

(6) 特許出願、外部発表等

		累計（件）	当該年度（件）
特許出願	国内出願	0	0
	外国出願	0	0
外部発表等	研究論文	0	0
	その他研究発表	9	5
	標準化提案・採択	0	0
	プレスリリース・報道	0	0
	展示会	0	0
	受賞・表彰	0	0

(7) 具体的な実施内容と最終成果

研究開発項目1：DeepProtect の高度化に関する研究

プライバシー保護連合学習技術「DeepProtect」を実運用に資する取引モニタリングシステムとするため、不正送金検知の現状を模擬したテスト環境下で、以下の3つの問題に対する要素技術の開発を行った。

- ① 不正取引データの合成による不均衡データ問題の緩和
- ② 安定した継続学習の実現
- ③ AIの不正判定を回避する敵対的サンプル攻撃への防御

- 1-1 不均衡データ問題の緩和を実現するため、敵対的生成ネットワークを用いた合成データの生成と確率モデルを用いた不正取引データの合成手法を提案した。組織ごとに保有しているデータのクラスタが一部異なっている状況を想定した人工データを作成し、これを使って不均衡データ問題の緩和を試みた。その結果、後者の確率モデルを用い、口座取引の合成データを生成する方法が有効であることを検証した。
- 1-2 連合学習と継続学習を同時に安定的に実施する手法を提案し、機構が提供する銀行のデータへの適用を通じて、その有用性を示した。機構が提供する4銀行取引データの分析を実施し、異なるデータカラムを有する環境で連合学習を実施する手法を提案した。その結果、目標とした再現率90%以上の検知を達成した。
- 1-3 犯罪者が銀行口座の入出金額や頻度、タイミングなどを意図的に変更して、AIによる検知を回避する敵対的サンプル攻撃が可能であることを、機構が提供する銀行取引データに対して実証した。また、この敵対的サンプル攻撃への防御が可能となるよう高度化する方法を検討した。

研究開発項目2：DeepProtect を用いた不正取引検知エンジンの開発

- 2-1 DeepProtect によって不正送金検知モニタリングを実施し、最新の犯罪手口を反映するために継続的な学習を実現するには、最新データを安定的に取り込むシステムが必要不可欠である。そこで、銀行における不正検知モニタリングの実施現場で学習実施を行うデータパイプラインと学習・推論を行うプロトタイプシステムを開発した。
- 2-2 銀行ローカル環境内のみで学習を実施するローカル学習実施機能、中央サーバーと通信しながら連合学習を実施する連合学習実施機能、新しい取引データを対象に推論を実行する推論機能を有するプロトタイプシステムを開発した。また、推論結果はデータパイプラインにフィードバックされ、人間系支援ツールに疑わしい取引データとして表示される。銀行担当者が疑わしい取引に対してラベルの修正を実施することで、正例として学習に取り込まれる機能をプロトタイプシステムに実現した。

研究開発項目3：継続実運用を想定した不正送金検知実証実験環境の整備

3-1 DeepProtect は水平型連合学習モデルであり、異なる銀行がもつ不正取引の特徴を連合学習スキームのもとで間接的に共有するため、不正検知実証実験を主導するEAGLYSが中心となって参加4銀行から提供された顧客口座の共通データ項目を分析した。その結果、不正送金検知特徴量の標準化は、銀行から提供を受けるデータ項目の共通化で実現するのではなく、共通のデータ項目をもつ銀行間で水平型連合学習モデルを複数個構築し、それらをアンサンブルすることで実質的な標準化を達成するアプローチを新たに提案した。

3-2 銀行内で継続的に DeepProtect を用いた不正検知や連合学習を行えるよう、不正送金監視者による運用を見据えて、人間系のフィードバックの取り込みを容易とする支援ツールの開発を実施した。具体的には、不正検知モニタリングシステムの運用を通じて、連合学習モデルの継続的学習と精度向上が容易にできるアノテーションシステムを開発した。

(8) 研究開発成果の展開・普及等に向けた計画・展望

DeepProtect は不正取引のデータ量が限られていても機械学習ベースの不正検知モニタリングが実施可能という利点があり、研究開発開始時には地銀など小規模な銀行におけるAIの導入への期待は大きかった。しかし、そこから2年以上が経過し、システムベンダーによるAIモニタリングサービスの導入検討が進んでいる。そこで、小規模な組織が多い信用金庫を初期顧客ターゲットとして、以下の活動を通じて本研究成果の社会導入を行う予定である。

(a) フェーズ1：プロダクトマーケットフィットを目的としたPOCフェーズ（～2026年度末）

信用金庫をターゲットとした、不正取引モニタリングの状況調査、および現場で生じるニーズのヒアリングを実施する。ヒアリングを通じて、本研究開発で開発したプロトタイプを紹介を行い、POCの実施提案を行う。2026年末までに5件のPOC実施を目標とする。POCにおける運用を通じて、不正検知モニタリングサービス、およびシステム要件整理を行う。またPOC実施金融機関で勘定系システムや既存のモニタリングシステムを提供するシステムベンダーにヒアリングを実施し、既存のモニタリングシステムへの組み込みを模索することで、ビジネスモデル仮説を構築する。

(b) フェーズ2：商用プロダクト開発（～2027年度末）

POCを通じて得られた要件をもとに、サービス提供として必要な機能要件、非機能要件を元としたプロダクト開発を行う。プロダクトは他のシステムと独立して不正検知モニタリングを行うシステム開発を主として実施するとともに、データパイプライン・学習実施モジュールを既存のモニタリングシステムに組み込むことが可能なバージョンの2種類のシステム開発を開発候補とし、開発内容はPOCフェーズの活動を通じて決定する。

(c) フェーズ3：プロダクト展開（～2029年度末）

プロダクトの展開を行うためにチャンネル開拓を実施する。チャンネル開拓の活動候補は、プロダクト開発内容に応じて実施する。