

Spis treści

Przedmowa	15
-----------------	----

Część I. Podstawy uczenia maszynowego 25

1. Krajobraz uczenia maszynowego 27

Czym jest uczenie maszynowe?	28
Dlaczego warto korzystać z uczenia maszynowego?	28
Przykładowe zastosowania	31
Rodzaje systemów uczenia maszynowego	33
Nadzorowanie uczenia	33
Uczenie wsadowe i uczenie przyrostowe	40
Uczenie z przykładów i uczenie z modelu	44
Główne problemy uczenia maszynowego	49
Niedobór danych uczących	50
Niereprezentatywne dane uczące	50
Dane kiepskiej jakości	52
Nieistotne cechy	52
Przetrenowanie danych uczących	53
Niedotrenowanie danych uczących	55
Podsumowanie	55
Testowanie i ocenianie	56
Strojenie hiperparametrów i dobór modelu	56
Niegodność danych	57
Ćwiczenia	59

2. Nasz pierwszy projekt uczenia maszynowego 61

Praca z rzeczywistymi danymi	61
Przeanalizuj całokształt projektu	63
Określ zakres problemu	63
Wybierz wskaźnik wydajności	65
Sprawdź założenia	67

Zdobądź dane	67
Uruchom przykładowy kod w serwisie Google Colab	68
Zapisz zmiany w kodzie i w danych	70
Zalety i wady interaktywności	71
Kod w książce a kod w notatnikach Jupyter	72
Pobierz dane	72
Rzut oka na strukturę danych	73
Stwórz zbiór testowy	77
Odkrywaj i wizualizuj dane, aby zdobywać nowe informacje	81
Wizualizuj dane geograficzne	81
Poszukaj korelacji	84
Eksperymentuj z kombinacjami atrybutów	86
Przygotuj dane pod algorytmy uczenia maszynowego	87
Oczyść dane	88
Obsługa tekstu i atrybutów kategoryalnych	90
Skalowanie i przekształcanie cech	94
Niestandardowe transformatory	97
Potoki transformujące	101
Wybierz i wytrenuj model	105
Trenuj i oceń model za pomocą zbioru uczącego	106
Dokładniejsze ocenianie za pomocą sprawdzianu krzyżowego	107
Wyreguluj swój model	109
Metoda przeszukiwania siatki	109
Metoda losowego przeszukiwania	111
Metody zespołowe	112
Analizowanie najlepszych modeli i ich błędów	112
Oceń system za pomocą zbioru testowego	113
Uruchom, monitoruj i utrzymuj swój system	114
Teraz Twoja kolej!	117
Ćwiczenia	117
3. Klasyfikacja	119
Zbiór danych MNIST	119
Uczenie klasyfikatora binarnego	122
Miary wydajności	122
Pomiar dokładności za pomocą sprawdzianu krzyżowego	123
Macierz pomyłek	124
Precyzja i pełność	125
Kompromis pomiędzy precyzją a pełnością	127
Wykres krzywej ROC	130

Klasyfikacja wieloklasowa	134
Analiza błędów	136
Klasyfikacja wieloetykietowa	140
Klasyfikacja wielowyjściowa	141
Ćwiczenia	143
4. Uczenie modeli	145
Regresja liniowa	146
Równanie normalne	148
Złożoność obliczeniowa	151
Gradient prosty	151
Wsadowy gradient prosty	154
Stochastyczny spadek wzdłuż gradientu	157
Schodzenie po gradiencie z minigrupami	160
Regresja wielomianowa	161
Krzywe uczenia	163
Regularyzowane modele liniowe	167
Regresja grzbietowa	167
Regresja metodą LASSO	170
Regresja metodą elastycznej siatki	172
Wczesne zatrzymywanie	173
Regresja logistyczna	174
Szacowanie prawdopodobieństwa	175
Funkcje ucząca i kosztu	176
Granice decyzyjne	177
Regresja softmax	180
Ćwiczenia	183
5. Maszyny wektorów nośnych	185
Liniowa klasyfikacja SVM	185
Klasyfikacja miękkiego marginesu	186
Nieliniowa klasyfikacja SVM	188
Jądro wielomianowe	189
Cechy podobieństwa	190
Gaussofskie jądro RBF	191
Klasy SVM i złożoność obliczeniowa	192
Regresja SVM	193
Mechanizm działania liniowych klasyfikatorów SVM	195
Problem dualny	198
Kernelizowane maszyny SVM	199
Ćwiczenia	202

6. Drzewa decyzyjne	203
Uczenie i wizualizowanie drzewa decyzyjnego	203
Wyliczanie prognoz	204
Szacowanie prawdopodobieństw przynależności do klas	206
Algorytm uczący CART	207
Złożoność obliczeniowa	208
Wskaźnik Giniego czy entropia?	208
Hiperparametry regularyzacyjne	209
Regresja	211
Wrażliwość na orientację osi	213
Drzewa decyzyjne mają znaczną wariancję	214
Ćwiczenia	215
7. Uczenie zespołowe i losowe lasy	217
Klasyfikatory głosujące	217
Agregacja i wklejanie	221
Agregacja i wklejanie w module Scikit-Learn	222
Ocena OOB	223
Rejony losowe i podprzestrzenie losowe	224
Losowe lasy	225
Zespół Extra-Trees	225
Istotność cech	226
Wzmacnianie	227
AdaBoost	227
Wzmacnianie gradientowe	231
Wzmacnianie gradientu w oparciu o histogram	234
Kontaminacja	235
Ćwiczenia	238
8. Redukcja wymiarowości	240
Kłątwa wymiarowości	241
Główne strategie redukcji wymiarowości	242
Rzutowanie	242
Uczenie rozmaitościowe	244
Analiza PCA	245
Zachowanie wariancji	246
Główne składowe	247
Rzutowanie na d wymiarów	248
Implementacja w module Scikit-Learn	249
Współczynnik wariancji wyjaśnionej	249
Wybór właściwej liczby wymiarów	249

Algorytm PCA w zastosowaniach kompresji	251
Losowa analiza PCA	252
Przyrostowa analiza PCA	252
Rzutowanie losowe	254
Algorytm LLE	256
Inne techniki redukowania wymiarowości	258
Ćwiczenia	259

9. Techniki uczenia nienadzorowanego261

Analiza skupień: algorytm centroidów i DBSCAN	262
Algorytm centroidów	264
Granice algorytmu centroidów	273
Analiza skupień w segmentacji obrazu	274
Analiza skupień w uczeniu półnadzorowanym	276
Algorytm DBSCAN	279
Inne algorytmy analizy skupień	282
Mieszanki gaussowskie	283
Wykrywanie anomalii za pomocą mieszanin gaussowskich	287
Wyznaczanie liczby skupień	289
Bayesowskie modele mieszane	291
Inne algorytmy służące do wykrywania anomalii i nowości	292
Ćwiczenia	293

Część II. Sieci neuronowe i uczenie głębokie 295

10. Wprowadzenie do sztucznych sieci neuronowych i ich implementacji z użyciem interfejsu Keras297

Od biologicznych do sztucznych neuronów	298
Neurony biologiczne	299
Operacje logiczne przy użyciu neuronów	300
Perceptron	301
Perceptron wielowarstwowy i propagacja wsteczna	306
Regresyjne perceptrony wielowarstwowe	309
Klasyfikacyjne perceptrony wielowarstwowe	311
Implementowanie perceptronów wielowarstwowych za pomocą interfejsu Keras	313
Tworzenie klasyfikatora obrazów za pomocą interfejsu sekwencyjnego	314
Tworzenie regresyjnego perceptronu wielowarstwowego za pomocą interfejsu sekwencyjnego	323
Tworzenie złożonych modeli za pomocą interfejsu funkcyjnego	324
Tworzenie modeli dynamicznych za pomocą interfejsu podklasowego	329
Zapisywanie i odczytywanie modelu	331

Stosowanie wywołań zwrotnych	332
Wizualizacja danych za pomocą narzędzia TensorBoard	333
Dostrajanie hiperparametrów sieci neuronowej	337
Liczba warstw ukrytych	341
Liczba neuronów w poszczególnych warstwach ukrytych	342
Współczynnik uczenia, rozmiar grupy i pozostałe hiperparametry	343
Ćwiczenia	345
11. Uczenie głębokich sieci neuronowych	348
Problemy zanikających/eksplodujących gradientów	348
Inicjalizacje wag Glorota i He	349
Lepsze funkcje aktywacji	351
Normalizacja wsadowa	358
Obcinanie gradientu	363
Wielokrotne stosowanie gotowych warstw	363
Uczenie transferowe w interfejsie Keras	365
Nienadzorowane uczenie wstępne	366
Uczenie wstępne za pomocą dodatkowego zadania	367
Szybsze optymalizatory	368
Optymalizacja momentum	369
Przyspieszony spadek wzdłuż gradientu (algorytm Nesterova)	370
AdaGrad	371
RMSProp	373
Optymalizator Adam	373
AdaMax	374
Nadam	375
AdamW	375
Harmonogramowanie współczynnika uczenia	377
Regularyzacja jako sposób zapobiegania przetrenowaniu	381
Regularyzacja ℓ_1 i ℓ_2	381
Porzucanie	382
Regularyzacja typu Monte Carlo (MC)	385
Regularyzacja typu max-norm	387
Podsumowanie i praktyczne wskazówki	388
Ćwiczenia	390
12. Modele niestandardowe i uczenie za pomocą modułu TensorFlow	391
Krótkie omówienie modułu TensorFlow	391
Korzystanie z modułu TensorFlow jak z biblioteki NumPy	394
Tensory i operacje	395
Tensory a biblioteka NumPy	396

Konwersje typów	397
Zmienne	397
Inne struktury danych	398
Dostosowywanie modeli i algorytmów uczenia	399
Niestandardowe funkcje straty	399
Zapisywanie i wczytywanie modeli zawierających elementy niestandardowe	400
Niestandardowe funkcje aktywacji, inicjalizatory, regularizatory i ograniczenia	402
Niestandardowe wskaźniki	403
Niestandardowe warstwy	405
Niestandardowe modele	408
Funkcje straty i wskaźniki oparte na elementach wewnętrznych modelu	410
Obliczanie gradientów za pomocą różniczkowania automatycznego	411
Niestandardowe pętle uczenia	415
Funkcje i grafy modułu TensorFlow	417
AutoGraph i kreślenie	419
Reguły związane z funkcją TF	421
Ćwiczenia	422
13. Wczytywanie i wstępne przetwarzanie danych za pomocą modułu TensorFlow	424
Interfejs tf.data	425
Łączenie przekształceń	426
Tasowanie danych	428
Przeplatanie wierszy z różnych plików	429
Wstępne przetwarzanie danych	430
Składanie wszystkiego w całość	431
Pobieranie wstępne	432
Stosowanie zestawu danych z interfejsem Keras	434
Format TFRecord	435
Skompresowane pliki TFRecord	436
Wprowadzenie do buforów protokołów	436
Bufory protokołów w module TensorFlow	438
Wczytywanie i analizowanie składni obiektów Example	439
Obsługa list list za pomocą bufora protokołów SequenceExample	440
Warstwy przetwarzania wstępnego Keras	441
Warstwa Normalization	441
Warstwa Discretization	444
Warstwa CategoryEncoding	444
Warstwa StringLookup	446
Warstwa Hashing	447
Kodowanie cech kategoryalnych za pomocą wektorów właściwościowych	447
Wstępne przetwarzanie tekstu	451

Korzystanie z wytrenowanych składników modelu językowego	453
Warstwy wstępnego przetwarzania obrazów	454
Projekt TensorFlow Datasets (TFDS)	455
Ćwiczenia	456
14. Głębokie widzenie komputerowe za pomocą splotowych sieci neuronowych	458
Struktura kory wzrokowej	459
Warstwy splotowe	460
Filtry	462
Stosy map cech	463
Implementacja warstw splotowych w interfejsie Keras	465
Zużycie pamięci operacyjnej	468
Warstwa łącząca	469
Implementacja warstw łączących w interfejsie Keras	471
Architektury splotowych sieci neuronowych	473
LeNet-5	475
AlexNet	476
GoogLeNet	479
VGGNet	482
ResNet	482
Xception	486
SENet	487
Inne interesujące struktury	489
Wybór właściwej struktury CNN	491
Implementacja sieci ResNet-34 za pomocą interfejsu Keras	492
Korzystanie z gotowych modeli w interfejsie Keras	493
Gotowe modele w uczeniu transferowym	494
Klasyfikowanie i lokalizowanie	497
Wykrywanie obiektów	499
W pełni połączone sieci splotowe	501
Sieć YOLO	503
Śledzenie obiektów	506
Segmentacja semantyczna	507
Ćwiczenia	510
15. Przetwarzanie sekwencji za pomocą sieci rekurencyjnych i splotowych	512
Neurony i warstwy rekurencyjne	513
Komórki pamięci	515
Sekwencje wejść i wyjść	515
Uczenie sieci rekurencyjnych	517

Prognozowanie szeregów czasowych	517
Rodzina modeli ARMA	522
Przygotowywanie danych dla modeli uczenia maszynowego	525
Prognozowanie za pomocą modelu liniowego	528
Prognozowanie za pomocą prostej sieci rekurencyjnej	529
Prognozowanie za pomocą głębokich sieci rekurencyjnych	530
Prognozowanie wielowymiarowych szeregów czasowych	531
Prognozowanie kilka taktów w przód	532
Prognozowanie za pomocą modelu sekwencyjnego	534
Obsługa długich sekwencji	537
Zwalczanie problemu niestabilnych gradientów	537
Zwalczanie problemu pamięci krótkotrwałej	540
Ćwiczenia	547
16. Przetwarzanie języka naturalnego	
 za pomocą sieci rekurencyjnych i mechanizmów uwagi	549
Generowanie tekstów szekspirowskich za pomocą znakowej sieci rekurencyjnej	550
Tworzenie zestawu danych uczących	551
Budowanie i uczenie modelu char-RNN	553
Generowanie sztucznego tekstu szekspirowskiego	554
Stanowe sieci rekurencyjne	555
Analiza opinii	558
Maskowanie	560
Korzystanie z gotowych reprezentacji właściwościowych i modeli językowych	563
Sieć typu koder – dekodek służąca do neuronowego tłumaczenia maszynowego	565
Dwukierunkowe sieci rekurencyjne	571
Przeszukiwanie wiązkowe	572
Mechanizmy uwagi	574
Liczy się tylko uwaga: pierwotna architektura transformatora	578
Zatrzęsienie modeli transformatorów	587
Transformatory wizualne	592
Biblioteka Transformers firmy Hugging Face	596
Ćwiczenia	599
17. Autokodery, generatywne sieci przeciwstawne i modele rozpraszające	601
Efektywne reprezentacje danych	602
Analiza PCA za pomocą niedopełnionego autokodera liniowego	604
Autokodery stosowe	605
Implementacja autokodera stosowego za pomocą interfejsu Keras	606
Wizualizowanie rekonstrukcji	607
Wizualizowanie zestawu danych Fashion MNIST	608

Nienadzorowane uczenie wstępne za pomocą autokoderów stosowych	609
Wiązanie wag	610
Uczenie autokoderów pojedynczo	611
Autokodery spłotowe	613
Autokodery odszumiające	614
Autokodery rzadkie	615
Autokodery wariacyjne	618
Generowanie obrazów Fashion MNIST	622
Generatywne sieci przeciwstawne	623
Problemy związane z uczeniem sieci GAN	627
Głębokie spłotowe sieci GAN	628
Rozrost progresywny sieci GAN	631
Sieci StyleGAN	634
Modele rozpraszające	636
Ćwiczenia	642
18. Uczenie przez wzmacnianie	644
Uczenie się optymalizowania nagród	645
Wyszukiwanie strategii	646
Wprowadzenie do narzędzia OpenAI Gym	648
Sieci neuronowe jako strategie	652
Ocenianie czynności: problem przypisania zasługi	654
Gradyenty strategii	655
Procesy decyzyjne Markowa	659
Uczenie metodą różnic czasowych	663
Q-uczenie	664
Strategie poszukiwania	665
Przybliżający algorytm Q-uczenia i Q-uczenie głębokie	666
Implementacja modelu Q-uczenia głębokiego	667
Odmiany Q-uczenia głębokiego	671
Ustalone Q-wartości docelowe	671
Podwójna sieć DQN	672
Odtwarzanie priorytetowych doświadczeń	672
Walcząca sieć DQN	673
Przegląd popularnych algorytmów RN	674
Ćwiczenia	678
19. Wielkoskalowe uczenie i wdrażanie modeli TensorFlow	679
Eksploracja modelu TensorFlow	680
Korzystanie z systemu TensorFlow Serving	680
Tworzenie usługi predykcyjnej na platformie Vertex AI	688
Wykonywanie zadań predykcji wsadowych w usłudze Vertex AI	695

Wdrażanie modelu na urządzeniu mobilnym lub wbudowanym	697
Przetwarzanie modelu na stronie internetowej	699
Przyspieszanie obliczeń za pomocą procesorów graficznych	701
Zakup własnej karty graficznej	702
Zarządzanie pamięcią operacyjną karty graficznej	704
Umieszczanie operacji i zmiennych na urządzeniach	706
Przetwarzanie równoległe na wielu urządzeniach	708
Uczenie modeli za pomocą wielu urządzeń	710
Zrównoleglanie modelu	710
Zrównoleglanie danych	712
Uczenie wielkoskalowe za pomocą interfejsu strategii rozpraszania	718
Uczenie modelu za pomocą klastra TensorFlow	720
Realizowanie dużych grup zadań uczenia za pomocą usługi Vertex AI	723
Strojenie hiperparametrów w usłudze Vertex AI	725
Ćwiczenia	728
Dziękuję!	729
A Lista kontrolna projektu uczenia maszynowego	731
B Różniczkowanie automatyczne	736
C Specjalne struktury danych	744
D Grafy TensorFlow	750
Skorowidz	757