Focus on the News

暗号化したままディープラーニングの標準的な学習処理ができる秘密計算技術を世界で初めて実現

NTTは，データを暗号化したまま一度も元データに戻さずに，ソフトマックス関数やAdam（adaptive moment estimation）と呼ばれる最適化処理を含む標準的なディープラーニングの学習処理を行う技術を，世界で初めて実現しました．

通常，データを利活用するためには，通信時や保管時に暗号化していたとしても，処理を行う際には元データに戻して処理する必要があります．このことは，データ所有者からすると情報漏洩のリスクを感じることから，企業秘密や個人のプライバシにかかわるデータの利活用に抵抗感を持つユーザや組織が少なくありません．特に所有者から他者，または同一組織内であっても，データ提供して積極的に利活用したい場合には，このことは大きな障害だと考えられます．

今回開発した技術を用いることで，企業秘密や個人のプライバシにかかわるデータをディープラーニングで活用する際に，サーバではデータを暗号化したまま一度も元データに戻さずに処理することが可能となります．つまり，ディープラーニングでのデータ活用に必要な①データ提供，②データの保管，③学習処理，④予測処理，のすべてのステップを暗号化した状態で行えます．サーバでは常にデータは暗号化されたままであり一度も元データに戻すことがないため，従来よりもユーザや組織が安心してデータを提供でき，学習に利用できるデータ量や種類が増え，精度の高いAIの実現が可能になると考えています．

■背　景

昨今，さまざまな分野のデジタルトランスフォーメーションが進んでおり，分野横断的なデータの蓄積やAIなどの高度分析がイノベーションを促進し，さまざまな産業・サービスが発展すると期待されています．その一方で，情報漏洩や不正利用の懸念がデータの提供・利活用促進を阻害する要因となっています．

NTTはそのような要因の解消に貢献するため，データを暗号化したまま一度も元データに戻さずに処理ができる秘密計算技術の研究開発を世界に先駆けて取り組んできました．NTTが取り組む秘密計算技術はISO国際標準である秘密分散技術を利用して暗号化されたデータを，一度も元データに戻さずに分析を行うため，企業の秘密情報や個人のプライバシにかかわる情報などの情報を安全に，安心して提供し利活用できる社会の実現に貢献すると期待されています．

今回，NTTはAIの中でも活用が進み始めているディープラーニングの標準的なアルゴリズムについて，暗号化したまま一度も元データに戻さずに処理できる技術を世界で初めて実現しました．この技術によって，企業秘密や個人のプライバシにかかわるさまざまな情報を活用する際，データ所有者からのデータ提供の安心感を高めることができ，データの量や種類の増加や，これに伴う精度向上・高度分析の実現につながると考えます．例えば，個人の位置情報やスケジュールを暗号化したまま，天気や企業のイベント情報などと併せて学習することで，最適な飲食店の仕入れや人員リソースの配備を予測することが考えられます．また，さらにこの技術を応用すれば，レントゲン写真，MRI，CTスキャン，顕微鏡写真などの医療データを秘匿しつつ学習し，検査結果に悪性腫瘍があるかなどを高速かつ精度良く判定することが可能になると期待されます．

■技術のポイント・特徴

秘密計算ではデータを暗号化したまま一度も元データに戻すことなく処理を行うため，その処理方法は通常の処理方法とは大きく異なります．そのため，暗号化していないデータでは簡単に処理できても秘密計算では実現が難しいという，不得手な処理がありました．秘密計算でこのような処理を行うには新しい技術や工夫が必要となります．

一般にディープラーニングの学習では，訓練データを入力として，複数の層を順番に処理して学習の途中結果を得て，その途中結果が十分に学習されたものであればそれを最終結果として出力し，そうでなければ途中結果を更新する処理（最適化処理）を行い，もう一度最初の層の処理から一連の処理を繰り返します．複数の層のうち最後の層は出力層と呼ばれ，標準的にはソフトマックス関数と呼ばれる数式を計算します．また，最適化処理ではSGD（Stochastic Gradient Descent）が原始的な手

法として知られていますが，繰り返しの回数が多いため，SGDを改善したAdamなどが主に用いられています．これらディープラーニングの標準的な学習処理で用いられるソフトマックス関数やAdamでは，割り算，指数，逆数，平方根を組み合わせた処理を行います（**図**）．

ディープラーニングの学習を秘密計算で行うとき，従来技術では上にあげた割り算，指数，逆数，平方根が秘密計算にとって不得手な処理なため，秘密計算でソフトマックス関数やAdamを計算することは困難でした．そのため，多くの先行研究は学習・予測のうちこれらを計算しなくても良い予測のみに絞ったものでした．また，いくつかは学習に取り組んだ先行研究もあるものの，ソフトマックス関数を非常に粗い精度で近似し，かつ最適化処理は原始的なSGDしか使用できませんでした．

今回，NTTは従来では計算が困難であったソフトマックス関数を高速かつ精度良く計算し，さらに主要な最適化処理であるAdamを利用できる秘密計算技術を開発しました．実現方法には２つの異なるアプローチがあり，それぞれの開発を行いました．１つはソフトマックス関数やAdamを計算するために，あらかじめ入出力の組を並べた対応表を用意し，入力と対応表を暗号化しつつ，入力に対応する出力が得られる秘匿写像と呼ばれる独自技術を利用するアプローチです．もう１つはソフトマックス関数やAdamを構成する割り算，指数，逆数，平方根それぞれについて専用の高速アルゴリズムを開発するアプローチです．

この技術を用いることで，データを暗号化したまま，標準的なディープラーニングのアルゴリズムを用いて学習することが可能となりました．例えば今回開発した割り算・指数・逆数・平方根の専用アルゴリズムを用いる行う方法では，６万件の手書き文字を判別するモデル学習において，１エポックの学習を２分程度で実行することができます．

■今後の展開

今後はAIの知見を持つパートナーと連携して実証実験等を行うことで，秘密計算を使ったディープラーニングの効果を実証していきたいと考えています．最終的には誰もが安心してデータの提供と利活用ができる環境を

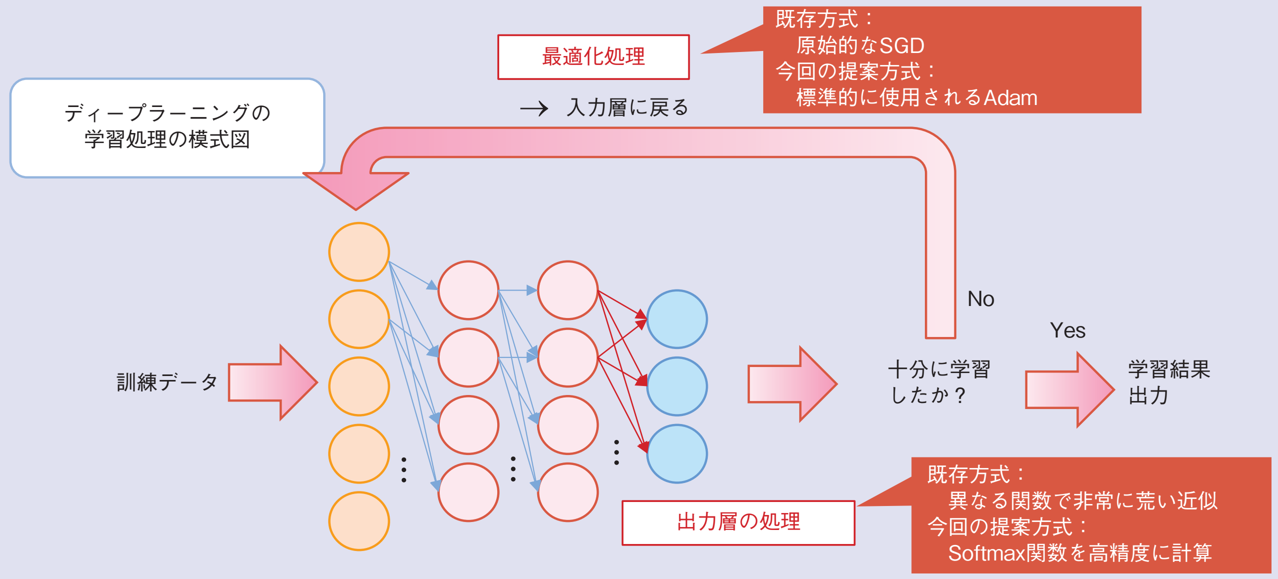

図　ディープラーニングの学習の模式図

提供していきたいと考えています．

◆問い合わせ先

NTTサービスイノベーション総合研究所
企画部広報担当
TEL　046-859-2032
E-mail　radnd-ml@hco.ntt.co.jp
URL　https://www.ntt.co.jp/news2019/1909/190902a.html

研究者紹介

データを暗号化したままAI分析を行う「秘密計算AI」の研究

三品　気吹
NTTセキュアプラットフォーム研究所
データセキュリティプロジェクト

私は昨年度NTTに入社し，「あらゆるデータ分析を秘密計算でもできるようにする」という野望を持ちながら，「秘密計算AI」の研究に取り組んでいます．

学生時代は生命情報科学を専門とし，ゲノム解析などに興味を持つうちに，究極の個人情報ともいわれるゲノムを安全に分析する技術として秘密計算に出会いました．もともと暗号理論は全くの専門外でしたが，「暗号化したまま計算ができるって，どういう仕組みだろう？」という好奇心や，「秘密計算でいろいろな分析をできるようにしたい！」という探求心の赴くまま，秘密計算で世界トップ技術を持つNTTの門をくぐり，現在に至ります．

秘密計算AIの研究は，無数に存在するAI手法のアルゴリズムを理解し，１つずつ秘密計算用アルゴリズムを設計・実装するという大掛かりなテーマです．通常のAIでは，すでに便利なライブラリが多数ありますが，秘密計算では四則演算など基本的な関数を組み合せて実装しており，そこが苦労する部分であると同時に，「私が秘密計算AIライブラリを，世界で初めてつくるんだ！」と感じるやりがいでもあります．

まだ秘密計算AIの研究は始まったばかりですが，今回のように世界初の成果を出すことができ，着々と手ごたえを感じています．誰もが秘密計算によって今よりもっと自由に，そして安心してデータ活用できる社会の実現をめざして，これからも研究に取り組んでいきます．