

Prof. dr hab. inż. Roman Słowiński
z Instytutu Informatyki Politechniki Poznańskiej

wygłosi odczyt pt.

"Wspomaganie decyzji oparte na wiedzy odkrytej z danych"

Komputerowe wspomaganie decyzji w oparciu o wiedzę odkrytą z danych jest jednym z najistotniejszych nurtów współczesnej informatyki. Obecne tendencje rozwojowe tej specjalności mają charakter metodologiczny i technologiczny. Pierwszy z nich wiąże się z wykorzystaniem teorii zbiorów przybliżonych, a drugi z implementacjami systemów wspomagania decyzji na platformach mobilnych i w Internecie.

Streszczenie:

Jednym z najistotniejszych nurtów współczesnej informatyki jest komputerowe wspomaganie decyzji w oparciu o wiedzę odkrytą z danych. Obecne tendencje rozwojowe tej specjalności wiążą się szczególnie z dwoma aspektami – metodologicznym i technologicznym.

Jeśli chodzi o aspekt metodologiczny, to przy odkrywaniu wiedzy z danych ważne jest, by uwzględniać semantykę danych oraz ich możliwe praktyczne „niedoskonałości”, jak niepewność, niedokładność, niespójność i niekompletność. W problemach decyzyjnych semantyka rozumiana jest dwójako: jako uporządkowanie danych według preferencji i jako monotoniczna zależność między oceną obiektów na poszczególnych atrybutach (kryteriach), a ich oceną globalną. Na przykład w problemie klasyfikacji obiektów semantyka ta oznacza, że atrybuty wg. których dokonywana jest ocena obiektów są wyposażone w skale preferencji (porządkowe, przedziałowe lub ilorazowe), a obiekty przydzielane są do klas uporządkowanych według preferencji, w ten sposób, że jeśli obiekt a ma ocenę nie gorszą od obiektu b na wszystkich atrybutach, to a nie powinien być przydzielony do klasy gorszej niż b . Powyższej semantyki nie można pominąć w analizie danych bez ryzyka wyindukowania z danych błędnych wzorców. Sprowadza to odkrywanie wiedzy do jednoczesnego modelowania preferencji i indukowania wzorców. Reprezentację tak rozumianej wiedzy umożliwiają reguły decyzyjne o specjalnej składni. Poza zdolnością wyrażania najbardziej skomplikowanych struktur preferencji (nieaddytywnych, niekompensacyjnych i nieprzechodnich), reprezentacja ta jest także bardziej odporna na wspomniane „niedoskonałości” danych. Reprezentacja wiedzy za pomocą reguł decyzyjnych spełnia ważny dla praktyki postulat czytelności, bowiem model ten „mówi” do użytkownika jego językiem i ma formę dekomponowalną na wyrażenia reprezentujące scenariusze współzależności między wartościami poszczególnych atrybutów, co nie jest możliwe w przypadku modelu syntetycznego, takiego jak funkcja wartości lub relacja matematyczna.

Ważnym elementem powyższej metodologii jest teoria zbiorów przybliżonych zaproponowana przez Zdzisława Pawlaka na początku lat osiemdziesiątych ubiegłego wieku. Okazało się, że wiele problemów praktycznych, w tym problemy odkrywania wiedzy z danych, nie da się sformułować i rozwiązać używając klasycznego, Cantorowskiego pojęcia zbioru. Teoria zbiorów przybliżonych jest jedną z propozycji odejścia od klasycznej definicji zbioru w celu modelowania pojęć nieostrych. Przedstawimy adaptację tej teorii do odkrywania wiedzy z danych z uwzględnieniem semantyki decyzyjnej.

Jeśli chodzi o aspekt technologiczny, to ważne jest, by współczesne systemy komputerowego wspomagania decyzji były implementowane na platformach mobilnych i w Internecie. Przez

platformy mobilne rozumie się komputery naręczne i tabliczkowe oraz telefony komórkowe. Dotarcie z narzędziem wspomagania decyzji na miejsce podejmowania decyzji stanowi przełom podobny do przejścia z przetwarzania wsadowego na komputery osobiste w latach osiemdziesiątych. Ponadto, implementowanie systemów wspomagania decyzji w Internecie ma duże znaczenie dla telekonsultacji w oparciu o dane multimedialne. W obu przypadkach znakomitą przykładową dziedziną zastosowań tych technologii jest medycyna (telemedycyna).