

# Spis treści

|                                                                                     |           |
|-------------------------------------------------------------------------------------|-----------|
| <b>Informacje o autorach</b>                                                        | <b>13</b> |
| <b>Informacje o recenzentach</b>                                                    | <b>15</b> |
| <b>Wstęp</b>                                                                        | <b>17</b> |
| <b>Rozdział 1. Umożliwianie komputerom uczenia się z danych</b>                     | <b>27</b> |
| <b>Tworzenie inteligentnych maszyn służących do przekształcania danych w wiedzę</b> | <b>28</b> |
| <b>Trzy różne rodzaje uczenia maszynowego</b>                                       | <b>28</b> |
| Przewidywanie przyszłości za pomocą uczenia nadzorowanego                           | 29        |
| Rozwiązywanie problemów interaktywnych<br>za pomocą uczenia przez wzmacnianie       | 32        |
| Odkrywanie ukrytych struktur za pomocą uczenia nienadzorowanego                     | 33        |
| <b>Wprowadzenie do podstawowej terminologii i notacji</b>                           | <b>35</b> |
| Notacja i konwencje używane w niniejszej książce                                    | 35        |
| Terminologia uczenia maszynowego                                                    | 37        |
| <b>Strategia tworzenia systemów uczenia maszynowego</b>                             | <b>38</b> |
| Wstępne przetwarzanie — nadawanie danym formy                                       | 38        |
| Trenowanie i dobór modelu predykcyjnego                                             | 39        |
| Ewaluacja modeli i przewidywanie wystąpienia nieznanych danych                      | 40        |
| <b>Wykorzystywanie środowiska Python do uczenia maszynowego</b>                     | <b>40</b> |
| Instalacja środowiska Python i pakietów z repozytorium Python Package Index         | 41        |
| Korzystanie z platformy Anaconda i menedżera pakietów                               | 41        |
| Pakiety przeznaczone do obliczeń naukowych,<br>analizy danych i uczenia maszynowego | 42        |
| <b>Podsumowanie</b>                                                                 | <b>42</b> |

|                                                                                                              |                |
|--------------------------------------------------------------------------------------------------------------|----------------|
| <b>Rozdział 2. Trenowanie prostych algorytmów uczenia maszynowego w celach klasyfikacji</b>                  | <b>45</b>      |
| <b>Sztuczne neurony — rys historyczny początków uczenia maszynowego</b>                                      | <b>46</b>      |
| Definicja formalna sztucznego neuronu                                                                        | 47             |
| Reguła uczenia perceptronu                                                                                   | 49             |
| <b>Implementacja algorytmu uczenia perceptronu w Pythonie</b>                                                | <b>51</b>      |
| Obiektowy interfejs API perceptronu                                                                          | 51             |
| Trenowanie modelu perceptronu na zestawie danych Iris                                                        | 54             |
| <b>Adaptacyjne neurony liniowe i zbieżność uczenia</b>                                                       | <b>60</b>      |
| Minimalizacja funkcji kosztu za pomocą metody gradientu prostego                                             | 61             |
| Implementacja algorytmu Adaline w Pythonie                                                                   | 63             |
| Usprawnianie gradientu prostego poprzez skalowanie cech                                                      | 67             |
| Wielkoskalowe uczenie maszynowe                                                                              |                |
| i metoda stochastycznego spadku wzdłuż gradientu                                                             | 69             |
| <b>Podsumowanie</b>                                                                                          | <b>73</b>      |
| <br><b>Rozdział 3. Stosowanie klasyfikatorów uczenia maszynowego za pomocą biblioteki scikit-learn</b>       | <br><b>75</b>  |
| <b>Wybór algorytmu klasyfikującego</b>                                                                       | <b>76</b>      |
| <b>Pierwsze kroki z biblioteką scikit-learn — uczenie perceptronu</b>                                        | <b>76</b>      |
| <b>Modelowanie prawdopodobieństwa przynależności do klasy za pomocą regresji logistycznej</b>                | <b>82</b>      |
| Regresja logistyczna i prawdopodobieństwo warunkowe                                                          | 82             |
| Wyznaczanie wag logistycznej funkcji kosztu                                                                  | 86             |
| Przekształcanie implementacji Adaline do postaci algorytmu regresji logistycznej                             | 88             |
| Uczenie modelu regresji logistycznej za pomocą biblioteki scikit-learn                                       | 92             |
| Zapobieganie przetrenowaniu za pomocą regularyzacji                                                          | 94             |
| <b>Wyznaczanie maksymalnego marginesu za pomocą maszyn wektorów nośnych</b>                                  | <b>97</b>      |
| Teoretyczne podłoże maksymalnego marginesu                                                                   | 98             |
| Rozwiązywanie przypadków nieliniowo rozdzielnych za pomocą zmiennych uzupełniających                         | 99             |
| Alternatywne implementacje w interfejsie scikit-learn                                                        | 101            |
| <b>Rozwiązywanie nieliniowych problemów za pomocą jądra SVM</b>                                              | <b>102</b>     |
| Metody jądrowe dla danych nierozdzielnych liniowo                                                            | 102            |
| Stosowanie sztuczki z funkcją jądra do znajdowania przestrzeni rozdzielających w przestrzeni wielowymiarowej | 104            |
| <b>Uczenie drzew decyzyjnych</b>                                                                             | <b>107</b>     |
| Maksymalizowanie przyrostu informacji — osiąganie jak największych korzyści                                  | 108            |
| Budowanie drzewa decyzyjnego                                                                                 | 112            |
| Łączenie wielu drzew decyzyjnych za pomocą modelu losowego lasu                                              | 115            |
| <b>Algorytm k-najbliższych sąsiadów — model leniwego uczenia</b>                                             | <b>119</b>     |
| <b>Podsumowanie</b>                                                                                          | <b>122</b>     |
| <br><b>Rozdział 4. Tworzenie dobrych zestawów danych uczących — wstępne przetwarzanie danych</b>             | <br><b>125</b> |
| <b>Kwestia brakujących danych</b>                                                                            | <b>126</b>     |
| Wykrywanie brakujących wartości w danych tabelarycznych                                                      | 126            |
| Usuwanie przykładów uczących lub cech niezawierających wartości                                              | 127            |

|                                                                                   |                |
|-----------------------------------------------------------------------------------|----------------|
| Wstawianie brakujących danych                                                     | 128            |
| Estymatory interfejsu scikit-learn                                                | 129            |
| <b>Przetwarzanie danych kategoryalnych</b>                                        | <b>130</b>     |
| Kodowanie danych kategoryalnych za pomocą biblioteki pandas                       | 131            |
| Mapowanie cech porządkowych                                                       | 131            |
| Kodowanie etykiet klas                                                            | 132            |
| Kodowanie „gorącojedynekowe” cech nominalnych<br>(z użyciem wektorów własnych)    | 133            |
| <b>Rozdzielanie zestawu danych na oddzielne podzbiory uczący i testowy</b>        | <b>136</b>     |
| <b>Skalowanie cech</b>                                                            | <b>138</b>     |
| <b>Dobór odpowiednich cech</b>                                                    | <b>140</b>     |
| Regularyzacje L1 i L2 jako kary ograniczające złożoność modelu                    | 141            |
| Interpretacja geometryczna regularyzacji L2                                       | 141            |
| Rozwiązania rzadkie za pomocą regularyzacji L1                                    | 143            |
| Algorytmy sekwencyjnego wyboru cech                                               | 146            |
| <b>Ocenianie istotności cech za pomocą algorytmu losowego lasu</b>                | <b>151</b>     |
| <b>Podsumowanie</b>                                                               | <b>154</b>     |
| <br><b>Rozdział 5. Kompresja danych poprzez redukcję wymiarowości</b>             | <br><b>155</b> |
| <b>Nienadzorowana redukcja wymiarowości za pomocą analizy głównych składowych</b> | <b>156</b>     |
| Podstawowe etapy analizy głównych składowych                                      | 156            |
| Wydobywanie głównych składowych krok po kroku                                     | 158            |
| Wyjaśniona wariancja całkowita                                                    | 160            |
| Transformacja cech                                                                | 161            |
| Analiza głównych składowych w interfejsie scikit-learn                            | 164            |
| <b>Nadzorowana kompresja danych za pomocą liniowej analizy dyskryminacyjnej</b>   | <b>167</b>     |
| Porównanie analizy głównych składowych z liniową analizą dyskryminacyjną          | 167            |
| Wewnętrzne mechanizmy działania liniowej analizy dyskryminacyjnej                 | 169            |
| Obliczanie macierzy rozproszenia                                                  | 169            |
| Dobór dyskryminant liniowych dla nowej podprzestrzeni cech                        | 171            |
| Rzutowanie przykładów na nową przestrzeń cech                                     | 173            |
| Implementacja analizy LDA w bibliotece scikit-learn                               | 174            |
| <b>Jądrowa analiza głównych składowych</b>                                        |                |
| <b>jako metoda odwzorowywania nierozdzielnych liniowo klas</b>                    | <b>176</b>     |
| Funkcje jądra oraz sztuczka z funkcją jądra                                       | 177            |
| Implementacja jądrowej analizy głównych składowych w Pythonie                     | 181            |
| Rzutowanie nowych punktów danych                                                  | 188            |
| Algorytm jądrowej analizy głównych składowych w bibliotece scikit-learn           | 191            |
| <b>Podsumowanie</b>                                                               | <b>192</b>     |
| <br><b>Rozdział 6. Najlepsze metody oceny modelu i strojenie parametryczne</b>    | <br><b>195</b> |
| <b>Usprawnianie cyklu pracy za pomocą kolejkowania</b>                            | <b>195</b>     |
| Wczytanie zestawu danych Breast Cancer Wisconsin                                  | 196            |
| Łączenie funkcji transformujących i estymatorów w kolejce czynności               | 197            |
| <b>Stosowanie k-krotnego sprawdzianu krzyżowego w ocenie skuteczności modelu</b>  | <b>198</b>     |
| Metoda wydzielania                                                                | 199            |
| K-krotny sprawdzian krzyżowy                                                      | 200            |

|                                                                                      |            |
|--------------------------------------------------------------------------------------|------------|
| <b>Sprawdzanie algorytmów za pomocą krzywych uczenia i krzywych walidacji</b>        | <b>204</b> |
| Diagnozowanie problemów z obciążeniem i wariancją                                    |            |
| za pomocą krzywych uczenia                                                           | 204        |
| Rozwiązywanie problemów przetrenowania i niedotrenowania                             |            |
| za pomocą krzywych walidacji                                                         | 208        |
| <b>Dostrajanie modeli uczenia maszynowego za pomocą metody przeszukiwania siatki</b> | <b>209</b> |
| Strojenie hiperparametrów przy użyciu metody przeszukiwania siatki                   | 210        |
| Dobór algorytmu poprzez zagnieżdżony sprawdzian krzyżowy                             | 211        |
| <b>Przegląd wskaźników oceny skuteczności</b>                                        | <b>213</b> |
| Odczytywanie macierzy pomyłek                                                        | 213        |
| Optymalizacja precyzji i pełności modelu klasyfikującego                             | 215        |
| Wykres krzywej ROC                                                                   | 217        |
| Wskaźniki zliczające dla klasyfikacji wieloklasowej                                  | 220        |
| Kwestia dysproporcji klas                                                            | 220        |
| <b>Podsumowanie</b>                                                                  | <b>223</b> |
| <b>Rozdział 7. Łączenie różnych modeli w celu uczenia zespołowego</b>                | <b>225</b> |
| <b>Uczenie zespołów</b>                                                              | <b>225</b> |
| <b>Łączenie klasyfikatorów za pomocą algorytmu głosowania większościowego</b>        | <b>229</b> |
| Implementacja prostego klasyfikatora głosowania większościowego                      | 230        |
| Stosowanie reguły głosowania większościowego do uzyskiwania prognoz                  | 235        |
| Ewaluacja i strojenie klasyfikatora zespołowego                                      | 237        |
| <b>Agregacja — tworzenie zespołu klasyfikatorów za pomocą próbek początkowych</b>    | <b>242</b> |
| Agregacja w pigułce                                                                  | 243        |
| Stosowanie agregacji do klasyfikowania przykładów z zestawu Wine                     | 244        |
| <b>Usprawnianie słabych klasyfikatorów za pomocą wzmocnienia adaptacyjnego</b>       | <b>248</b> |
| Wzmacnianie — mechanizm działania                                                    | 248        |
| Stosowanie algorytmu AdaBoost za pomocą biblioteki scikit-learn                      | 252        |
| <b>Podsumowanie</b>                                                                  | <b>255</b> |
| <b>Rozdział 8. Wykorzystywanie uczenia maszynowego w analizie sentymentów</b>        | <b>257</b> |
| <b>Przygotowywanie zestawu danych IMDb movie review do przetwarzania tekstu</b>      | <b>258</b> |
| Uzyskiwanie zestawu danych IMDb                                                      | 258        |
| Przetwarzanie wstępne zestawu danych IMDb do wygodniejszego formatu                  | 259        |
| <b>Wprowadzenie do modelu worka słów</b>                                             | <b>260</b> |
| Przekształcanie słów w wektory cech                                                  | 261        |
| Ocena istotności wyrazów za pomocą ważenia częstości termów                          |            |
| — odwrotnej częstości w tekście                                                      | 262        |
| Oczyszczanie danych tekstowych                                                       | 264        |
| Przetwarzanie tekstu na znaczniki                                                    | 266        |
| <b>Uczenie modelu regresji logistycznej w celu klasyfikowania tekstu</b>             | <b>268</b> |
| <b>Praca z większą ilością danych — algorytmy sieciowe i uczenie pozardzeniowe</b>   | <b>270</b> |
| <b>Modelowanie tematyczne za pomocą alokacji ukrytej zmiennej Dirichleta</b>         | <b>273</b> |
| Rozkładanie dokumentów tekstowych za pomocą analizy LDA                              | 274        |
| Analiza LDA w bibliotece scikit-learn                                                | 274        |
| <b>Podsumowanie</b>                                                                  | <b>277</b> |

|                                                                                               |            |
|-----------------------------------------------------------------------------------------------|------------|
| <b>Rozdział 9. Wdrażanie modelu uczenia maszynowego do aplikacji sieciowej</b>                | <b>279</b> |
| Serializacja wyuczonych estymatorów biblioteki scikit-learn                                   | 280        |
| Konfigurowanie bazy danych SQLite                                                             | 283        |
| Tworzenie aplikacji sieciowej za pomocą środowiska Flask                                      | 285        |
| Nasza pierwsza aplikacja sieciowa                                                             | 285        |
| Sprawdzanie i wyświetlanie formularza                                                         | 287        |
| Przekształcanie klasyfikatora recenzji w aplikację sieciową                                   | 293        |
| Pliki i katalogi — wygląd drzewa katalogów                                                    | 295        |
| Implementacja głównej części programu w pliku app.py                                          | 296        |
| Konfigurowanie formularza recenzji                                                            | 298        |
| Tworzenie szablonu strony wynikowej                                                           | 299        |
| Umieszczanie aplikacji sieciowej na publicznym serwerze                                       | 301        |
| Tworzenie konta w serwisie PythonAnywhere                                                     | 301        |
| Przesyłanie aplikacji klasyfikatora filmowego                                                 | 302        |
| Aktualizowanie klasyfikatora recenzji filmowych                                               | 303        |
| Podsumowanie                                                                                  | 305        |
| <b>Rozdział 10. Przewidywanie ciągłych zmiennych docelowych za pomocą analizy regresywnej</b> | <b>307</b> |
| Wprowadzenie do regresji liniowej                                                             | 308        |
| Prosta regresja liniowa                                                                       | 308        |
| Wielowymiarowa regresja liniowa                                                               | 308        |
| Zestaw danych Housing                                                                         | 310        |
| Wczytywanie zestawu danych Housing do obiektu DataFrame                                       | 310        |
| Wizualizowanie ważnych elementów zestawu danych                                               | 312        |
| Analiza związków za pomocą macierzy korelacji                                                 | 313        |
| Implementacja modelu regresji liniowej wykorzystującego zwykłą metodę najmniejszych kwadratów | 315        |
| Określanie parametrów regresyjnych za pomocą metody gradientu prostego                        | 316        |
| Szacowanie współczynnika modelu regresji za pomocą biblioteki scikit-learn                    | 319        |
| Uczenie odpornego modelu regresyjnego za pomocą algorytmu RANSAC                              | 321        |
| Ocenianie skuteczności modeli regresji liniowej                                               | 324        |
| Stosowanie regularyzowanych metod regresji                                                    | 327        |
| Przekształcanie modelu regresji liniowej w krzywą — regresja wielomianowa                     | 329        |
| Dodawanie członów wielomianowych za pomocą biblioteki scikit-learn                            | 329        |
| Modelowanie nieliniowych zależności w zestawie danych Housing                                 | 331        |
| Analiza nieliniowych relacji za pomocą algorytmu losowego lasu                                | 334        |
| Regresja przy użyciu drzewa decyzyjnego                                                       | 334        |
| Regresja przy użyciu losowego lasu                                                            | 336        |
| Podsumowanie                                                                                  | 339        |
| <b>Rozdział 11. Praca z nieoznakowanymi danymi — analiza skupień</b>                          | <b>341</b> |
| Grupowanie obiektów na podstawie podobieństwa przy użyciu algorytmu centroidów                | 342        |
| Algorytm centroidów w bibliotece scikit-learn                                                 | 342        |
| Inteligentniejszy sposób dobierania pierwotnych centroidów za pomocą algorytmu k-means++      | 346        |
| Twarda i miękka analiza skupień                                                               | 347        |

|                                                                        |            |
|------------------------------------------------------------------------|------------|
| Stosowanie metody łokcia do wyszukiwania optymalnej liczby skupień     | 349        |
| Ujęcie ilościowe jakości analizy skupień za pomocą wykresu profilu     | 351        |
| <b>Organizowanie skupień do postaci drzewa skupień</b>                 | <b>355</b> |
| Oddolne grupowanie skupień                                             | 356        |
| Przeprowadzanie hierarchicznej analizy skupień na macierzy odległości  | 357        |
| Dołączanie dendrogramów do mapy cieplnej                               | 360        |
| Aglomeracyjna analiza skupień w bibliotece scikit-learn                | 361        |
| <b>Wyznaczanie rejonów o dużej gęstości za pomocą algorytmu DBSCAN</b> | <b>363</b> |
| <b>Podsumowanie</b>                                                    | <b>368</b> |

## Rozdział 12. Implementowanie wielowarstwowej sieci neuronowej od podstaw 369

|                                                                               |            |
|-------------------------------------------------------------------------------|------------|
| <b>Modelowanie złożonych funkcji przy użyciu sztucznych sieci neuronowych</b> | <b>370</b> |
| Jednowarstwowa sieć neuronowa — powtórzenie                                   | 371        |
| Wstęp do wielowarstwowej architektury sieci neuronowych                       | 373        |
| Aktywacja sieci neuronowej za pomocą propagacji w przód                       | 376        |
| <b>Klasyfikowanie pisma odręcznego</b>                                        | <b>379</b> |
| Zestaw danych MNIST                                                           | 379        |
| Implementacja perceptronu wielowarstwowego                                    | 385        |
| <b>Trenowanie sztucznej sieci neuronowej</b>                                  | <b>395</b> |
| Obliczanie logistycznej funkcji kosztu                                        | 395        |
| Wyjaśnienie algorytmu wstecznej propagacji                                    | 398        |
| Uczenie sieci neuronowych za pomocą algorytmu propagacji wstecznej            | 399        |
| <b>Zbieżność w sieciach neuronowych</b>                                       | <b>402</b> |
| <b>Jeszcze słowo o implementacji sieci neuronowej</b>                         | <b>404</b> |
| <b>Podsumowanie</b>                                                           | <b>404</b> |

## Rozdział 13. Równoległe przetwarzanie sieci neuronowych za pomocą biblioteki TensorFlow 407

|                                                                                            |            |
|--------------------------------------------------------------------------------------------|------------|
| <b>Biblioteka TensorFlow a skuteczność uczenia</b>                                         | <b>408</b> |
| Wyzwania związane z wydajnością                                                            | 408        |
| Czym jest biblioteka TensorFlow?                                                           | 409        |
| W jaki sposób będziemy poznawać bibliotekę TensorFlow?                                     | 411        |
| <b>Pierwsze kroki z biblioteką TensorFlow</b>                                              | <b>411</b> |
| Instalacja modułu TensorFlow                                                               | 411        |
| Tworzenie tensorów w TensorFlow                                                            | 412        |
| Manipulowanie typem danych i rozmiarem tensora                                             | 413        |
| Przeprowadzanie operacji matematycznych na tensorach                                       | 414        |
| Dzielenie, nawarstwianie i łączenie tensorów                                               | 415        |
| <b>Tworzenie potoków wejściowych za pomocą tf.data, czyli interfejsu danych TensorFlow</b> | <b>416</b> |
| Tworzenie obiektów Dataset z istniejących tensorów                                         | 417        |
| Łączenie dwóch tensorów we wspólny zestaw danych                                           | 418        |
| Potasuj, pogrupuj, powtórz                                                                 | 419        |
| Tworzenie zestawu danych z plików umieszczonych w lokalnym magazynie dyskowym              | 422        |
| Pobieranie dostępnych zestawów danych z biblioteki tensorflow_datasets                     | 425        |
| <b>Tworzenie modelu sieci neuronowej za pomocą modułu TensorFlow</b>                       | <b>430</b> |
| Interfejs Keras (tf.keras)                                                                 | 430        |
| Tworzenie modelu regresji liniowej                                                         | 431        |

|                                                                                                            |            |
|------------------------------------------------------------------------------------------------------------|------------|
| Uczenie modelu za pomocą metod <code>.compile()</code> i <code>.fit()</code>                               | 435        |
| Tworzenie perceptronu wielowarstwowego klasyfikującego kwiaty z zestawu danych Iris                        | 436        |
| Ocena wytrenowanego modelu za pomocą danych testowych                                                      | 439        |
| Zapisywanie i wczytywanie wyuczonego modelu                                                                | 440        |
| <b>Dobór funkcji aktywacji dla wielowarstwowych sieci neuronowych</b>                                      | <b>440</b> |
| Funkcja logistyczna — powtórzenie                                                                          | 441        |
| Szacowanie prawdopodobieństw przynależności do klas w klasyfikacji wieloklasowej za pomocą funkcji softmax | 443        |
| Rozszerzanie zakresu wartości wyjściowych za pomocą funkcji tangensa hiperbolicznego                       | 444        |
| Aktywacja za pomocą prostowanej jednostki liniowej (ReLU)                                                  | 446        |
| <b>Podsumowanie</b>                                                                                        | <b>448</b> |
| <b>Rozdział 14. Czas na szczegóły — mechanizm działania biblioteki TensorFlow</b>                          | <b>449</b> |
| <b>Cechy kluczowe TensorFlow</b>                                                                           | <b>450</b> |
| <b>Grafy obliczeniowe TensorFlow: migracja do wersji TensorFlow 2</b>                                      | <b>451</b> |
| Grafy obliczeniowe                                                                                         | 451        |
| Tworzenie grafu w wersji TensorFlow 1.x                                                                    | 452        |
| Migracja grafu do wersji TensorFlow 2                                                                      | 453        |
| Wczytywanie danych wejściowych do modelu: TensorFlow 1.x                                                   | 453        |
| Wczytywanie danych wejściowych do modelu: TensorFlow 2                                                     | 454        |
| Poprawianie wydajności obliczeniowej za pomocą dekoratorów funkcji                                         | 455        |
| <b>Obiekty <code>Variable</code> służące do przechowywania i aktualizowania parametrów modelu</b>          | <b>456</b> |
| <b>Obliczanie gradientów za pomocą różniczkowania automatycznego i klasy <code>GradientTape</code></b>     | <b>460</b> |
| Obliczanie gradientów funkcji straty w odniesieniu do zmiennych modyfikowalnych                            | 460        |
| Obliczanie gradientów w odniesieniu do tensorów niemodyfikowalnych                                         | 462        |
| Przechowywanie zasobów na obliczanie wielu gradientów                                                      | 462        |
| <b>Upraszczenie implementacji popularnych struktur za pomocą interfejsu Keras</b>                          | <b>463</b> |
| Rozwiązywanie problemu klasyfikacji XOR                                                                    | 466        |
| Zwiększenie możliwości budowania modeli za pomocą interfejsu funkcyjnego Keras                             | 471        |
| Implementowanie modeli bazujących na klasie <code>Model</code>                                             | 472        |
| Pisanie niestandardowych warstw Keras                                                                      | 473        |
| <b>Estymatory TensorFlow</b>                                                                               | <b>476</b> |
| Praca z kolumnami cech                                                                                     | 477        |
| Uczenie maszynowe za pomocą gotowych estymatorów                                                           | 481        |
| Stosowanie estymatorów w klasyfikacji zestawu pisma odręcznego MNIST                                       | 484        |
| Tworzenie niestandardowego estymatora z istniejącego modelu Keras                                          | 486        |
| <b>Podsumowanie</b>                                                                                        | <b>488</b> |
| <b>Rozdział 15. Klasyfikowanie obrazów za pomocą głębokich spłotowych sieci neuronowych</b>                | <b>489</b> |
| <b>Podstawowe elementy spłotowej sieci neuronowej</b>                                                      | <b>490</b> |
| Spłotowe sieci neuronowe i hierarchie cech                                                                 | 490        |
| Spłot dyskretny                                                                                            | 492        |
| Warstwy podpróbkowania                                                                                     | 501        |

|                                                                                                                       |                |
|-----------------------------------------------------------------------------------------------------------------------|----------------|
| <b>Implementowanie sieci CNN</b>                                                                                      | <b>502</b>     |
| Praca z wieloma kanałami wejściowymi/barw                                                                             | 503            |
| Regularyzowanie sieci neuronowej metodą porzucania                                                                    | 506            |
| Funkcje straty w zadaniach klasyfikacji                                                                               | 509            |
| <b>Implementacja głębokiej sieci splotowej za pomocą biblioteki TensorFlow</b>                                        | <b>511</b>     |
| Architektura wielowarstwowej sieci CNN                                                                                | 511            |
| Wczytywanie i wstępne przetwarzanie danych                                                                            | 512            |
| Implementowanie sieci CNN za pomocą interfejsu Keras                                                                  | 513            |
| <b>Klasyfikowanie płci na podstawie zdjęć twarzy za pomocą sieci splotowej</b>                                        | <b>518</b>     |
| Wczytywanie zestawu danych CelebA                                                                                     | 519            |
| Przekształcanie obrazów i dogenerowanie danych                                                                        | 520            |
| Uczenie modelu CNN jako klasyfikatora płci                                                                            | 525            |
| <b>Podsumowanie</b>                                                                                                   | <b>530</b>     |
| <br><b>Rozdział 16. Modelowanie danych sekwencyjnych za pomocą rekurencyjnych sieci neuronowych</b>                   | <br><b>533</b> |
| <b>Wprowadzenie do danych sekwencyjnych</b>                                                                           | <b>534</b>     |
| Modelowanie danych sekwencyjnych — kolejność ma znaczenie                                                             | 534            |
| Przedstawianie sekwencji                                                                                              | 535            |
| Różne kategorie modelowania sekwencji                                                                                 | 536            |
| <b>Sieci rekurencyjne służące do modelowania sekwencji</b>                                                            | <b>537</b>     |
| Mechanizm zapętlania w sieciach rekurencyjnych                                                                        | 537            |
| Obliczanie aktywacji w sieciach rekurencyjnych                                                                        | 539            |
| Rekurencja w warstwie ukrytej a rekurencja w warstwie wyjściowej                                                      | 542            |
| Problemy z uczeniem długofalowych oddziaływań                                                                         | 544            |
| Jednostki LSTM                                                                                                        | 546            |
| <b>Implementowanie wielowarstwowej sieci rekurencyjnej przy użyciu biblioteki TensorFlow do modelowania sekwencji</b> | <b>548</b>     |
| Pierwszy projekt — przewidywanie sentymentów na recenzjach z zestawu danych IMDb                                      | 548            |
| Drugi projekt — modelowanie języka na poziomie znaków w TensorFlow                                                    | 561            |
| <b>Przetwarzanie języka za pomocą modelu transformatora</b>                                                           | <b>572</b>     |
| Wyjaśnienie mechanizmu samouwagi                                                                                      | 573            |
| Wieloblokowy mechanizm uwagi i komórka transformatora                                                                 | 575            |
| <b>Podsumowanie</b>                                                                                                   | <b>577</b>     |
| <br><b>Rozdział 17. Generatywne sieci przeciwstawne w zadaniach syntetyzowania nowych danych</b>                      | <br><b>579</b> |
| <b>Wprowadzenie do generatywnych sieci przeciwstawnych</b>                                                            | <b>580</b>     |
| Autokodery                                                                                                            | 580            |
| Modele generatywne syntetyzujące nowe dane                                                                            | 582            |
| Generowanie nowych prób za pomocą sieci GAN                                                                           | 584            |
| Funkcje straty generatora i dyskryminatora w modelu GAN                                                               | 585            |
| <b>Implementowanie sieci GAN od podstaw</b>                                                                           | <b>587</b>     |
| Uczenie modeli GAN w środowisku Google Colab                                                                          | 587            |
| Implementacja sieci generatora i dyskryminatora                                                                       | 590            |
| Definiowanie zestawu danych uczących                                                                                  | 593            |
| Uczenie modelu GAN                                                                                                    | 595            |

|                                                                         |            |
|-------------------------------------------------------------------------|------------|
| <b>Poprawianie jakości syntetyzowanych obrazów za pomocą sieci GAN:</b> | <b>603</b> |
| <b>splotowej i Wassersteina</b>                                         | <b>603</b> |
| Splot transponowany                                                     | 603        |
| Normalizacja wsadowa                                                    | 605        |
| Implementowanie generatora i dyskryminatora                             | 607        |
| Wskaźniki odmienności dwóch rozkładów                                   | 613        |
| Praktyczne stosowanie odległości EM w sieciach GAN                      | 616        |
| Kara gradientowa                                                        | 617        |
| Implementacja sieci WGAN-DP służącej do uczenia modelu DCGAN            | 617        |
| Załamanie modu                                                          | 622        |
| <b>Inne zastosowania modeli GAN</b>                                     | <b>623</b> |
| <b>Podsumowanie</b>                                                     | <b>624</b> |

## **Rozdział 18. Uczenie przez wzmacnianie jako mechanizm podejmowania decyzji w skomplikowanych środowiskach** **625**

|                                                                             |            |
|-----------------------------------------------------------------------------|------------|
| <b>Wprowadzenie: uczenie z doświadczenia</b>                                | <b>626</b> |
| Filozofia uczenia przez wzmacnianie                                         | 626        |
| Definicja interfejsu agent-środowisko w systemie uczenia przez wzmacnianie  | 628        |
| <b>Podstawy teoretyczne uczenia przez wzmacnianie</b>                       | <b>629</b> |
| Procesy decyzyjne Markowa                                                   | 629        |
| Wyjaśnienie matematyczne procesów decyzyjnych Markowa                       | 630        |
| Terminologia uczenia przez wzmacnianie: zwrot, strategia i funkcja wartości | 633        |
| Programowanie dynamiczne za pomocą równania Bellmana                        | 636        |
| <b>Algorytmy uczenia przez wzmacnianie</b>                                  | <b>637</b> |
| Programowanie dynamiczne                                                    | 638        |
| Uczenie przez wzmacnianie metodą Monte Carlo                                | 641        |
| Uczenie metodą różnic czasowych                                             | 643        |
| <b>Implementacja naszego pierwszego algorytmu uczenia przez wzmacnianie</b> | <b>646</b> |
| Wprowadzenie do pakietu OpenAI Gym                                          | 646        |
| Rozwiązywanie problemu świata blokowego za pomocą Q-uczenia                 | 654        |
| Krótko o algorytmie Q-uczenia głębokiego                                    | 658        |
| <b>Podsumowanie rozdziału i książki</b>                                     | <b>665</b> |