

Statystyczna analiza dużych zbiorów danych

Małgorzata Bogdan

Zasadniczym celem statystycznej analizy dużych zbiorów danych jest identyfikacja mechanizmu generującego dane, co następnie można wykorzystać do przewidywania nowych wartości a także do optymalizacji procesów decyzyjnych. Zadanie to jest trudne ze względu na dużą liczbę parametrów opisujących złożone mechanizmy generujące duże dane. Problem ten zwykle rozwiązuje się poprzez odpowiednią redukcję wymiaru modeli statystycznych. W czasie wykładu omówimy dwa podstawowe zagadnienia statystyczne - identyfikację powiązań między zmiennymi w dużej bazie danych (uczenie bez nadzoru) oraz identyfikację zmiennych wpływających na zadane cechy (uczenie z nadzorem). Zaprezentujemy wybrane współczesne metody statystyczne służące do rozwiązywania tych problemów a także przedstawimy konkretne przykłady zastosowań do analizy danych medycznych i finansowych.

M. Bogdan, UNIWERSYTET WROCLAWSKI, PLAC UNIWERSYTECKI 1, WROCLAW 50-137
Adres e-mail: malgorzata.bogdan@uwr.edu.pl