

令和 3 年 5 月 1 日現在

機関番号：12701

研究種目：基盤研究(B)（一般）

研究期間：2018～2020

課題番号：18H03305

研究課題名（和文）浸透学習法とその応用に関する研究

研究課題名（英文）Percolative Learning and its applications

研究代表者

長尾 智晴（Nagao, Tomoharu）

横浜国立大学・大学院環境情報研究院・教授

研究者番号：10180457

交付決定額（研究期間全体）：（直接経費） 13,500,000円

研究成果の概要（和文）：本研究では、申請者らが先に開発してPCT出願した独自の深層回路である浸透学習法（percolative learning method）と呼ぶ階層型神経回路網の学習手法の基礎理論、実現方法と複数分野への応用について研究を行なった。浸透学習は学習時のみ利用できる入力情報（Aux）を、学習時・運用時の両方で利用できる入力情報（Main）に“浸透”させることで、Mainだけを学習・運用に用いた場合より高精度の認識・分類を実現する手法である。応用例として時系列予測とソフトセンサを扱い、浸透学習が従来手法より高い精度を実現することを確認し、浸透学習が様々な問題に適用できることを示した。

研究成果の学術的意義や社会的意義

これまでの神経回路網や深層学習では、学習時のみ利用できるデータは、学習しても実際の運用の際に使いえなくなるので、結局利用されてこなかった。これに対して、我々が開発した浸透学習を使うことで、そのようなデータを有効に活用することができるようになった。本事業において浸透学習の理論・方法・応用について具体的な検討を行い、浸透学習が有効であることを示すことができた。今後、様々な分野で浸透学習を利用することが考えられ、その効果と波及効果は大きいと考えられる。

研究成果の概要（英文）：We previously developed "PLM: Percolative Learning Method" which is a kind of learning method for layered deep neural networks. In this project, we studied theory, methods and applications of PLM. In PLM, we can "percolate" Aux data which are used only for learning into Main data which are used for learning and testing. We proved that PLM could achieve the precision rate higher than a conventional deep neural network experimentally. We dealt with time dependence data prediction and realization of software-sensor, and we showed that PLM is effective for various fields.

研究分野：知能情報学

キーワード：機械学習 深層学習 ニューラルネットワーク 進化計算法

1. 研究開始当初の背景

現在、機械学習の分野で注目されている深層学習(ディープラーニング)は、階層型ニューラルネットワークモデルである。深層学習、およびその基礎となっている階層型ニューラルネットワークでは、仮に、学習時だけに使える情報があっても、その情報が運用時には使えないとすれば、学習に用いることができなかつた。これはある意味では当然のことであつた。これに対して研究代表者は、ニューラルネットワークの入力情報を、Aux(学習時は使えるが、運用時は使えなくなる情報)と Main(学習時・運用時の両方で使える情報)に分けて考え、学習時に Aux データを利用することで、学習・運用ともに Main データだけを利用する従来型のニューラルネットワークより高精度の出力を行うことができる浸透学習法(PLM:Percolative Learning Method)を先に開発して PCT 出願した((国際出願番号 PCT/JP2018/028633)。本研究では、この浸透学習法(以下、PLM と略記)の理論・方法論・応用について研究を行なつた。このような機能をもつたニューラルネットワークは他に存在しない。なお、PLM の原理は PCT 出願済みで実施例も既にあつたが、PLM の理論・方法論・応用についての検討が充分とは言えない状況であり、研究を進める必要があつた。

2. 研究の目的

上記を背景にして、本研究では、PLM を汎用性のある深層学習法として確立することを目的として、PLM の理論・方法論・応用について研究を行なつた。

3. 研究の方法

はじめに、本研究のベースとなる PLM について簡単に示す。図 1 に PLM の基本構造を示す。まず始めに、回路の最終出力である「認識結果」を出力として、通常確率的勾配降下法によって回路網の全ての結合荷重を決定する。次に、Aux からの影響である「浸透データ」の部分の出力が変わらないことを条件として、Aux からのピンク色の信号線(点線部分)の結合荷重を強制的に減らしつつ、Main からの赤い信号線の結合荷重を最適化する。この際、青い信号線の結合荷重は変化させない。これによって、最終的にピンク色の信号線を削除し、Aux からの信号が入力されない回路を作る。これを運用に用いる際は、Main データだけを利用するが、Aux からの影響が自然に考慮された出力を行うことができるようになる。これが PLM の原理である。

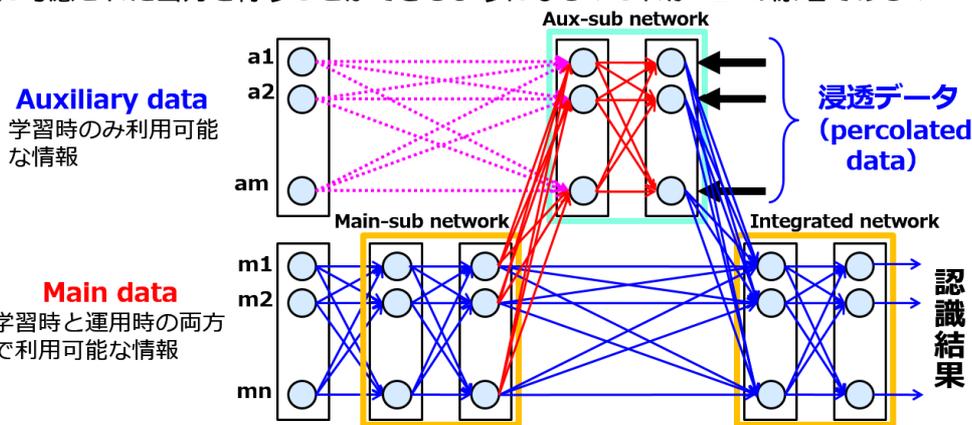


図 1 浸透学習法(PLM)の基本構造

次に本研究の方法について示す。

(1)理論 当初、仮想的な誤差関数を想定して PLM の理論を作ること検討したが、PLM の応用対象である具体的な問題の誤差関数は非常に複雑であることと、問題によってその様相が大きく異なることから、誤差関数を多変量の連続関数として定義してアプローチを見出すことはあまり有効ではないと考えられた。そこで、理論については方法論や応用を進めた後で整理すべきものと考え、方法論を開発する中で、改良の際にその原理について考えることで理論的な知見を深めることとした。

(2)方法論 PLM では、Aux データの減衰方法と、浸透データの維持が最も重要である。前述の基礎的な学習法で成功する問題もあるが、中には、この基礎的な方法では、浸透データが消えてしまう場合や、Aux からの影響を十分に出力に対して考慮することができない場合があつた。そこで、学習方法について様々な面から考察を行なつた。

(3)応用 これまでに手書き文字認識のためのデータベース(MNIST)などへ PLM を適用した際の有効性は確認されていたが、より現実味のあるデータに対しての有効性は未知数であつた。そこで、いわゆるベンチマーク問題だけでなく、より具体的な問題として、時系列信号の将来変動予測の問題とソフトセンサの問題を扱つた。ソフトセンサとは、あるセンサの情報・影響を別のセンサでソフトウェアによって代替しようとするものである。本研究では、これらの問題に対し

て PLM を適用することでその有効性を確認した。

4. 研究成果

本研究によって得られた成果を次に示す。

(1) PLM の方法・理論について

初めに, Aux からの入力への減衰方法について検討した。始めに開発した方法は, 前述のように結合荷重を強制的に減らすのではなく, Aux からの信号の間引き率を徐々に高める手法である。この手法は, 結果的には基礎的な学習法と同等の性能を示すことができることを確認した。

次に, 図 1 に示した基本的な構造では, Aux-sub-network への Main からの入力が Main の情報を加工した後のものになっているが, 場合によっては Main データを直接入力して Aux と関係づけた方がよい場合も散見された。このため, ネットワーク構造そのものを進化計算法によって最適化する方法について検討した。その結果, 問題によっては構造を変化させた方が有効であることを確認するとともに, PLM のための効果的な構造最適化法を開発した。

さらに, 応用問題へ PLM を適用するにあたって, 精度の逐次上昇へのニーズがあることがわかった。下記に示す応用などでは, 出力に影響すると考えられる信号が, 画像認識の場合のように充分とは言えないケースが多々ある。例えば, ある企業の株価を予測する場合, その企業の過去の株価はもちろん重要であるが, その他, 株価に影響を与える情報が何であるのか, 何を留意すれば予測に必要な十分な情報なのかは誰にもわからない。世の中の全ての情報を入力することができれば良いがそれは不可能である。このため, まずは関係が想定できるような株価, 為替データ, ニュースデータなどで基礎回路を作り, 出力の精度が充分ではない場合は, 入力する情報を増やすアプローチが一般的である。そこで, 図 2 に示すように, 基礎回路の入力データを Main (顕在変数), 追加するデータを Aux (潜在変数) とみなして浸透学習を行うことを考えた。最初の顕在変数群を入力する回路に, 随時, 精度向上を図ることができる潜在変数 a_1, a_2, \dots を追加して学習する。この過程で, 最終的にはその結線を浸透学習で除去して運用時は不要にする。これによって, 精度は高まるが, 構造は元のままに保つことができる。これにより, 精度と説明性を両立させている, と考えることもできる。本方式は, 時系列予測だけでなく, 様々な問題に適用することができる。



図 2 精度逐次上昇法 (仮称・特願 2021-032097)

(2) PLM の応用について

時系列信号の将来変動予測問題への適用結果

ここでは時系列信号の将来変動を予測する問題に PLM を適用した。図 3 に考え方を示す。例えば, ある経済指数の 1 週間後の値を予測する際, 図 3 の上側に示すように従来手法では現在 (t) より過去のデータを入力して, 将来 ($t+d$) の値を出力するように学習を行う。これに対して提案する PLM では, Aux として未来の情報を学習に用いることができる。なお, この際, 予測対象の値を入れてしまうと, その信号が強過ぎるため, 学習過程でそこから出力に直結する回路が出来てしまい, その信号線の影響を排除することができなくなるため, 敢えて正解に当たる予測データは入力せず, その周辺のデータを入力する必要がある。

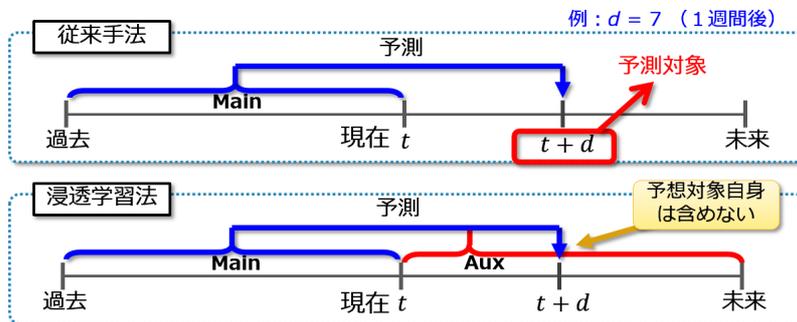


図 3 PLM による将来変動予測の考え方

実験結果の例を表 1 に示す。同表から, PLM を用いる方が従来型の深層回路より高精度に予測することができていることがわかる。これは, PLM が敢えて $t+d$ ではない, 予測精度が高い未来の

値を求めてから時間軸を逆方向に予測する計算をしているためと考えることができる。そのような回路を学習で自然に獲得することができることは PLM の優秀さを表していると考えられる。

表 1 浸透学習法(PLM)による将来変動予測結果の例

	方向的中率 [%]	MAE (Mean Absolute Error)
当日変動率	58.08	0.353
従来のNN	65.21	0.318
浸透学習法	68.20	0.304

ソフトセンサの問題への適用結果

次の応用として、ソフトセンサに PLM を利用してみた。ここでは、ある機械装置のセンサ値 x_A を PLM の Aux, 別のセンサ値 x_M を PLM の Main と考え、 x_A を x_M で再現することを試みた。その際、回路構造を図 1 の基本構造から若干修正して図 4 に示すように変更した。これは、浸透特徴(h_{perc})を表現するための回路(Main sub Net)と、最終出力 y を表現するための回路(Main Net)を明確に区別するための工夫である。図 5 にこの回路網構造による実験結果の例を示す。

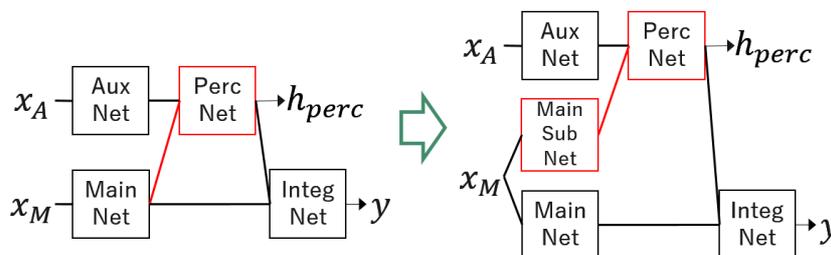


図 4 回路網に関する改良

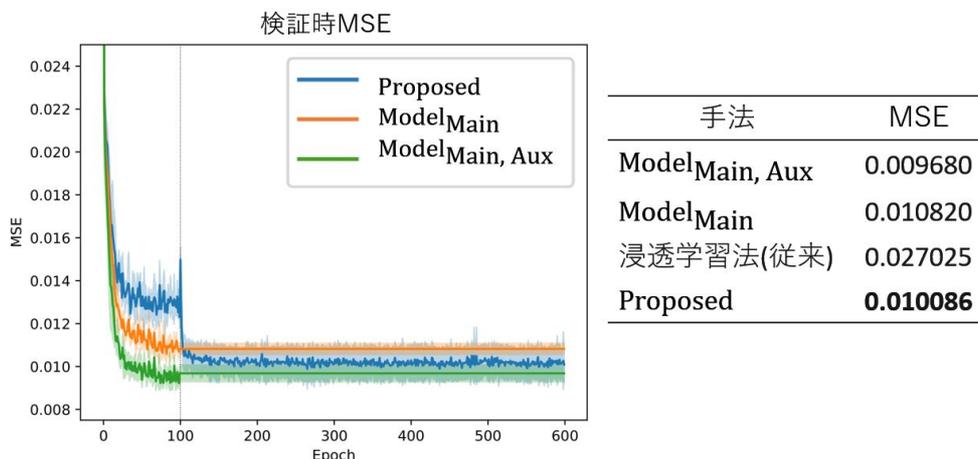


図 5 改良した回路による実験結果の例

図 5 から、今回用いた回路網によって、従来法(ModelMain), および従来の浸透学習法より大幅な精度改善を実現することができたことがわかる。

以上のように、本研究では、先に提案した画期的な深層学習モデルである浸透学習法(PLM)の理論・方法論・応用について研究を行なった。その結果、特に時系列信号の将来変動予測と、ソフトセンサの応用についての PLM の有効性を検証することができたとともに、浸透学習を成功させるための回路網構造や学習方法についての有効な知見を得ることができた。

浸透学習法の設定は、現実世界の問題において頻繁に登場する問題設定であり、浸透学習法は学術的に新規性が高いだけでなく、極めて実用性が高い手法であると考えられる。今後は、学習方法についての理論的考察を深めることや、今回扱わなかった問題に対する PLM の有効性の検証などを進める予定である。

本研究は極めて有益な研究成果を得ることができ、順調かつ当初予定通りの成果を挙げて実施することができたと考えている。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 1件）

1. 著者名 Kazuki Takaishi, Masayuki Kobayashi, Miku Yanagimoto and Tomoharu Nagao	4. 巻 1
2. 論文標題 Percolative Learning: Time-Series Prediction from Future Tendencies	5. 発行年 2018年
3. 雑誌名 Conference: 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)	6. 最初と最後の頁 1643-1648
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/SMC.2018.00285	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計1件（うち招待講演 1件 / うち国際学会 0件）

1. 発表者名 長尾智晴
2. 発表標題 深層学習から浸透学習へ
3. 学会等名 応用脳科学コンソーシアム 応用脳科学アカデミー（招待講演）
4. 発表年 2019年

〔図書〕 計0件

〔出願〕 計1件

産業財産権の名称 ニューラルネットワークシステム、学習制御装置、演算方法、学習制御方法およびプログラム	発明者 長尾智晴, 小林雅幸	権利者 横浜国立大学
産業財産権の種類、番号 特許、2021-032097	出願年 2021年	国内・外国の別 国内

〔取得〕 計0件

〔その他〕

6. 研究組織

氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------