

STATYSTYCZNE SYSTEMY UCZĄCE SIĘ

Prof. dr hab. inż. Jacek Koronacki

Celem kursu jest wprowadzenie słuchaczy w szeroki wachlarz zagadnień z zakresu odpowiadającego tytułowi kursu. Statystyczne systemy uczące się wyrosły z wielowymiarowej analizy statystycznej i stanowią dziś najszerzej stosowany dział uczenia maszynowego. Skupimy się na wyjaśnieniu najważniejszych pojęć i metod, podaniu ich intuicyjnej interpretacji, a nie na szczegółach technicznych lub matematycznych. Nie wystarczy niestety czasu na zilustrowanie metod dogłębną analizą przykładów ich zastosowań. Tym niemniej, całość powinna wskazać, dlaczego analizy statystyczne są dziś tak powszechnie stosowane w badaniach naukowych oraz we wszystkich dziedzinach gospodarki i biznesu.

Zakładać będziemy, że słuchacze znają podstawy statystyki.

Plan kursu:

1. Preliminaria
- 2 i 3. Liniowa analiza regresji
4. Liniowa analiza dyskryminacyjna – (uogólnione) podejście Fishera; dyskryminacja oparta na regresji liniowej oraz dyskryminacja logistyczna.
5. Klasyfikator bayesowski i metoda największej wiarygodności w analizie dyskryminacyjnej.
6. Zasady nieparametrycznej estymacji gęstości w klasach, metoda najbliższych sąsiadów i metody pokrewne.
7. Drzewa klasyfikacyjne.
8. Rodziny klasyfikatorów.
9. Maszyny wektorów podpierających.
10. Analiza skupień – metoda k-średnich i metody pokrewne (np. SOM).
11. Hierarchiczna analiza skupień.
12. Redukcja wymiarowości – analiza składowych głównych, analiza czynnikowa i skalowanie wielowymiarowe; wzmianka o analizie składowych niezależnych.
13. Nieparametryczna analiza regresji.
- 14 i 15. Problemy nieklasyczne – analiza skupień na podzbiorach atrybutów; problem uwzględnienia własności geometrycznych skupień; klasyfikacja pod częściowym nadzorem.

Literatura:

J. Koronacki i J. Ćwik, *Statystyczne systemy uczące się*, WNT 2005, 2008.