

■ 刊行にあたって

人工知能が大きな話題になっている。これは、社会と文明に大きな変革を及ぼす第4次産業革命をもたらすかもしれない。このとき、人工知能の仕組みをよく理解した万全の心構えが我々の側に必要である。

人工知能はコンピュータと共に歩んだ60年を超す歴史を持っている。しかし、社会を揺るがすほどにその性能が上がったのは、ここ10年ほどのことである。その中心になった技術が深層学習である。これは、脳の神経回路にヒントを得たもので、さまざまなデータを与えれば、その中の隠れた仕組みや構造を自動的に学習してシステムを築き上げるという、まさに人間の名人がやってのける技を目指している。

ただ、これはまだ限られた範囲の問題にしか適用できず、人間のような自由自在の汎用性はない。しかし、人間ではとても扱えないほどの超多量のデータから規則を抜き出すとなれば、決して侮ることのできない新しい技術である。これを可能にしたのは、コンピュータやインターネットを始めとする情報技術の驚くべき進歩である。

人工知能は、人間の知能の領域に進出し、そのもとになる心にまで踏み込もうとしている。人間が今日あるのは永年の進化によるものであり、心や知能はその結果生まれた。これが今日の文明と社会を生み、我々の生活の基盤となっている。人工知能は社会や文明を変えようとしているのであろうか。技術的特異点が2045年に訪れ、人工知能は人類の知能を完全に凌駕し、人間はその保護のもとに生活するという恐ろしい話までである。

私はこの話にくみしないが、人工知能が社会と文明に与える影響は甚大であることは間違いない。こうした事態を論じるときに、人工知能技術、特に今の中核にある深層学習についてのしっかりとした理解が必要である。

人工知能は人間が作ったものである。だから、我々はその隅々までも熟知している。とはいえ、何故そんなにうまくいえるのか、その根本はまだ理解できていない。学習の結果うまくいき、囲碁では人間の名人をも凌駕するといわれてもそれだけでは困る。これは脳についても同じである。進化の結果、こんな素晴らしい脳が出来上がった。それがうまく働く原理はどこ

にあるのか、我々はまだ理解できていない。脳の仕組みと人工知能は、共通する情報の基本原理を使い、これを別々の仕方でも実現したのかもしれない。

人工知能は便利である。ツールとして多くのソフトやライブラリが利用できるし、これをうまく使えば役に立つだろう。しかし、それでは人工知能を理解したことにはならない。その本質に迫る理解が必要である。世の中に人工知能の解説書は多いが、基礎から始め、その仕組みを理論的に明快に説明したのは本書が初めてといってよい。

深層学習にしても永い歴史を持っている。学習の基本的な仕組みを初学者にもわかるように数理的に明らかにするとともに、最新の入り組んだ工夫に至るまでを懇切丁寧に説明してある。また、アルファ碁などの仕組みについても詳しい説明がみられるのも嬉しい。ところどころに挟んであるエピソードも楽しい。

深層学習を学びこれを理解したいという学生、これをさらに発展させようと意気込んでいる研究者、またこれを使ってさらに有用な技術を作り出そうとしている技術者にとって、必読の文献と言える。基礎のしっかりした理解があれば、今後の発展にも十分に対応ができるからである。

本書が上梓されたことを喜びたい。

2017年3月

理化学研究所脳科学総合研究センター 特別顧問
甘利俊一

■ まえがき

紛れもなく深層学習は今世紀の科学技術における革新ですし、今後もその評価は変わらないでしょう。同時代において科学技術のエキサイティングな進展と遭遇できることはとても幸運なことです。それだけを理由にしても、深層学習をじっくり学んでみる価値があるのではないのでしょうか。

深層学習はしばしば、「理論的な仕組みはまったく分かっていないにもかかわらず、発見法的な手法を積み上げるだけでうまくいっている」といわれます。このようなイメージはある面では実状を反映していますが、実際には研究開発の現状を正しく捉えていません。深層学習の研究開発はまぐれ当たりなわけではなく、慎重な数理解析に基づいてデザインされています。つまり相応の根拠があるのです。最先端のレベルで深層学習のデザインを行おうとするならば、このような理論的背景を押さえておく必要があるでしょう。そこで本書ではできる限りこれまでの研究の理論的な側面を紹介していきます。その一方で深層学習がなぜ極めて高い性能を実現できるのかに関しては、いまだに多くのことが分かっていません。特に訓練可能性、表現能力、汎化可能性という3つの理論上の大きな謎があります。入門書としての性格上、今現在も世界中で進行中の理論研究に関しては解説できませんが、いくつかの重要な問題意識については繰り返し紹介します。ぜひさらなる理解への端緒としてください。

深層学習や機械学習の分野では数多くの素晴らしいフレームワークとチュートリアルが無料で提供されており、その解説や実装結果もすでにたくさん公開されています。したがって、本書では実装に関する内容は一切省きました。しかし理論的な詳細を理解するうえでも、TensorFlowやChainerなどで実装・実験をすることは有用です。読者の皆様は、フレームワークでの実験をしながら学ばれるのもよいと思います。

著者の専門は素粒子理論・超弦理論ですが、所属する理論科学連携研究推進グループ(iTHES)主催の講演会などで深層学習の面白さに触発されて以来、個人的に勉強を続けてきました。勉強ノートを蓄えているのを知った周囲から、書籍にしたほうがよいと後押しされ、講談社サイエンティフィクを

紹介していただきました。このような経緯から、本書は著者のこれまでの勉強の記録をもとに、特に基礎的で重要であろう話題を選んで執筆したものです。そのため、これまでの碩学らの論文や素晴らしい教科書の数々に多くを負っています。しかしそのうえであえて1冊を加えるのは、著者が勉強し始めたときに「もしこんな教科書があったらよかったのに」と空想する内容をまとめることにも価値があると思ったからです。そこで、予備知識がない状態から研究の最前線を垣間みれるレベルまで読者を引き上げるような教科書を目指しました。その試みが成功したか否かは、真摯に読者の皆様の判断を仰ぎます。

ニューラルネットワーク研究の立役者である甘利俊一先生には、本書の素稿すべてに目を通して詳細なコメントをいただいたのみならず、身にあまる素晴らしい巻頭言まで書いていただきました。感謝の言葉もありません。また同僚である小川軌明博士、田中章詞博士、日高義将博士および本郷優博士には原稿に対し多くの有益なコメントをいただいたほか、普段から刺激的な議論を通じて著者の理解を鍛えていただいています。本書に他書より優れたところがあるならば、それは各氏のおかげです。もちろん本書における無理解は著者の責任であることはいうまでもありません。また講談社サイエンティフィクの横山真吾氏には、不慣れな著者の執筆を辛抱強く鼓舞していただきました。そして理化学研究所の理論科学連携研究推進グループ (iTHES)、および数理創造プログラム (iTHEMS) には、「科学者の自由な楽園」を地でいく理想的な研究環境を提供していただいています。改めて、これまでの継続的な支援を感謝します。

最後に、本書が、本邦からたくさんの若い深層学習の研究開発者が誕生することの一助となれば幸いです。

2017年4月

瀧 雅人

■ 目 次

■ 刊行にあたって	iii
■ まえがき	v

第 1 章 はじめに	1
------------	---

Chapter 2	第 2 章 機械学習と深層学習	4
	2.1 なぜ深層学習か？	4
	2.2 機械学習とは何か	6
	2.2.1 代表的なタスク	7
	2.2.2 さまざまなデータセット	8
	2.3 統計入門	10
	2.3.1 標本と推定	10
	2.3.2 点推定	12
	2.3.3 最尤推定	17
	2.4 機械学習の基礎	19
	2.4.1 教師あり学習	20
	2.4.2 最小二乗法による線形回帰	21
	2.4.3 線形回帰の確率的アプローチ	24
	2.4.4 最小二乗法と最尤法	26
	2.4.5 過適合と汎化	26
	2.4.6 正則化	28
2.4.7 クラス分類	29	
2.4.8 クラス分類へのアプローチ	31	
2.4.9 ロジスティック回帰	32	
2.4.10 ソフトマックス回帰	34	
2.5 表現学習と深層学習の進展	36	
2.5.1 表現学習	36	
2.5.2 深層学習の登場	37	

Chapter 3	第 3 章 ニューラルネットワーク	41
	3.1 神経細胞のネットワーク	41
	3.2 形式ニューロン	44
	3.3 パーセプトロン	47
	3.3.1 形式ニューロンによるパーセプトロン	47
	3.3.2 パーセプトロンとミンスキー	48
3.4 順伝播型ニューラルネットワークの構造	50	

3.4.1	ユニットと順伝播型ニューラルネットワーク	50
3.4.2	入力層	52
3.4.3	中間層	52
3.4.4	出力層	54
3.4.5	関数	55
3.5	ニューラルネットによる機械学習	55
3.5.1	回帰	56
3.5.2	2 値分類	57
3.5.3	多クラス分類	58
3.6	活性化関数	59
3.6.1	シグモイド関数とその仲間	60
3.6.2	正規化線形関数	60
3.6.3	マックスアウト	62
3.7	なぜ深層とすることが重要なのか	63

第 4 章	勾配降下法による学習	65
4.1	勾配降下法	65
4.1.1	勾配降下法	67
4.1.2	局所的最小値の問題	68
4.1.3	確率的勾配降下法	69
4.1.4	ミニバッチの作り方	71
4.1.5	収束と学習率のスケジューリング	72
4.2	改良された勾配降下法	73
4.2.1	勾配降下法の課題	73
4.2.2	モーメンタム法	75
4.2.3	ネステロフの加速勾配法	77
4.2.4	AdaGrad	77
4.2.5	RMSprop	78
4.2.6	AdaDelta	80
4.2.7	Adam	81
4.2.8	自然勾配法	83
4.3	重みパラメータの初期値の取り方	84
4.3.1	LeCun の初期化	84
4.3.2	Glorot の初期化	85
4.3.3	He の初期化	85
4.4	訓練サンプルの前処理	87
4.4.1	データの正規化	87
4.4.2	データの白色化	88
4.4.3	画像データの局所コントラスト正規化*	91

第 5 章	深層学習の正則化	93
5.1	汎化性能と正則化	93
5.1.1	汎化誤差と過学習	93

5.1.2	正則化	96
5.2	重み減衰	97
5.2.1	重み減衰の効果	97
5.2.2	スパース正則化と不良条件問題	98
5.3	早期終了	99
5.3.1	早期終了とは	99
5.3.2	早期終了と重み減衰の関係	100
5.4	重み共有	101
5.5	データ拡張とノイズの付加	101
5.5.1	データ拡張と汎化	101
5.5.2	ノイズの付加とペナルティ項	102
5.6	バギング	103
5.7	ドロップアウト	104
5.7.1	ドロップアウトにおける学習	105
5.7.2	ドロップアウトにおける推論	107
5.7.3	ドロップアウトの理論的正当化	108
5.8	深層表現のスパース化	111
5.9	バッチ正規化	111
5.9.1	内部共変量シフト	112
5.9.2	バッチ正規化	112

Chapter 6	第 6 章 誤差逆伝播法	114
6.1	パーセプトロンの学習則とデルタ則*	114
6.2	誤差逆伝播法	117
6.2.1	パラメータ微分の複雑さとトイモデル	118
6.2.2	誤差関数の勾配計算	120
6.2.3	逆伝播計算の初期値	123
6.2.4	勾配計算	124
6.2.5	デルタの意味	125
6.3	誤差逆伝播法はなぜ早いのか	126
6.4	勾配消失問題, パラメータ爆発とその対応策	128
6.4.1	事前学習	129
6.4.2	ReLU 関数	130

Chapter 7	第 7 章 自己符号化器	132
7.1	データ圧縮と主成分分析	132
7.2	自己符号化器	136
7.2.1	砂時計型ニューラルネット	137
7.2.2	再構成誤差による学習	139
7.2.3	符号化器の役割	140
7.2.4	自己符号化器と主成分分析	141
7.3	スパース自己符号化器	142
7.3.1	自己符号化器のスパース化	142

	7.3.2 スパース自己符号化器の誤差逆伝播法	143
7.4	積層自己符号化器と事前学習	146
	7.4.1 積層自己符号化器	146
	7.4.2 事前学習	148
7.5	デノイジング自己符号化器	149
7.6	収縮自己符号化器*	150
	7.6.1 収縮自己符号化器と多様体学習	150
	7.6.2 他の自己符号化器との関係	151

	第 8 章 畳み込みニューラルネット	153
8.1	一次視覚野と畳み込み	153
	8.1.1 ヒューベル・ウィーゼルの階層仮説	153
	8.1.2 ニューラルネットと畳み込み	156
8.2	畳み込みニューラルネット	157
	8.2.1 画像データとチャネル	157
	8.2.2 畳み込み層	158
	8.2.3 1×1 畳み込み*	162
	8.2.4 因子化した畳み込み*	162
	8.2.5 ストライド	163
	8.2.6 パディング	165
	8.2.7 プーリング層	166
	8.2.8 局所コントラスト正規化層*	168
	8.2.9 局所応答正規化層*	168
	8.2.10 ネットワーク構造	169
8.3	CNN の誤差逆伝播法	170
	8.3.1 畳み込み層	170
	8.3.2 プーリング層	172
8.4	学習済みモデルと転移学習	172
8.5	CNN はどのようなパターンを捉えているのか	173
8.6	脱畳み込みネットワーク*	174
8.7	インセプションモジュール*	175

	第 9 章 再帰型ニューラルネット	177
9.1	時系列データ	177
9.2	再帰型ニューラルネット	178
	9.2.1 ループと再帰	178
	9.2.2 実時間リカレント学習法	181
	9.2.3 ネットワークの展開	183
	9.2.4 通時的誤差逆伝播法	184
9.3	機械翻訳への応用	186
9.4	RNN の問題点	186
9.5	長・短期記憶	187
	9.5.1 メモリー・セル	188

9.5.2	ゲート	189
9.5.3	LSTM	189
9.5.4	LSTM の順伝播	190
9.5.5	LSTM の逆伝播	192
9.5.6	ゲート付き再帰的ユニット*	195
9.6	再帰型ニューラルネットと自然言語処理*	196
9.6.1	Seq2Seq 学習	198
9.6.2	ニューラル会話モデル	199

第 10 章	ボルツマンマシン	200
10.1	グラフィカルモデルと確率推論	200
10.1.1	有向グラフィカルモデル*	201
10.1.2	無向グラフィカルモデル*	205
10.2	ボルツマンマシン	211
10.2.1	隠れ変数なしのボルツマンマシン	211
10.2.2	隠れ変数ありのボルツマンマシン	213
10.3	ボルツマンマシンの学習と計算量爆発	215
10.3.1	隠れ変数のない場合	216
10.3.2	対数尤度関数の凸性	219
10.3.3	勾配上昇法と計算量	221
10.3.4	ダイバージェンスによる学習	222
10.3.5	隠れ変数のある場合	223
10.4	ギブスサンプリングとボルツマンマシン	227
10.4.1	マルコフ連鎖	228
10.4.2	Google とマルコフ連鎖	229
10.4.3	定常分布	231
10.4.4	マルコフ連鎖モンテカルロ法	233
10.4.5	ギブスサンプリングとボルツマンマシン	234
10.5	平均場近似	240
10.6	制限付きボルツマンマシン	246
10.6.1	制限付きボルツマンマシンの学習	248
10.6.2	ブロック化ギブスサンプリング	250
10.7	コントラストティブダイバージェンス法とその理論	252
10.7.1	コントラストティブダイバージェンス法はなぜうまくいくのか*	256
10.7.2	コントラストティブダイバージェンスの最小化*	261
10.7.3	持続的コントラストティブダイバージェンス法 (PCD 法)	263
10.8	ディープビリーフネットワーク	265
10.8.1	DBN の事前学習	267
10.8.2	DBN の微調整	271
10.8.3	DBN からのサンプリング	273
10.8.4	DBN での推論	273
10.9	ディープボルツマンマシン	274
10.9.1	DBM の事前学習	275
10.9.2	DBM の微調整	278

■	10.9.3 順伝播型ニューラルネットへの変換	281
Chapter 11	第 11 章 深層強化学習	283
	11.1 強化学習	283
	11.1.1 マルコフ決定過程	284
	11.1.2 ベルマン方程式と最適政策	287
	11.1.3 TD 誤差学習	293
	11.1.4 Q 学習	295
	11.2 関数近似と深層 Q ネット	297
	11.2.1 Q 学習と関数近似	297
	11.2.2 深層 Q 学習	301
	11.3 アタリゲームと DQN	304
	11.4 方策学習	308
	11.4.1 勾配上昇法による方策学習	308
	11.4.2 方策勾配定理の証明	310
	11.5 アルファ碁	311
	11.5.1 モンテカルロ木探索の考え方	311
	11.5.2 SL 方策ネットワーク P_{σ}	312
	11.5.3 ロールアウト方策 P_{π}	313
	11.5.4 LR 方策ネットワーク P_{ρ}	313
	11.5.5 価値ネットワーク v	314
	11.5.6 方策と価値ネットワークによるモンテカルロ木探索	315
■	付録 A 確率の基礎	318
	A.1 確率変数と確率分布	318
	A.2 連続確率変数と確率質量関数	321
	A.3 期待値と分散	324
	A.4 情報量とダイバージェンス	326
■	付録 B 変分法	328
	B.1 汎関数	328
	B.2 オイラー・ラグランジュ方程式	329
■	参考文献	331
■	索引	337