

今泉 允聡 著

『深層学習の原理に迫る』

岩波書店、2021年4月刊、B6並製、126ページ、1,200円+税

最近の日本社会では画像認識、自動運転、自動翻訳などAI (Artificial Intelligence) の話題がTVコマーシャルを含めしばしばメディアに登場している。こうした人々が注目する新たな実用技術の背後に「深層学習 (Deep Learning)」という機械学習 (Machine Learning)、統計的学習 (Statistical Learning) が関係しているらしい。本書を読み終わっての評者の感想は、「深層学習」について理論家からの説明を聞く機会があることは、現代社会を生きていく為の教養として有益であった、ことである。著者である今泉允聡氏は統計科学、特に統計的学習理論における新進気鋭の若手理論家である。本書は著者がしばらく前に属していた統計数理研究所が開催した一般向けの公開講座、チュートリアルの内容をベースに比較的わかりやすく書き下ろした小品である。

まずは本書の概略を紹介しておこう。第1章「深層学習の登場」、第2章「深層学習とは何か」、第3章「なぜ多層が必要なのか」、第4章「膨大なパラメータの謎」、第5章「なぜパラメータの学習ができる?」、第6章「原理を知ることには価値はあるか?」という全6章、116項から本書は構成されている。

第1章は2016年に起きた衝撃的なニュース、グーグル・アルファ碁による世界トップクラスの棋士の撃破、それに遡ること4年、2012年ヒントン教授 (トロント大学) の学習アルゴリズム AlexNet の発表から近年における「深層学習」

の進化の動向である。第2章はニューラルネットと「深層学習」について、過度には数学を利用しないようかなり配慮された解説である。次に第3章では実はかなりの歴史を持っている単純なニューラルネット・モデルから始まり、従来の数学理論、関数近似理論における常識からの逸脱と不連続関数の表現や膨大なデータ構造における「深層学習」の優位性の説明の試み、などが紹介されている。第4章では「深層学習」モデルが含む膨大なパラメータ数の謎について従来の数学・統計科学では説明できない重要な論点、その解釈が述べられている。例えば評者が大学時代に統計学で学んだパラメータ (母数) の数についての「バイアスとばらつきのトレードオフ」についての常識が応用上で根本的に覆されているようだ、という論点は重要である。すなわち「深層学習」が統計学の利用者や統計科学の関係者には見過ごせない挑戦的課題を投げかけているのである。次に第5章では「深層学習」モデルにおける膨大な数のパラメータについての評価関数の数値的評価を巡る問題を解説している。どうやら普通の数理統計学の教科書の説明で出てくる滑らかな凸関数の評価関数を最適化する、などとは様相が異なり、97項の印象的な図で示されている多数の極値が存在するギザギザした形状の高次元関数の最適化問題、という説明は、深層学習がどうして役に立つか、なぜ数理的に難しいのか、素朴な疑問についてある種の納得感が得られる。最後の第6章では

「深層学習」の理論的研究の重要性を述べる、ことで全体を締めくくっている。

普段はAIや機械学習などにはあまり縁がない評者はまだ一読しただけであるが、日本だけではなく世界のメディアがこしばらく注目している深層学習、機械学習、統計的学習を巡る様々な論点について前より理解が深まり、重要性についての認識を改めることができた。本書への感想を幾つか述べておこう。

本書のテーマである「深層学習」はごく最近になり様々な実用化が実現しつつあるが、実は根本的に理論的には「なぜ応用がうまく行くのか分かっていない」という深層学習の若手研究者である著者の説明は実に興味深い。応用上で画期的な性能があることが分かってからまだ約10年ほどであり、有用性についてはかなりの人々が認めてはいるものの、現在でも数学・物理・工学・統計学をはじめ様々な分野における世界のトップクラスの研究者が研究を続けているらしい。難しい数理的議論はともかく、そうした理由の一端について有益な情報が得られた感がある。著者を含めた若手研究者たちの今後の研究成果を期待、応援したい。

ただし本書は元々、専門家ではなく一般読者向けに書かれ、数学・数理的議論をあまり利用しないという著者の方針の為と思われるが、数理的内容の説明には十分には理解できない箇所が少なくない。例えば「二重降下 (double descent)」の説明を例として挙げておくが、これから本書を読もうとする読者はこうした点について了とする必要があるだろう。(なお日本統計学会誌 (日本語) 第50巻2 (2021年3月)号には本テーマについての特集記事があり、数学・数理的側面により関心のある読者は今泉氏の論文を含め無料ダウンロードで読むことは可

能である。) むろん、2023年初頭の段階においてもまだまだ「深層学習」は未開の分野であるから、専門的研究論文はさておき、本書で説明を試みている「深層学習」の理論についてどこまで正確に理解できるかは疑問となることは仕方がない、とは評者の意見である。

しかし、統計学の教育を受けた旧世代にとっては良く知らない間に社会は劇的に変化、現代の統計科学では「深層学習」を含む統計的学習理論 (statistical learning theory) の役割は大きくなっている。例えば読者は、樹木モデル (Trees-based methods)、ラッソ (Lasso)、正則化 (Regularization)、リサンプリング (Resampling)、サポートベクター分類器 (Support Vector Machine) などに加えて深層学習 (Deep Learning) など理解されているだろうか。

本書を一読して深層学習について優れた研究業績を重ねている若手研究者が書いた書籍という感想を持ったが、理論的内容の細部はともかく、内容を分かりやすく解説できる説明力には感心した場面が少なくない。例えば統計的データ解析や学習理論ではよく知られている「過適合 (overfitting)、過学習」をめぐる「訓練データと予測力」を巡る説明では「受験生がある大学の過去の問題の解答をすべて丸暗記しても、実は本番ではあまり振るわない」という身近な類推などには感心させられた。

既に日本を含め現代社会ではAIや学習理論の応用から無縁でいることは困難になっている。統計科学やデータ・サイエンスに関心のある読者が「深層学習」や「統計的学習」などを考察するきっかけとなればと考え、本書の一読を推奨したい。

評者：国友 直人・くにとも なおと
(統計数理研究所特任教授、東京大学名誉教授)