A method to translate privacy requirements into a configuration for privacy-preserving machine learning applied to multi-centric studies

Moitree Basu[1,2], Jan Ramon[1]

[1] MAGNET, Inria Lille, [2] University of Lille, France

Abstract

Recently, due to the potential of machine learning approaches and the increased awareness of privacy risks, there is an increased interest in **privacy-preserving machine learning**. However, real-world medical applications are often complex. Our research, therefore, focuses on two objectives: we want to develop algorithms that make privacy-preserving machine learning more **interpretable** for non-experts and which automatically optimize the parameters and strategy of a machine learning solution to **respect privacy requirements** while **optimizing utility**, i.e., **maximizing precision** and **minimizing cost**.

At its core, our methodology is based on a **constraint programming** approach. It is known that non-experts can relatively easily learn to express requirements in the form of **constraints**. At the same time, this approach allows us to use a wide range of publicly available constraint program solvers.

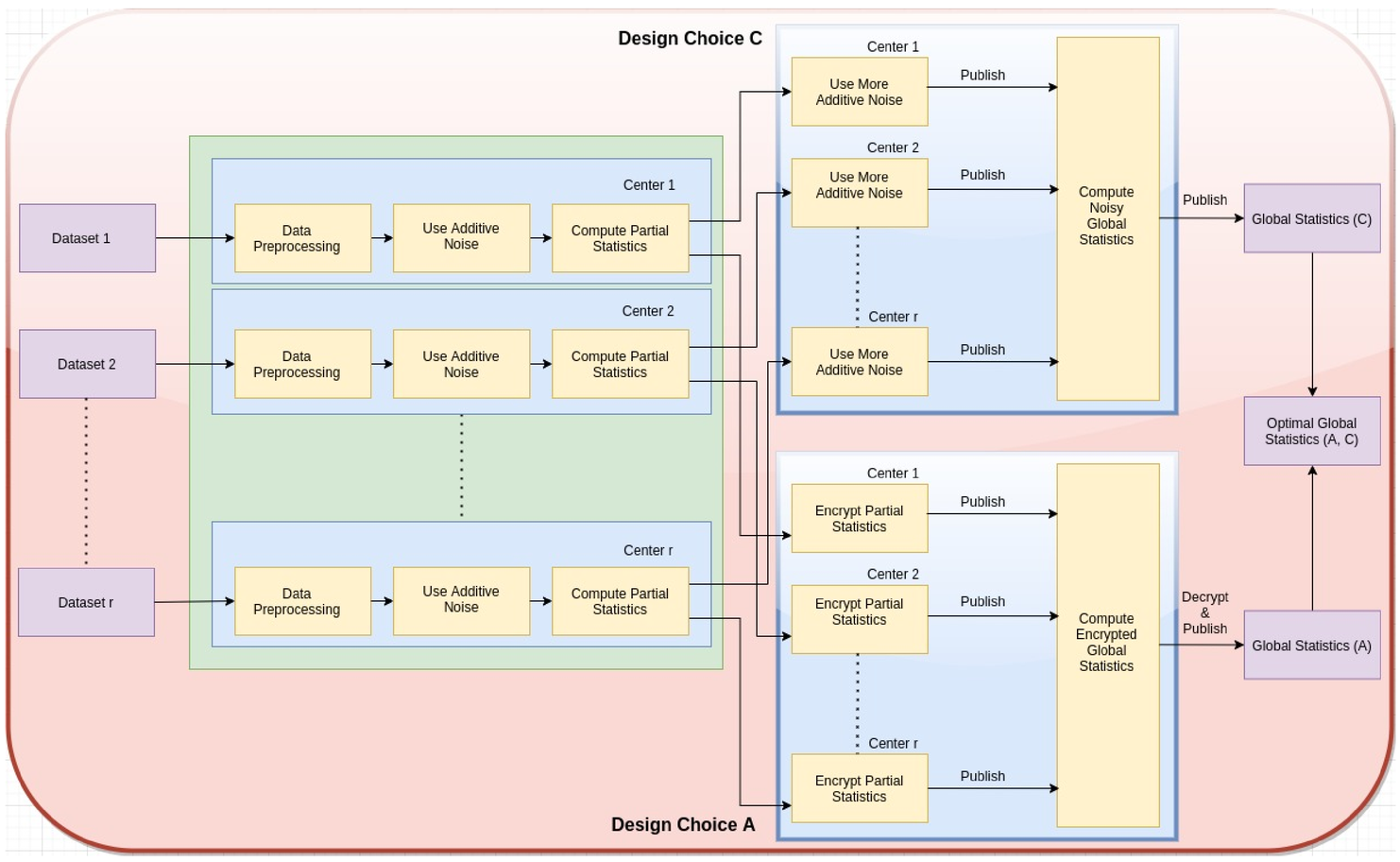
Background

Differential Privacy:

Let, $\epsilon > 0, \delta \geq 0$. A (randomized) protocol $\mathcal{A}$ is (ϵ, δ)-differentially private if for all neighboring datasets X, X', i.e., datasets differing only in a single data point, and for all sets of possible outputs $\mathcal{O}$, we have:

$$Pr(\mathcal{A}(X) \in \mathcal{O}) \leq e^{\varepsilon} Pr(\mathcal{A}(X') \in \mathcal{O}) + \delta.$$

Example Problem Scenario



Specification:

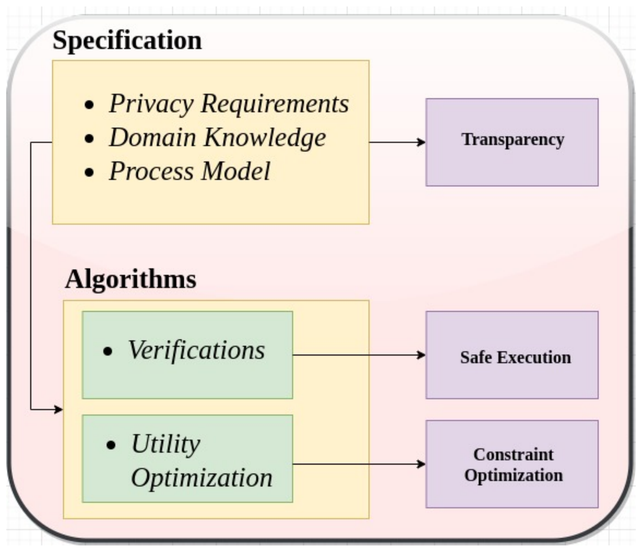
- **Problem Statement**
  - Problem setting: Multi-centric study
  - Goal: Compute linear regression, achieve privacy
  - N patients, r centers, U_i set of patients in center i
  - Every patient j has q features, $x_{j,k}$ with $k = 1...q$
    - $x_{j,0} = 1$ for bias
    - $x_{j,1} =$ some specific health parameter (F) of interest
  - Output: Compute the queries $Q_{k,l} = \sum_{j=1}^{N} x_{j,k} x_{j,l}$ for all $k, l \in 0...q$
    - compute the counts, sums, sum-of-squares, co-variances
- **Design Choices**
  - "A" uses costly encryption for fully private, accurate computation of global sum $Q_{k,l}$
  - "C" adds more noise to these partial sums and publishes them for public computation of global sum $Q_{k,l}$

Objective Function Derived for the Example Scenario:

- $\min \lambda Obj_{\mathrm{A}} + (1-\lambda) Obj_{\mathrm{C}}$, where $0 \leq \lambda \leq 1$
- $Obj_{\mathrm{A}} = C_{encrypt} + \sum_{i=1}^{r} |U_i| \sigma^2_{\mathrm{A},i}$
- $Obj_{\mathrm{C}} = \sum_{i=1}^{r} |U_i| \sigma^2_{\mathrm{C},i}$

Proposed Method: Problem Formulation

Block Diagram:



Privacy Requirements:

- Individual patients need privacy
  - add noise $\eta_{v,j,k} \sim \mathcal{N}\left(0, \sigma^2_{v,i}\right)$ to $x_{j,k}$ with $j \in U_i$ where $v \in \{\mathrm{A}, \mathrm{C}\}$
  - all $x_{j,:}$ with $k = 1 \ldots q$ should be $(\epsilon^{(pp)}, \delta^{(pp)})$-DP in any published model
  - all $x_{j,:}$ with $k = 1 \ldots q$ should be $(\epsilon^{(pl)}, \delta^{(pl)})$-DP in partial statistics seen by researchers locally in the hospital
  - $(\epsilon^{(pp)}, \delta^{(pp)}) \ll (\epsilon^{(pl)}, \delta^{(pl)})$ as every center trusts its own researchers better
- Center-wise partial average needs privacy
  - avoid bad reputation: disclosed aberrant statistics of the patient population can be wrongly interpreted as incorrect care
  - compute local statistics: $Q_{v,i,k,l} = \sum_{j \in U_i} (x_{j,k} + \eta_{v,j,k})(x_{j,l} + \eta_{v,j,l})$
  - for all i and k, $\sum_{j \in U_i} x_{j,1} x_{j,k}$ should be $(\epsilon^{(u)}, \delta^{(u)})$-DP (for every center, for the specific medical parameter F)
- Aggregation
  - the centers make design choice from $v \in \{\mathrm{A}, \mathrm{C}\}$
  - either publish the $Q_{\mathrm{C},i,k,l}$ and aggregate publicly, or use a seure private aggregation for $Q_{\mathrm{A},i,k,l}$
- We can easily adapt to other requirements

Constraint Optimization

- Patient privacy:
  - $b^\top([D^{(1)} \ldots D^{(r)}\ I]\mathrm{diag}(\sigma_{\mathrm{A}})^2[D^{(1)} \ldots D^{(r)}\ I]^\top)^{-1} b \leq t_{GM}(\epsilon^{(pp)}, \delta^{(pp)})$: In design A, only multi-centric aggregate statistics are published.
  - For $i : 1 \ldots r$: $b^\top([D^{(i)}\ I]\sigma^2_{\mathrm{C},i}[D^{(i)}\ I]^\top)^{-1} b \leq t_{GM}(\epsilon^{(pp)}, \delta^{(pp)})$: In design C statistics of every unit are published.
  - For $i : 1 \ldots r, v \in \{\mathrm{A}, \mathrm{C}\}$: $b^\top([D^{(i)}\ I]\sigma^2_{v,i}[D^{(i)}\ I]^\top)^{-1} b \leq t_{GM}(\epsilon^{(pl)}, \delta^{(pl)})$: Statistics of every unit are available to the local researchers.
- Intra-center confidentiality of computed statistic on specific feature F:
  - For $i : 1 \ldots r$: $(b')^\top ([D^{(i)\prime}\ I]\sigma^2_{\mathrm{C},i}[D^{(i)\prime}\ I]^\top)^{-1} b' \leq t_{GM}(\epsilon^{(u)}, \delta^{(u)})$: In scenario C, the local statistics on specific feature F are published.
  - $(b')^\top ([D^{(1)} \ldots D^{(r)}\ I]\mathrm{diag}(\sigma_{\mathrm{A}})^2[D^{(1)} \ldots D^{(r)}\ I]^\top)^{-1} b' \leq t_{GM}(\epsilon^{(u)}, \delta^{(u)})$: In scenario A, the local statistics on specific feature F are part of the eventually published overall aggregates.

**pp: inter-center global publish, pl: intra-center local publish, u: F feature specific calculation, b, b': $f(\epsilon, \delta, D)$, D: feature weight matrix,
$t_{GM}(\epsilon, \delta)$: minimal variance of additive Gaussian noise needed to achieve (ϵ, δ)-differential privacy, σ^2_A, σ^2_C: variances of noise terms in design choice A & C respectively

Discussion

Summary: We hope this idea can facilitate the collaboration between medical and machine learning experts, and focus on the problem at hand rather than on technical issues. We apply our techniques to the problem of organizing **multi-centric studies** in a **secure and privacy-preserving** way. The goal is to develop an algorithm allowing the user to specify a wide range of constraints, e.g., requirements on the **privacy of patients**, requirements on the **confidentiality of information on participating centers**, the **trust centers have in different involved parties**, the **statistics** which need to be computed, the **procedures** which must be followed, the required **precision** etc. These specifications are then used as **constraints** while **optimizing the utility**, which in turn **preserves privacy** and helps us to **automate** the process to a great extent. The result of optimizing such a constraint program (if feasible) is then a **federated, secure, privacy-preserving machine learning strategy** satisfying all requirements.