

連載解説「Deep Learning (深層学習)」にあたって

神畠 敏弘

(産業技術総合研究所)

松尾 豊

(東京大学)

Deep Learning を皆さんは知っているだろうか？

2012 年は、機械学習分野にとって、まさしく deep learning の年であったといえよう。Langford の機械学習関連のブログ [Langford] など、2012 年の顕著な成果として取り上げられるのは当然として、一般紙である New York Times [Markoff 12] にまで記事が掲載された。新しい機械学習手法がこれほど話題になったことは、サポートベクタマシンやノンパラメトリックベイズなど最近のどの手法でもなかったことである。

なぜこれほど話題になったのだろうか。2011 年に音声認識の分野で注目され始めてはいた [Interspeech 2011] のだが、2012 年になってから、いろいろな分析・予測・認識処理のコンペティションで連戦連勝しているからである。しかも 2 位以下を置き去りにした圧勝である。例えば、この Large Scale Visual Recognition Challenge 2012 [ILSVRC 2012] は画像の分類問題を対象としたコンペティションである。結果を見ていただくと、2 位以下の予測誤差は 10^{-2} のオーダーの差だが、1 位の Hinton らのグループ SuperVision は 2 位と約 10^{-1} の差を付けている。この強さは、言葉どおり桁違いである。その他のコンペティションでの様子も、Kurzweil のブログ [Kurzweil] で取り上げられているように圧倒的である。このような大きな差が付くことは非常にまれであり、この deep learning は数十年に一度のブレイクスルーといってよいだろう。そこで、遅ればせながら、この deep learning について連載解説を企画した。

deep learning は新しい機械学習手法であると述べたが、これは正しくもあり、誤りでもある。なぜなら、deep learning の手法は人工知能研究の初期からあるニューラルネットワークの一つであるからである。ここで少しニューラルネットワークの歴史に触れておこう。1940 年代に、シナプスに信号が入力されると、その信号に応じて新たな信号の発火が起こり、それが次のシナプスに伝播するという、人間の脳の神経ネットワークのモデルとして、McCulloch-Pitts モデルが提案された。このモデルは 1950 年代には、Rosenblatt をはじめとする研究者によりパーセプトロンなどが開発され、計算機上でシミュレートされるようになった。しかし、1969

年に Minsky と Papert によりパーセプトロンの限界が示され、その研究は下火になった。この限界は、1980 年代には Rumelhart らの効率的な多層ニューラルネットワークの開発により解消され、本格的に非線形の識別問題が扱えるようになり、その研究と実問題への応用は発展を遂げた。だが、大域的な最適解は計算できないという問題は残っており、この弱点を解消したサポートベクタマシンが 1990 年代に登場したことにより、ニューラルネットは機械学習世界の主役の座を譲ることとなった。再び冬の時代を迎えたニューラルネット研究ではあるが、Hinton や Bengio らは研究を根気強く続け、新たな手法を発表し続けた。特に、並列計算技術とともに用いることで、超多層ニューラルネットの学習を可能にした、2000 年代中期の contrastive divergence と pre-training は重要な技術である。そして、現在、冒頭で述べたような画期的な成果を示し、ニューラルネットは二度目の復活を成し遂げた。不遇の時代にも忍耐強く改良を続け、現在の成功につなげたニューラルネット研究者の高い見識や不断の努力と、時流に流されず優れた研究結果を見逃さなかった査読者の先見の明に対してはただ畏敬の念を抱くのみである。

本連載解説では、deep learning の訳語として『深層学習』を用いる。その理由は、“deep” には技術的な側面と内容的な側面の二つの意味が込められていると考えからである。技術的な側面から見ると、深層学習のニューラルネットは、非常に多くの層で構成されている、すなわち『深い層構造』をもっている。一方で、内容的な側面から見ると、深層学習は、表面的な識別関数を獲得するのではなく、より多様な入力のコマンドによる複雑な意味表現、すなわち『深層にある表現』を獲得することを目的としている。なお、Bengio は深層学習の本質を表現学習 (representation learning) であるとし、今年、2013 年 4 月に第 1 回を開催した深層学習関連の国際会議の名称も International Conference on Learning Representation [ICLR 2013] としている。

連載解説『Deep Learning (深層学習)』は全 6 回の予定である。第 1 回と第 2 回は深層学習のモデルと学習手法の基本で、今号 (2013 年 5 月号) では、安田宗

樹がボルツマンマシン型のモデルを、次号(2013年7月号)では、麻生英樹がフィード・フォワード型のモデルを中心に、その歴史的な発展なども踏まえて解説する。2013年9月号掲載の第3回は、岡野原大輔が実装面について解説する。モデルとともに、データの規模も、深層学習の重要な要素であるため、並列計算技術も必要となる。後半3回(2013年11月号～2014年3月号)は、各応用分野において深層学習をどのように適用するかを扱う。第4回は岡谷貴之が画像認識分野について、第5回は久保陽太郎が音声認識分野について、そして第6回はボレガラ・ダヌシカが自然言語処理分野についてそれぞれ解説する。

深層学習についていくつか関連資料をあげておこう。最も著名なサイトは deeplearning.net [DLnet] であり、各種の資料、ソフトウェア、その他の関連情報などを網羅している。2012年は深層学習のチュートリアルがさまざまな国際会議で開催された。深層学習の中心研究者である Bengio は機械学習の国際会議 ICML[Bengio] で、Hinton はサマースクール [Hinton] にてチュートリアル資料を公開しており、これらを最初に参照すると良いであろう。深層学習の応用についても、画像認識の国際会議 CVPR[Fergus]、自然言語処理の国際会議 ACL[Socher]、そして音声認識の ICASSP[Yu] などでもチュートリアル講演が行われた。

冒頭に述べたような圧倒的な性能を示した深層学習ではあるが、課題は山積している。まず、理論面的に解明されていない部分が数多くある。pre-training を採用したことによる誤差はどれくらいか？ ほかに効率の良い学習法はあるのか？ pre-training が特徴の学習であるとするならば、既存の非線形次元削減や特徴生成とは何が違ったのか？ その違いの本質が解明され、大規模データとそれにみあう複雑度のモデルを扱うことが可能であるならば、ニューラルネット以外でも深層学習と同等のモデルが獲得できるのか？ 一方で、実用面でも、現状では調整すべきパラメータが多く性能を発揮するには忍耐強い調整を必要とするともいわれている。複雑な

モデルの獲得には、大規模データを処理できることが必須となるが、並列計算手法の改良はまだ必要であろう。これらの問題に対し、多くの研究者が取り組み始めており、向こう数年から数十年にわたって、画期的な成果が次々とこの分野では発表されるであろう。以上、現在の機械学習とその関連分野で最もホットな話題である deep learning (深層学習) について、本連載解説を楽しんでいただきたい。

◇ 参考文献 ◇

- [Bengio] Bengio, Y.: Representation learning tutorial, *ICML 2012 Tutorial*, <http://www.iro.umontreal.ca/~bengioy/talks/deep-learning-tutorial-2012.html>
- [DLnet] Deep Learning, <http://deeplearning.net/>
- [Fergus] Fergus, R., et al.: Deep learning methods for vision, *CVPR 2012 Tutorial*, http://cs.nyu.edu/~fergus/tutorials/deep_learning_cvpr12/
- [Hinton] Hinton, G., et al.: Graduate summer school: Deep learning, feature learning, <https://www.ipam.ucla.edu/schedule.aspx?pc=gss2012>
- [ICLR 2013] *International Conference on Learning Representations 2013*, <https://sites.google.com/site/representationlearning2013/>
- [ILSVRC 2012] *Large Scale Visual Recognition Challenge 2012*, <http://www.image-net.org/challenges/LSVRC/2012/results.html>
- [Interspeech2011] *Speech Recognition Leaps Forward*, <http://research.microsoft.com/en-us/news/features/speechrecognition-082911.aspx>
- [Kurzweil] How bio-inspired deep learning keeps winning competitions, *Kurzweil Accelerating Intelligence*, <http://www.kurzweilai.net/how-bio-inspired-deep-learning-keeps-winning-competitions>
- [Langford] Langford, J.: Deep learning 2012, *Machine Learning (Theory)*, <http://hunch.net/?p=2609>
- [Markoff 12] Markoff, J.: Scientists See Promise in Deep-Learning Programs, *NY Times* (2012), <http://nyti.ms/SgcVec>
- [Socher] Socher, R., et al.: Deep learning for NLP (without Magic), *ACL 2012 Tutorial*, <http://www.socher.org/index.php/DeepLearningTutorial/DeepLearningTutorial>
- [Yu] Yu, D., et al.: Deep learning and its applications in signal processing, *ICASSP2012 Tutorial*, http://www.icassp2012.com/Tutorial_09.asp