

Autor: Andrzej Wójtowicz

Tytuł: Grupowa klasyfikacja danych niekompletnych – podejście nieimputacyjne z zastosowaniem we wspomaganie diagnostyki guzów jajnika

Streszczenie:

W niniejszej pracy doktorskiej zająłem się problemem klasyfikacji danych niekompletnych. Motywacja do podjęcia badań ma swoje źródło w medycynie, gdzie bardzo często występuje zjawisko braku danych. Najpopularniejszą metodą radzenia sobie z tym problemem jest imputacja danych, będąca uzupełnieniem brakujących wartości na podstawie statystycznych zależności między cechami. W moich badaniach przyjąłem inną strategię rozwiązania tego problemu. Wykorzystując opracowane wcześniej klasyfikatory można przekształcić je do formy, która zwraca przedział możliwych predykcji. Następnie, poprzez zastosowanie operatorów agregacji oraz metod progowania, można dokonać finalnej klasyfikacji. W niniejszej pracy pokazuję jak dokonać ww. przekształcenia klasyfikatorów oraz jak wykorzystać strategie agregacji danych przedziałowych do klasyfikacji. Opracowane przeze mnie metody podnoszą jakość klasyfikacji danych niekompletnych w problemie wspomaganie diagnostyki guzów jajnika. Dodatkowa analiza wyników na zewnętrznych zbiorach danych z repozytorium uczenia maszynowego Uniwersytetu Kalifornijskiego w Irvine (UCI) wskazuje, że przedstawione metody są komplementarne z imputacją.

