

Spis treści

Przedmowa 9

Podziękowania 11

O książce 13

O autorach 17

CZĘŚĆ I. PODSTAWY UCZENIA GŁĘBOKIEGO 19

Rozdział 1. Czym jest uczenie głębokie? 21

1.1. Sztuczna inteligencja, uczenie maszynowe i uczenie głębokie 22

1.1.1. Sztuczna inteligencja 22

1.1.2. Uczenie maszynowe 22

1.1.3. Formy danych umożliwiające uczenie 24

1.1.4. "Głębia" uczenia głębokiego 26

1.1.5. Działanie uczenia głębokiego przedstawione na trzech rysunkach 27

1.1.6. Co dotychczas osiągnięto za pomocą uczenia głębokiego? 29

1.1.7. Nie wierz w tymczasową popularność 30

1.1.8. Nadzieje na powstanie sztucznej inteligencji 31

1.2. Zanim pojawiło się uczenie głębokie: krótka historia uczenia maszynowego 32

1.2.1. Modelowanie probabilistyczne 32

1.2.2. Wczesne sieci neuronowe 33

1.2.3. Metody jądrowe 33

1.2.4. Drzewa decyzyjne, lasy losowe i gradientowe wzmocnianie maszyn 35

- 1.2.5. Powrót do sieci neuronowych 35
- 1.2.6. Co wyróżnia uczenie głębokie? 36
- 1.2.7. Współczesne uczenie maszynowe 37
- 1.3. Dlaczego uczenie głębokie? Dlaczego teraz? 38
 - 1.3.1. Sprzęt 38
 - 1.3.2. Dane 39
 - 1.3.3. Algorytmy 40
 - 1.3.4. Nowa fala inwestycji 40
 - 1.3.5. Demokratyzacja uczenia głębokiego 41
 - 1.3.6. Co dalej? 41

Rozdział 2. Matematyczne podstawy sieci neuronowych 43

- 2.1. Pierwszy przykład sieci neuronowej 44
- 2.2. Reprezentacja danych sieci neuronowych 47
 - 2.2.1. Skalary (tensory zerowymiarowe) 48
 - 2.2.2. Wektory (tensory jednowymiarowe) 48
 - 2.2.3. Macierze (tensory dwuwymiarowe) 48
 - 2.2.4. Trójwymiarowe tensory i tensory o większej liczbie wymiarów 49
 - 2.2.5. Główne atrybuty 49
 - 2.2.6. Obsługa tensorów R 50
 - 2.2.7. Wsad danych 50
 - 2.2.8. Prawdziwe przykłady tensorów danych 51
 - 2.2.9. Dane wektorowe 51
 - 2.2.10. Dane szeregu czasowego i dane sekwencyjne 52
 - 2.2.11. Dane w postaci obrazów 52
 - 2.2.12. Materiały wideo 53
- 2.3. Koła zębate sieci neuronowych: operacje na tensorach 53

- 2.3.1. Operacje wykonywane element po elemencie 54
- 2.3.2. Operacje na tensorach o różnych wymiarach 55
- 2.3.3. Iloczyn tensorowy 55
- 2.3.4. Zmiana kształtu tensora 57
- 2.3.5. Geometryczna interpretacja operacji tensorowych 58
- 2.3.6. Interpretacja geometryczna uczenia głębokiego 59
- 2.4. Silnik sieci neuronowych: optymalizacja gradientowa 60
 - 2.4.1. Czym jest pochodna? 61
 - 2.4.2. Pochodna operacji tensorowej: gradient 62
 - 2.4.3. Stochastyczny spadek wzdłuż gradientu 63
 - 2.4.4. Łączenie pochodnych: algorytm propagacji wstecznej 66
- 2.5. Ponowna analiza pierwszego przykładu 67
- 2.6. Podsumowanie rozdziału 68

Rozdział 3. Rozpoczynamy korzystanie z sieci neuronowych 71

- 3.1. Anatomia sieci neuronowej 72
 - 3.1.1. Warstwy: podstawowe bloki konstrukcyjne uczenia głębokiego 72
 - 3.1.2. Modele: sieci warstw 73
 - 3.1.3. Funkcja straty i optymalizatory: najważniejsze elementy konfiguracji procesu uczenia 74
- 3.2. Wprowadzenie do pakietu Keras 75
 - 3.2.1. Keras, TensorFlow, Theano i CNTK 76
 - 3.2.2. Instalowanie pakietu Keras 77
 - 3.2.3. Praca z pakietem Keras: krótkie wprowadzenie 77
- 3.3. Przygotowanie stacji roboczej do uczenia głębokiego 79
 - 3.3.1. Dwie opcje uruchamiania pakietu Keras 79
 - 3.3.2. Wady i zalety uruchamiania uczenia głębokiego w chmurze 80
 - 3.3.3. Jaki procesor graficzny najlepiej nadaje się do uczenia głębokiego? 80

3.4. Przykład klasyfikacji binarnej: klasyfikacja recenzji filmów	81
3.4.1. Zbiór danych IMDB	81
3.4.2. Przygotowywanie danych	82
3.4.3. Budowa sieci neuronowej	83
3.4.4. Walidacja modelu	87
3.4.5. Używanie wytrenowanej sieci do generowania przewidywań dotyczących nowych danych	90
3.4.6. Dalsze eksperymenty	90
3.4.7. Wnioski	91
3.5. Przykład klasyfikacji wieloklasowej: klasyfikacja krótkich artykułów prasowych	91
3.5.1. Zbiór danych Agencji Reutera	91
3.5.2. Przygotowywanie danych	93
3.5.3. Budowanie sieci	93
3.5.4. Walidacja modelu	94
3.5.5. Generowanie przewidywań dotyczących nowych danych	96
3.5.6. Inne sposoby obsługi etykiet i funkcji straty	97
3.5.7. Dlaczego warto tworzyć odpowiednio duże warstwy pośrednie?	97
3.5.8. Dalsze eksperymenty	98
3.5.9. Wnioski	98
3.6. Przykład regresji: przewidywanie cen mieszkań	99
3.6.1. Zbiór cen mieszkań w Bostonie	99
3.6.2. Przygotowywanie danych	100
3.6.3. Budowanie sieci	100
3.6.4. K-składowa walidacja krzyżowa	101
3.6.5. Wnioski	105
3.7. Podsumowanie rozdziału	105

Rozdział 4. Podstawy uczenia maszynowego	107
--	-----

- 4.1. Cztery rodzaje uczenia maszynowego 108
 - 4.1.1. Uczenie nadzorowane 108
 - 4.1.2. Uczenie nienadzorowane 108
 - 4.1.3. Uczenie częściowo nadzorowane 109
 - 4.1.4. Uczenie przez wzmacnianie 109
- 4.2. Ocena modeli uczenia maszynowego 109
 - 4.2.1. Zbiory treningowe, walidacyjne i testowe 111
 - 4.2.2. Rzeczy, o których warto pamiętać 114
- 4.3. Wstępna obróbka danych, przetwarzanie cech i uczenie cech 114
 - 4.3.1. Przygotowywanie danych do przetwarzania przez sieci neuronowe 115
 - 4.3.2. Przetwarzanie cech 116
- 4.4. Nadmierne dopasowanie i zbyt słabe dopasowanie 118
 - 4.4.1. Redukcja rozmiaru sieci 119
 - 4.4.2. Dodawanie regularyzacji wag 121
 - 4.4.3. Porzucanie - technika dropout 123
- 4.5. Uniwersalny przepływ roboczy uczenia maszynowego 125
 - 4.5.1. Definiowanie problemu i przygotowywanie zbioru danych 125
 - 4.5.2. Wybór miary sukcesu 126
 - 4.5.3. Określanie techniki oceny wydajności modelu 127
 - 4.5.4. Przygotowywanie danych 127
 - 4.5.5. Tworzenie modeli lepszych od linii bazowej 128
 - 4.5.6. Skalowanie w górę: tworzenie modelu, który ulega nadmiernemu dopasowaniu 129
 - 4.5.7. Regularyzacja modelu i dostrajanie jego hiperparametrów 129
- 4.6. Podsumowanie rozdziału 130

CZĘŚĆ II. UCZENIE GŁĘBOKIE W PRAKTYCE 131

Rozdział 5. Uczenie głębokie i przetwarzanie obrazu 133

5.1. Wprowadzenie do konwolucyjnych sieci neuronowych 134

5.1.1. Działanie sieci konwolucyjnej 136

5.1.2. Operacja max-pooling 141

5.2. Trenowanie konwolucyjnej sieci neuronowej na małym zbiorze danych 143

5.2.1. Stosowanie uczenia głębokiego w problemach małych zbiorów danych 144

5.2.2. Pobieranie danych 144

5.2.3. Budowa sieci neuronowej 147

5.2.4. Wstępna obróbka danych 148

5.2.5. Stosowanie techniki augmentacji danych 151

5.3. Korzystanie z wcześniej wytrenowanej konwolucyjnej sieci neuronowej 155

5.3.1. Ekstrakcja cech 155

5.3.2. Dostrajanie 163

5.3.3. Wnioski 167

5.4. Wizualizacja efektów uczenia konwolucyjnych sieci neuronowych 168

5.4.1. Wizualizacja pośrednich aktywacji 169

5.4.2. Wizualizacja filtrów konwolucyjnych sieci neuronowych 175

5.4.3. Wizualizacja map ciepła aktywacji klas 181

5.5. Podsumowanie rozdziału 185

Rozdział 6. Uczenie głębokie w przetwarzaniu tekstu i sekwencji 187

6.1. Praca z danymi tekstowymi 188

6.1.1. Kodowanie słów i znaków metodą gorącej jedynki 189

6.1.2. Osadzanie słów 192

6.1.3. Łączenie wszystkich technik: od surowego tekstu do osadzenia słów 197

6.1.4. Wnioski 203

6.2. Rekurencyjne sieci neuronowe 203

6.2.1. Warstwa rekurencyjna w pakiecie Keras	206
6.2.2. Warstwy LSTM i GRU	209
6.2.3. Przykład warstwy LSTM zaimplementowanej w pakiecie Keras	212
6.2.4. Wnioski	213
6.3. Zaawansowane zastosowania rekurencyjnych sieci neuronowych	214
6.3.1. Problem prognozowania temperatury	214
6.3.2. Przygotowywanie danych	217
6.3.3. Punkt odniesienia w postaci zdrowego rozsądku	220
6.3.4. Podstawowe rozwiązanie problemu przy użyciu techniki uczenia maszynowego	221
6.3.5. Punkt odniesienia w postaci pierwszego modelu rekurencyjnego	223
6.3.6. Stosowanie rekurencyjnego porzucania w celu zmniejszenia nadmiernego dopasowania	225
6.3.7. Tworzenie stosów warstw rekurencyjnych	226
6.3.8. Korzystanie z dwukierunkowych rekurencyjnych sieci neuronowych	228
6.3.9. Kolejne rozwiązania	232
6.3.10. Wnioski	233
6.4. Konwolucyjne sieci neuronowe i przetwarzanie sekwencji	234
6.4.1. Przetwarzanie sekwencji za pomocą jednowymiarowej sieci konwolucyjnej	234
6.4.2. Jednowymiarowe łączenie danych sekwencyjnych	235
6.4.3. Implementacja jednowymiarowej sieci konwolucyjnej	235
6.4.4. Łączenie sieci konwolucyjnych i rekurencyjnych w celu przetworzenia długich sekwencji	237
6.4.5. Wnioski	241
6.5. Podsumowanie rozdziału	242
Rozdział 7. Najlepsze zaawansowane praktyki uczenia głębokiego	245
7.1. Funkcjonalny interfejs programistyczny pakietu Keras: wykraczanie poza model sekwencyjny	246
7.1.1. Wprowadzenie do funkcjonalnego interfejsu API	247

7.1.2. Modele z wieloma wejściami	249
7.1.3. Modele z wieloma wyjściami	251
7.1.4. Skierowane acykliczne grafy warstw	254
7.1.5. Udostępnianie wag warstwy	258
7.1.6. Modele pełniące funkcję warstw	259
7.1.7. Wnioski	260
7.2. Monitorowanie modeli uczenia głębokiego przy użyciu wywołań zwrotnych pakietu Keras i narzędzia TensorBoard	260
7.2.1. Używanie wywołań zwrotnych w celu sterowania procesem trenowania modelu	260
7.2.2. Wprowadzenie do TensorBoard - platformy wizualizacji danych pakietu TensorFlow	264
7.2.3. Wnioski	268
7.3. Korzystanie z pełni możliwości modeli	268
7.3.1. Konstrukcja zaawansowanych architektur	269
7.3.2. Optymalizacja hiperparametru	272
7.3.3. Składanie modeli	274
7.3.4. Wnioski	276
7.4. Podsumowanie rozdziału	276
 Rozdział 8. Stosowanie uczenia głębokiego w celu generowania danych	 279
8.1. Generowanie tekstu za pomocą sieci LSTM	281
8.1.1. Krótka historia generatywnych sieci rekurencyjnych	281
8.1.2. Generowanie danych sekwencyjnych	282
8.1.3. Dlaczego strategia próbkowania jest ważna?	282
8.1.4. Implementacja algorytmu LSTM generującego tekst na poziomie liter	285
8.1.5. Wnioski	289
8.2. DeepDream	290
8.2.1. Implementacja algorytmu DeepDream w pakiecie Keras	291
8.2.2. Wnioski	296

8.3. Neuronowy transfer stylu 297

8.3.1. Strata treści 298

8.3.2. Strata stylu 298

8.3.3. Implementacja neuronowego transferu stylu przy użyciu pakietu Keras 299

8.3.4. Wnioski 304

8.4. Generowanie obrazów przy użyciu wariacyjnych autoenkoderów 306

8.4.1. Próbkowanie z niejawnej przestrzeni obrazów 306

8.4.2. Wektory koncepcyjne używane podczas edycji obrazu 307

8.4.3. Wariacyjne autoenkodery 308

8.4.4. Wnioski 314

8.5. Wprowadzenie do generatywnych sieci z przeciwnikiem 315

8.5.1. Schematyczna implementacja sieci GAN 316

8.5.2. Zbiór przydatnych rozwiązań 317

8.5.3. Generator 318

8.5.4. Dyskryminator 319

8.5.5. Sieć z przeciwnikiem 320

8.5.6. Trenowanie sieci DCGAN 320

8.5.7. Wnioski 322

8.6. Podsumowanie rozdziału 323

Rozdział 9. Wnioski 325

9.1. Przypomnienie najważniejszych koncepcji 326

9.1.1. Sztuczna inteligencja 326

9.1.2. Co sprawia, że uczenie głębokie to wyjątkowa dziedzina uczenia maszynowego? 326

9.1.3. Jak należy traktować uczenie głębokie? 327

9.1.4. Najważniejsze technologie 328

9.1.5. Uniwersalny przepływ roboczy uczenia maszynowego 329

9.1.6. Najważniejsze architektury sieci	330
9.1.7. Przestrzeń możliwości	334
9.2. Ograniczenia uczenia głębokiego	336
9.2.1. Ryzyko antropomorfizacji modeli uczenia maszynowego	337
9.2.2. Lokalne uogólnianie a ekstremalne uogólnianie	339
9.2.3. Wnioski	340
9.3. Przyszłość uczenia głębokiego	341
9.3.1. Modele jako programy	342
9.3.2. Wykraczanie poza algorytm propagacji wstecznej i warstwy różniczkowalne	343
9.3.3. Zautomatyzowane uczenie maszynowe	344
9.3.4. Nieustanne uczenie się i wielokrotne używanie modułowych procedur składowych	345
9.3.5. Przewidywania dotyczące dalekiej przyszłości	346
9.4. Bycie na bieżąco z nowościami związanymi z szybko rozwijającą się dziedziną	348
9.4.1. Zdobywaj wiedzę praktyczną, pracując z prawdziwymi problemami przedstawianymi w serwisie Kaggle	348
9.4.2. Czytaj o nowych rozwiązaniach w serwisie arXiv	348
9.4.3. Eksploruj ekosystem związany z pakietem Keras	349
9.5. Ostatnie słowa	349
 DODATKI	 351
 Dodatek A. Instalowanie pakietu Keras i innych bibliotek niezbędnych do jego działania w systemie Ubuntu	 353
 Dodatek B. Uruchamianie kodu w środowisku RStudio Server przy użyciu zdalnej instancji procesora graficznego EC2	 359
 Skorowidz	 365