

Spis treści

Wprowadzenie 9

1. Podstawowe zagadnienia 15

Funkcje 16

Matematyka 16

Diagramy 16

Kod 17

Pochodne 20

Matematyka 20

Diagramy 20

Kod 21

Funkcje zagnieżdżone 22

Diagram 22

Matematyka 22

Kod 23

Inny diagram 23

Reguła łańcuchowa 24

Matematyka 24

Diagram 24

Kod 25

Nieco dłuższy przykład 27

Matematyka 27

Diagram 27

Kod 28

Funkcje z wieloma danymi wejściowymi 29

Matematyka 30

Diagram 30

Kod 30

Pochodne funkcji z wieloma danymi wejściowymi 31

Diagram 31

Matematyka 31

Kod 32

Funkcje przyjmujące wiele wektorów jako dane wejściowe 32

Matematyka 33

Tworzenie nowych cech na podstawie istniejących 33

Matematyka 33

Diagram 33

Kod 34

Pochodne funkcji z wieloma wektorami wejściowymi 35

Diagram 35

Matematyka 36

Kod 36

Następny etap - funkcje wektorowe i ich pochodne 37

Diagram 37

Matematyka 37

Kod 38

Funkcje wektorowe i ich pochodne w kroku wstecz 38

Grafy obliczeniowe z danymi wejściowymi w postaci dwóch macierzy dwuwymiarowych 40

Matematyka 41

Diagram 43

Kod 43

Ciekawa część - krok wstecz 43

Diagram 44

Matematyka 44

Kod 46

Podsumowanie 50

2. Wprowadzenie do budowania modeli 51

Wstęp do uczenia nadzorowanego 52

Modele uczenia nadzorowanego 53

Regresja liniowa 55

Regresja liniowa - diagram 55

Regresja liniowa - bardziej pomocny diagram (i obliczenia matematyczne) 57

Dodawanie wyrazu wolnego 58

Regresja liniowa - kod 58

Uczenie modelu 59

Obliczanie gradientów - diagram 59

Obliczanie gradientów - matematyka (i trochę kodu) 60

Obliczanie gradientów - (kompletny) kod 61

Używanie gradientów do uczenia modelu 62

Ocena modelu - testowe i treningowe zbiory danych 63

Ocena modelu - kod 63

Analizowanie najważniejszej cechy 65

Budowanie sieci neuronowych od podstaw 66

Krok 1. Zestaw regresji liniowych 67

Krok 2. Funkcja nieliniowa 67

Krok 3. Inna regresja liniowa 68

Diagramy 68

Kod 70

Sieci neuronowe - krok wstecz 71

Uczenie i ocena pierwszej sieci neuronowej 73

Dwa powody, dla których nowy model jest lepszy 74

Podsumowanie 75

3. Deep learning od podstaw 77

Definicja procesu deep learning - pierwszy krok 77

Elementy sieci neuronowych - operacje 79

Diagram 79

Kod 80

Elementy sieci neuronowych - warstwy 82

Diagramy 82

Elementy z elementów 84

Wzorzec warstwy 86

Warstwa gęsta 88

Klasa NeuralNetwork (i ewentualnie inne) 89

Diagram 89

Kod 90

Klasa Loss 90

Deep learning od podstaw 92

Implementowanie treningu na porcjach danych 92

Klasa NeuralNetwork - kod 93

Nauczyciel i optymalizator 95

Optymalizator 95

Nauczyciel 97

Łączenie wszystkich elementów 98

Pierwszy model z dziedziny deep learning (napisany od podstaw) 99

Podsumowanie i dalsze kroki 100

4. Rozszerzenia 101

Intuicyjne rozważania na temat sieci neuronowych 102

Funkcja straty - funkcja softmax z entropią krzyżową 104

Komponent nr 1. Funkcja softmax 104

Komponent nr 2. Entropia krzyżowa 105

Uwaga na temat funkcji aktywacji 108

Eksperymenty 111

Wstępne przetwarzanie danych 111

Model 112

Eksperyment: wartość straty z użyciem funkcji softmax z entropią krzyżową 113

Współczynnik momentum 113

Intuicyjny opis współczynnika momentum 114

Implementowanie współczynnika momentum w klasie Optimizer 114

Eksperyment - algorytm SGD ze współczynnikiem momentum 116

Zmniejszanie współczynnika uczenia 116

Sposoby zmniejszania współczynnika uczenia 116

Eksperymenty - zmniejszanie współczynnika uczenia 118

Inicjowanie wag 119

Matematyka i kod 120

Eksperymenty - inicjowanie wag 121

Dropout 122

Definicja 122

Implementacja 122

Eksperymenty - dropout 123

Podsumowanie 125

5. Konwolucyjne sieci neuronowe 127

Sieci neuronowe i uczenie reprezentacji 127

Inna architektura dla danych graficznych 128

Operacja konwolucji 129

Wielokanałowa operacja konwolucji 131

Warstwy konwolucyjne 131

Wpływ na implementację 132

Różnice między warstwami konwolucyjnymi a warstwami gęstymi 133

Generowanie predykcji z użyciem warstw konwolucyjnych - warstwa spłaszczenia 134

Warstwy agregujące 135

Implementowanie wielokanałowej operacji konwolucji 137

Krok w przód 137

Konwolucja - krok wstecz 140

Porcje danych, konwolucje dwuwymiarowe i operacje wielokanałowe 144

Konwolucje dwuwymiarowe 145

Ostatni element - dodawanie kanałów 147

Używanie nowej operacji do uczenia sieci CNN 150

Operacja Flatten 150

Kompletna warstwa Conv2D 151

Eksperymenty 152

Podsumowanie 153

6. Rekurencyjne sieci neuronowe 155

Najważniejsze ograniczenie - przetwarzanie odgałęzień 156

Automatyczne różniczkowanie 158

Pisanie kodu do akumulowania gradientów 158

Powody stosowania sieci RNN 162

Wprowadzenie do sieci RNN 163

Pierwsza klasa dla sieci RNN - RNNLayer 164

Druga klasa dla sieci RNN - RNNNode 165

Łączenie obu klas 166

Krok wstecz 167

Sieci RNN - kod 169

Klasa RNNLayer 170

Podstawowe elementy sieci RNNNode 172

Zwykłe węzły RNNNode 173

Ograniczenia zwykłych węzłów RNNNode 175

Pierwsze rozwiązanie - węzły GRUNode 176

Węzły LSTMNode 179

Reprezentacja danych dla opartego na sieci RNN modelu języka naturalnego na poziomie znaków 182

Inne zadania z obszaru modelowania języka naturalnego 182

Łączenie odmian warstw RNNLayer 183

Łączenie wszystkich elementów 184

Podsumowanie 185

7. PyTorch 187

Typ Tensor w bibliotece PyTorch 187

Deep learning z użyciem biblioteki PyTorch 188

Elementy z biblioteki PyTorch - klasy reprezentujące model, warstwę, optymalizator i wartość straty 189

Implementowanie elementów sieci neuronowej za pomocą biblioteki PyTorch - warstwa DenseLayer 190

Przykład - modelowanie cen domów w Bostonie z użyciem biblioteki PyTorch 191

Elementy oparte na bibliotece PyTorch - klasy optymalizatora i wartości straty 192

Elementy oparte na bibliotece PyTorch - klasa nauczyciela 193

Sztuczki służące do optymalizowania uczenia w bibliotece PyTorch 195

Sieci CNN w bibliotece PyTorch 196

Klasa DataLoader i transformacje 198

Tworzenie sieci LSTM za pomocą biblioteki PyTorch 200

Postscriptum - uczenie nienadzorowane z użyciem autoenkoderów 202

Uczenie reprezentacji 203

Podejście stosowane w sytuacjach, gdy w ogóle nie ma etykiet 203

Implementowanie autoenkodera za pomocą biblioteki PyTorch 204

Trudniejszy test uczenia nienadzorowanego i rozwiązanie 209

Podsumowanie 210

A. Skok na głęboką wodę 211

Reguła łańcuchowa dla macierzy 211

Gradient dla wartości straty względem wyrazu wolnego 215

Konwolucje z użyciem mnożenia macierzy 215