

# STATYSTYCZNE SYSTEMY UCZĄCE SIĘ

## Przedmowa do wydania drugiego

Pierwsze wydanie *Statystycznych systemów uczących się* rozeszło się już dość dawno temu. Od tego czasu zaczęliśmy być pytani, najczęściej via poczta elektroniczna, czy planujemy wznowienie książki. Dlatego zdecydowaliśmy się takie wznowienie - wszakże unacześnione - przygotować.

Co do istoty, niniejsze wydanie jest identyczne z wydaniem pierwszym. Jednak jest to zarazem wydanie rozszerzone o nieco głębsze potraktowanie wybranych zagadnień klasycznych, lub o dodanie nowego materiału, opartego na wynikach badań z ostatnich lat.

Gdy mówić o rozszerzeniach pierwszego rodzaju, wymienić należy dokładniejsze omówienie - w szczególności - problemu diagnostyki modelu logistycznego (w punkcie 1.3.4) oraz - w ogólności - problemów estymacji i diagnostyki uogólnionych modeli liniowych (w podrozdz. 5.1), jak też nowy podrozdział poświęcony efektom losowym i modelom mieszanym (podrozdz. 5.3). W punkcie 2.3.2, dotyczącym wyboru i oceny klasyfikatora, zostało dodanych trochę informacji niesłusznie pominiętych w wydaniu wcześniejszym.

Zagadnieniom nieklasycznym poświęcony jest nowy podrozdz. 9.4, w którym jest mowa o algorytmach analizy skupień na podzbiorach atrybutów, spektralnej analizie skupień oraz uczeniu pod częściowym nadzorem. Ponadto, w podrozdz. 7.2 dodana została między innymi informacja o drzewach ortogonalnych.

Wśród nowinek dodanych w tym wydaniu, w pewnym sensie miejsce pośrednie zajmuje algorytm hierarchicznej analizy skupień uwzględniający odległości między obserwacjami wewnątrz tego samego skupienia (w podrozdz. 9.2) oraz nowy podrozdz. 8.7, poświęcony nieklasycznym metodom jądrowym w systemach uczących się. Wspomniany algorytm dlatego zajmuje miejsce pośrednie między metodami klasycznymi i nowymi, że jakkolwiek sam jest algorytmem stosunkowo nowym, to jest kolejną propozycją rozwiązania problemu postawionego dziesiątki lat temu. Z kolei nieklasyczne metody jądrowe były już przedmiotem analizy w pierwszym wydaniu książki, ale nie miały specjalnego podrozdziału, który pozwolił wzbogacić materiał ich dotyczący. Takie pośrednie miejsce zajmuje także krótka informacja o lemacie Johnsona-Lindenstraussa, który został opublikowany w roku 1984, ale jego zastosowania oraz idee związane z rzutowaniem danych na losowe podprzestrzenie znalazły należne im miejsce w uczeniu maszynowym znacznie później (p. podrozdz. 8.6).

Oczywiście, naturalną kolejną rzeczą, w różnych miejscach tekstu usunęliśmy zauważone błędy, wprowadziliśmy minimalne zmiany, mające na celu polepszenie wykładu, wykaz książek cytowanych wzbogaciliśmy o kilka pozycji oraz dodaliśmy nowe odwołania do szczególnie ciekawych artykułów czasopiśmienniczych z ostatnich lat (np. dotyczących rodzin klasyfikatorów i w szczególności ich regularyzacji).

*Jacek Koronacki i Jan Ćwik*

Warszawa, w sierpniu 2008