

Przedmowa

Niniejszy podręcznik poświęcony jest wybranym zagadnieniom eksploracji danych. W dobie bujnego rozwoju technik komputerowego przetwarzania ogromnych baz danych problem optymalności metod przetwarzania nabiera szczególnego znaczenia. Praca wychodzi naprzeciw tym potrzebom, przedstawiając syntetyczne ujęcie współczesnych metod eksploracji mających największe znaczenie i zastosowanie w praktyce.

Zawartość podręcznika została podzielona na 16 rozdziałów. Rozdział 1 stanowi wprowadzenie w tematykę eksploracji danych, wyjaśniające podstawowe pojęcia tego działu wiedzy, w tym dotyczące statystyki, technik i algorytmów optymalizacji, analizy sygnałów, rozpoznawania wzorców, sztucznej inteligencji oraz technik wizualizacji danych. Rozdział 2 traktuje o podstawowych metodach globalnej optymalizacji, w tym algorytmu symulowanego wyżarzania oraz algorytmów genetycznych i ewolucyjnych.

Rozdziały 3 i 4 dotyczą regresji, czyli zadania aproksymacji funkcji na podstawie zbioru danych pomiarowych. Rozdziały 5, 6, 7, 8 przedstawiają różne sposoby klasyfikacji danych liczbowych, reprezentowanych zwykle w postaci wektorowej, a także problemów predykcji (rozdział 9).

Rozdział 10 omawia metody oceny jakości wyniku klasyfikacji lub predykcji. Rozdziały 11 i 12 dotyczą zadań pozyskiwania i selekcji cech numerycznych najlepiej charakteryzujących analizowany proces. Rozdział 13 opisuje zadanie klasteryzacji danych, czyli ich grupowania metodami nienadzorowanego uczenia.

W szerszym pojęciu klasteryzacja jest jednym ze sposobów poszukiwania asocjacji między danymi, co zostało szerzej naświetlone w rozdziale 14. Rozdział 15 stanowi przegląd narzędzi wizualizacji graficznej dostępnej w Matlabie. Na koniec w rozdziale 16 przedstawiono dwa przykłady zastosowania metod eksploracji danych w medycynie.

Rozważania teoretyczne poparte są licznymi przykładami eksploracji danych zaimplementowanymi w postaci odpowiednich programów w Matlabie. Pozwala to czytelnikowi na łatwą weryfikację nabytej wiedzy teoretycznej.

Książka stanowi unikalne źródło wiedzy dotyczącej metod eksploracji danych dla studentów na studiach II i III stopnia. Może być polecana jako podstawowy materiał nie tylko do wykładów z eksploracji danych i pozyskiwania wiedzy, ale także jako materiał rozszerzający do innych klasycznych wykładów, jak sztuczna inteligencja, rozpoznawanie wzorców czy inżynieria biomedyczna na kierunku informatyka, automatyka czy na kierunkach pokrewnych. Może służyć pracownikom firm, tzw. analitykom danych, specjalizującym się w życiu zawodowym w analizie danych i odkrywaniu wiedzy z baz i hurtowni danych.