Privacy Preserving Neural Network Classification

Beyza Bozdemir[1], Gamze Tillem[2], Melek Önen[1], Orhan Ermis[1]
[1]EURECOM, France
[2]Cyber Security Group, Delft University of Technology, The Netherlands

Privacy goals in PAPAYA

PAPAYA Objectives

Privacy by Design

- **PP analytics**: **Neural Networks (NN),** clustering, statistics

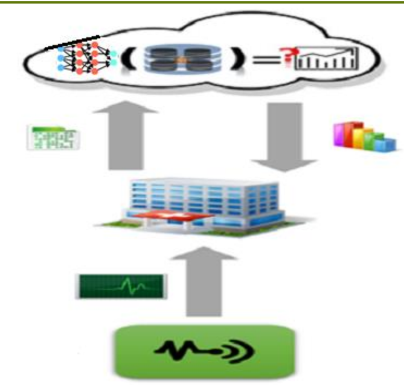
Different Settings

- One data owner vs. multiple data owners
- One querier vs. multiple queriers

Use Cases

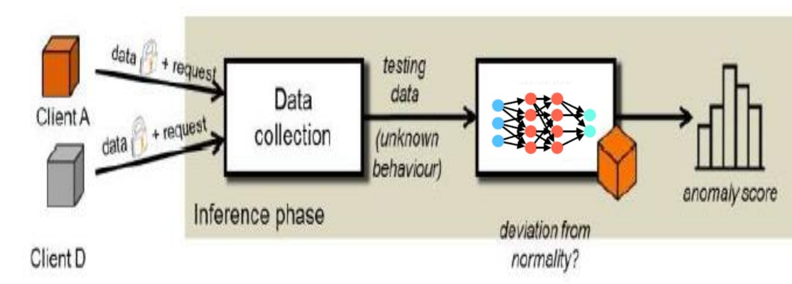
healthcare

Arrhythmia Detection



mobile and phone usage

Threat Detection



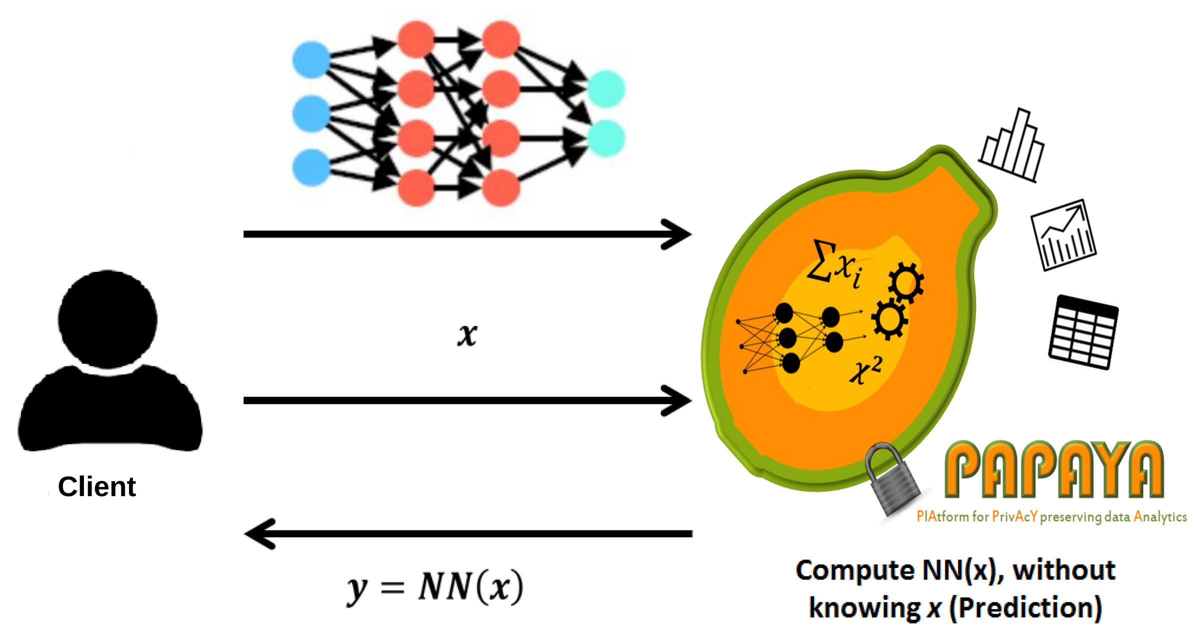
Analytics example – Neural Network classification

NN Layers and Operations

- **Input Layer**
- **Hidden Layer**

 - Convolutional layer (matrix multiplications)
 - Activation layer (sigmoid, tanh, etc.)
 - Pooling layer
 - Fully connected layer

- **Output Layer** (softmax, etc.)



Privacy by Design Challenges

- **Privacy vs. efficiency**

 Deep NN ⇒ Significant overhead with cryptographic tools

- **Privacy vs. accuracy**

 Complex operations (sigmoid, softmax, etc.) ⇒ Not suitable to crypto tools

- **Real Numbers**

Privacy Preserving Neural Network Classification – Existing solutions

with Homomorphic Encryption [1]

- Non-interactive
- Only linear operations (eg. AF is approximated to x^2)
- Expensive in computation cost
- No communication cost

with Secure Two-party Computation [2]

- Interactive – Client is involved
- Linear operations and comparisons
- Efficient in computation cost
- Expensive in communication cost

Privacy Preserving Neural Network Classification: A Hybrid Solution

Hybrid Solution

Features

- Paillier for linear operations

 ⇒ Optimized computational overhead

 ⇒ Less computation time compared to [1]

- Paillier for x^2

 ⇒ New interactive protocol to compute x^2

- 2PC for comparison only (ReLU case)

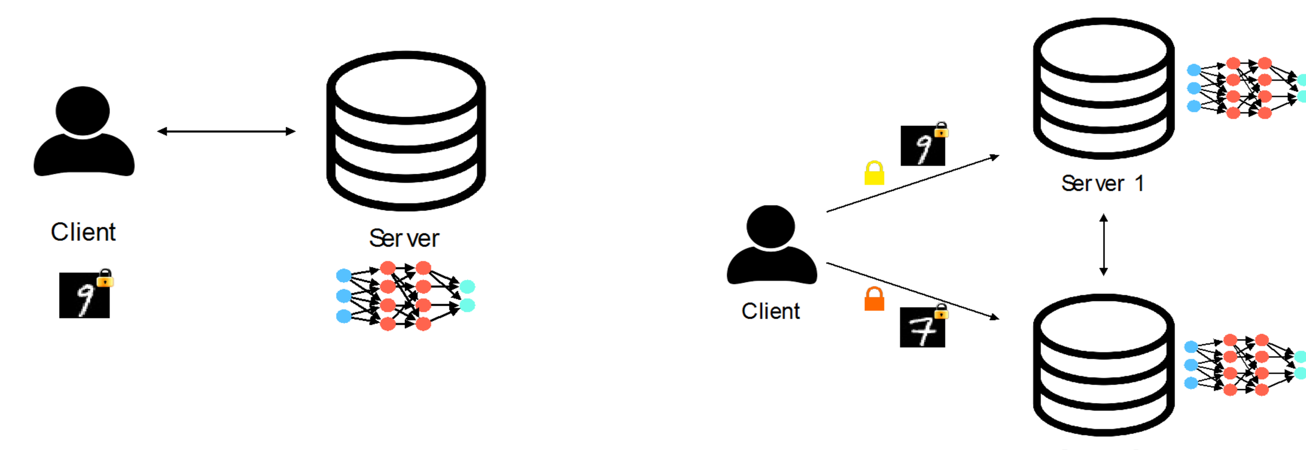
 ⇒ Optimized communication overhead

 ⇒ Less bandwidth usage compared to [2]

- Similar level of accuracy as in [1, 2]

Input
Convolution layer
Activation layer (x^2)
Pooling layer (avg)
Convolution layer
Pooling layer (avg)
Fully Connected layer
Activation layer (x^2)
Fully Connected layer
Output

Flexible solution: 2 settings



Results

- Computation cost 30-fold better than [1]
- Communication cost 27-fold better than [2]

Technique	Computation Cost (s)	Communication Cost (MB)
HE [1]	297	372.2
2PC [2]	1.2	47.6
Hybrid Solution	10	1.73

References:
[1] Gilad-Bachrach, Ran, et al. "Cryptonets: Applying neural networks to encrypted data with high throughput and accuracy." *International Conference on Machine Learning*. 2016.
[2] Liu, Jian, *et al. "Oblivious neural network predictions via minionn transformations." Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. ACM, 2017.*

EURECOM
www.eurecom.fr
www.papaya-project.eu
@ProjectPapaya
PAPAYA Project
papaya@eurecom.fr

The research leading to these results has received funding from the European Union's Horizon 2020 Research and Innovation Programme, through the PAPAYA project, under Grant Agreement No. 786767.
This poster reflects the view of the consortium only. The Research Executive Agency is not responsible for any use that may be made of the information it contains.