2024年09月24日　技術情報

ネットワーク分散AIの実現に向けた連合学習

#AI・生成AI　#IoT・データ利活用　#研究開発

執筆者

谷村 崇仁 (Tanimura Takahito)
日立製作所 研究開発グループ デジタルサービスプラットフォームイノベーションセンタ エッジインテリジェンス研究部 所属
現在，ネットワークおよび分散AIの研究開発に従事
博士（工学）
電子情報通信学会シニア会員，日本物理学会会員

垂水 信二 (Tarumi Shinji)
日立製作所 研究開発グループ ヘルスケアイノベーションセンタ デジタルヘルスケア研究部 所属
現在，ヘルスケアデータ解析および分散AIの研究開発に従事
人工知能学会会員

北川 雄一 (Kitagawa Yuichi)
日立製作所 研究開発グループ デジタルサービスプラットフォームイノベーションセンタ エッジインテリジェンス研究部 所属
現在，分散AIの研究開発に従事
電子情報通信学会会員

礒田 有哉 (Isoda Yuya)
日立製作所 研究開発グループ デジタルサービスプラットフォームイノベーションセンタ データマネジメント研究部 所属
現在，データ管理およびアプリケーション管理の研究開発に従事
情報処理学会会員

中野 和香子 (Nakano Wakako)
日立製作所 研究開発グループ デジタルサービスプラットフォームイノベーションセンタ エッジインテリジェンス研究部 所属
現在，コンピュータビジョンおよび分散AIの研究開発に従事
電子情報通信学会会員，情報処理学会会員

高瀬 誠由 (Takase Masayuki)
日立製作所 研究開発グループ デジタルサービスプラットフォームイノベーションセンタ エッジインテリジェンス研究部 所属
現在，ネットワークおよび分散AIの研究開発に従事
電子情報通信学会会員

ハイライト

生成AIをはじめとする深層学習ベースAIの構築とドメイン特化には膨大な業種特化データが不可欠であるが，このような特化データの収集は一般に困難である。このようなデータ確保の課題に対し，連合学習は一つの解決策を提示している。連合学習は，複数の組織に分散する機密性の高いデータを，機密を保護しつつ効果的に利用する技術である。

本稿では，連合学習の基本的な手法と日立の取り組みを解説するとともに，近年注目を集めている生成AI への適用可能性について述べる。

1. はじめに

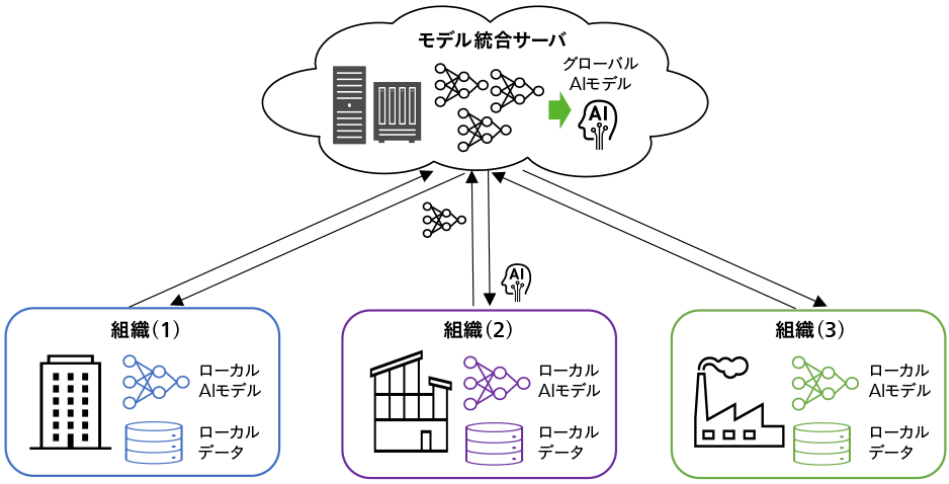
現代におけるAI（Artificial Intelligence）技術の飛躍的な進歩は，学習に用いる膨大なデータに依存している。そのため，ドメインや業種に特化した高度なAIサービスを実現するには，質の高い大量のドメイン／業種特化データの確保が，必要不可欠である。

連合学習[1)]は，このデータ確保に関する課題の一部を解決するため，複数の組織に分散したデータを効果的に利用する技術である。従来の方法でデータを一か所に集めるだけでは，データの機密性やプライバシーの問題が発生する可能性があった。しかし，連合学習を用いることで，データの機密を保護したまま，分散したデータをAIの学習に利用することが可能となる。本稿では，連合学習の基本的な方法とその課題を紹介し，さらに生成AIへの適用可能性について考察する。

2. 連合学習の概要

連合学習は，分散型のAI学習技術の一つであり，複数箇所に分散する学習データを物理的に集積することなく，すべてのデータを用いてAIが学習することを可能にする。連合学習の基本的な枠組みを図1に示す。

図1｜連合学習の概念図



注：略語説明　AI（Artificial Intelligence）

組織の持つ学習データ（ローカルデータ）を各組織に留め置いたままAIモデルを学習することで，データのプライバシーを保護しながら，複数の組織間で協調的にAIモデルを学習する。

連合学習におけるAI訓練プロセスは，以下のステップを繰り返すことで行われる。

(1) 各参加者は，自身が保有する学習データ（ローカルデータ）を用いて，参加者ごとのAIモデル（ローカルAIモデル）を学習する。

(2) 各参加者は，学習したローカルAIモデルをモデル統合サーバに送信する。この際，学習データ自体は共有されない。

(3) モデル統合サーバは，受信した各参加者のローカルAIモデルを統合し，すべてのローカルAIモデルの特徴を引き継いだAIモデル（グローバルAIモデル）を生成する。

(4) 生成されたグローバルAIモデルは，モデル統合サーバから各参加者に配布される。配布されたグローバルAIモデルは，各参加者の新たなローカルAIモデルとして利用される。

(5) 上記の1から4のプロセスを繰り返し，モデルをアップデートする。

この方法の特筆すべき点は，各参加者のローカルデータをサーバに集めることなく，すべての参加者のデータを反映したグローバルAIモデルを作成できることである。このようにして作られたグローバルAIモデルは，個々の参加者が保有するデータのみで学習したモデルよりも，高い性能を示すことが期待される。

3. 連合学習のユースケース

連合学習は，複数の組織が共通する課題を持ち，これを協力して解決したい場合に特に有効である。

連合学習は，複数組織間で知見を共有することで，参加組織全体の利益が増加する分野での応用が想定される。医療分野はその代表的な例である。複数の医療事業者が連携することで，患者の診察データを外部に出すことなく，高度な医療診断モデルを共同で構築することが可能となる。これにより，個々の医療機関では得られない規模のデータを活用しつつも，患者のプライバシーを保護することができる。

日立の取り組みの一つとして，代理店を介した営業手法が見られる損害保険業界において，異なる組織に分散して保存されている文章およびデータを連合学習で疑似的に統合することで，総合的な知見を引き出す事例がある。

他にも，金融業界における不正検知モデルの共同開発や，製造業における故障リスク診断ソリューションなど，さまざまな分野で連合学習の活用が期待されている。

4. 連合学習の技術的課題と日立の技術

4.1　組織間のデータ不均衡性とその解決

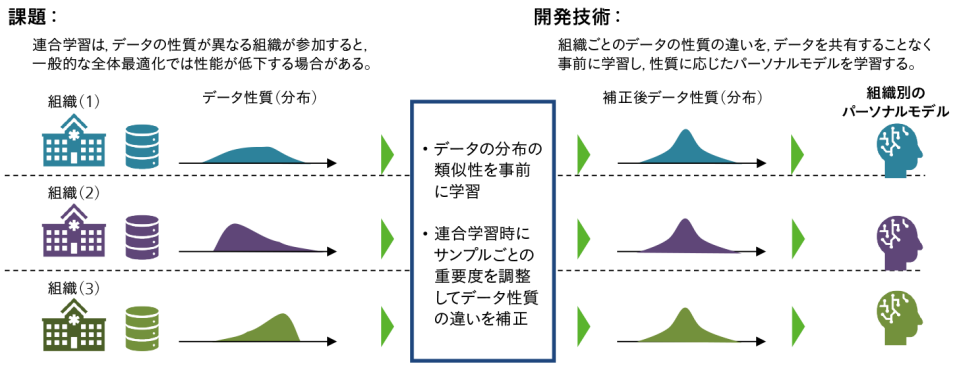
連合学習における大きな課題の一つがデータの不均衡性である。参加者間でデータの量や質に差がある場合，連合学習で得られるモデルの性能に偏りが生じることが知られている。この問題に対して，日立は参加者のデータの性質に応じたモデルを構築するパーソナライズド連合学習技術[2)]を開発した。この技術の特徴は，一般的な連合学習が単一のグローバルモデルを生成するのに対し，各参加者に応じたパーソナルモデルを生成する点にある。

技術の詳細を説明する。まず各参加者が保有するデータの類似度を，参加者間で直接データを開示せずに推定する。次に，推定した類似度の情報を用いて，各データサンプルの重要度（どれだけ学習に寄与させるか）を柔軟に変化させながら連合学習を行う（図2参照）。この方式により，参加者それぞれのデータと近い性質を持つデータを優先的に学習し，パーソナルモデルとして提供することが可能になる。

公開医療データセットを用いて，複数の病院のデータに基づくパーソナライズド連合学習によって生存予測モデルを生成し，評価を行った。その結果，各病院に分散したデータに一般的な連合学習を適用した場合と比較して，パーソナライズド連合学習によって生成されたモデルはより高い性能を得た。さらにすべての病院データを直接統合して機械学習する理想的な状態と比較しても遜色ない結果が得られた。

本技術の活用により，連合学習のすべての参加者に対して性能の高いモデルを安定して提供することが可能となる。このため，データの不均衡性が課題となるさまざまな分野での応用が期待される。

図2｜パーソナライズド連合学習の概念図



一般的な連合学習が単一のグローバルモデルを生成するのに対し，各参加者に応じたパーソナルモデルを学習することでユーザーに合わせたモデル生成を可能にする。

4.2　ネットワーク負荷の増大とその解決

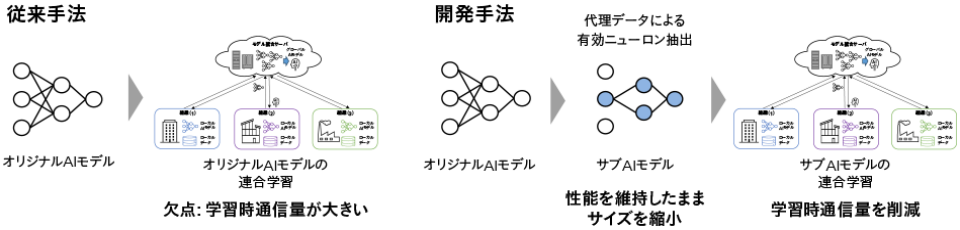
連合学習では，各参加者がモデル情報を統合サーバに転送し，かつ，統合サーバがモデル情報を各参加者に配信するため，学習時の通信量が増大しネットワーク負荷が大きくなるという技術課題がある。特に参加者数が多い場合や，モデルの規模が大きい場合に，この課題は顕著となる。これに対し日立では，宝くじ仮説（Lottery Ticket Hypothesis）※）と代理データセットを用いてオリジナルのAIモデルから小規模なサブAIモデルを抽出し，通信に用いることで，ネットワーク負荷の小さい連合学習を実現している[3)]。

連合学習ではモデル統合サーバが各組織の機密データを保有できないため，データ駆動型のサブAIモデル抽出法の適用が難しかった。日立は，組織外部に持ち出せない機密データに代えて，統合サーバで合成可能な代理データを用いてサブAIを抽出する方式を開発し，実際の機密データを用いる場合と遜色ない結果を得ることができた（図3参照）。

本技術の活用により連合学習時の通信負荷が削減されるため，モバイルネットワークを含むさまざまなネットワーク環境での応用が期待される。

※）密な深層ニューラルネットワークモデルには，「当選チケット」と呼ばれる，元のモデルと同等の性能を達成できる，よりパラメータ数の少ないサブモデルが含まれているという仮説。

図3｜代理データによるサブAIモデル抽出の概念図



ネットワークを介するモデル情報交換をサブAIモデルで代替することで，連合学習時の通信量を削減する。

5. 生成AIと連合学習の融合に向けて

5.1　連合学習による独自生成AI構築

自組織のナレッジを学習した独自の生成AIは，組織の業務フロー変革に大きな可能性をもたらす。独自生成AIを作成する一般的な手法として，ベースとなる生成AIモデルを基に，それぞれの利用者の目的に応じ，各自の保有するデータによって事後調整（ファインチューニング）を行う方法がある。事後調整に用いるデータが高品質かつ大量であれば，より専門的な内容に特化した有用な独自生成AIが得られる可能性が高まる。しかしながら，一般的に事後調整に用いる訓練データは組織の機密データである場合が多く，容易に他の組織と共有することができない場合が多い。

ここで，連合学習の知見を生成AIの事後調整に活用することで，訓練データを秘匿したまま生成AIに組み込むことが可能となる。このような共同ファインチューニングにより，AIモデルが一般的な知識に加え，各組織の専門性や経験を反映した高度な生成能力を獲得することが期待される。

5.2　生成AIエージェントと連合学習の連携

生成AIエージェントとは，生成AIに対して，状況に応じて適切なツールを自律的に活用して問題を解決する能力を付加したものである。この生成AIエージェントが呼び出すツールの一つとして，連合学習により獲得した予測型AIモデルを設定することで，生成AIの持つ優れた一般推論能力や対人インタフェース能力と，業務データに特化した予測型AIモデルの予測能力を分離しながら，その双方を効果的に活用することが可能となる。

例えば，金融アドバイザリーサービスにおける活用が考えられる。生成AIベースの対話エージェントが顧客とのコミュニケーションを担当し，具体的な金融予測や分析が必要な場合には連合学習で構築された高度な金融予測モデルを呼び出すことで，顧客対応の柔軟性と専門的な分析能力を両立させることができる。

さらにこうした連携により，連合学習モデルの出力を自然言語で説明することや，生成AIの推論プロセスに連合学習モデルの予測結果を組み込むことも可能となる。これにより，ユーザーにとってより理解しやすく，信頼性の高いAIシステムを構築することができると期待される。

6. おわりに

本稿では，連合学習の概要と特徴を解説し，その課題と，生成AIへの適用可能性について考察した。連合学習は，データの機密性を保ちながら分散データを効果的に利用できる革新的な技術であり，AIのさらなる発展に大きく寄与する可能性を秘めている。

参考文献など

1） H. B. McMahan et al., ”Communication-efficient learning of deep networks from decentralized data,” Proc. of AISTATS 2017, pp. 1273-1282（2017）

2） S. Tarumi et al., ”Personalized Federated Learning for Institutional Prediction Model using Electronic Health Records: A Covariate Adjustment Approach,” Annu Int Conf IEEE Eng Med Biol Soc. 2023 Jul:2023:1-4（2023.7）

3） T. Tanimura et al., “Compressing Model before Federated Learning by Transferrable Surrogate Lottery Ticket,” IEEE Consumer Communications & Networking Conference （CCNC） 2023, pp. 620-623（2023.1）

日立評論

日立評論は、イノベーションを通じて社会課題に応える
日立グループの取り組みを紹介する技術情報メディアです。

日立評論Webサイトでは、日立の技術者・研究者自身の執筆による論文や、
対談やインタビューなどの企画記事、バックナンバーを掲載しています。ぜひご覧ください。

日立評論(日本語) Webサイト

https://www.hitachihyoron.com/jp/

Hitachi Review(英語) Webサイト

https://www.hitachihyoron.com/rev/

日立評論メールマガジン

Webサイトにてメールマガジンに登録いただきますと、
記事の公開をはじめ日立評論に関する最新情報をお届けします。