

Spis treści

Wprowadzenie	7
1. Podstawy języka R	11
2. Techniki programowania w systemie GNU R	27
2.1. Rozkład współczynnika determinacji R^2 z próby	28
2.2. Dokładność oszacowań parametrów regresji liniowej	33
2.3. Graficzna prezentacja zależności w danych	38
2.4. Rozkład łączny dwóch cech	49
2.5. Wielokrotne stosowanie funkcji	54
2.6. Ocena mocy testów na normalność rozkładu	60
2.7. Przekształcanie danych	68
2.8. Manipulowanie ciągami znaków	75
3. Procedury budowy modeli predykcyjnych	79
3.1. Porównanie dwóch modeli predykcyjnych	80
3.2. Przydatność zbioru testowego	88
3.3. Walidacja krzyżowa a wybór postaci modelu	92
3.4. Optymalizacja progu odcięcia	100
3.5. Graficzna ocena klasyfikatora	108
3.6. Odwzorowanie <i>score</i> w prawdopodobieństwo	118
4. Algorytmy analizy danych	127
4.1. Selekcja zmiennych i metody regularyzacyjne	128
4.2. Uogólnione modele addytywne	141
4.3. Drzewa klasyfikacyjne i lasy losowe	146
4.4. Sieci neuronowe	157
4.5. Bagging modeli predykcyjnych	166
4.6. Analiza skupień i redukcja wymiaru danych	173
5. Zakończenie	181