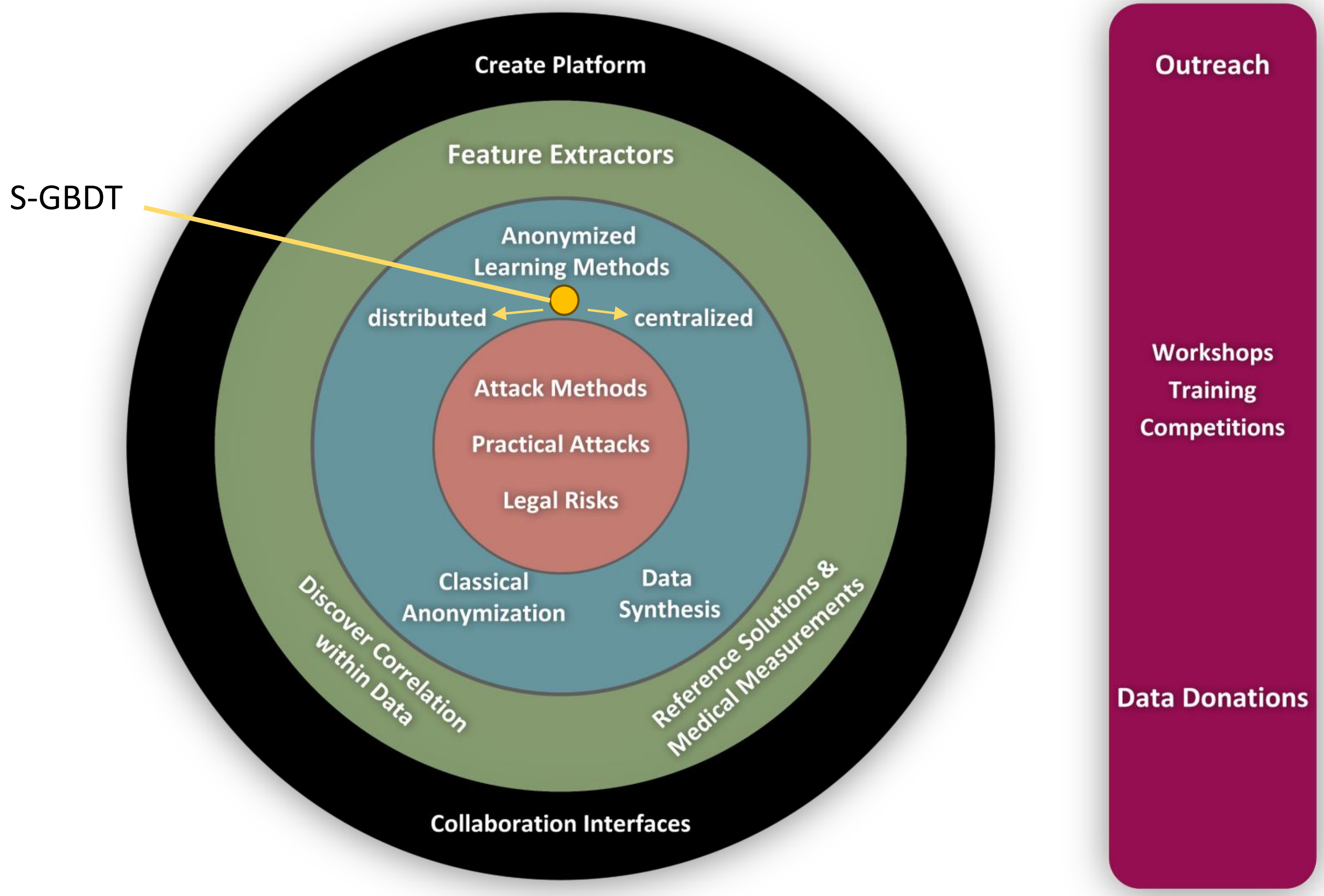
regular training



(a) **Abalone**

learning on streams

Contribution to AnoMed



Full version on arXiv