

Tytuł rozprawy doktorskiej:

„Metody stosowania wiedzy dziedzinowej do poprawiania jakości klasyfikatorów”

Streszczenie

Rozprawa dotyczy metod pozwalających na zastosowania wiedzy dziedzinowej do poprawiania jakości klasyfikatorów, przy czym poprawa jakości dotyczy zarówno metod ekstrakcji cech, metod konstrukcji klasyfikatorów oraz metod stosowania ich do predykcji wartości decyzji dla nowych obiektów. W szczególności, zaproponowano następujące metody poprawiania jakości klasyfikatorów: definiowanie eksperckich cech (atrybutów) z wykorzystaniem wiedzy dziedzinowej wyrażonej w języku wykorzystującym logikę temporalną, nową metodę mierzenia jakości cięć podczas dyskretyzacji z nadzorem wykorzystującą zdefiniowaną na podstawie wiedzy dziedzinowej macierzy odległości pomiędzy wartościami atrybutu decyzyjnego, nowe drzewo decyzyjne wykorzystujące nadmiarowe cięcia weryfikujące podział w danym węźle drzewa, nową metodę liczenia podobieństwa pomiędzy obiektami (np. pacjentami) wykorzystującą zdefiniowaną przez eksperta ontologię pojęć wraz z jej zastosowaniem do tworzenia klasyfikatora metodą k najbliższych sąsiadów oraz nową metodę generowania reguł krzyżowych opisujących wpływ czynnika zakłócającego percepcję za pomocą klasyfikatora.

Wszystkie wspomniane wyżej metody zostały zaimplementowane w bibliotece oprogramowania CommoDM, która jest jednym z rozszerzeń biblioteki RSES-lib.

Implementacje opracowanych metod zostały przetestowane na rzeczywistych zbiorach danych. Były to dane porównawcze znane z literatury, ale również własne dane medyczne zebrane w związku z powstaniem rozprawy. Te ostatnie zbiory danych wiążą się z medycznym aspektem rozprawy, który dotyczy wspomagania leczenia pacjentów ze stabilną chorobą niedokrwienną serca, a główny z problemów medycznych rozważany w pracy to problem przewidywania obecności istotnych zwężeń w tętnicach wieńcowych serca w oparciu o nieinwazyjne monitorowanie serca metodą Holtera.

Wyniki eksperymentów potwierdzają skuteczność zastosowania dodatkowej wiedzy dziedzinowej w zadaniu tworzenia i testowania klasyfikatorów, gdyż po zastosowaniu nowych metod jakość klasyfikatorów znacznie się podwyższyła. Jednocześnie interpretacja kliniczna uzyskanych wyników jest bardziej zgodna z wiedzą medyczną.

Prezentowane w rozprawie badania naukowe były wspierane przez granty o numerach DEC-2013/09/B/ST6/01568 oraz DEC-2013/09/B/NZ5/00758 Narodowego Centrum Nauki, a ich wyniki opublikowano w 10 publikacjach, w tym w 3 publikacjach z ministerialnej listy A, w 3 publikacjach indeksowanych w Web of Science, w jednym rozdziale w monografii zagranicznej i w 3 publikacjach pokonferencyjnych.