

Spis treści

Wstęp 9

1. Wprowadzenie do uczenia głębokiego 11

Uczenie maszynowe pożera informatykę 11

Podstawowe elementy uczenia głębokiego 12

W pełni połączona warstwa 13

Warstwa splotowa 13

Warstwy rekurencyjnej sieci neuronowej 14

Komórki LSTM 15

Architektury uczenia głębokiego 15

LeNet 16

AlexNet 16

ResNet 17

Automatyczne generowanie opisów 18

Neuronowe tłumaczenie maszynowe firmy Google 18

Modele jednorazowe 19

AlfaGo 21

Generatywne sieci kontradiktoryjne 22

Neuronowe maszyny Turinga 23

Środowiska uczenia głębokiego 23

Ograniczenia TensorFlow 24

Podsumowanie 25

2. Wprowadzenie do podstawowych elementów TensorFlow 27

Poznajemy tensory 27

Skalary, wektory i macierze 28

Algebra macierzy 31

Tensory	33
Tensory w fizyce	34
Dygresje matematyczne	35
Proste obliczenia w TensorFlow	36
Instalacja TensorFlow i rozpoczęcie pracy	36
Inicjalizacja stałych tensorów	37
Próbkowanie losowych tensorów	38
Dodawanie i skalowanie tensorów	39
Operacje na macierzach	39
Typy tensorów	41
Manipulacje kształtem tensora	41
Wprowadzenie do rozgłaszania	42
Programowanie imperatywne i deklaratywne	43
Grafy TensorFlow	44
Sesje TensorFlow	45
Zmienne TensorFlow	45
Podsumowanie	47
3. Regresja liniowa i logistyczna z TensorFlow	49
Przegląd matematyczny	49
Funkcje i różniczkowalność	49
Funkcje straty	51
Metoda gradientu prostego	55
Systemy automatycznego różniczkowania	57
Uczenie z TensorFlow	59
Tworzenie ćwiczebnych zbiorów danych	59
Nowe koncepcje TensorFlow	64
Uczenie modeli liniowych i logistycznych w TensorFlow	68

Regresja liniowa w TensorFlow	68
Regresja logistyczna w TensorFlow	75
Podsumowanie	80
4. W pełni połączone sieci głębokie	81
Czym jest w pełni połączona sieć głęboka?	81
"Neurony" w sieciach w pełni połączonych	83
Uczenie w pełni połączonych sieci z propagacją wsteczną	85
Twierdzenie o uniwersalnej zbieżności	86
Dlaczego głębokie sieci?	87
Szkolenie w pełni połączonych sieci neuronowych	88
Reprezentacje możliwe do uczenia	88
Aktywacje	89
Zapamiętywanie w sieciach w pełni połączonych	89
Regularyzacja	90
Szkolenie sieci w pełni połączonych	93
Implementacja w TensorFlow	93
Instalacja DeepChem	93
Zbiór danych Tox21	94
Przyjmowanie minigrup węzłów zastępczych	95
Implementacja warstwy ukrytej	95
Dodawanie porzucania do warstwy ukrytej	96
Implementacja minigrup	97
Ocena dokładności modelu	97
Korzystanie z TensorBoard do śledzenia zbieżności modeli	98
Podsumowanie	99
5. Optymalizacja hiperparametrów	101

Ewaluacja modelu i optymalizacja hiperparametrów 102

Wskaźniki, wskaźniki, wskaźniki 103

Wskaźniki klasyfikacji binarnej 103

Wskaźniki klasyfikacji wieloklasowej 106

Wskaźniki regresji 107

Algorytmy optymalizacji hiperparametrów 108

Ustalenie linii bazowej 108

Metoda spadku studenta 110

Metoda przeszukiwania siatki 111

Losowe wyszukiwanie hiperparametrów 112

Zadanie dla czytelnika 113

Podsumowanie 113

6. Splotowe sieci neuronowe 115

Wprowadzenie do architektur splotowych 116

Lokalne pola recepcyjne 116

Jądra splotowe 118

Warstwy łączące 120

Tworzenie sieci splotowych 120

Rozszerzone warstwy splotowe 121

Zastosowania sieci splotowych 122

Wykrywanie i lokalizacja obiektów 122

Segmentacja obrazu 123

Sploty grafowe 123

Generowanie obrazów przy użyciu autokoderów wariacyjnych 124

Trenowanie sieci splotowej w TensorFlow 129

Zbiór danych MNIST 129

Wczytywanie zbioru MNIST 130

Podstawowe elementy sieci splotowych w TensorFlow 132

Architektura splotowa 134

Ewaluacja trenowanych modeli 137

Zadanie dla czytelnika 139

Podsumowanie 139

7. Rekurencyjne sieci neuronowe 141

Przegląd architektur rekurencyjnych 142

Komórki rekurencyjne 144

Długa pamięć krótkoterminowa (LSTM) 144

Bramkowane jednostki rekurencyjne (GRU) 146

Zastosowania modeli rekurencyjnych 146

Generowanie danych przez sieci rekurencyjne 146

Modele seq2seq 147

Neuronowe maszyny Turinga 149

Praca z rekurencyjnymi sieciami neuronowymi w praktyce 150

Przetwarzanie korpusu językowego Penn Treebank 151

Kod przetwarzania wstępnego 152

Wczytywanie danych do TensorFlow 154

Podstawowa architektura rekurencyjna 155

Zadanie dla czytelnika 157

Podsumowanie 157

8. Uczenie przez wzmacnianie 159

Procesy decyzyjne Markowa 163

Algorytmy uczenia przez wzmacnianie 164

Q-uczenie 165

Uczenie się polityki 166

Szkolenie asynchroniczne 168

Ograniczenia uczenia przez wzmacnianie 168

Gra w kółko i krzyżyk 170

Obiektowość 170

Abstrakcyjne środowisko 171

Środowisko gry w kółko i krzyżyk 171

Abstrakcja warstwowa 174

Definiowanie grafu warstw 176

Algorytm A3C 180

Funkcja straty A3C 183

Definiowanie wątków roboczych 185

Trenowanie polityki 187

Zadanie dla czytelnika 188

Podsumowanie 189

9. Szkolenie dużych głębokich sieci 191

Specjalistyczny sprzęt dla głębokich sieci 191

Szkolenie z użyciem CPU 192

Szkolenie z użyciem GPU 193

Procesory tensorowe 194

Bezpośrednio programowalne macierze bramek 195

Układy neuromorficzne 196

Rozproszone szkolenie głębokich sieci 197

Równoległość danych 197

Równoległość modeli 198

Szkolenie na równoległych danych z użyciem wielu układów GPU na zbiorze CIFAR10 199

Pobieranie i wczytywanie danych 201

Głębokie zanurzenie w architekturę 202

Szkolenie na wielu układach GPU 204

Zadanie dla czytelnika 206

Podsumowanie 207

10. Przyszłość głębokiego uczenia 209

Głębokie uczenie poza branżą techniczną 209

Głębokie uczenie w przemyśle farmaceutycznym 210

Głębokie uczenie w prawie 211

Głębokie uczenie w robotyce 211

Głębokie uczenie w rolnictwie 212

Etyczne wykorzystanie głębokiego uczenia 212

Czy uniwersalna sztuczna inteligencja jest nieuchronna? 214

Co dalej? 214

Skorowidz 216