

## マルチエージェントによる Deep Learning の提案

## Deep Learning by Multi Agent System

黒滝 紘生      中山 浩太郎      松尾 豊  
Hiroki Kurotaki      Kotaro Nakayama      Yutaka Matsuo

東京大学工学系研究科 技術経営戦略学専攻

The Department of Technology Management for Innovation, Graduate School of Engineering, The University of Tokyo

We reconstructed Stacked Denoising Autoencoder, a novel Deep Learning model, with the methodology of multi-agent system. We redesigned behaviors of each neuron as autonomous agents. In our model, each agent represents one feature and has locally limited views of other agent's information, thus the computation is principally done without explicit global level manipulations. This brand new basal model, which is more similar to human brain, can help establishing connection between Deep Learning and multi-agent, and developing systems that has more representational ability with its potential to expansion of many aspects. We theoretically proved the equivalence of the ordinary SDA and the proposed multi-agent system. We also verified this equivalence empirically with XOR toy dataset and MNIST image recognition. In addition, we proposed Sparse Connect SDA, a variant of SDA whose connections are pruned randomly, and tested its behavior with the datasets.

## 1. 概要

機械学習の一分野である深層学習は、近年目覚ましい発展を遂げている。特に画像・音声・文書などの認識タスクで、大きな性能向上をもたらしている。従来の深層学習は、ユニット間の接続を行列で実装した、大域的なアルゴリズムによって記述されてきた。しかしこの方法では、モデル構造における制約が大きく、特に、モデルを構成する個々のユニットを、アルゴリズムの操作対象にすることは難しかった。また、ユニットを徐々に増やしたり、つなぎ替えるようなアプローチにも、制限があった。さらに、行列を用いることで、メモリの消費量や、学習・利用時の計算量が増大してしまっていた。

本研究では、ユニットを自律したエージェントと捉え、マルチエージェントシステムの手法を取り入れて、深層学習のアルゴリズムを再構成することで、この問題点を解決した。具体的には、ニューラルネットワークを、行列として捉えるのではなく、独立したエージェントの通信で再定義した。このとき、ノード同士のローカルな通信のみで、深層学習の代表的手法の一つである、積層雑音除去自己符号化器 (Stacked Denoising Autoencoder, 以下 SDA) [Vincent 08] が実装できることを示した。また、このマルチエージェントによる SDA 実装が、従来の SDA と等価な計算を行っていることを、理論・実験の両面から示した。

これらのエージェントとしてのユニットは、アルゴリズムを変更することなく、自由に追加・削除できるようになっている。これによって、ユニットを徐々に増やしたり、つなぎ替えるようなアプローチによる、新たなニューラルネットワークの構成を行うための基礎を築くことができた。このマルチエージェントシステムの応用として、SDA から自由にエッジを削除できて、メモリ効率や実行効率に優れたモデルを構築できることが挙げられる。そこで、エッジを削除したネットワークの基礎的な振る舞いを調べるため、筆者らは、SDA からランダムにエッジを削除して得られる、“Sparse Connect SDA”を提案して、その挙動を調べた。これらの実験は、人工的な XOR

連絡先: 黒滝 紘生, 東京大学工学系研究科, 〒 113-0033 東京都文京区本郷 7-3-1 工学部 2 号館, kurotaki@weblab.t.u-tokyo.ac.jp

データセット、および、MNIST 手書き文字認識データセットにて行われた。

深層ニューラルネットとマルチエージェントシステムを合わせた研究は、筆者の知る限り先例がない。今後、提案手法を推し進めることで、将来的に、行列をベースにした大域的なモデルでは実現できない、ユニットの混合方法や、より自由なネットワーク構造が、できるようになり、深層学習の発展の上で、大きな貢献をもたらすと考えている。また将来的には、マルチエージェントの特性である、局所的な記述のみで複雑なシステムを組める点を活かして、複雑化する深層学習モデルを、より少ない知識で扱えるような枠組みの開発も目指していきたいと考えている。

## 2. 背景

深層学習は、人間の脳にある神経細胞ネットワークを、近似的に数式モデルへ落とし込んだ、ニューラルネットワークと呼ばれる学習モデルの一種である。深層学習モデルは、画像認識・音声認識・自然言語処理などの分野で、必要不可欠な技術の一つになっている。深層学習を応用した例には、画像や動画の説明文の自動生成、化合物結合などのグラフネットワーク予測、ゲームの自動プレイ、プログラムの解釈と実行など、他にも様々な応用ソフトウェアが提案されており、その中には、自動運転やロボットの自動学習、あるいは IBM の質問応答システム Watson のように、実社会での応用にむけた研究が進んでいる分野もある。これらの応用ソフトウェアが実現されれば、産業界に与えるインパクトも大きいと考えられる。

従来の深層学習のアルゴリズムでは、ユニットをレイヤーと呼ばれる単位でグループ化することが事実上の標準になっていた。例えば、SDA では、複数のレイヤーユニットによる出力を、1次元のベクトルとして一まとめに扱っている。しかし、本来ニューラルネットワークに、「レイヤーでユニットを区切り、有限次元の空間に格子状に並べる」といった制約は存在しない。この制約は、人間の脳で言えば、神経細胞がいくつかの層に分かれて一列に並んでいるのに相当する。現在の深層学習計算モデルにおいて、レイヤーを用いた手法が中心となっているのは、計算を行うハードウェアの都合が大きい。具体的

は、GPU やマルチコア CPU などの並列計算ハードウェアを最大限に使うため、また CUDA や BLAS といった並列演算ソフトウェアの恩恵を受けるために、計算モデルの大部分を行列積で済ませている。しかし、実際の人間の脳では、神経細胞は網目状のネットワークを構成している。モデルを行列の形に単純化したことで、モデルの表現能力を制限してしまっていると考えられる。レイヤーに囚われない、より自由で柔軟なネットワーク構造を得るための仕組みが求められている。

一方、マルチエージェントシステムは、自律動作する多数の知的なエージェントの集まりとして、システムを構成して、複雑な問題を扱う方法である。エージェントが満たすべき要件については、様々なバリエーションがあるが、一例として、

- 自律した動作
- 局所的な情報による動作
- 通信による協調、分散知能

などが挙げられる。マルチエージェントシステムは、多数の意思決定者がいるような環境をシミュレートする、実験要素の強い経済学のような分野でも使われている。マルチエージェントシステムは、複雑な環境を複雑なままモデリングする手法と考えられ、単一のシステムやアルゴリズムでは記述が難しい問題も、複数の独立システムと、その通信という形で捉えることによって、容易に定式化・実装できるようになることが特徴の一つである。この特徴は、個別のユニットの独立性を高めるという今回の課題に適合している。加えて、マルチエージェントシステムの特長は、深層学習で近年課題となっている、複数種データを取り扱うマルチモーダル学習の課題にも適していると考えている。

以上を踏まえて、「深層学習の発展のため、マルチエージェントシステムの手法と知見を取り入れたい」と、本研究の目的はまとめられる。深層学習のアルゴリズムを、マルチエージェントの手法で再構成することで、従来のレイヤーベースの深層学習に比べて、より自由なネットワークやアルゴリズムを構築するための、新たな基礎を構築できると考えている。

### 3. 関連研究

ここではまず、本研究の提案手法がベースにしている。深層学習の SDA モデルについて述べる。その後、ユニットの自律性や、ユニット間の柔軟な接続関係を重視したり、学習中に形状を探索できるような、ニューラルネットワークの先行研究について述べる。

#### 3.1 深層学習と Stacked Denoising Autoencoder

深層学習は、ニューラルネットワークのうち、普通隠れレイヤーが 2 つ以上のものをいう。深層学習では、従来よりも多くのレイヤーを扱うことで、より複雑な関数を学習できるようになったと考えられている。2006 年に Hinton らによって、レイヤーの多いニューラルネットワークを効率よく学習させる方法が発見され、発展への道が開けた。本研究で扱う Stacked Denoising Autoencoder(SDA) は、2008 年に発表された [1]。基本的なアイデアは、各レイヤーの重みをあらかじめ教師無し学習で調整するところにある。これにより、隠れ層が効率の良い素性を学習できるようになった。この教師無し学習のことを、事前学習 (以下 pretraining) と呼び、全体をつなげて学習する段階の方は、finetuning と呼ぶ。SDA では、pretraining 段階で、雑音自己符号化器 (Denoising Autoencoder、以下 DA) というモデルによる教師無し学習をさせている。DA では、入力

データの一部を隠し、それを再構成する課題を与える。この過程で、入力=再構成データをよく表現するような特徴量が学習されている、と考えられている。SDA は、DA によって得られた特徴量を積み重ねることによって、複数層の隠れレイヤーの学習に成功しており、深層学習の代表的モデルの一つとなっている。本研究でも、この SDA をベースにして、SDA と等価な計算を行えるようなアルゴリズムを提案している。

#### 3.2 学習中に変化するネットワーク

この SDA も含めて、現在深層学習で広く使われているニューラルネットワークでは、大半の場合、ユニット間の接続構成や形状は固定されている。接続するエッジの重みを変更することで、擬似的にユニットのネットワーク構成を変更していると捉えることもできるが、その全体的な形が、行列状に制約されていることは変わらない。ネットワーク形状も探索することで、より様々なモデルを学習できると考えられるが、深層学習と結びついた形で行う方法は確立していない。

ネットワーク内でユニットをグループ化する方法として、[Warde-Farley 14] では、学習中にネットワークの隠れユニットのグループ化を進める手法が提案されている。しかし隠れレイヤーの数を学習することはできず、隠れレイヤー数は一層に限定されている。

Neuro Evolution と呼ばれる一群のアプローチでは、ネットワークの形状を初めから固定せず、進化的アルゴリズム (EA) によってニューラルネットの構造を探索している。HyperNEAT[Stanley 09] は、Neuro Evolution の代表的手法の一つであり、ネットワークの重みを出力する高次のネットワークを用いることで、効率良い学習を行っている。しかし、深層学習のネットワークに近い構成が得られるよう促したり、深層学習に組み入れるような手法は、確立していない。

枝刈り (Pruning) と呼ばれる方法では、ネットワークを構成するエッジやユニットを、一定の条件に従って削除している。その方法としては、1 次/2 次微分の情報による感度分析が代表的である。本研究では、この手法を応用したアプローチを前提に、基本的な性質を実験で明らかにしている。

接続制約としてのレイヤーを考えたとき、生物の脳の神経回路において、ニューロンの接続は、必ずしも厳密なレイヤー構造を取っているわけではない。そこで、現在のニューラルネットよりも、より脳の神経回路に近いニューロンやシナプスのモデルを用いた、スパイクニューラルネット (Spiking Neural Net) と呼ばれる研究も盛んになされている。しかし、これらの研究は、より脳に近いシミュレーションや、特殊なハードウェアでの実装による低電力消費などが主目的となっており、深層学習モデルを拡張するという本研究の目的にはそぐわない。深層学習の学習パラメータ数を削減するための手法は、現在も盛んに研究が行われている分野である。本研究で実験した、エッジ削減ネットワークも、この研究の一種だと考えている。学習パラメータ数を削減することで、より多くの回数の学習が出来ると共に、学習後のモデルを表現するのに必要なメモリサイズも減少するため、スマートフォンや家電の組み込み環境のような、メモリが十分に搭載できない環境でも、深層学習モデルを利用できる可能性が高まると考えられる。一方で、一般的にはモデルのパラメータ数が減少すれば、モデルの表現力も低くなるため、必要な学習性能を維持できることが重要である。深層学習との関連では、重み行列の一部を共有する、ハッシュ (Hashing) と呼ばれる手法や、重み行列のパラメータの有効桁数を量子化や 2 値化で減少させる手法も研究されている。

## 4. マルチエージェントによる深層学習

ここでは、マルチエージェントの手法により、深層学習を構成するためのアルゴリズムについて、概要を述べる。アルゴリズムや証明の詳細は、紙面の都合上割愛する。

### 4.1 システムの概要

提案手法がマルチエージェントシステムであるために満たすべき条件として、本研究では、

- システムは、一つの環境と、複数のユニットから成り立つ。
- 各ユニットは、独立して行動を行っている。
- 各ユニットは、局所的な情報しか持たない。

という3つの条件を設定した。これらは、深層学習の再構成という本研究の目的から、考えられたものである。

マルチエージェント (MA) 版アルゴリズムの系は、ユニットと環境から構成される。ユニットは、ノードとエッジから成っている。エッジは、ノードとノードの間を結んでいる。

ノードは、エッジでつながった他のノードと、メッセージの送受信によって、情報をやり取りできる。この情報には、変数が示す値や、変数に蓄積された誤差が含まれる。またノードは環境からデータを受け取ることがある。

各ユニットのノードは、メッセージかデータを受け取ったときのみ、これに反応して動作する (イベントドリブン)。このとき、あらかじめ定められたアルゴリズムを始動して、変数を更新したり、自らもメッセージを送信できる。逆に、データやメッセージを受け取るまでは、原則として、演算や別のメッセージ送信を行わない。

ノードがもつ変数は、データや特徴量の値そのものに限られない。ノードは、隣接するエッジの重みパラメータや、特徴量の計算に必要な中間変数なども含むことができる。これらの変数も、更新の対象となる。

以上の条件を元に、筆者はマルチエージェント深層学習モデルを構成する、各種類のユニットについて、具体的なアルゴリズムを書き下した。それらが通常のマルチエージェントでないネットワークのアルゴリズムと、数式レベルで一致することを証明した。具体的には、各数式によって変数が更新される値と、その更新順序が、通常版と MA 版とで一対一対応していることを示した。この証明は、Autoencoder の前身となる、隠れレイヤー1層のネットワーク (MLP) から始まり、単層の Autoencoder、Denoising Autoencoder、多層の SDA と進んだ。

### 4.2 Sparse Connect SDA

筆者はさらに、得られた MA 版の SDA を拡張することで、学習時にネットワークから不要なエッジを削除できるようなアルゴリズムのネットワークを組むことができることを示した。具体的には、一定の接続率を定め、各レイヤー間のエッジ接続を確率的に削減したネットワークを生成した。本マルチエージェントシステムでは、学習中 (オンライン) にエッジを削減することも可能だと考えられるが、今回はエッジを削減した場合の基本的な挙動を調べるのが目的のため、より実験時の解釈が容易な実験設定として、レイヤー・ユニット接続生成時に一括してエッジを削減することにした。このモデルは今後の研究にて様々なノード・エッジの追加・削減パターンを与える上で、基礎的な意味合いを持っている。このアルゴリズムでは、マルチエージェントシステムによって深層学習を再構成したことによって得られた特徴である、各ユニットが独立して動作し

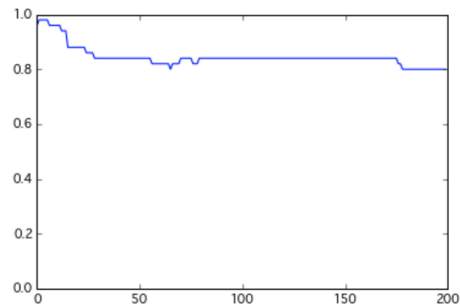


図 1: 平均識別エラー率の比較 (SDA、MNIST、通常版)

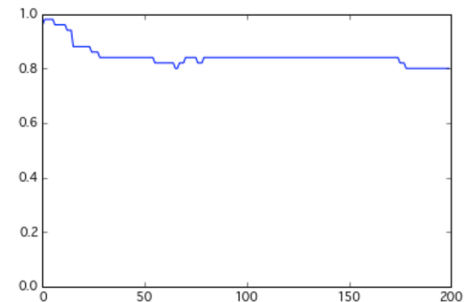


図 2: 平均識別エラー率の比較 (SDA、MNIST、MA 版)

ており、アルゴリズムの全体を書き換える必要なく、それぞれのユニットを操作でき、特にユニットの数を増減できるような性質を、活用している。このことによって、深層学習のアルゴリズムに、エッジ削減の手法をより自然な形で取り入れ、また実装することができている。例えば、通常のニューラルネットワークモデルでは、行列のサイズを学習中に変更したり、あるいはメモリや計算時間を消費して、不必要に大きな行列をあらかじめ確保しておく必要があるが、マルチエージェントシステムでは、必要なときに必要な分だけのエッジ接続を生成することができる。本マルチエージェントシステムによるエッジ削減方法では、エッジを削減した分だけ学習時間も削減されると予想できる。このことは、実験の章で確認する。

## 5. 実験と考察

筆者は、マルチエージェントによる深層学習のアルゴリズムを実装して、その性能を確認した。特に、従来の深層学習の実装と比べて、提案手法によって再構成された深層学習の実装が、数式上だけでなく、実験的にも同等の性能を示していることを示した。

さらに、3章で述べた、エッジを確率的に削減するネットワークを実装して、接続率を変化させたときに、学習性能や学習時間がどのように変化するかを調べた。

データセットには、人工的な XOR 関数のデータセットと、MNIST 手書き文字認識データセットの一部を用いた。

### 5.1 マルチエージェント版と通常版の比較

図 1,2 は、通常版と MA 版の SDA モデルによる、MNIST データセットに対する識別エラー率の比較で、5回の実験の平均を取っている。縦軸がエラー率、横軸が学習の進行度 (エポック数) である。2つのグラフがほぼ同じ曲線を描いていることがわかる。なお、MLP モデル、XOR データセットの組み合わせでも同様の実験を行っており、pretraining や fine tuning

表 1: 重み行列の誤差の比較 (MLP、MNIST)

グループ	絶対誤差の最大値
入力→隠れレイヤー行列	$1.08 \times 10^{-19}$
隠れレイヤーバイアス	0.0
隠れ→出力レイヤー行列	$8.7 \times 10^{-19}$
出力レイヤーバイアス	0.0

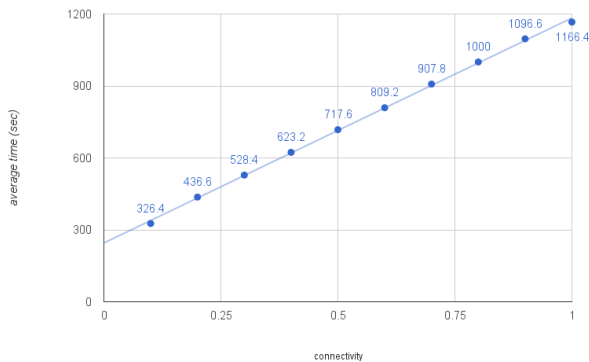


図 3: 接続率変更による平均実験時間の変化 (MNIST)

時の各訓練コストについても比較を行っている。どの組み合わせでも、同様にグラフの形状が等しくなっている。このことから、理論面だけでなく、実験的にも、2つのモデルが同じ内容を計算していることが示唆される。

このことをさらに詳しく確かめるため、通常版と MA 版とで、データを 1 つ与えたときの、各対応するエッジ重みの変化を比較して、最大でどの程度の誤差が生じているか調べた (表 1)。この誤差は、Python 実装により生じる一般的な計算誤差よりも 3 桁以上小さい。今回実装した MA 版と通常版の誤差は、深層学習の上では無視できる程度の差であると考えている。

## 5.2 Sparse Connect SDA の性質

我々は、Sparse Connect SDA の接続確率を変化させて、MNIST データセットにおける、最終テストエラーおよび実験時間を比較した。

図 3 では、縦軸が学習の実験時間、横軸が接続率を表している。接続率が小さいほど、実験時間が短くなっていることがわかる。これは、学習に必要なパラメータを削減したことによる効果だと考えられる。

図 4 では、縦軸が学習完了後の最終テストエラー率の平均、横軸が接続率を表している。接続率を下げても、エラー率が上昇 (悪化) しないケースがあるとわかる。この結果は前章での予想とは異なっている。理由は調査中だが、一つの仮説として、「接続率が高い複雑なモデルほど、学習が困難である」と考えている。また、エッジの削減が、疎なネットワークモデルの学習を促したとも考えられる。

## 6. まとめ

本研究では、マルチエージェントシステムの考えを深層学習に取り入れて、深層学習のアルゴリズムを再構成する、マルチエージェント深層学習を提案した。特に、深層学習の代表的

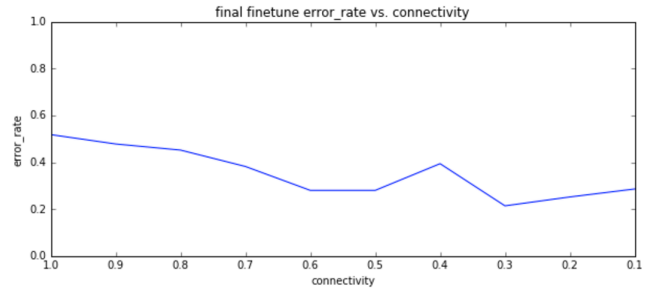


図 4: 接続率変更による平均エラー率の変化 (MNIST)

なモデルの一つである、積層雑音除去自己符号化器 (Stacked Denoising Autoencoder, SDA) を、提案手法によって再構成できて、通常版の SDA と同等のアルゴリズムを計算できることを示した。深層学習をマルチエージェント化することにより、深層ニューラルネットの長期的な課題である、ネットワーク構成やアルゴリズムにおける自由度の低さを、解決するための基礎を築くことができたと考えている。

さらに、提案手法によるマルチエージェント版の SDA を拡張して、エッジを削除した場合のネットワークを実装できることを示した。最も簡単なエッジ削除パターンとして、ランダムに削除する Sparse Connect SDA を提案して、その挙動を調べた。このとき、最終的な識別性能を大きく落とさず、学習や識別にかかる時間を削減できる可能性を示した。この新しいネットワークは、マルチエージェント版深層学習を応用することで、従来よりも自由なネットワークを構成できる一つの例となっている。

今後の研究課題としては、より純粋な分散エージェント知能として、モデルを洗練していくと共に、様々なユニットを混合したネットワークの構成などが考えられる。さらに、将来的には、マルチエージェントの記述の特性を活かして、誰でも深層学習を扱えるような枠組みの整備を目指している。

## 参考文献

- [Vincent 08] P. Vincent, H. Larochelle, Y. Bengio, and P. A. Manzagol. Extracting and Composing Robust Features with Denoising Autoencoders. In Proceedings of the 25th International Conference on Machine Learning, pp. 1096-1103, 2008
- [Warde-Farley 14] D. Warde-Farley, A. Rabinovich, and D. Anguelov. Self-informed Neural Network Structure Learning. arXiv:1412.6563 [cs, stat], 2014.
- [Stanley 09] K. O. Stanley, D. B. D' Ambrosio, and J. Gauci. A Hypercube-based Encoding for Evolving Large-scale Neural Networks. Artificial Life, Vol. 15, No. 2, pp. 185-212, Apr 2009.