組織間連合学習AIによる社会課題へのチャレンジ
銀行不正送金検知の取組み

小澤 誠一

国立大学法人 神戸大学
数理・データサイエンスセンター センター長
大学院工学研究科 電気電子工学専攻
株式会社 テラアクソン 代表取締役研究責任者

TelaAxon

KOBE UNIVERSITY
神戸大学

自己紹介

小澤 誠一 (ozawasei@kobe-u.ac.jp)
所　属： 神戸大学 数理・データサイエンスセンター・センター長
工学研究科電気電子工学専攻/未来医工学研究開発センター
株式会社テラアクソン・代表取締役研究責任者

◎ **研究内容**（第2次人工知能ブームのときから）
ニューラルネット，機械学習，追加学習、パターン認識，ビッグデータ解析，セキュリティ，Security for AI、画像認識，文書解析、プライバシー保護機械学習、スマートアグリなど

◎ **進行中の研究テーマ**

1) 機械学習のサイバーセキュリティへの応用
サイバー攻撃検知・観測・可視化，攻撃情報の収集, なりすまし検知、悪性サイト、悪性JaveScript判定、security for AI など

2) 機械学習の文書解析への応用
金融文書解析、SNS炎上検知

3) 深層ニューラルネットを使った画像認識
農作物の生育情報取得、磁場センシング画像による危険物検知

4) 暗号データに対する機械学習アルゴリズムの開発
プライバシー保護データマイングと異常検知、複数銀行間連合学習を用いた不正送金検知

研究開発の概要、目的、背景

■ **特殊詐欺による不正出金被害**

- R4年度　17,570件（一日当たり約48件）
  被害額は370.8億円（+88.8億円、+31.5%）
- オレオレ詐欺、預貯金詐欺、キャッシュカード詐欺、架空料金請求詐欺、還付金詐欺など
- 依然として、高齢者をターゲットとした被害が多い。

認知件数合計

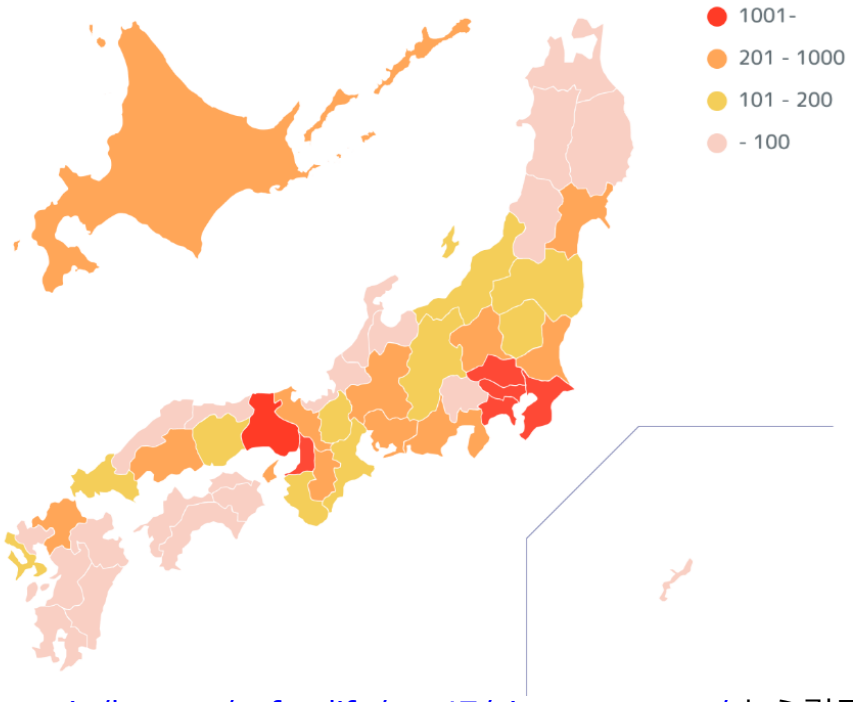
17,570件

合計被害額

370億8,135万5,000円

被害件数上位の厳重注意地域

1位	東京	3,218 件
2位	神奈川	2,090 件
3位	大阪	2,064 件
4位	千葉	1,457 件
5位	埼玉	1,387 件



警視庁特殊詐欺対策ページ： https://www.npa.go.jp/bureau/safetylife/sos47/circumstances/ から引用

研究開発の概要、目的、背景

■ マネーロンダリング（資金洗浄）及びテロ資金供与・拡散金融

- 特定事業者→所管行政庁に届けられた「疑わしい取引」
  R4年度　約58.3万件（一日当たり約1598件）
  警察庁「犯罪収益移転防止に関する年次報告書（令和４年）」
- 麻薬取引、脱税、粉飾決算、非合法ビジネス（ヤミ金融、賭博など）
  テロ資金供与、拡散金融（資産凍結等措置の対象となっている者に、資金または金融サービスの提供をする行為）
- FATFの監視対象国ではないものの「重点フォローアップ」(2021.8.30)が必要とされ、最終評価期限(5年後)までにFATFへ改善報告を3回程度必要

pwc: https://www.pwc.com/jp/ja/knowledge/prmagazine/pwcs-view/202111/35-07.html

財務省（2021.8）
- 「マネロン・テロ資金供与・拡散金融対策に関する行動計画」を発表
- 2024年春を期限として「取引時確認、顧客管理の強化および平準化の観点から、**取引スクリーニング、取引モニタリングの共同システムの実用化を図る**」としている。

全銀協：AML/CFT業務の高度化・共同化

■全国銀行協会（2022.10.13発表）

2022年10月13日

各　位

一般社団法人全国銀行協会

AML／CFT業務の高度化・共同化に係る新会社の設立について

一般社団法人 全国銀行協会（会長：半沢淳一　三菱UFJ銀行頭取）は、本日開催の理事会において、AML／CFT業務の高度化・共同化を図ることを目的とした株式会社（当協会が100%出資）を新たに設立することを決定いたしました。

新会社においては、「取引モニタリング等のAIスコアリングサービス」を提供予定であり、本年度中に準備会社を設立のうえ、2024年度以降の段階的なサービス提供に向けた準備を進めてまいります。

当協会は、2020年度、NEDO（国立研究開発法人 新エネルギー・産業技術総合開発機構）から実証事業を受託し、AI等の先端技術を活用した高度なシステムの共同化による効率的かつ実効的なマネー・ローンダリング対策の実現に必要な規制の精緻化の可能性、課題についての調査・整理を実施いたしました。その後、2021年度には「AML／CFT業務共同化に関するタスクフォース」を設置し、共同化の実現について検討を進めてまいりました。

今般、その検討の結果として新たに株式会社を設立することを決定したものです。

https://www.zenginkyo.or.jp/news/2022/n101302/ から一部を抜粋

株式会社マネー・ローンダリング対策共同機構
（2023.1.6 設立）

[業務内容]
- ❑ AIスコアリングサービス
- ❑ 業務高度化支援サービス

2024年度以降の段階的なサービス提供予定

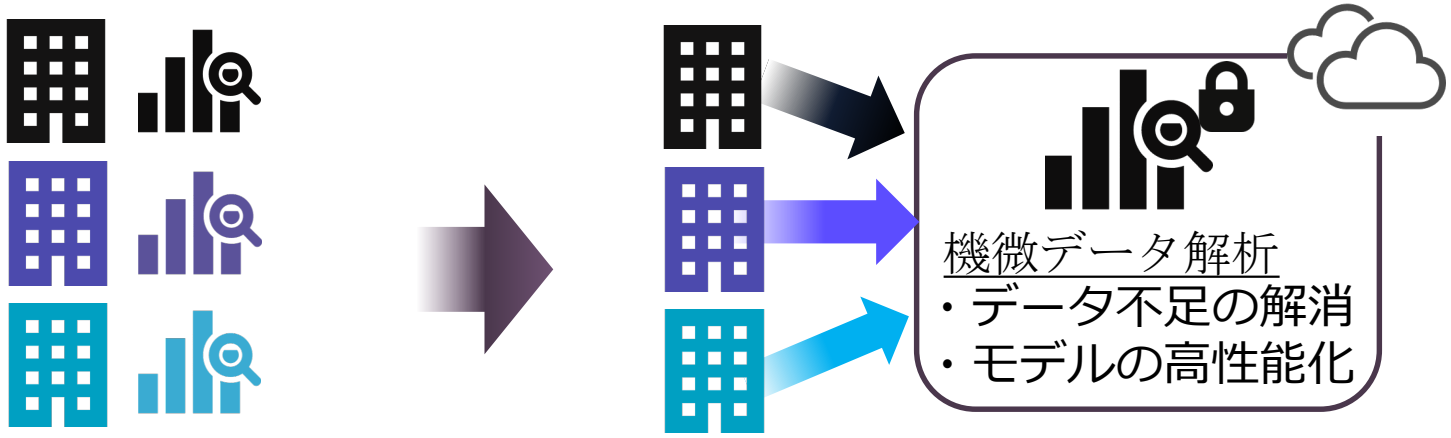
組織間プライバシー保護学習が求められる背景

~背景~　組織が保有するデータの多くは，組織を超えての利活用が容易でない．

組織間データ利活用
データ量の増大，多様化により，AIの予測精度向上が期待できる

個人の機微な情報
情報漏洩リスク，関連法による規制から，組織間データ利活用が制限される

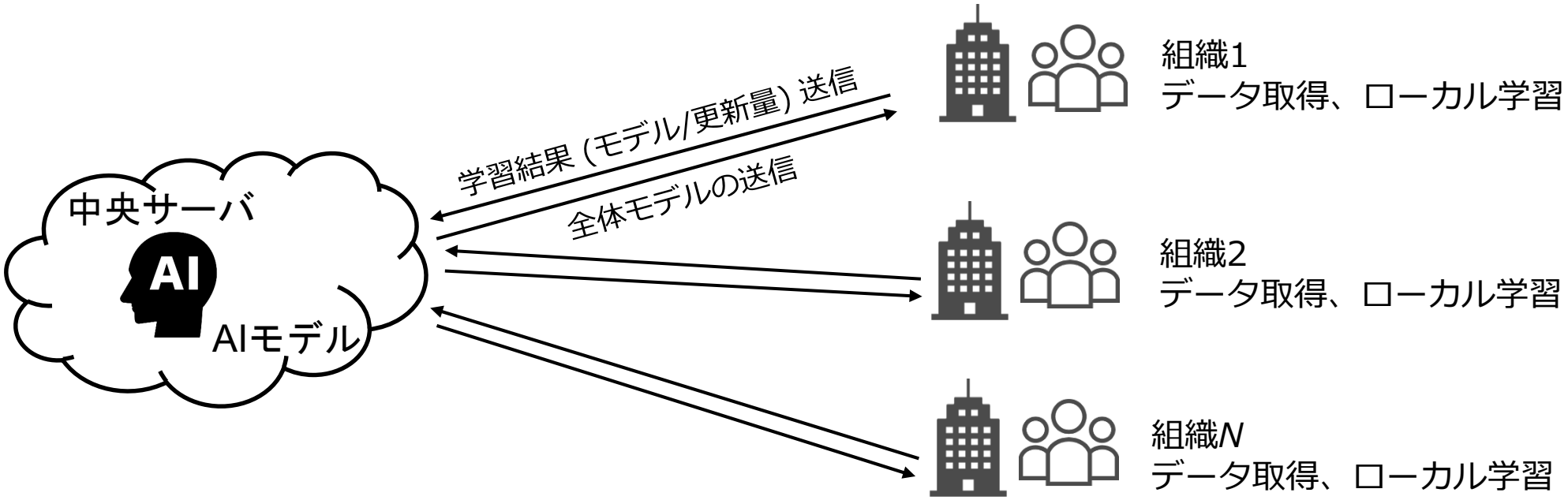
実社会に遍在するデータを有効に利活用するため，
組織間プライバシー保護データ解析手段を実現する必要がある



各組織は独自に解析　　各組織が協力して解析

組織間プライバシー保護連合学習

複数のデータ提供者（組織）が互いにデータを直接共有せず、連合学習する。



プライバシー保護機械学習モデル (PPML)

Type-1: 通信経路のみを暗号化

Type-2: 通信経路も中央サーバでも暗号化（準同型暗号が必要）

連合学習方式

水平型連合学習

銀行A

UID	年齢	性別	入金合計	出金合計	残高
0001	35	女	330,500	128,000	35,000
0002	28	男	280,000	250,000	2,000
0003	78	女	125,000	78,000	123,500

AI 異なる顧客の同種データを学習

銀行B

UID	年齢	性別	入金合計	出金合計	残高
0101	18	女	30,500	5,000	35,000
0102	58	男	580,000	450,000	542,000
0103	40	男	325,000	178,000	958,500

垂直型連合学習

銀行A

UID	年齢	性別	入金合計	出金合計	残高
0001	35	女	330,500	128,000	35,000
0002	28	男	280,000	250,000	2,000
0003	78	女	125,000	78,000	123,500

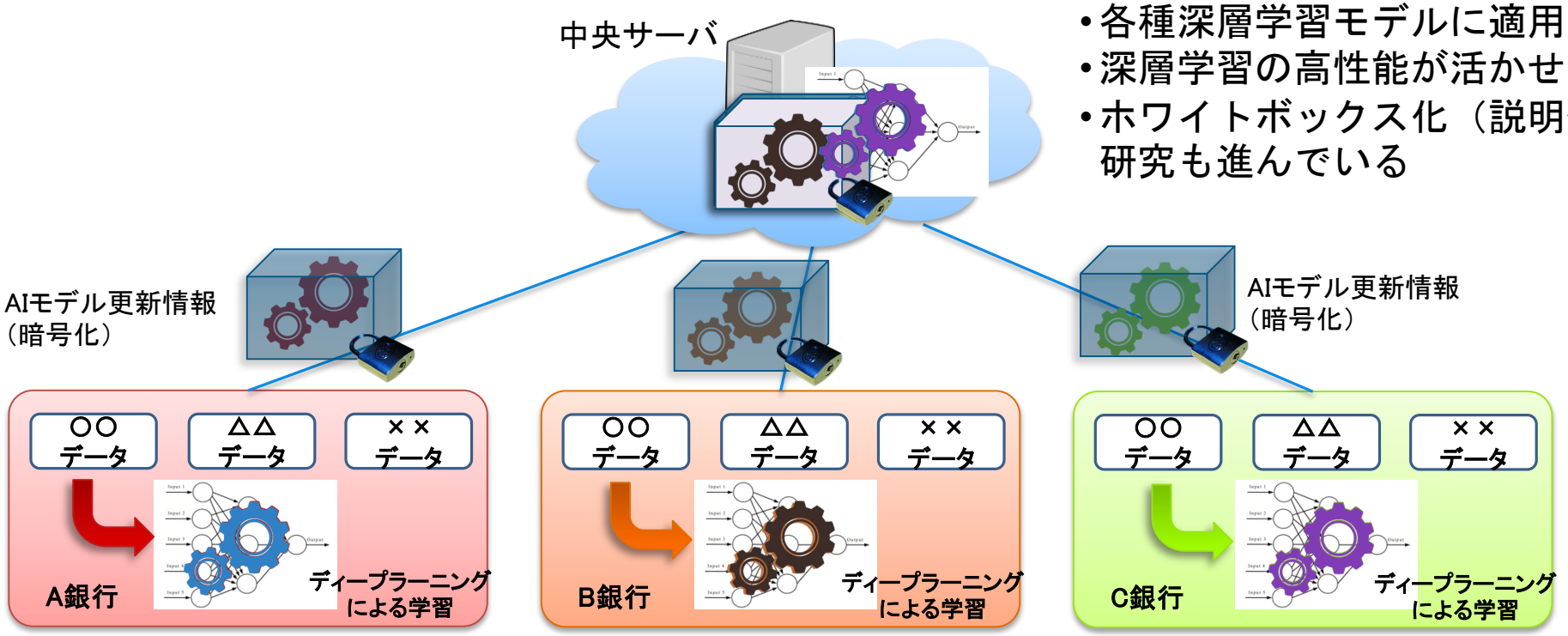
AI 同一顧客の異種データを学習

会社A

UID	年齢	購入品	金額	個数
0001	35	バッグ	2,500	2
0002	28	時計	30,000	1
0003	78	書籍	5,000	5

水平連合学習型深層学習（DeepProtect）

複数の組織が持つデータを外部に開示することなく深層学習を行うプライバシー保護深層学習システム



- 各種深層学習モデルに適用可能
- 深層学習の高性能が活かせる
- ホワイトボックス化（説明性向上）の研究も進んでいる

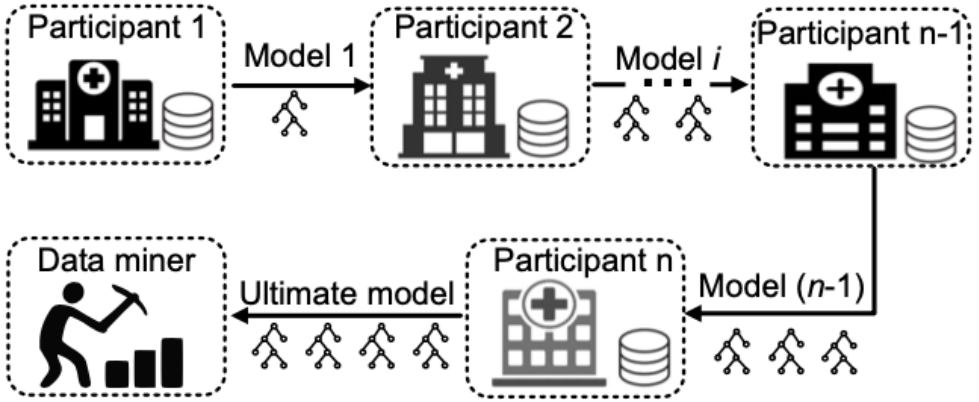
複数組織で連携した分散協調型の深層学習

特許第7058440号

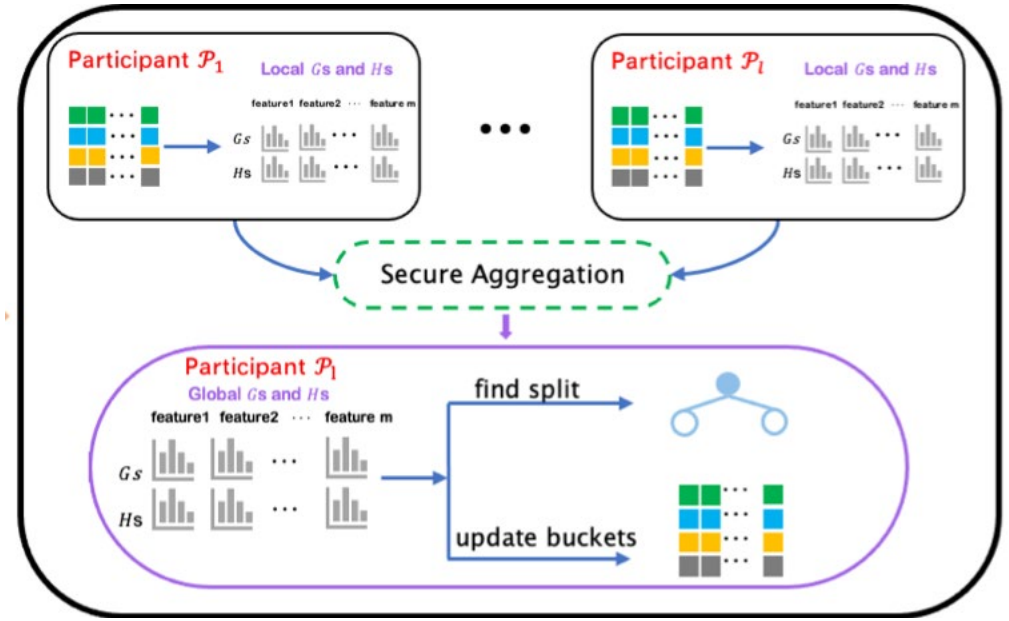
- Le Trieu Phong, Tran Thi Phuong: Privacy-Preserving Deep Learning via Weight Transmission. *IEEE Trans. Inf. Forensics Secur.* 14(11): 3003-3015 (2019)

組織間水平連合学習のアプローチ

✔ **TFL (Tree-based Federated Learning)**
(Zhao et al., 2018)
- 組織ごとに順次、決定木をローカル学習
- シンプルな実装、組織間通信は1回のみ
- 性能は低い

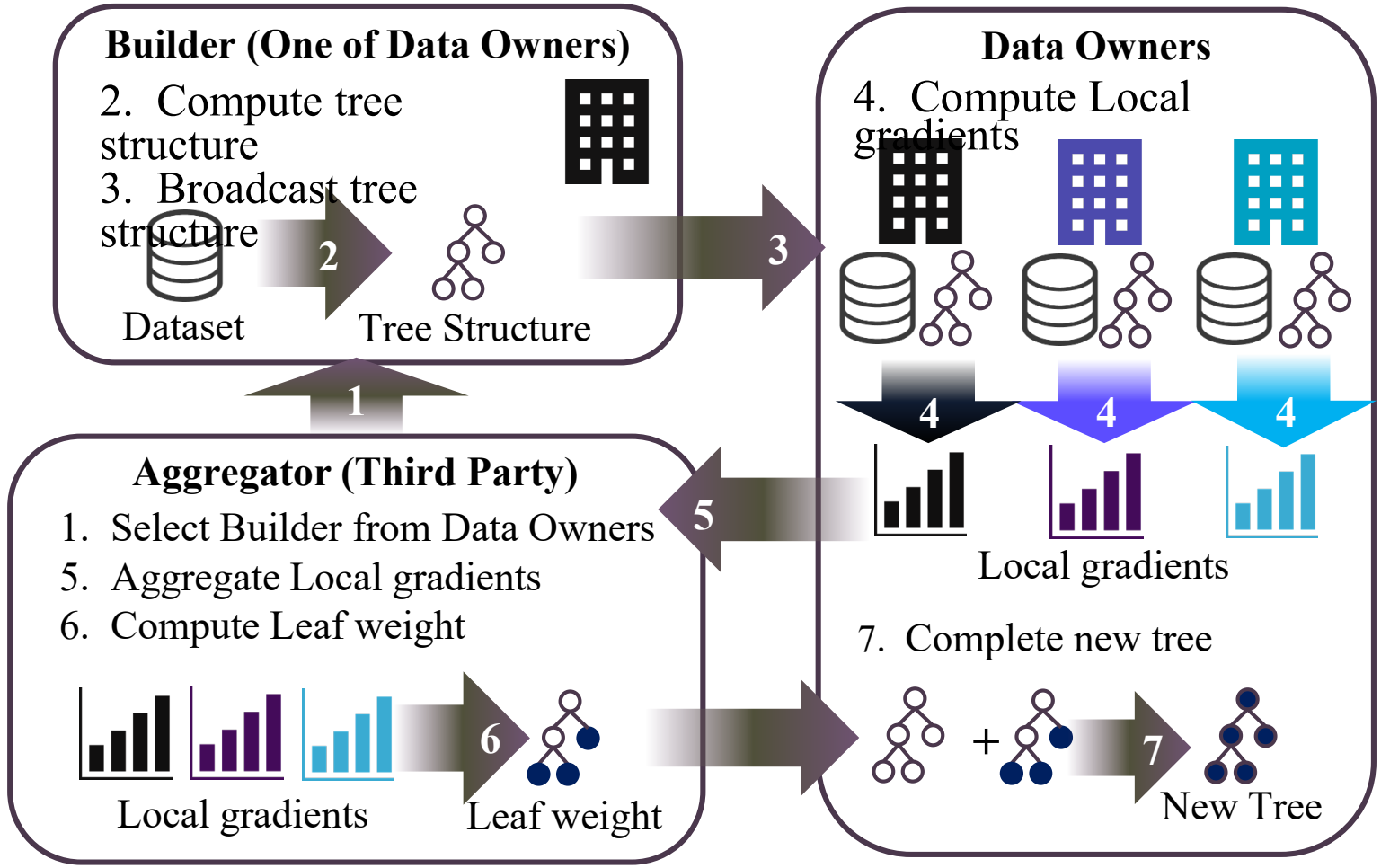


✔ **FederBoost** (Tian et al., 2020)
- 水平型FederBoostは通常のGBDTと同様の手順でグローバルに学習
- 性能は通常のGBDTと同じ
- レベル（深さ）ごとに組織間通信が必要
- ノードに割り当てられるデータ数が少ない場合、組織データの漏洩が問題



水平連合学習勾配ブースティング決定木アンサンブル

eFL-Boost (Efficient Federated Learning for Gradient Boosting Decision Trees)



【特徴】

通信コスト：低い

情報漏洩：他組織には木構造情報のみ

予測性能：水平型FederBoostと同等

F. Yamamoto, S. Ozawa, L. Wang, “eFL-Boost: Efficient Federated Learning for Gradient Boosting Decision Trees,” IEEE Access, vol. 10, pp. 43954-43963, 2022

銀行不正送金実証実験 1

CREST-AI /AIP加速課題 (2016-2024)

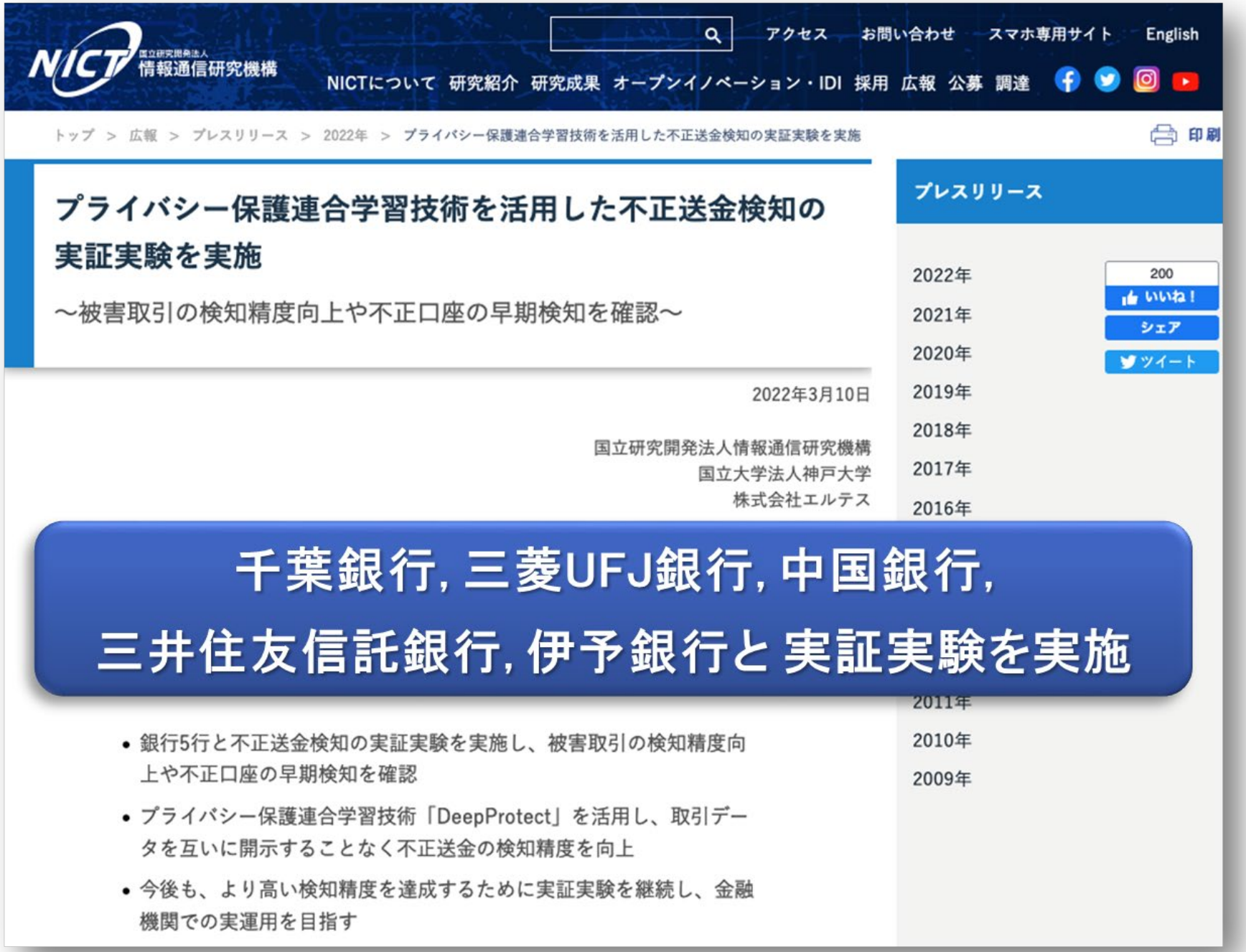
研究領域：

「イノベーション創発に資する人工知能基盤技術の創出と統合化」

(研究総括：栄藤稔(大阪大学))

戦略目標:

急速に高度化・複雑化が進む人工知能基盤技術 を用いて多種膨大な情報の利活用を可能とする統合化技術 の創出



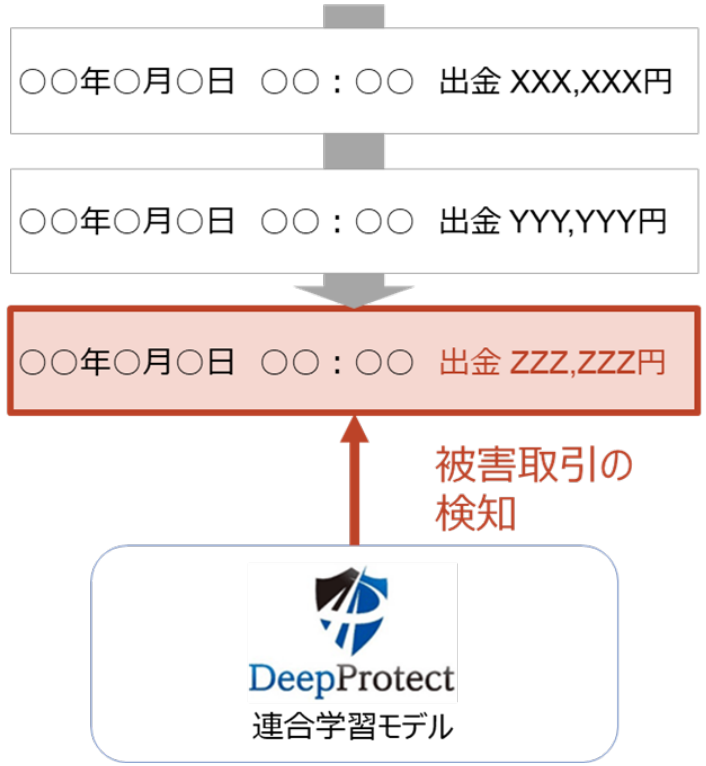
NICT
神戸大学
(株)エルテス

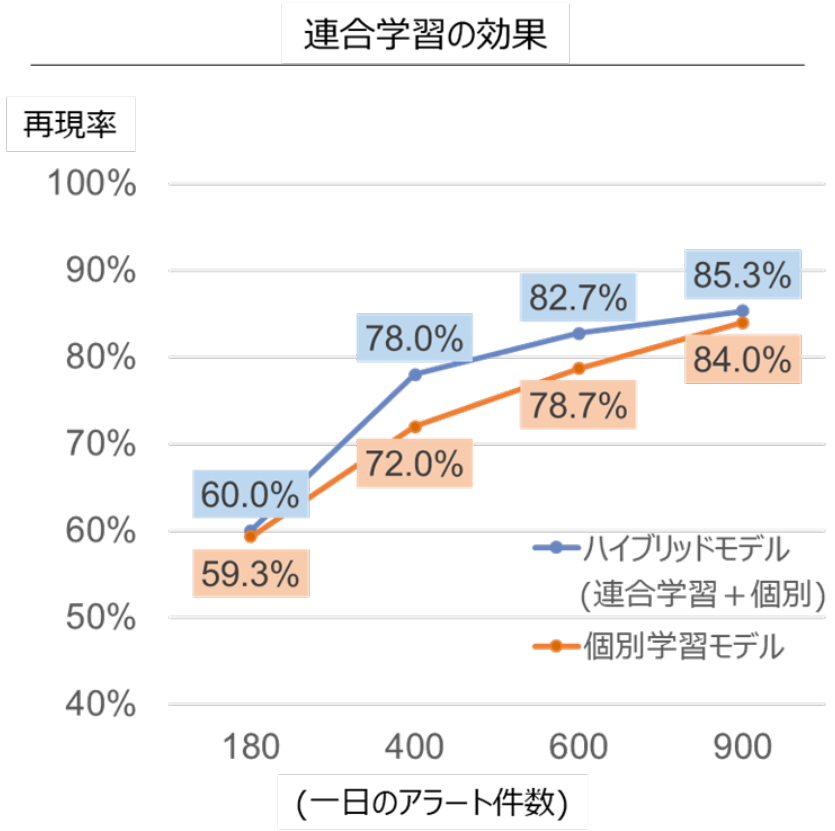
JST CREST 加速フェーズ実績

実証実験結果①：被害取引の検知

- 連合学習モデルにより検知精度が向上、検知精度は1日あたりのアラート件数が600件以上で当初目標の80%以上を達成
- 個別学習モデルでは検知できなかった不正取引が検知された事例も

口座A

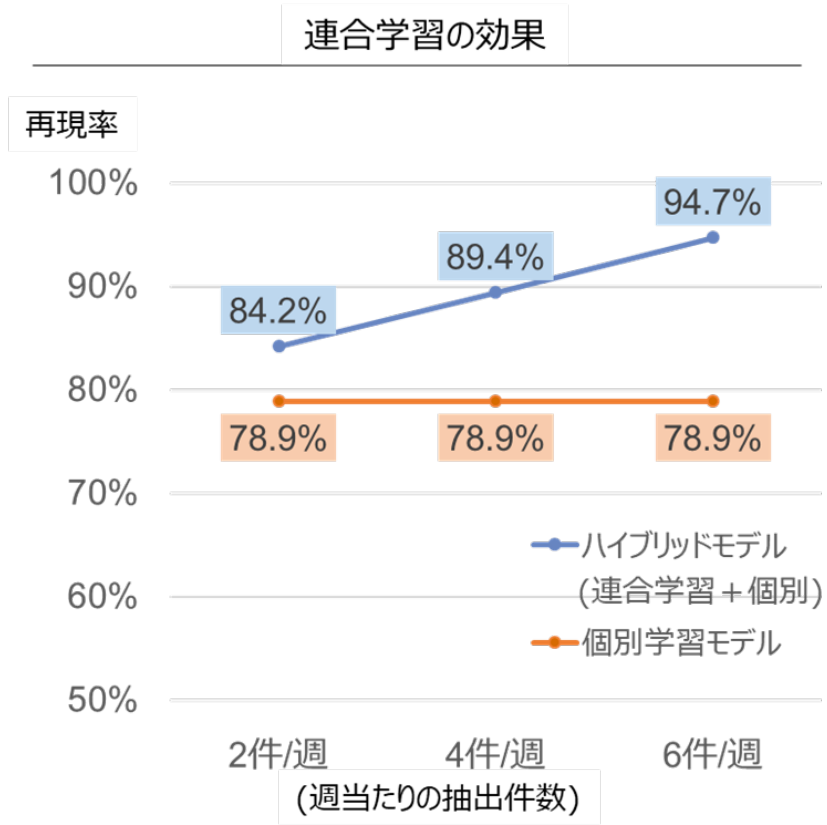




JST CREST 加速フェーズ実績

実証実験結果②：不正口座の検知

- 個別学習モデルと比較してハイブリッドモデル（個別学習モデルと連合学習モデルの組み合わせ）が高い性能、検知率80%以上を達成
- 実際の不正口座凍結よりも20～50週程度の早期検知が可能



連合学習の課題

（１）顧客口座・取引情報の組織間不整合

【理想】

- ✔ 取引・口座特徴量の標準化
- ✓ 判定基準の統一化

【対策】

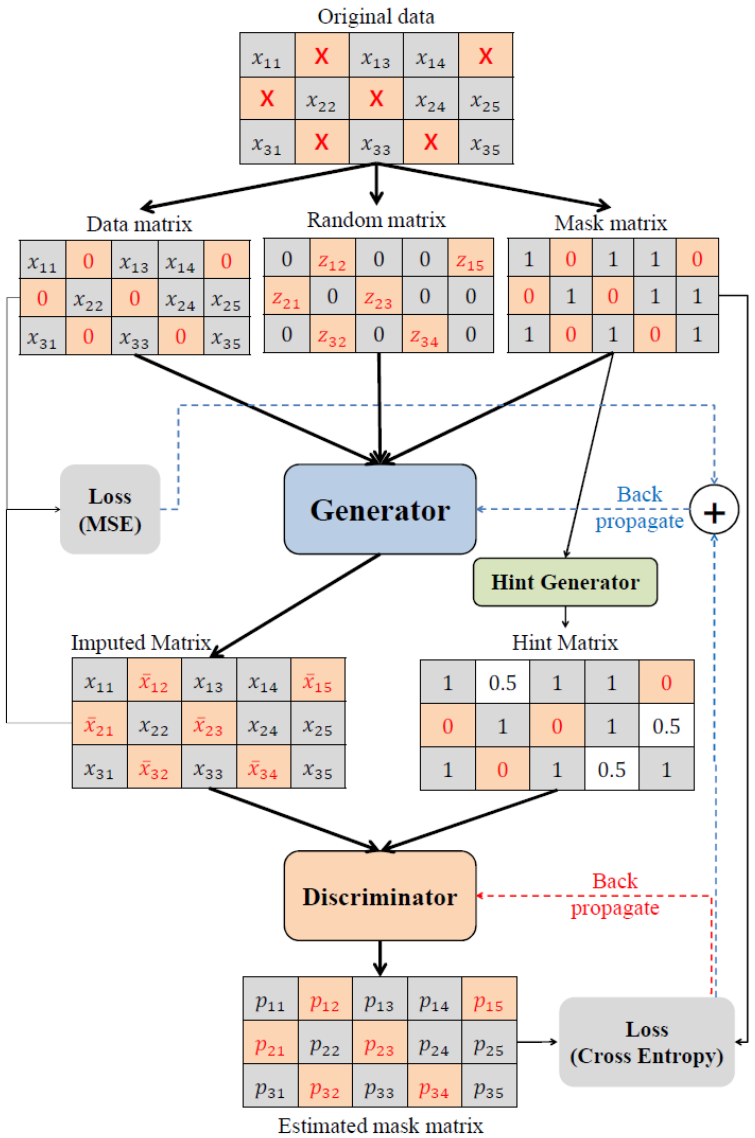
- ➢ 欠損値補完手法の導入
  - ○ 線形補間、GAN　など
  - ○ 非数値データ、質的データの取扱い
- ➢ 識別器のハイブリッド化
  - ○ アンサンブルモデル

  → 提供データ項目が多ければ得する
- ➢ データ提供者の選択・重み付け
- ➢ データ定型化（標準化）を推進

	特徴量	銀行1	銀行2	銀行3	銀行4
1	年齢	○	○	○	○
2	法人格（法人・個人）	○	○	○	×
3	口座開設期間	○	○	×	○
4	国籍	○	○	×	○
5	口座・取引場所距離	○	○	×	○
6	入出金時刻	○	○	×	○
7	入金額	○	○	○	○
8	出金額	○	○	○	○

※連合学習の対象は１つではなく、複数あってよい。

Generative Adversarial Imputation Nets (GAIN)



- 欠損値はある分布に従って観測される連続値
- 欠損箇所を'0' としたマスク行列を作成し、そこから一定確率で '1' と '0' を反転させたヒント行列を作成
- 実データ、欠損値に対してある分布に従って値を与える乱数行列、マスク行列をGeneratorに与えて、補完行列を計算
- Discriminatorは補完行列とヒント行列が与えられて、欠損データを当てる
- Generatorは実データとDiscriminatorが間違えるような欠損補完をしたい。対して、DiscriminatorはGeneratorの欠損補完を正しく当てたい。

銀行不正送金実証実験２
NICT高度通信・放送研究開発委託研究 (2022-2024)

【研究開発課題名】
プライバシー保護連合学習の高度化に関する研究開発

【副題】
継続実運用に資する不正取引モニタリングに向けたプライバシー保護連合学習の高度化

【受託者名】
代表研究者：国立大学法人神戸大学
研究分担者：EAGLYS株式会社

研究開発体制及び分担

代表研究者
（国立大学法人神戸大学）

- 研究開発項目１-１：実運用を模擬したテスト環境の開発と性能評価
- 研究開発項目 1-3：DeepProtect の敵対的サンプルとデータ汚染攻撃への耐性向上
- 研究開発項目 3-1：不正送金検知特徴量と不正判定基準の標準化

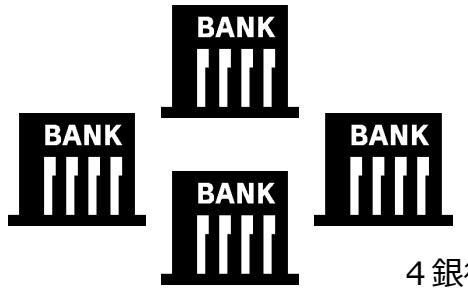
研究分担者
（ＥＡＧＬＹＳ株式会社）

- 研究開発項目 1-2：DeepProtect の継続学習化
- 研究開発項目 2-1：データパイプライン構築モジュールの開発
- 研究開発項目 2-2：DeepProtect 学習実施モジュールの開発
- 研究開発項目 3-2：人間系フィードバックを容易とする支援ツールの開発

銀行との連携
計算基盤の利用など

NICT
■ 実証実験計算基盤（銀行データ解析サーバ）
■ 口座取引・顧客口座データ

データ提供



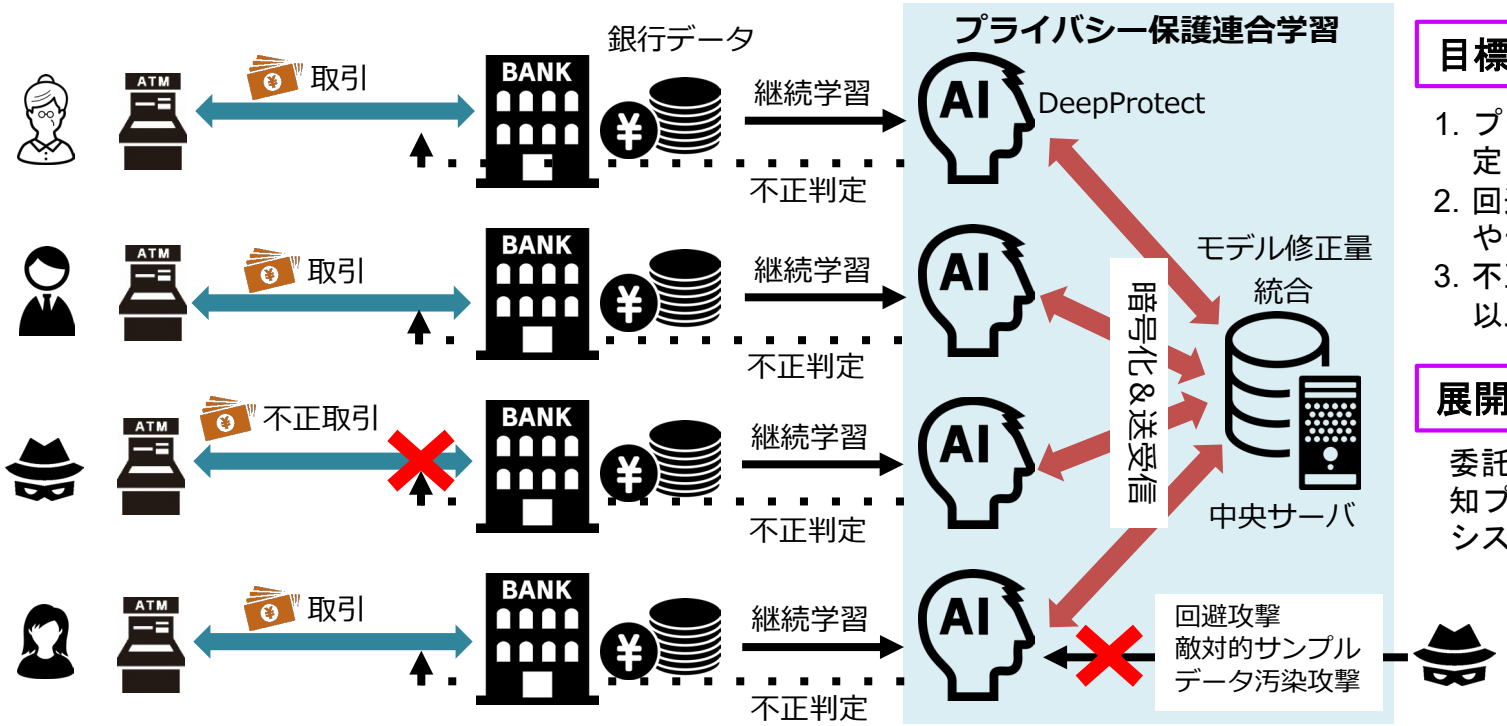
４銀行以上

課題229
採択番号22901

プライバシー保護連合学習の高度化に関する研究開発

継続実運用に資する不正取引モニタリングに向けたプライバシー保護連合学習の高度化

研究概要：組織の機微なデータを他組織と共有しなくても、高度なAIを協調して構築できる連合学習は、プライバシー保護を重視する社会実装に不可欠な技術となりつつある。本研究では、現状の連合学習で実運用上解決すべき問題である安定した継続学習および回避攻撃を意図した敵対的サンプルやデータ汚染攻撃への耐性向上を実現し、DeepProtectの高度化を行うことを目的とする。また、本技術を喫緊の社会課題であるマネーロンダリング対策に導入し、４行以上の金融機関との連携を通して不正送金検知実証実験を実施し、不正取引の検知再現率が継続的に90%以上維持されることを目標とする。



目標

1. プライバシー保護連合学習における安定した継続学習の実現
2. 回避攻撃を意図した敵対的サンプル・やデータ汚染攻撃への耐性向上
3. 不正取引の検知再現率が継続的に90%以上を達成

展開・普及計画

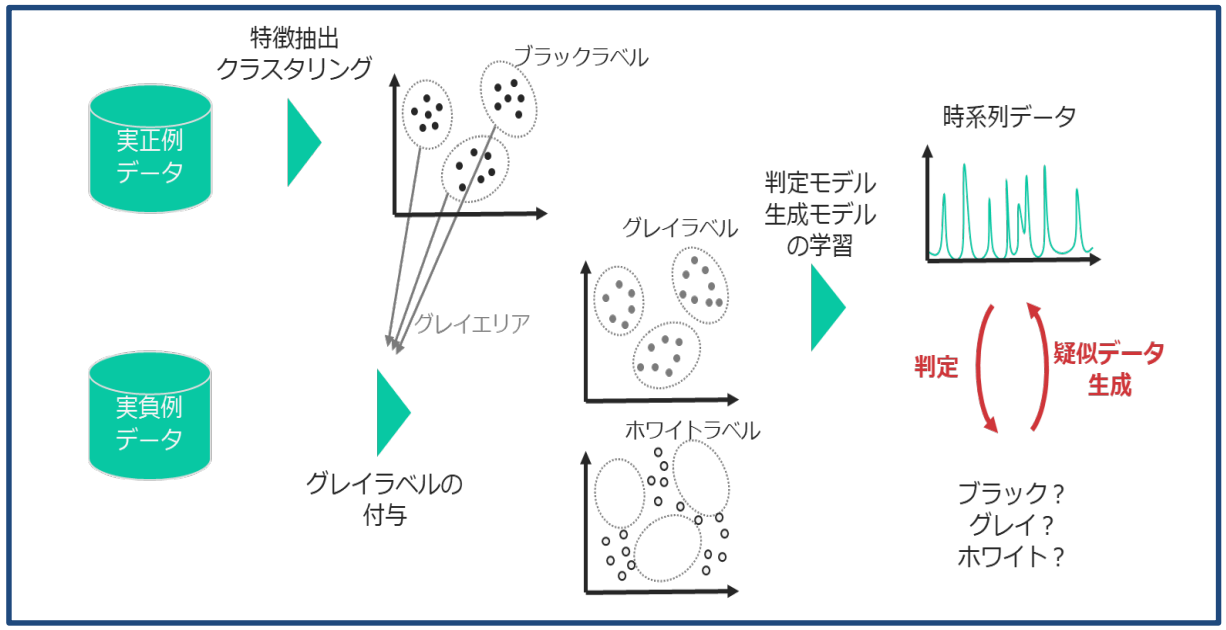
委託研究の終了後において、不正取引検知プロトタイプの取引モニタリング共同システムへの採用を目指す

DeepProtectの高度化に関する研究（1）

研究開発項目1-1　実運用を模擬したテスト環境の開発と性能評価（神戸大学）

■ データ加工やデータの可視化など、自由度の高いデータ分析とシステム開発が必要
■ 銀行不正送金検知の実運用を模擬したテスト環境を開発

→ **疑似データ生成モデルの構築**

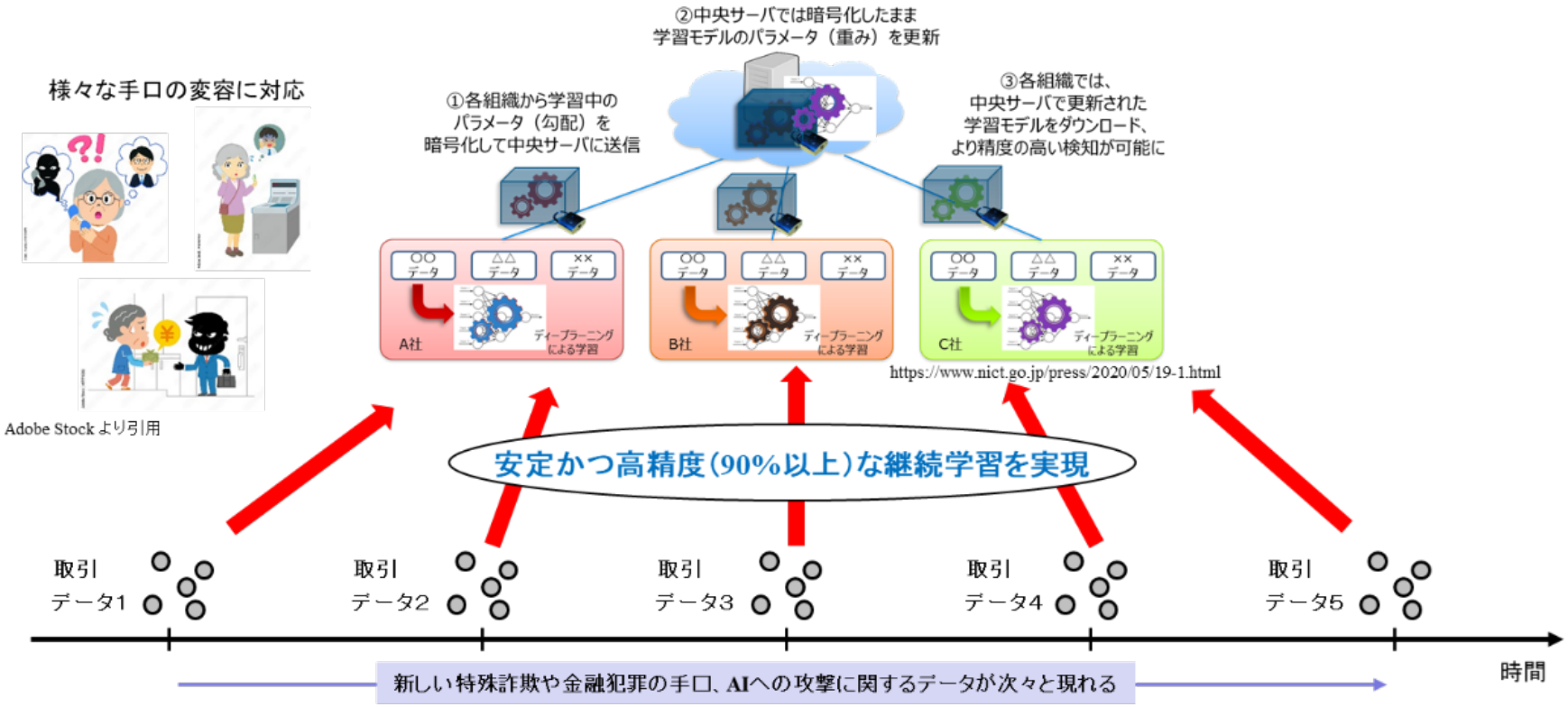


1. クラスタリングベースの生成手法
2. 敵対的生成ネットワーク（GAN）を用いた生成手法

DeepProtectの高度化に関する研究 (2)

研究開発項目1-2　Deep Protect の継続学習化（EAGLYS）

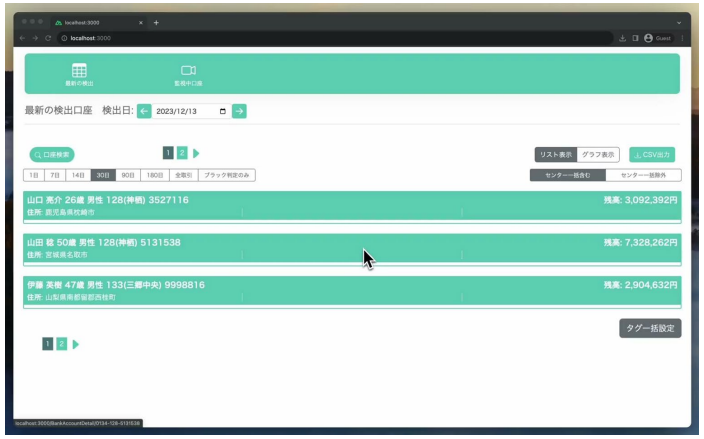
■忘却による精度劣化がなく、安定した継続学習手法を開発
■長期的な運用でも、現実的なリソースで高精度検知が持続運用されることを目指す。

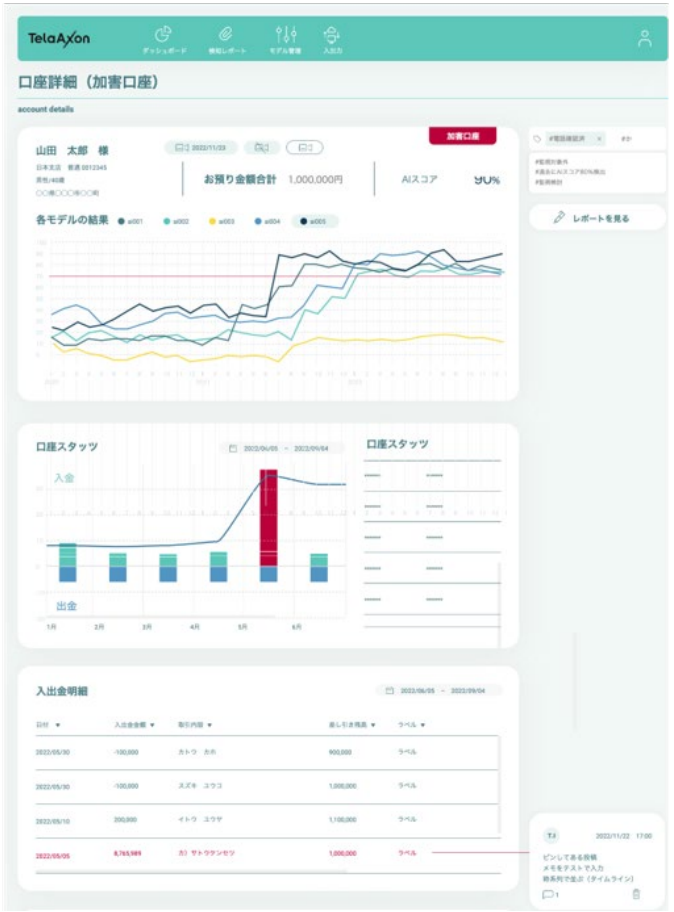


継続実運用を想定した不正検知実証実験環境の整備

研究開発項目3-2　人間系フィードバックを容易とする支援ツールの開発（EAGLYS）

実運用の過程において、人間系のフィードバックを取り込むために、検知結果の表示、結果修正、修正内容のデータソース反映を実施するシステムを開発





まとめ

- 特殊詐欺、マネーロンダリングの発生件数は依然増加している。
- 組織が互いにデータを共有しなくても、AIを学習できる仕組み、例えば「連合学習」が社会実装で求められている。
- 実問題では、データオーナによっては、すべてのデータ項目を出せないことが多く、完全な「水平型連合学習」にならない。
- 欠損値を補完する仕組みの導入が望ましいが、非数値データの生成が課題
- 継続的に学習・予測するしくみが必要
- 個々の銀行内に不正検知AIシステムを設置し、金融犯罪担当者のフィードバックをもらいながら、継続運用する「行内実証実験」の準備が整いつつある。