

DeepProtectの技術と実績

～ 機微情報の暗号化データを用いたAIモデルの連合学習とプライバシー保護 ～



概要

DeepProtectは、機微情報を外部に開示することなく、暗号化したデータにより複数組織で連携して深層学習を行い、高精度なAIモデルが得られるNICT独自のプライバシー保護連合学習技術です。金融分野での実証実験を経て、医療分野への応用にも着手します。

コア技術

- 暗号化された情報のみを中央サーバに送信することで、情報漏洩を防止
- 準同型暗号などの技術を連合学習に应用することで、プライバシー保護と実用性の両立を実現

社会実装業績

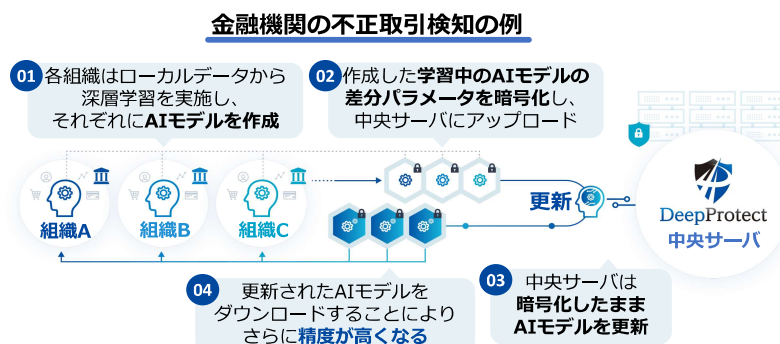
- (R4年度～R6年度)
 - 複数の銀行の実データで実証実験を行い、高い検知精度を達成できることを確認
- JST CREST(JPMJCR19F6) /AIP加速課題(JPMJCR22U5) や高度通信・放送研究開発委託研究(採択番号22901) を遂行
- (R7年度～)
 - JST経済安全保障重要技術育成プログラム(JPMJKP24U1)「高機能暗号を活用した連合学習技術の高度化と医療データへの応用」を開始予定

学界への影響

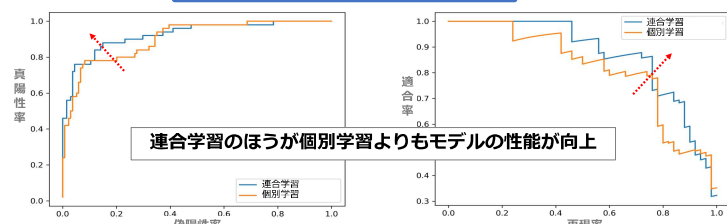
基本論文が2023 IEEE Signal Processing Society Best Paper Award 受賞、2025/5/7現在引用数1863回を達成

L. T. Phong, Y. Aono, T. Hayashi, L. Wang, and S. Moriai, "Privacy-Preserving Deep Learning via Additively Homomorphic Encryption", IEEE Trans. Inf. Forensics Secur., May 2018.

L. T. Phong, T. T. Phung, L. Wang, S. Ozawa, "Frameworks for Privacy-Preserving Federated Learning," IEICE TRANS. INF. & SYST., E107-D(1), 2-12 (2024) 【Invited Paper】



医療分野のMRIデータ解析の例



(a) ROC曲線*1

*1二値分類モデルの性能を評価するグラフであり、曲線が左上に近いほど、偽陽性率を抑えつつ真陽性率が高く、性能が優れていることを示す

(b)適合率・再現率PR曲線*2

*2偽りのあるデータに対するモデルの性能評価に有効なグラフであり、曲線が右上に近いほど、適合率と再現率が高く、性能が優れていることを示す
・適合率：モデルが陽性と予測したものうち、実際に陽性であるものの割合
・再現率：実際に陽性であるもののうち、モデルが正しく陽性と判定した割合

特徴

- ・ 連合学習と準同型暗号の融合
- ・ データセットは非公開のままプライバシーを守りながら深層学習
- ・ 個々のデータセットだけでの学習よりも性能向上



DeepProtect
Webサイト

ユースケース

- ・ 金融：他の銀行に顧客データを開示することなく複数銀行間で連携して不正取引を検知するシステム
- ・ 医療：他の医療機関に患者のデータを開示することなく共同で放射線画像診断支援を行うAI

今後の展開

- ・ 企業へ技術移転→ビジネス展開
- ・ 更なる社会実装の促進
- ・ 金融・医療に限らずさまざまな分野での活用の可能性を探索

【お問合せ先】

サイバーセキュリティ研究所 セキュリティ基盤研究室
Mail : security@ml.nict.go.jp

NICT オープンハウス 2025

Copyright © 2025 NICT All Rights Reserved.