Classification under local differential privacy [1]

Thomas Berrett, Cristina Butucea

Alexander Baumann

April 12, 2022

Table of Contents

1. Setting
2. Construction of the privacy mechanism
3. Excess risk of the classifier
   - Formulation of the main theorem
   - Outline of the proof
4. Minimax rate of convergence of the excess risk
5. Comparison to non-privatized setup
6. Numerical experiments

Table of Contents

1. Setting
2. Construction of the privacy mechanism
3. Excess risk of the classifier
   - Formulation of the main theorem
   - Outline of the proof
4. Minimax rate of convergence of the excess risk
5. Comparison to non-privatized setup
6. Numerical experiments

Setting

- Given original data $(X_i, Y_i) \in \mathbb{R}^d \times \{0, 1\}$ with $1 \leq i \leq n$
- Produce randomized variables Z_i on a measurable space $(\mathcal{Z}, \mathcal{B})$, which represent the privatized data.
- Here, we consider the non-interactive local case, i.e. $\{X_i, Y_i\} \to Z_i$. This is defined through a Markov kernel (Appendix Definition 6):

$$Q_i : \mathcal{B} \times \left(\mathbb{R}^d \times \{0, 1\}\right) \to [0, 1]$$

$$Q_i\left(\cdot | x_i, y_i\right) \sim Z_i | \{X_i = x_i, Y_i = y_i\}$$

Setting

Definition 1

Let $\alpha > 0$.
In the setting from before, a privacy mechanism is called $\alpha-$locally differentially private ($\alpha-$LDP) if:

$$\sup_{A_i \in \mathcal{B}} \sup_{\substack{(x_i, y_i) \\ (x_i', y_i')}} \frac{Q_i\left(A_i | x_i, y_i\right)}{Q_i\left(A_i | x_i', y_i'\right)} \leq e^{\alpha} \qquad \forall i \in \{1, \ldots, n\}$$

In this case, write $Q \in \mathcal{Q}_{\alpha}$.

Interpretation:
If $\alpha \to 0$ and hence $exp(\alpha) \to 1$, then the privatized data is almost untraceable.

Table of Contents

1. Setting
2. Construction of the privacy mechanism
3. Excess risk of the classifier
   - Formulation of the main theorem
   - Outline of the proof
4. Minimax rate of convergence of the excess risk
5. Comparison to non-privatized setup
6. Numerical experiments

Construction of the privacy mechanism

Given original data $(X_i, Y_i) \in \mathbb{R}^d \times \{0, 1\}$ with $1 \leq i \leq 2n$, bandwidth parameter $h \in (0, \infty)$ and $\alpha > 0$.

- For $j \in \mathbb{Z}^d$, define $x_j := h \cdot j$
- Define $B_i := \left(\mathbb{1}_{\{\|X_i - x_j\|_\infty < h\}}\right)_{j \in \mathbb{Z}^d} \in \mathbb{R}^{\mathbb{Z}^d}$
- Define $\epsilon_i := (\epsilon_{ij})_{j \in \mathbb{Z}^d} \in \mathbb{R}^{\mathbb{Z}^d}$, where $\epsilon_{ij} \sim \mathit{Laplace}\left(0, \frac{2^{d+1}}{\alpha}\right)$
- Define the privatized variables

$$Z_i := \begin{cases} B_i + \epsilon_i, & i \leq n \\ Y_i \cdot B_i + \epsilon_i, & n < i \leq 2n \end{cases}$$

$\Rightarrow$ non-interactive local privacy mechanism

Construction of the privacy mechanism

Theorem 1

The defined privacy mechanism is $\alpha-$LDP.

Table of Contents

1. Setting
2. Construction of the privacy mechanism
3. Excess risk of the classifier
    - Formulation of the main theorem
    - Outline of the proof
4. Minimax rate of convergence of the excess risk
5. Comparison to non-privatized setup
6. Numerical experiments

Excess risk of a classifier

- Define $\eta(x) := \mathbb{P}[Y = 1|X = x]$
- Bayes Classifier: $C^B(x) := \mathbb{1}_{\{\eta(x) \geq 0.5\}}$

$\Rightarrow$ Minimising the risk $R_P(C) := \mathbb{P}[C(X) \neq Y]$ over all classifiers C

- Excess risk:

$$\begin{aligned}\mathcal{E}_P(C) &:= R_P(C) - R_P\left(C^B\right) \\ &= \mathbb{E}\left[\left(\mathbb{P}[C(X) = 0|X] - \mathbb{1}_{\{\eta(X) < 0.5\}}\right) \cdot (2\eta(X) - 1)\right]\end{aligned}$$

Construction of the classifier

- Set $j^*(x) := \arg\min_{j \in \mathbb{Z}^d} \|x - x_j\|_\infty$
- Let $x_0 \in \mathbb{R}^d$ be the point to classify. Define

$$T_n = T_n(x_0) := \frac{1}{n} \sum_{i=n+1}^{2n} Z_{ij^*(x_0)} - \frac{1}{2n} \sum_{i=1}^{n} Z_{ij^*(x_0)}$$

- Classifier: $C_n(x_0) := \mathbb{1}_{\{T_n(x_0) \geq 0\}}$

Main theorem

Theorem 2

Assume $h \to 0$ *and* $n\alpha^2 h^{2d} \to \infty$ *as* $n \to \infty$.
Then, for every probability distribution P *on* $\mathbb{R}^d \times \{0,1\}$:

$$\mathcal{E}_P(C_n) \to 0 \quad (n \to \infty)$$

Outline of the proof

- Show:
  $\mathcal{E}_P(C_n) = \mathbb{E}\left[\left(\mathbb{P}\left[T_n(X) < 0|X\right] - \mathbb{1}_{\{\eta(X)<0.5\}}\right) \cdot (2\eta(X) - 1)\right] \to 0$

  $\Rightarrow$ *Idea:* Use dominated convergence

Outline of the proof: Step 1

- $\Omega_0 := \left\{ x_0 \in \mathbb{R}^d \,:\, \liminf_{h \longrightarrow 0} \frac{\mathbb{P}[\|X - x_0\|_\infty < h]}{h^d} > 0 \right\}$

Lemma 1

$$\mathbb{P}\left[X \in \Omega_0\right] = 1$$

Idea of proof.

- With the notation $\kappa(A) := \mathbb{P}\left[X \in A\right]$, we have:

$$\Omega_\kappa := \Omega_0 = \left\{ x_0 \in \mathbb{R}^d \,:\, \liminf_{h \longrightarrow 0} \frac{\kappa\left(B_{x_0}(h)\right)}{h^d} > 0 \right\}$$

Outline of the proof: Step 1

Idea of proof.

- Use Lebesgue's decomposition theorem (see Appendix Theorem 4) applied to κ and the Lebesgue measure λ
  $\Rightarrow$ Obtain measures μ,ν such that:
  - $\kappa = \mu + \nu$
  - $\mu \ll \lambda$ and $\nu \perp \lambda$
- Obtain the sets Ω_μ and Ω_ν and the decomposition $\Omega_\kappa = \Omega_\mu \cup \Omega_\nu$

  $\Rightarrow \mathbb{P}\left[X \notin \Omega_0\right] = \kappa\left(\Omega_\kappa^c\right) \leq \mu\left(\Omega_\mu^c\right) + \nu\left(\Omega_\nu^c\right)$
- Use the properties of μ and ν to show:
  - $\mu\left(\Omega_\mu^c\right) = 0$
  - $\nu\left(\Omega_\nu^c\right) = 0$

  $\Rightarrow \mathbb{P}\left[X \in \Omega_0\right] = 1$

$\square$

Outline of the proof: Step 2

Recall $\Omega_0 = \left\{ x_0 \in \mathbb{R}^d \; : \; \liminf_{h \to 0} \frac{\mathbb{P}[\|X - x_0\|_\infty < h]}{h^d} > 0 \right\}$.

- $\mathcal{X}^*_{x_0} := \left\{ x \in \mathbb{R}^d \; : \; \|x - x_{j^*(x_0)}\|_\infty < h \right\} = B_{x_{j^*(x_0)}}(h)$
- $x_0 \in \Omega_0 \; \Rightarrow \mathbb{P}\left[X \in \mathcal{X}^*_{x_0} \right] > 0$
- Hence, one can define:

$$\Omega_1 := \left\{ x_0 \in \Omega_0 \; : \; \limsup_{h \to 0} \left| \frac{\mathbb{E}\left[T_n\right]}{\mathbb{P}\left[X \in \mathcal{X}^*_{x_0}\right]} - \left(\eta(x_0) - \frac{1}{2} \right) \right| = 0 \right\}$$

- $\mathbb{E}[T_n] = \mathbb{E}\left[\mathbb{1}_{\{X \in \mathcal{X}^*_{x_0}\}} \cdot \left(\eta(X) - \frac{1}{2} \right) \right]$

By the Lebesgue differentiation theorem (see Appendix Theorem 5), one obtains $\mathbb{P}\left[X \in \Omega_1\right] = 1$.

Outline of the proof: Step 3

Let $x_0 \in \Omega_1 \subset \Omega_0$ such that $\eta(x_0) \neq \frac{1}{2}$. Define

$$\delta := \frac{1}{2} \cdot \left| \eta(x_0) - \frac{1}{2} \right| \cdot \mathbb{P}\left[X \in \mathcal{X}^*_{x_0}\right]$$

Lemma 2

1) $exp\left(-\frac{n\alpha^2\delta^2}{2^{2d+6}}\right) \to 0$

2) $\frac{\mathbb{E}[T_n(x_0)]}{\delta} \to 2 \cdot sign\left(\eta(x_0) - \frac{1}{2}\right)$

Outline of the proof: Step 3

Recall $\Omega_0 = \left\{ x_0 \in \mathbb{R}^d \ : \ \liminf_{h \longrightarrow 0} \frac{\mathbb{P}[\|X - x_0\|_\infty < h]}{h^d} > 0 \right\}$.

Proof.

1) We have:

$$\frac{n\alpha^2\delta^2}{2^{2d+6}} \sim n\alpha^2 \mathbb{P}\left[X \in \mathcal{X}^*_{x_0}\right]^2 \geq n\alpha^2 \left(\frac{h}{2}\right)^{2d} \left(\frac{\mathbb{P}\left[\|X - x_0\|_\infty < \frac{h}{2}\right]}{\left(\frac{h}{2}\right)^d}\right)^2 \to \infty$$

The final convergence can be shown by using $x_0 \in \Omega_0$ and the assumption $n\alpha^2 h^{2d} \to \infty$.

2) $$\frac{\mathbb{E}[T_n]}{\delta} = 2 \cdot \underbrace{\frac{\mathbb{E}[T_n]}{\mathbb{P}\left[X \in \mathcal{X}^*_{x_0}\right]}}_{\xrightarrow{x_0 \in \Omega_1} \eta(x_0) - \frac{1}{2}} \cdot \frac{1}{\left|\eta(x_0) - \frac{1}{2}\right|} = 2 \cdot sign\left(\eta(x_0) - \frac{1}{2}\right)$$

Let $x_0 \in \Omega_1$.

$$\left(\mathbb{P}\left[T_n(x_0) < 0\right] - \mathbb{1}_{\{\eta(x_0)<1/2\}}\right) \cdot (2\eta(x_0) - 1)$$

$$\leq 2 \cdot \underbrace{exp\left(-\frac{n\alpha^2\delta^2}{2^{2d+6}}\right)}_{\xrightarrow{Lemma\ 2.1)} 0} \cdot \mathbb{1}_{\left\{\eta(x_0)\neq\frac{1}{2}\right\}} + 2 \cdot \underbrace{\mathbb{1}_{\left\{\frac{\mathbb{E}[T_n(x_0)]}{\delta\, sign(2\eta(x_0)-1)} \leq 1;\ \eta(x_0)\neq\frac{1}{2}\right\}}}_{\xrightarrow{Lemma\ 2.2)} 0} + \underbrace{\mathbb{1}_{\left\{\frac{\mathbb{E}[T_n(x_0)]}{\delta} < 1;\ \eta(x_0)\neq\frac{1}{2}\right\}}}_{\xrightarrow{Lemma\ 2.2)} 0}$$

The same is also true almost surely when replacing x_0 by X since $\mathbb{P}\left[X \in \Omega_1\right] = 1$.
One can now finish the proof by applying dominated convergence theorem, so we have:

$$\mathcal{E}_P(C_n) = \mathbb{E}\left[\left(\mathbb{P}\left[T_n(X) < 0\right] - \mathbb{1}_{\{\eta(X)<1/2\}}\right) \cdot (2\eta(X) - 1)\right] \rightarrow 0.$$

$\square$

Table of Contents

1 Setting

2 Construction of the privacy mechanism

3 Excess risk of the classifier
- Formulation of the main theorem
- Outline of the proof

4 Minimax rate of convergence of the excess risk

5 Comparison to non-privatized setup

6 Numerical experiments

Minimax rate

- Let $\alpha > 0$ be the privacy parameter and $n \in \mathbb{N}$ the sample size.
- We want to examine the minimax excess risk over all privacy mechanisms $Q \in \mathcal{Q}_\alpha$ and all classifier depending only on the privatized data, i.e.

$$\mathcal{R}_{n,\alpha}(\mathcal{P}) := \inf_{\substack{Q \in \mathcal{Q}_\alpha \\ C_n}} \sup_{P \in \mathcal{P}} \mathcal{E}_P(C_n)$$

where $\mathcal{P}$ is a class of distributions.

The class of distributions

Let $\beta \in (0, 1]$, $\gamma \in [0, \infty)$ and $L, C_0, r_0, c_0, \mu \in (0, \infty)$ and define $\theta := (\beta, L, \gamma, C_0, r_0, c_0, \mu)$.

We consider the class $\mathcal{P}(\theta)$ of distributions which satisfy the following three conditions:

- (β, L)-Hölder smoothness condition
- (γ, C_0)-margin condition
- (c_0, r_0, μ)-strong density assumption

The class of distributions

Definition 2

Let $\beta \in (0,1]$ and $L > 0$.
We call a distribution P (β, L)-Hölder if $\eta(x) = \mathbb{P}[Y = 1 | X = x]$ is $\beta-$Hölder with constant L, i.e.

$$\left|\eta(x) - \eta(x')\right| \leq L \cdot \left|x - x'\right|^{\beta} \qquad \forall x, x' \in [0,1]^d$$

Definition 3

Let $\gamma \geq 0$ and $C_0 > 0$.
A distribution P satisfies the (γ, C_0)-margin condition if:

$$\mathbb{P}\left(0 < \left|\eta(X) - \frac{1}{2}\right| \leq t\right) \leq C_0 \cdot t^{\gamma} \qquad \forall t > 0$$

The class of distributions

Definition 4

Let $c_0, r_0 > 0$ and λ the Lebesgue measure.
A Lebesgue-measurable set $A \subset [0,1]^d$ is called (c_0, r_0)-regular if:

$$\lambda\left(A \cap B_x(r)\right) \geq c_0 \cdot \lambda\left(B_x(r)\right) \qquad \forall r \in (0, r_0],\ x \in A$$

Definition 5

Let $c_0, r_0, \mu > 0$.
A distribution P satisfies the (c_0, r_0, μ)-strong density assumption if X has a density f such that:

- supp(f) is (c_0, r_0)-regular
- $f(x) \geq \mu$ for all $x \in$ supp(f)

Minimax rate

Theorem 3

Let $\theta = (\beta, L, \gamma, C_0, r_0, c_0, \mu)$ *as before such that* $\beta\gamma \leq d$.
Then there exist constants c *and* C *such that:*

$$c \cdot \left(n\alpha^2\right)^{-\frac{\beta(1+\gamma)}{2\beta+2d}} \leq \mathcal{R}_{n,\alpha}\left(\mathcal{P}(\theta)\right) \leq C \cdot \left(n\alpha^2\right)^{-\frac{\beta(1+\gamma)}{2\beta+2d}} \qquad \forall n \in \mathbb{N}, \alpha \in (0,1]$$

Table of Contents

1. Setting
2. Construction of the privacy mechanism
3. Excess risk of the classifier
   - Formulation of the main theorem
   - Outline of the proof
4. Minimax rate of convergence of the excess risk
5. **Comparison to non-privatized setup**
6. Numerical experiments

Comparison to non-privatized setup

	Privatized	**Non-privatized**
Convergence assumption for excess risk	$n\alpha^2 h^{2d}$	nh^d
Minimax rate	$\left(n\alpha^2\right)^{-\frac{\beta(1+\gamma)}{2\beta+2d}}$	$n^{-\frac{\beta(1+\gamma)}{2\beta+d}}$

Table: Comparison of the privatized and non-privatized setting

Table of Contents

1 Setting

2 Construction of the privacy mechanism

3 Excess risk of the classifier
- Formulation of the main theorem
- Outline of the proof

4 Minimax rate of convergence of the excess risk

5 Comparison to non-privatized setup

6 Numerical experiments

Data

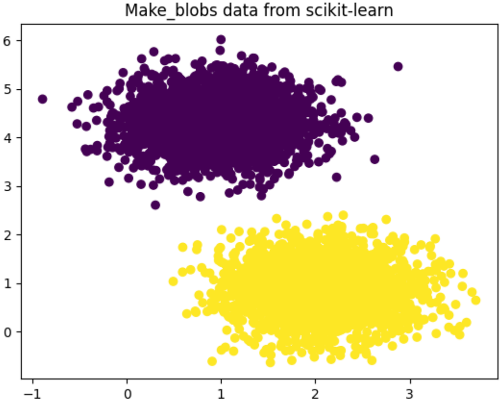

Figure: Artificial data from sklearn to classify

Numerical results

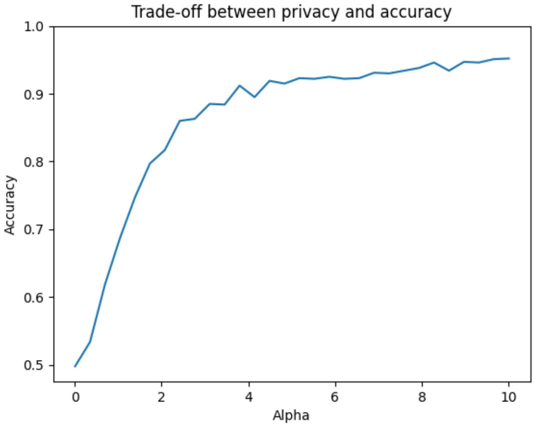

Figure: Accuracy with respect to the parameter α

References

[1] Thomas Berrett and Cristina Butucea. "Classification under local differential privacy". In: *arXiv preprint arXiv:1912.04629* (2019).

[2] Achim Klenke. *Probability theory: a comprehensive course*. Springer Science & Business Media, 2013.

[3] Elias M Stein and Rami Shakarchi. *Real Analysis, Princeton Lectures in Analysis III*. 2005.

Appendix

Appendix

Definition 6 (Markov Kernel, Definition 8.25 in [2])

Let $(\Omega_1, \mathcal{A}_1)$ and $(\Omega_2, \mathcal{A}_2)$ be two measureable spaces.
A map $\kappa : \mathcal{A}_2 \times \Omega_1 \to [0, 1]$ is called a Markov Kernel if:

- For any $A_2 \in \mathcal{A}_2$ the map $\omega_1 \mapsto \kappa(A_2, \omega_1)$ is $\mathcal{A}_1$-measureable.
- For any $\omega_1 \in \Omega_1$ the map $A_2 \mapsto \kappa(A_2, \omega_1)$ is a probability measure $(\Omega_2, \mathcal{A}_2)$.

Appendix

Theorem 4 (Lebesgue's decomposition theorem, Theorem 7.33 in [2])

Let μ, ν be σ-finite measures on a measureable space $(\Omega, \mathcal{A})$.
Then ν can be uniquely decomposed into measures ν_1 and ν_2 such that:

- $\nu = \nu_1 + \nu_2$
- $\nu \ll \mu$
- $\nu \perp \mu$

Appendix

Theorem 5 (Lebesgue differentiation theorem, Theorem 1.3 in [3])

Let λ be the Lebesgue measure.
If a function f on $\mathbb{R}^d$ is integrable, then we have:

$$\lim_{\substack{x \in B \\ \lambda(B) \to 0}} \frac{1}{\lambda(B)} \int_B f(y)\, dy = f(x) \qquad \text{for a.e. } x$$

In particular, the set

$$E := \left\{ x \in \mathbb{R}^d : \limsup_{\substack{x \in B \\ \lambda(B) \to 0}} \left| \frac{1}{\lambda(B)} \int_B f(y)\, dy - f(x) \right| > 0 \right\}$$

has measure zero.