

Spis treści

1. Wprowadzenie do systemów uczących się	9
1.1. Uczenie maszynowe	9
1.2. Uczenie maszynowe a sztuczna inteligencja	11
1.3. Problemy, które rozwiązuje uczenie maszynowe	12
1.4. System informacyjny w sensie Pawlaka	12
1.5. System decyzyjny w sensie Pawlaka	13
1.6. Idea uczenia maszynowego	14
1.6.1. Problem uczenia maszynowego	14
1.6.2. Uczenie klasyfikatora	16
1.6.3. Wymiar Vapnika-Chervonenkisa	17
1.6.4. Metoda PAC	18
1.7. Uczenie nadzorowane	19
1.8. Uczenie nienadzorowane	21
1.9. Uczenie półnadzorowane	22
1.10. Uczenie przez wzmacnianie	22
1.11. Twierdzenie <i>No free lunches</i>	23
Literatura	24
2. Metody wstępnego przetwarzania danych	25
2.1. Popularne zbiory rzeczywistych danych	25
2.2. Typy wartości w atrybutach	26
2.2.1. Typ numeryczny	26
2.2.2. Typ kategorialny	26
2.3. Oczyszczanie danych	27
2.3.1. Praca z nieprawidłowymi wartościami	27
2.3.2. Uzupełnianie brakujących wartości	33
2.4. Redukcja wymiarowości systemu	36
2.4.1. Analiza głównych składowych	36
2.4.2. Nadzorowana redukcja wymiarowości za pomocą liniowej analizy dyskryminacyjnej	39
2.4.3. Ocena istotności atrybutów za pomocą korelacji	42
2.4.4. Inne metody redukcji wymiarowości danych	45
Literatura	45
3. Uczenie modeli i ich metryki	47
3.1. Parametry wydajności	47
3.1.1. Dokładność globalna	47
3.1.2. Dokładność zbalansowana	47
3.1.3. Stopień prawidłowej trafności w klasę decyzyjną	48
3.1.4. Stopień prawidłowej nietrafności w klasę decyzyjną	48
3.1.5. Stopień błędnej trafności w klasę decyzyjną	49
3.1.6. Stopień błędnej nietrafności w klasę decyzyjną	49
3.1.7. Precyzja	50
3.1.8. Pełność	50
3.1.9. Wskaźnik F_1	50
3.1.10. Dobór precyzji i pełności dla modelu	51
3.1.11. Krzywa ROC	53

3.2. Funkcje kosztu	54
3.2.1. Pojęcie funkcji kosztu	54
3.2.2. Błąd średniokwadratowy	56
3.2.3. Funkcja Hubera	57
3.2.4. Binarna entropia krzyżowa	57
3.2.5. Kategorialna entropia krzyżowa	58
3.2.6. Funkcja zawiasowa	58
3.3. Metody oceny wydajności modeli	58
3.3.1. Metoda „trenuj i testuj”	58
3.3.2. <i>k</i> -krotna walidacja krzyżowa	60
3.3.3. Walidacja krzyżowa Monte Carlo	61
3.3.4. Wewnętrzna walidacja krzyżowa	61
3.3.5. <i>Leave one out</i>	62
3.3.6. <i>Bootstrap</i>	62
3.3.7. <i>Bagging</i>	63
3.3.8. Losowe lasy	63
3.4. Zjawisko przetrenowania i niedotrenowania modelu	65
3.4.1. Problem przetrenowania i niedotrenowania	65
3.4.2. Przyczyny niedotrenowania i przetrenowania modeli	66
3.4.3. Wykorzystanie podzbiorów walidacyjnych	67
3.4.4. Ewaluacja stopnia wytrenowania modelu na podstawie metody T&T	67
3.4.5. Obciążenie modelu	68
3.4.6. Wariancja modelu	69
3.5. Regularyzacja	69
3.5.1. Regularyzacja grzbietowa	69
3.5.2. Regularyzacja lasso	70
3.5.3. Regularyzacja <i>ElasticNet</i>	70
Literatura	71
4. Wprowadzenie do sztucznych sieci neuronowych	72
4.1. Perceptron – podstawowa jednostka sieci neuronowej	72
4.2. Sztuczna sieć neuronowa	74
4.3. Funkcja aktywacji	76
4.3.1. Idea funkcji aktywacji	76
4.3.2. Popularne funkcje aktywacji	76
4.4. Metody optymalizacyjne	80
4.4.1. Metoda gradientu prostego	80
4.4.2. Momentum	84
4.4.3. Przyspieszony spadek wzdłuż gradientu	85
4.4.4. RMSProp	86
4.4.5. <i>Adaptive moment estimation</i> (Adam)	87
4.5. Uczenie sztucznej sieci neuronowej	88
4.6. Metody projektowania sieci neuronowych do klasyfikacji i regresji	90
Literatura	92
5. Problemy uczenia głębokiego	93
5.1. Niestabilne gradienty	93
5.1.1. Metody inicjalizacji wag sieci neuronowych	93
5.1.2. Dobór funkcji aktywacji	94
5.1.3. Normalizacja wsadowa	95
5.1.4. Obcinanie gradientu	96

5.2. Regularyzacja	97
5.2.1. Wczesne zatrzymywanie	97
5.2.2. Dropout	98
5.3. Regularyzacja metodą Monte Carlo	99
5.4. Regularyzacja metodą max-norm	100
5.5. Metody projektowania sieci neuronowych odpornych na problemy uczenia głębokiego	100
Literatura	101
6. Implementacja sieci neuronowych z zastosowaniem biblioteki TensorFlow	102
6.1. Keras	102
6.1.1. Instalacja w środowisku języka Python	102
6.1.2. Weryfikacja poprawności instalacji oraz wersji biblioteki	103
6.1.3. Wbudowane warstwy modeli sieci neuronowych	103
6.1.4. Podstawowe metody i atrybuty warstw	104
6.2. Tworzenie modeli z zastosowaniem interfejsu Keras	105
6.2.1. Sequential API	105
6.2.2. Functional API	107
6.2.3. Subclassing API	111
6.3. Wizualizacja, trening, ocena wydajności i zapisywanie modelu	113
6.4. Studium przypadku – budowa modelu decyzyjnego diagnozującego schorzenia układu krążenia	118
Literatura	123
7. Wdrażanie wytrenowanych modeli w chmurze publicznej	125
7.1. Chmura publiczna	125
7.2. Usługi bezserwerowe	126
7.3. Google Cloud Platform	126
7.4. Wdrażanie modeli za pomocą Google Cloud Platform	127
7.5. Wykonywanie predykcji za pomocą wdrożonego modelu	137
Literatura	142
Streszczenie	144