

# HETAL: Efficient Privacy-preserving Transfer Learning with Homomorphic Encryption

---

Seewoo Lee, Garam Lee, Jung Woo Kim, Junbum Shin, Mun-Kyu Lee  
40th International Conference on Machine Learning, Honolulu, Hawaii



Berkeley  
UNIVERSITY OF CALIFORNIA



인하대학교  
INHA UNIVERSITY

# Solving privacy issues in machine learning

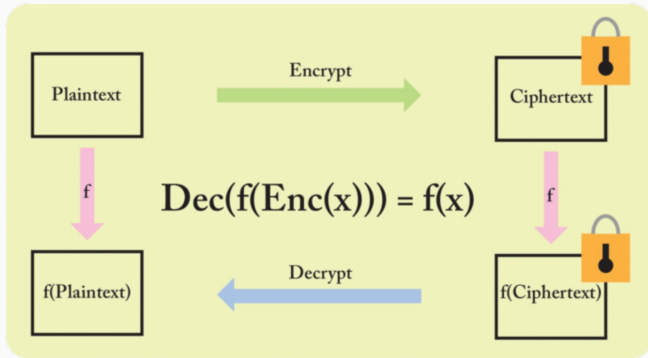
## Question

How can we create / use machine learning models *without leaking private data*?

Privacy-preserving Machine Learning (PPML): prevent privacy leakage in machine learning. For example,

- Federated Learning
- Differential Privacy
- Trusted Execution Environment
- Secure Multi-Party Computation
- **Homomorphic Encryption**

# Homomorphic Encryption (HE)



HE can perform computation on encrypted data *without having to decrypt it*.

# Cheon-Kim-Kim-Song (CKKS) Scheme

- CKKS scheme supports unlimited number of SIMD operations over real and complex numbers, which is appropriate for machine learning applications
  - Addition / Multiplication
  - Rotation / Complex conjugation
  - Bootstrapping
- Lattice-based scheme, quantum-safe
- Needs to approximate non-polynomial functions

- Biggest challenge of HE: computationally expensive

## Question

Can we make HE-based PPML practical?

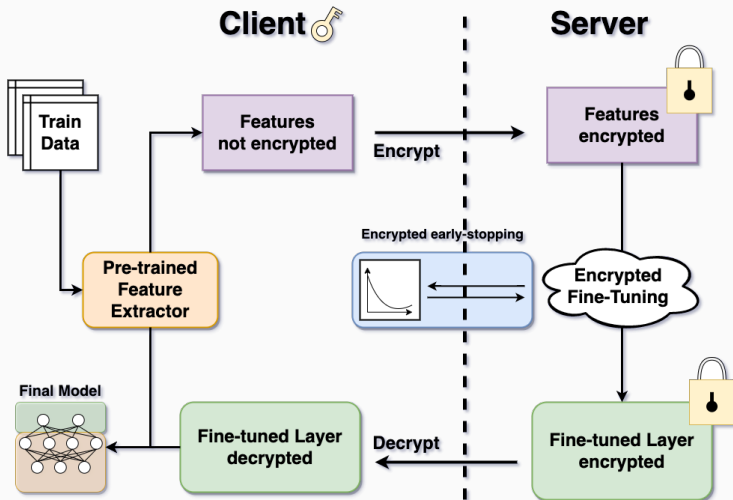


Figure 1: HETAL protocol

- We propose **HETAL**, an efficient and practical **H**omomorphic **E**ncryption based **Tr**Ansfer **L**earning algorithm for privacy-preserving transfer learning.
- Fine-tune classifiers with *encrypted* features extracted from pre-trained models
- *Encrypted early-stopping*
  - Simple communication protocol to determine whether to stop the training or not
- *Encrypted fine-tuning*
  - Softmax approximation on wide interval
  - Efficient encrypted matrix multiplication

# Softmax approximation

- We combine **domain extension** [1] and **normalization** (via comparison [2]) to approximate softmax on a wide range.
- We prove that approximation error is bounded by an explicit constant that does not depend on the domain extension index.
- Our softmax approximation covers inputs in range  $[-128, 128]$  with estimated maximum / average errors of **0.0037**  $\sim$  **0.0224** and **0.0022**  $\sim$  **0.0046**.
  - For input dimensions 3, 5, 7, and 10.

---

[1] Cheon et al., *Efficient homomorphic evaluation on large intervals*, IEEE Transaction on Information Forensics and Security, 2022

[2] Cheon et al., *Efficient homomorphic comparison methods with optimal complexity*, International Conference on the Theory and Application of Cryptology and Information Security, 2020



# Softmax approximation

## Theorem

Let  $p : \mathbb{R}^c \rightarrow \mathbb{R}^c$  be an approximation of the softmax on  $[-R, R]^c$  satisfying

$$\|\text{Softmax}(\mathbf{x}) - p(\mathbf{x})\|_\infty < \epsilon.$$

Then for  $\mathbf{x} \in [-\frac{1}{2}L^n R, \frac{1}{2}L^n R]^c$ , we have

$$\|\text{Softmax}(\mathbf{x}) - p(D_n(\text{Norm}(\mathbf{x})))\|_\infty < \beta + \epsilon,$$

where  $\beta = \beta(\delta, c, r, L, d)$  is a constant that depends only on  $\delta, c, r, L, d$ . Here  $D_n$  is a domain extension polynomial and Norm is a normalization map.

## Encrypted matrix multiplication

- We propose **DiagABT** and **DiagATB**, which are optimized matrix multiplication algorithms of the form  $AB^T$  and  $A^TB$ .
- These algorithms are optimized with techniques including
  - off-diagonal masking
  - tiling
  - complexification
  - partial rotation

# Encrypted matrix multiplication

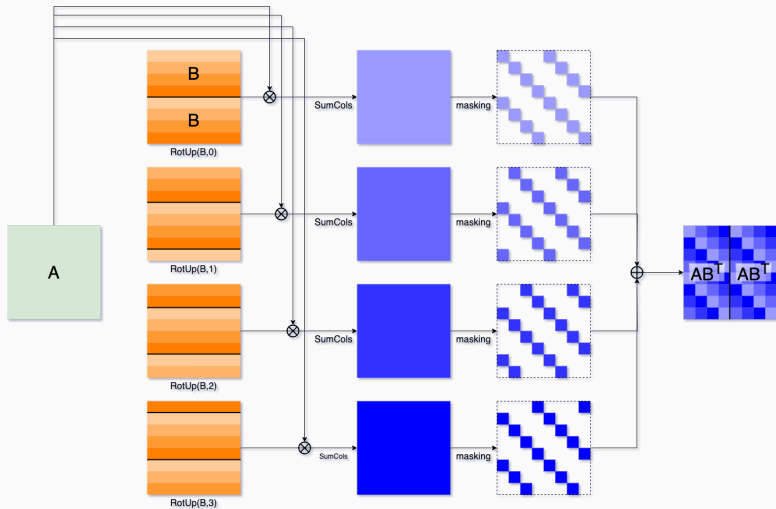


Figure 2: DiagABT algorithm for  $A \in \mathbb{R}^{8 \times 8}$  and  $B \in \mathbb{R}^{4 \times 8}$

# Encrypted matrix multiplication

Our matrix multiplication algorithms show **1.8 to 324 times** speed up compared to baseline algorithms.

| $(a, b, c)$     | $AB^T (A \in \mathbb{R}^{a \times b}, B \in \mathbb{R}^{c \times b})$ |                       |               |         |       | $A^T B (A \in \mathbb{R}^{a \times c}, B \in \mathbb{R}^{a \times b})$ |                       |               |         |      |
|-----------------|---|-----------------------|---------------|---------|-------|--|-----------------------|---------------|---------|------|
|                 | (Jin et al.)*   | ColMajor <sup>†</sup> | DiagABT       | Speedup |       | (Jin et al.)*  | RowMajor <sup>†</sup> | DiagATB       | Speedup |      |
| (128, 128, 4)   | 0.8192  | 0.1104                | <b>0.0601</b> | 13.63   | 1.84  | 10.0352  | 0.1171                | <b>0.0415</b> | 241.81  | 2.82 |
| (256, 256, 8)   | 3.2768  | 0.3203                | <b>0.1211</b> | 27.06   | 2.64  | 40.1408  | 0.3167                | <b>0.1239</b> | 323.98  | 2.56 |
| (512, 769, 4)   | 4.9216  | 0.7609                | <b>0.1223</b> | 40.24   | 6.22  | 60.2896  | 0.7176                | <b>0.3343</b> | 180.35  | 2.15 |
| (1024, 769, 8)  | 9.8432  | 3.0428                | <b>0.3710</b> | 26.53   | 8.20  | 120.5792   | 2.8546                | <b>1.2558</b> | 96.02   | 2.27 |
| (2048, 769, 16) | 19.6864   | 12.6251               | <b>1.2376</b> | 15.91   | 10.20 | 241.1584   | 11.8220               | <b>4.9970</b> | 48.26   | 2.37 |

**Table 1:** Comparison with baseline matrix multiplication algorithms.

\* Jin et al., *Secure transfer learning for machine fault diagnosis under different operating conditions*, International Conference on Provable Security, 2020

<sup>†</sup> Crockett, *A low-depth homomorphic circuit for logistic regression model training*, Workshop on Encrypted Computing & Applied Homomorphic Cryptography, 2020

# Results

With CryptoLab's HEaaS library and a single NVIDIA Ampere A40 GPU, we fine-tune classifiers with encrypted features **in an hour** with **at most 0.51% accuracy drop**.

| dataset             | encrypted    |                 | not encrypted |         |                      |
|---------------------|--------------|-----------------|---------------|---------|----------------------|
|                     | Running time |                 | ACC (a)       | ACC (b) | ACC loss ((b) - (a)) |
|                     | Total (s)    | Time / Iter (s) |               |         |                      |
| MNIST               | 3442.29      | 9.46            | 96.73%        | 97.24%  | 0.51%                |
| CIFAR-10            | 3114.30      | 15.72           | 96.57%        | 96.62%  | 0.05%                |
| Face Mask Detection | 566.72       | 4.29            | 95.46%        | 95.46%  | 0.00%                |
| DermaMNIST          | 1136.99      | 7.06            | 76.06%        | 76.01%  | -0.05%               |
| SNIPS               | 1264.27      | 6.95            | 95.00%        | 94.43%  | -0.57%               |

**Table 2:** Transfer learning results on 5 benchmark datasets.

## Summary

- We present an efficient HE-based transfer learning algorithm called HETAL.
- Efficiency of HETAL is based on two main components:
  - Highly precise softmax approximation algorithm with wide domain of approximation;
  - Efficient encrypted matrix multiplication algorithms, DiagABT and DiagATB.
- We trained classifiers with encrypted features extracted from five real-world datasets via pre-trained transformers in an hour.



paper



code