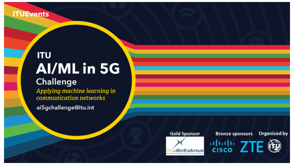
Imperial College London



Privacy-Preserving AI/ML in 5G Networks for Healthcare Applications (ITU-ML5G-PS-022)

Dopamine: Differentially Private Secure Federated Learning on Medical Data

Team: I****L diagnostics**

Mohammad Malekzadeh, Burak Hasircioglu, Nitish Mital, Kunal Katarya, Mehmet Emre Ozfatura

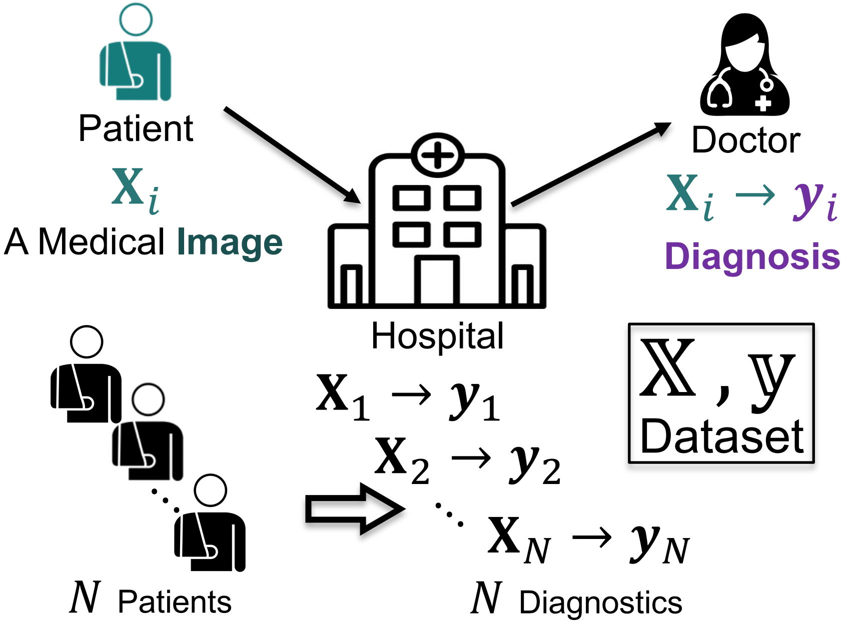
Supervisor: Prof. Deniz Gündüz

Information Processing and Communications Lab (IPC-Lab)

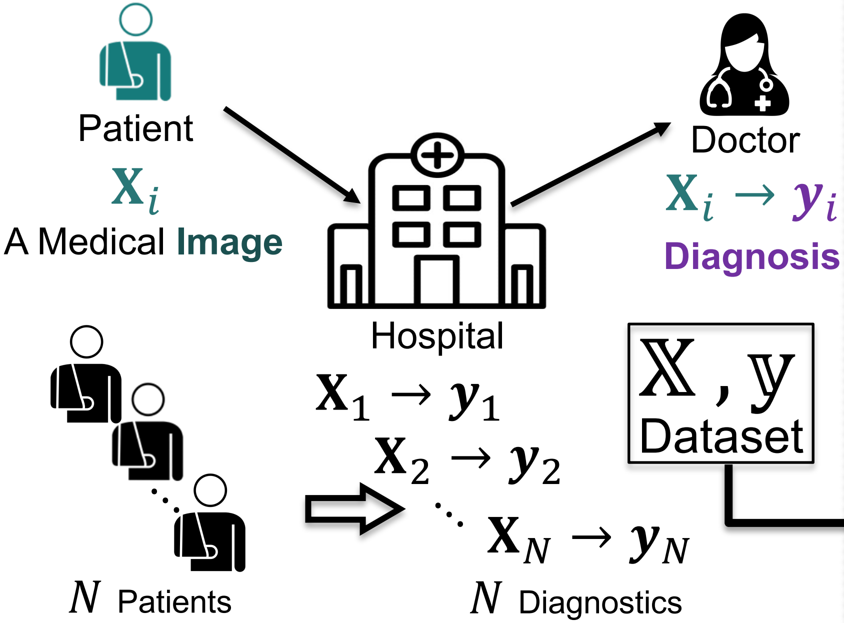
Department of Electrical and Electronic Engineering

Imperial College London

Problem Setting

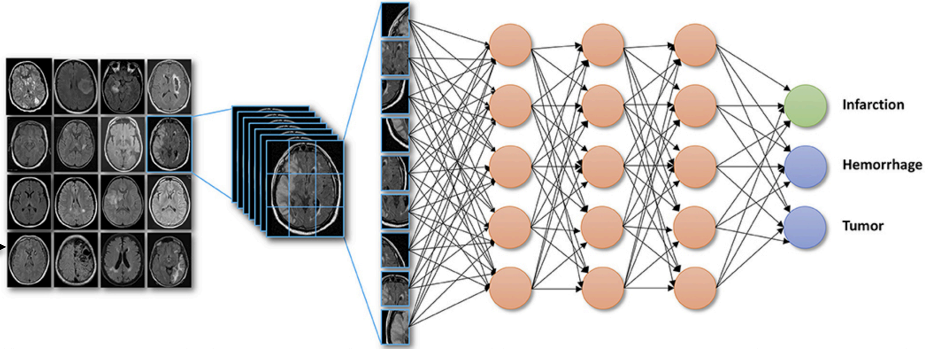


Motivation



Taking Advantage of Medical Data

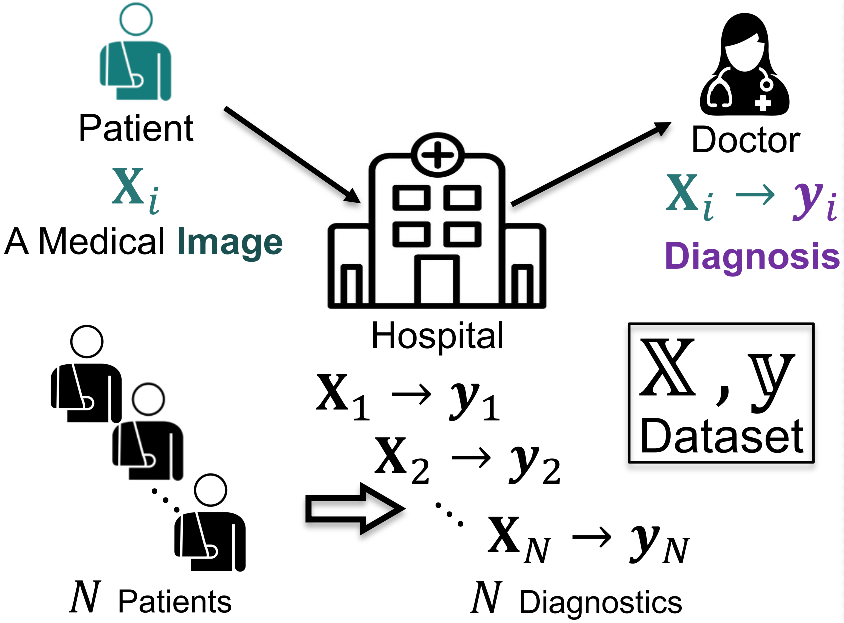
Deep Neural Networks



[1]

[1] Zhu, Guangming, et al. "Applications of deep learning to neuro-imaging techniques." *Frontiers in Neurology* 10 (2019): 869.

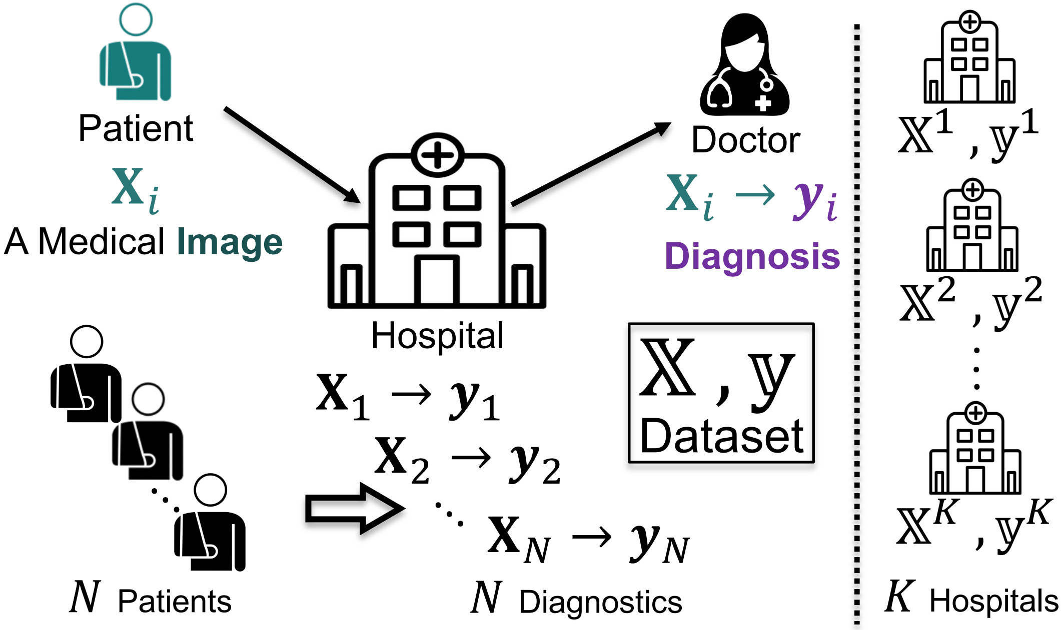
Motivation



Pervasive **Connectivity** enables **Automated Diagnosis**.

- Coverage
  - More Patients
  - Rural Area & Developing Countries
- Efficiency:
  - Faster
  - Cheaper
- Lower Burden on Healthcare System
  - Decision for further examination?
  - Giving Short-Term Advice

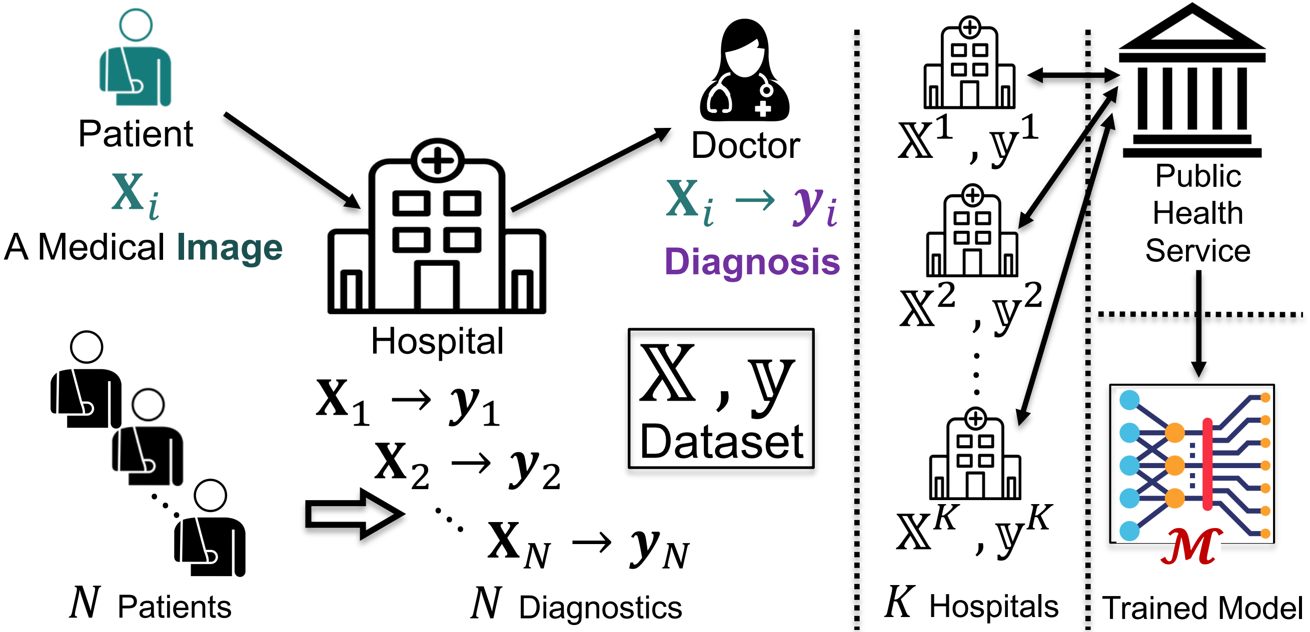
Challenge



Medical Dataset are Distributed and Kept Private.

Patients' **Privacy** is as important as Patient's **Health**

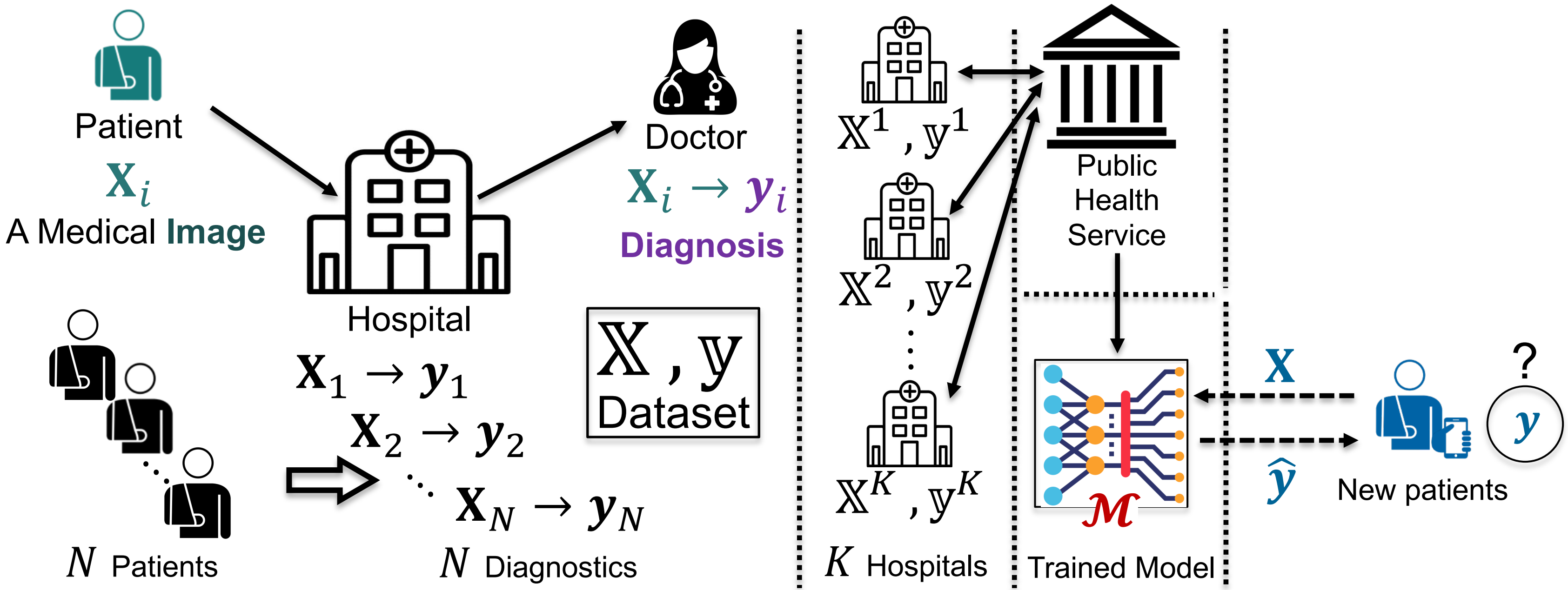
Solution



To **Train a DNN** on Distributed Datasets using **Federated Learning**[2]

[2] Truex, Stacey, et al. "A hybrid approach to privacy-preserving federated learning." Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security. 2019.

Solution(cont.)

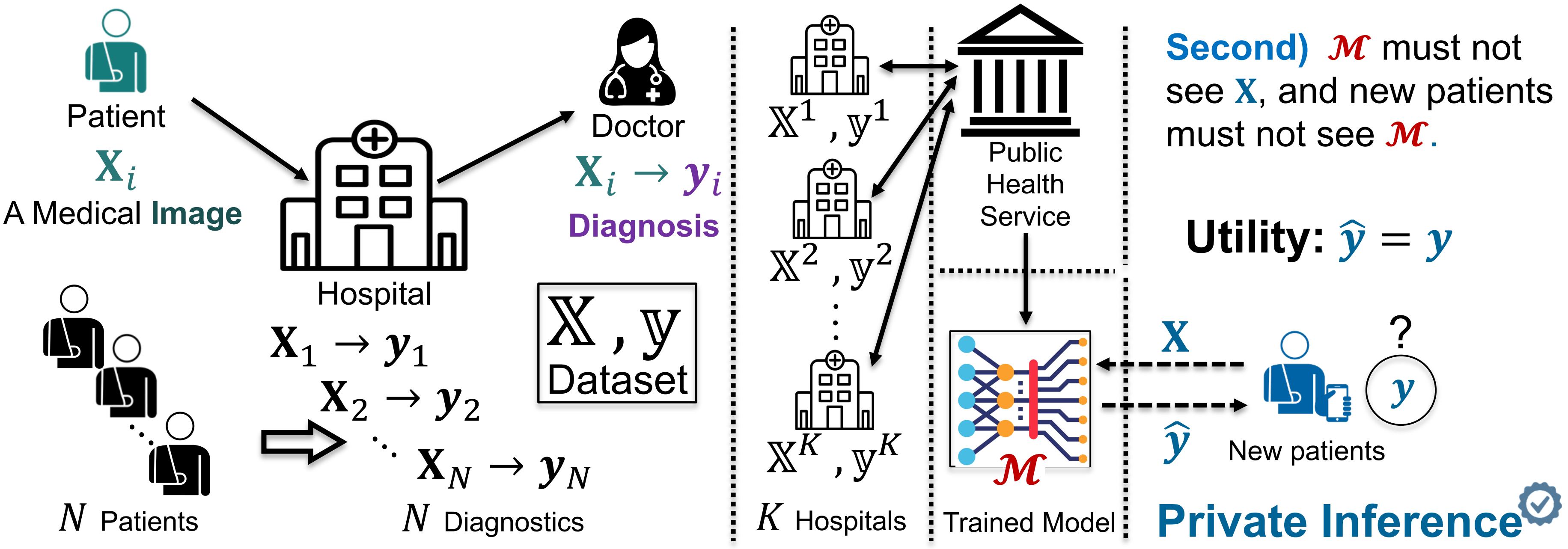


Requirements

Privacy?

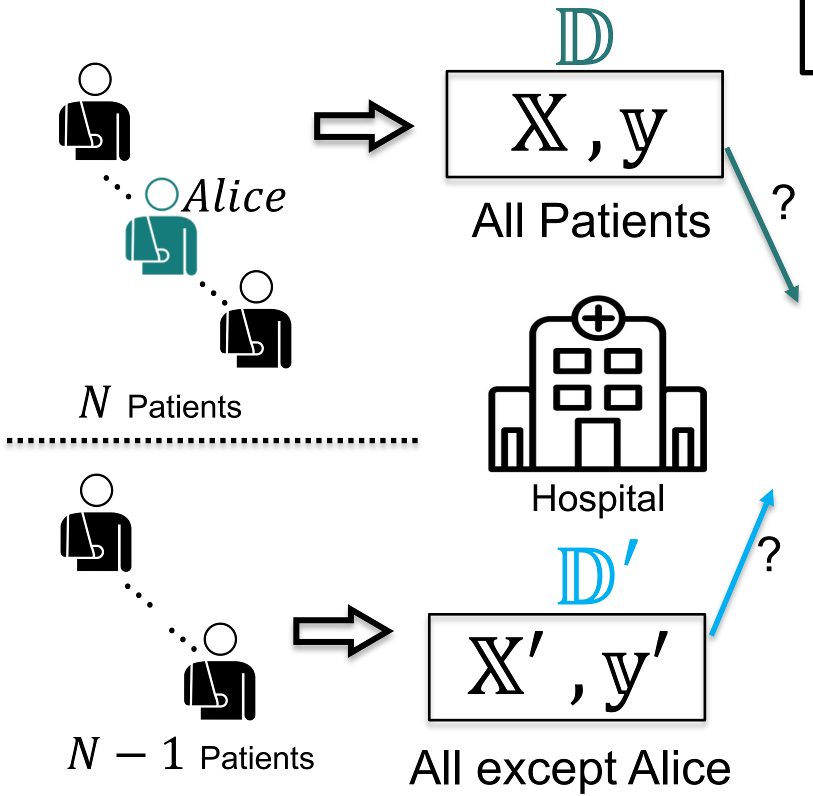
Differential Privacy

First) $\mathcal{M}$ must not reveal the presence (or absence) of any patient i in the training set.



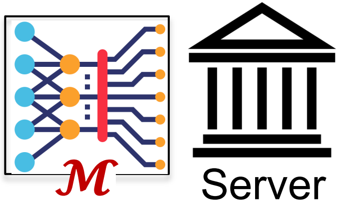
Differential Privacy

$$e^{-\epsilon} \leq \frac{\Pr(\mathcal{M} \textit{ is trained on } \mathbb{D})}{\Pr(\mathcal{M} \textit{ is trained on } \mathbb{D}')} \leq e^{\epsilon}$$



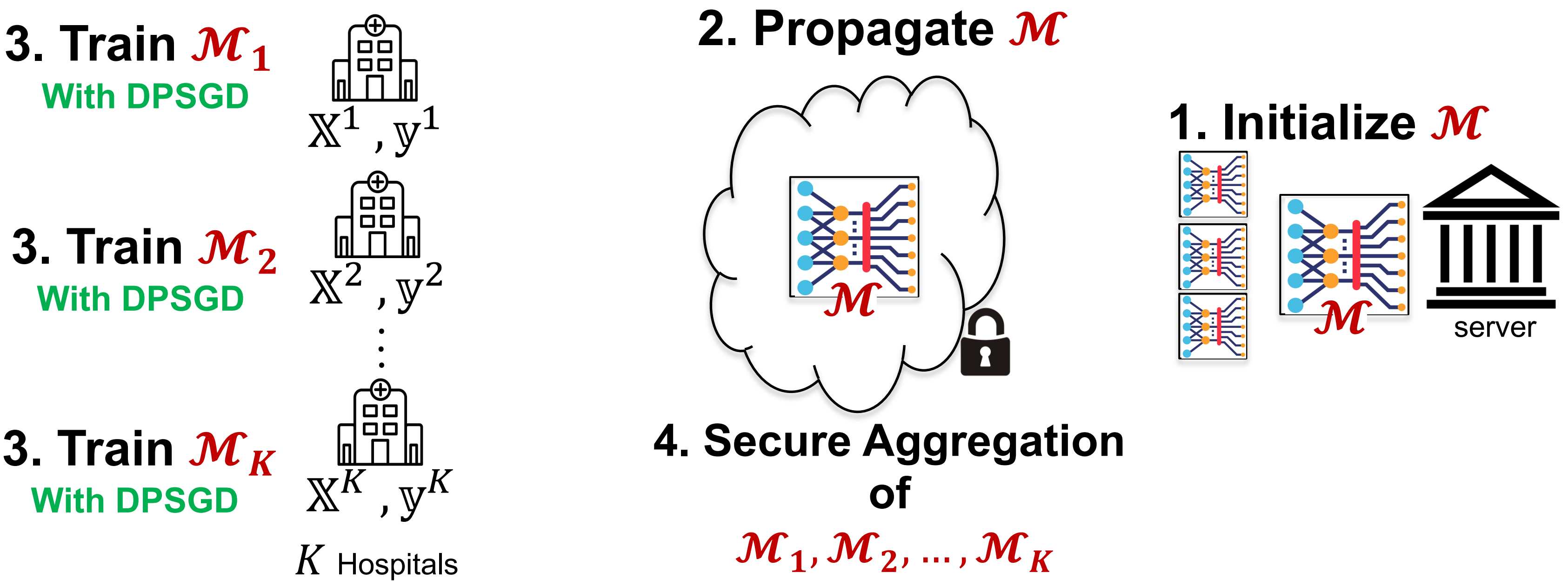
"Deep learning with differential privacy"[3]

Known as DPSGD

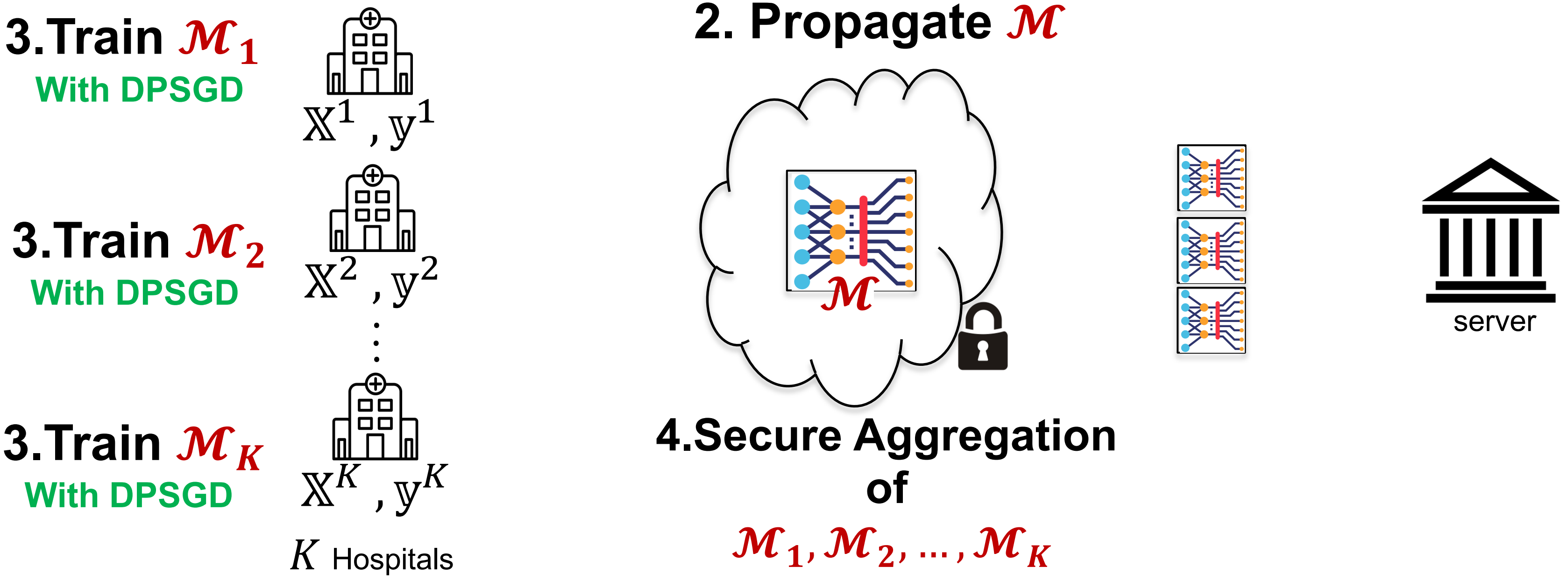


[3] Abadi, Martin, et al. "Deep learning with differential privacy." Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. 2016.

Implementation



Implementation(cont.)



Why Secure Aggregation?

Better Accuracy while keeping the same DP Privacy Guarantee.

Adding Gaussian Noise for DP

$$\mathcal{F}(x)$$

$$\mathcal{M}(x) = \mathcal{F}(x) + \mathcal{N}\left(\mu = 0, \sigma^2 = \frac{2\ln(1.25/\delta)(\Delta_2\mathcal{F})^2}{\epsilon^2 \boldsymbol{K}}\right)$$

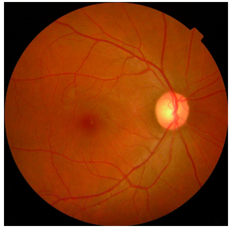
Dopamine's Training Algorithm

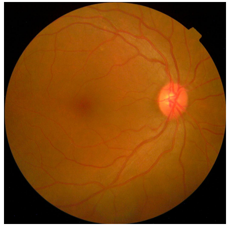
Algorithm 1 *Dopamine*'s Training

Input: K: number of hospitals, $\mathbb{D}$: distributed dataset, $\mathbf{w}$: model's trainable parameters, $\mathcal{L}(\cdot,\cdot)$: loss function, q: sampling probability, σ: noise scale, C: gradient norm bound, η: learning rate, β: momentum, T: number of rounds, (ϵ, δ): bounds on the patient-level differential privacy loss.
Output: $\mathbf{w}$: final parameters.
$\mathbf{w}_G^0$ = random initialization.
$\hat{\epsilon} = 0$
for $t : 1, \ldots, T$ **do**
  for $k : 1, \ldots, K$ **do**
    $\mathbf{w}_k^t = \mathbf{w}_G^{t-1}$
    $\mathbb{D}_k^t = Sampling(\mathbb{D}_k)$ // by uniformly sampling each item in $\mathbb{D}_k$ independently with probability q.
    for $\mathbf{x}_i \in \mathbb{D}_k^t$ **do**
      $\mathbf{g}^t(\mathbf{x}_i) = \nabla_{\mathbf{w}} \mathcal{L}(\mathbf{w}_k^t, \mathbf{x}_i)$
      $\bar{\mathbf{g}}^t(\mathbf{x}_i) = \mathbf{g}^t(\mathbf{x}_i) / \max\left(1, \frac{||\mathbf{g}^t(\mathbf{x}_i)||_2}{C}\right)$
    end for
    $\tilde{\mathbf{g}}_k^t = \frac{1}{|\mathbb{D}_k^t|}\left(\sum_i \bar{\mathbf{g}}^t(\mathbf{x}_i) + \mathcal{N}(0, \frac{\sigma^2 \cdot C^2 \cdot \mathbf{I}}{K})\right)$
    $\hat{\mathbf{g}}_k^t = \tilde{\mathbf{g}}_k^t + \beta \hat{\mathbf{g}}_k^{t-1}$ $//\hat{\mathbf{g}}_k^0 = 0$
    $\mathbf{w}_k^t = \mathbf{w}_k^t - \eta \hat{\mathbf{g}}_k^t$
  end for
  $\hat{\epsilon} = CalculatePrivacyLoss(\delta, q, \sigma, t)$ // by Moments Accountant (Abadi et al. 2016)
  if $\hat{\epsilon} > \epsilon$ **then**
    return $\mathbf{w}_G^{t-1}$
  end if
  $\mathbf{w}_G^t = \frac{1}{K}\left(SecureAggregation(\sum_k \mathbf{w}_k^t)\right)$
end for

Evaluation

- **Dataset:**
  - **Diabetic Retinopathy**[4]
  - **Five** Classes: *normal*, *mild*, *moderate*, *severe*, and *proliferative*.
  - **3662** images: 2931 for training, 731 for testing.
  - Dimensions: **224×224**





- **Deep Neural Network:**
  - **SqueezeNet**[5]
  - 50x **fewer** parameters than the famous AlexNet.
  - Yet, achieves the **same** level of AlexNet's accuracy on ImageNet.

- **Simulation:**
  - **10** hospitals and **1** server.
  - data distributed **i.i.d** and **equal**.

[4] Choi, J. Y.; et. al. . 2017. Multi-categorical deep learning neural network to classify retinal images: A pilot study employing small database. PLOS ONE12: 1–16.

[5] Iandola, Forrest N., et al. "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size." *arXiv preprint arXiv:1602.07360* (2016).

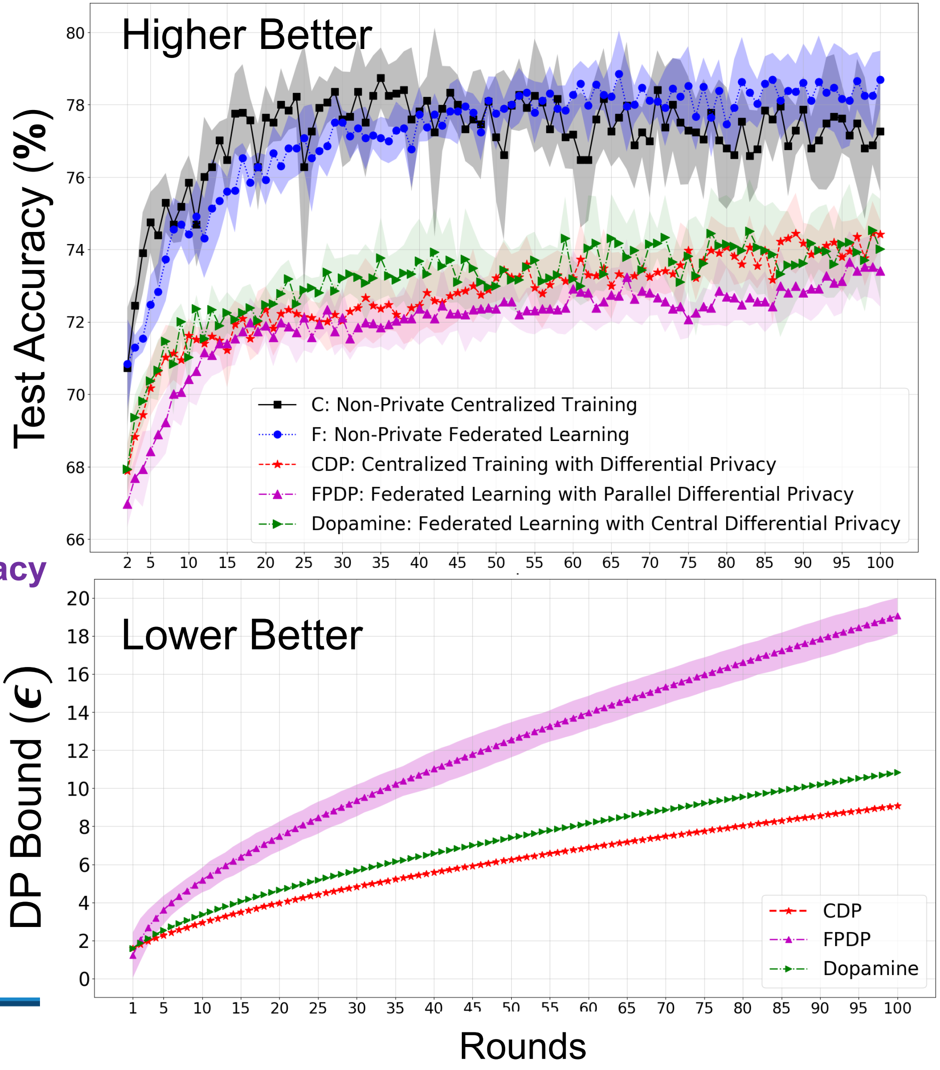
Experimental Results

Baselines:

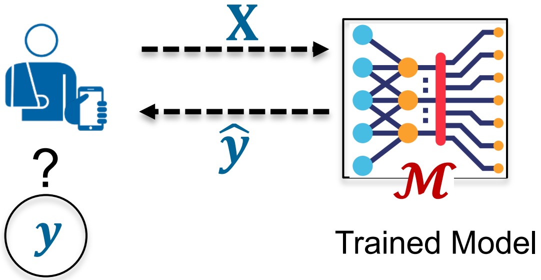
1) **Centralized Learning without Privacy**
2) **Federated Learning without Privacy**
3) **Centralized Learning with Differential Privacy**
4) **Federated Learning with Parallel Differential Privacy**
5) **Our Solution**

(1) & (3)
are not achievable in practice!

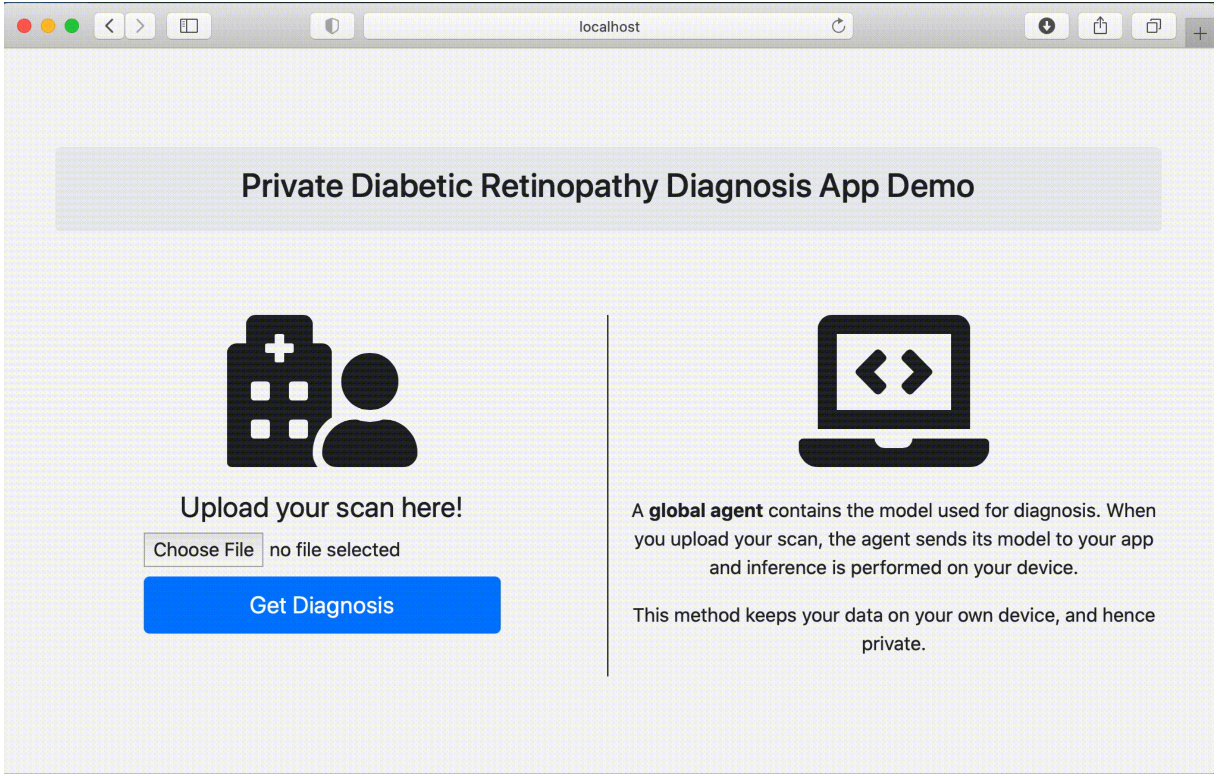
(2)
is not an acceptable alternative!



Private Inference



Patients don't share their $\mathbf{X}$,
but
$\mathcal{M}$ is sent to the **patients'** devices.



https://imperial-diagnostics.herokuapp.com

Contributions

https://github.com/ipc-lab/private-ml-for-health

1. First to implement **Federated Learning** on **DNNs** with **Patient-Level DP** on a **Medical Dataset**
2. First to use **Momentums** in **Federated DP-SGD** achieving **Better Accuracy & Stable Training**

In Progress

1. End-to-end Secure Aggregation Using Homomorphic encryption
2. Further Evaluation: Other datasets --- Other DNNs.
3. Keeping the trained Model Private at the Server's Side.

Q/A

Open Questions

1. More accurate and efficient FL algorithms with DP.
2. When patients could have more than one sample data.