



Spis treści

Przedmowa	XIII
Podziękowania	XIV
O tej książce	XV
O autorze	XVIII

1. Wprowadzenie do głębokiego uczenia:	
dlaczego warto się tego nauczyć	1
W tym rozdziale	1
Zapraszamy do <i>grokowania</i> głębokiego uczenia	1
Dlaczego warto opanować głębokie uczenie się	2
Czy trudno będzie się tego nauczyć?	3
Dlaczego powinieneś przeczytać tę książkę	3
Co jest potrzebne, aby zacząć	5
Zapewne potrzebna jest jakaś znajomość Pythona	6
Podsumowanie	7
2. Podstawowe koncepcje: jak maszyny się uczą?	9
W tym rozdziale	9
Czym jest głębokie uczenie?	9
Czym jest uczenie się maszyn?	10
Nadzorowane uczenie się maszyn	11
Nienadzorowane uczenie się maszyn	12
Uczenie parametryczne kontra nieparametryczne	13
Nadzorowane uczenie parametryczne	14
Parametryczne uczenie nienadzorowane	16

Uczenie nieparametryczne	18
Podsumowanie	19
3. Wprowadzenie do prognozowania neuronowego: propagacja w przód	21
W tym rozdziale	21
Krok 1: Prognoza	22
Wykonywanie prognozy przez prostą sieć neuronową	24
Czym jest sieć neuronowa?	25
Co robi ta sieć neuronowa?	26
Wykonywanie prognozy dla wielu danych wejściowych	28
Wiele danych wejściowych: Co robi ta sieć neuronowa?	30
Wykonywanie prognozy z wieloma danymi wyjściowymi	37
Prognozowanie z wieloma danymi wejściowymi i wyjściowymi	39
Wiele wejść i wyjść: jak to działa?	41
Prognozowanie na podstawie prognoz	43
Elementarz NumPy	45
Podsumowanie	48
4. Wprowadzenie do prognozowania neuronowego: metoda gradientowa	49
W tym rozdziale	49
Porównywanie	50
Nauka	51
Po co mierzyć błędy?	53
Jaka jest najprostsza forma uczenia się sieci neuronowej?	54
Nauka metodą ciepło-zimno	56
Charakterystyka uczenia metodą ciepło-zimno	57
Obliczanie jednocześnie kierunku i wielkości na podstawie błędu	58
Jedna iteracja metody gradientowej	60
Uczenie się jest po prostu redukowaniem błędu	62
Przyjrzyjmy się kilku krokom uczenia się	64
Dlaczego to działa? Czym naprawdę jest <code>weight_delta</code> ?	66
Skupmy całą uwagę na jednej idei	68
Pudełko z wystającymi drążkami	69
Pochodne: Podejście drugie	70
Co naprawdę trzeba wiedzieć	71
Czego nie musimy naprawdę wiedzieć	72

Jak użyć pochodnych do uczenia się	72
Psucie metody gradientowej	75
Wizualizacja nadmiernych korekt	76
Rozbieżność	77
Wprowadzenie parametru alpha	78
Alpha w kodzie źródłowym	78
Zapamiętywanie	79
5. Uczenie się dla wielu wag jednocześnie: uogólnienie metody gradientowej	81
W tym rozdziale	81
Uczenie się metodą gradientową przy wielu danych wejściowych	82
Prosty gradient z wieloma danymi wejściowymi – objaśnienie	84
Przyjrzyjmy się kilku krokom uczenia się	88
Zamrożenie jednej wagi: co to daje?	91
Uczenie się metodą gradientową z wieloma wyjściami	93
Metoda gradientowa z wieloma wejściami i wyjściami	96
Czego uczą się te wagi?	98
Wizualizacja wartości wag	100
Wizualizowanie iloczynów skalarnych (sum ważonych)	101
Podsumowanie	102
6. Budowanie pierwszej głębokiej sieci neuronowej: wprowadzenie do propagacji wstecznej	103
W tym rozdziale	103
Problem światła ulicznych	104
Przygotowywanie danych	106
Macierze i relacje macierzowe	107
Tworzenie macierzy w Pythonie	110
Budowanie sieci neuronowej	111
Uczenie się całego zbioru danych	112
Pełna, wsadowa i stochastyczna metoda gradientowa	113
Sieci neuronowe uczą się korelacji	114
Presja w górę i w dół	115
Przypadek brzegowy: przeuczenie (overfitting)	117
Przypadek brzegowy: sprzeczne presje	118
Uczenie się pośredniej korelacji	120
Tworzenie korelacji	121

Układanie sieci neuronowych w stos: przegląd	122
Propagacja wsteczna: długodystansowy wkład błędu	123
Propagacja wsteczna: dlaczego to działa?	124
Liniowe czy nieliniowe	125
Dlaczego sieć neuronowa nadal nie działa	126
Tajemnica korelacji warunkowej	127
Krótką przerwa	128
Nasza pierwsza „głęboka” sieć neuronowa	129
Propagacja wsteczna w kodzie	130
Jedna iteracja propagacji wstecznej	132
Zebrańie wszystkiego razem	134
Dlaczego głębokie sieci są ważne?	135
7. Jak obrazować sieci neuronowe:	
w myślach i na papierze	137
W tym rozdziale	137
Czas na uproszczenie	137
Korelacja zbiorcza	139
Wcześniejsza, zbyt złożona wizualizacja	140
Uproszczona wizualizacja	141
Jeszcze większe upraszczanie	142
Zobaczmy, jak prognozuje ta sieć	144
Wizualizacja przy użyciu liter zamiast obrazów	145
Łączenie zmiennych	146
Wszystko obok siebie	146
Ważność narzędzi wizualizacji	147
8. Uczenie się sygnałów i ignorowanie szumu:	
wprowadzenie do regularyzacji i przetwarzania wsadowego ...	149
W tym rozdziale	149
Trójwarstwowa sieć dla danych MNIST	149
No dobrze, to było łatwe	151
Zapamiętywanie a uogólnianie	153
Przeuczenie w sieciach neuronowych	154
Skąd bierze się przeuczenie	155
Najprostsza regularyzacja: wczesne zatrzymanie	156
Branżowy standard regularyzacji: wykluczanie (dropout)	157
Dlaczego wykluczanie działa: działanie zespołowe	158

Wykluczanie w kodzie	159
Przetwarzanie z wykluczeniem dla danych MNIST	161
Mini-wsadowa metoda gradientowa	162
Podsumowanie	165
9. Modelowanie prawdopodobieństwa i nieliniowość: funkcje aktywacji	167
W tym rozdziale	167
Czym jest funkcja aktywacji?	168
Standardowe funkcje aktywacji dla ukrytych warstw	171
Standardowe funkcje aktywacji dla warstwy wyjściowej	172
Kluczowy problem: Dane wejściowe zawierają podobieństwa	175
Obliczanie funkcji softmax	176
Stosowanie funkcji aktywacji	177
Mnożenie delta przez nachylenie	179
Przełożenie wielkości wyjścia na nachylenie (pochoďną)	181
Aktualizowanie sieci MNIST	182
10. Neuronowe uczenie się krawędzi i narożników: wprowadzenie do konwolucyjnych sieci neuronowych	185
W tym rozdziale	185
Wielokrotne wykorzystanie tych samych wag w wielu miejscach	185
Warstwa konwolucyjna	187
Prosta implementacja w NumPy	189
Podsumowanie	193
11. Sieci neuronowe rozumiejące język: king-man+woman ==?	195
W tym rozdziale	195
Czym jest rozumienie języka?	195
Przetwarzanie języka naturalnego (NLP)	196
Nadzorowane NLP	197
Zbiór danych recenzji filmowych z IMDB	198
Przechwytywanie korelacji słów w danych wejściowych	200
Prognozowanie recenzji filmów	201
Wprowadzenie do warstw osadzających	202
Interpretowanie wyjścia	204
Architektura sieci neuronowej	205

Porównywanie osadzeń słów	207
Czym jest znaczenie neuronu?	209
Wypełnianie pustych miejsc	210
Znaczenie wynika ze straty	212
King – Man + Woman $\approx$ Queen	215
Analogie słów	216
Podsumowanie	218
12. Sieci neuronowe piszące jak Szekspir: warstwy rekurencyjne dla danych o zmiennej długości	219
W tym rozdziale	219
Wyzwanie: dane o dowolnej długości	220
Czy porównywanie naprawdę ma znaczenie?	221
Zaskakująca siła uśrednionych wektorów słów	222
Jak informacja jest przechowywana w tych osadzeniach?	223
Jak sieć neuronowa wykorzystuje osadzenia?	224
Ograniczenia wektorów „worka słów”	225
Wykorzystanie macierzy jednostkowych do sumowania osadzeń słów	226
Macierze, które niczego nie zmieniają	228
Uczenie się macierzy przejść	229
Uczenie się w celu tworzenia użytecznych wektorów zdań	230
Propagacja w przód w Pythonie	231
Jak zastosować do tego propagację wsteczną?	232
Nauczmy tę sieć!	233
Przygotowania	234
Propagacja w przód przy danych dowolnej długości	235
Propagacja wsteczna dla danych o dowolnej długości	236
Aktualizowanie wag przy danych dowolnej długości	237
Uruchomienie i analiza wyjścia	238
Podsumowanie	241
13. Przedstawiamy automatyczną optymalizację: budujemy platformę głębokiego uczenia	243
W tym rozdziale	243
Czym jest platforma głębokiego uczenia?	244
Wprowadzenie do tensorów	245
Wprowadzenie do automatycznego obliczania gradientu (autograd) ..	246

Małe podsumowanie	248
Tensory używane wielokrotnie	249
Ulepszenie funkcji autograd w celu obsługi wielokrotnie używanych tensorów	250
Jak działa wsteczna propagacja dodawania?	252
Dodawanie obsługi zmiany znaku	254
Dodawanie obsługi dodatkowych funkcji	255
Wykorzystanie autograd do uczenia sieci neuronowej	259
Dodanie automatycznej optymalizacji	261
Dodanie wsparcia dla różnych typów warstw	262
Warstwy zawierające warstwy	263
Warstwy funkcji straty	264
Jak uczyć platformę	265
Warstwy nieliniowości	266
Warstwa osadzająca	268
Dodawanie indeksowania do mechanizmu autograd	269
Warstwa osadzająca powraca	270
Warstwa entropii krzyżowej	271
Warstwa rekurencyjna	273
Podsumowanie	277
 14. Uczenie się pisania jak Szekspir:	
długa pamięć krótkotrwała (LSTM)	279
W tym rozdziale	279
Modelowanie znaków języka	279
Potrzeba obciążenia propagacji wstecznej	280
Obciążona propagacja wsteczna	282
Próbka wyjścia	285
Zanikające i eksplodujące gradienty	286
Mały przykład propagacji wstecznej w RNN	287
Komórki LSTM	288
Intuicyjne rozumienie bramek LSTM	290
Warstwa LSTM	291
Ulepszanie modelu znaków języka	292
Uczenie LSTM modelu znaków języka	293
Dostrajanie modelu LSTM znaków języka	294
Podsumowanie	295

15. Głębokie uczenie się na niewidzianych danych:	
wprowadzenie do uczenia sfederowanego	297
W tym rozdziale	297
Problem prywatności w głębokim uczeniu się	298
Sfederowane uczenie się	299
Uczenie się wykrywania spamu	300
Stwórzmy federację	302
Włamywanie się do sfederowanego uczenia się	303
Bezpieczne agregowanie	304
Szyfrowanie homomorficzne	305
Homomorficznie szyfrowane sfederowane uczenie się	307
Podsumowanie	308
16. Co dalej:	
krótki przewodnik	311
W tym rozdziale	311
Gratulacje!	312
Krok 1: Początek nauki PyTorch	312
Krok 2: Udział w innych kursach głębokiego uczenia się	313
Krok 3: Matematyczne podręczniki głębokiego uczenia się	313
Krok 4: Utworzenie bloga i nauczanie głębokiego uczenia się	314
Krok 5: Twitter	315
Krok 6: Stosowanie wiedzy z artykułów naukowych	315
Krok 7: Zdobądź dostęp do GPU	316
Krok 8: Zarabianie przy zdobywaniu praktyki	316
Krok 9: Dołączenie do projektu open source	317
Krok 10: Rozwijanie lokalnej społeczności	317
Indeks	319