

Politechnika Poznańska
Wydział Informatyki
Instytut Informatyki

Streszczenie rozprawy doktorskiej

**ZASTOSOWANIE TEORII ZBIORÓW
PRZYBLIŻONYCH Z RELACJĄ DOMINACJI
DO PROBLEMÓW PORZĄDKOWANIA
I KLASYFIKACJI NA PODSTAWIE PODOBIEŃSTWA**

Marcin Szeląg

Promotor
Prof. dr hab. inż. Roman Słowiński

Poznań, 2015

Przedmowa

Aktywność człowieka w wielu dziedzinach życia jest nieodłącznie związana z rozwiązywaniem *problemów decyzyjnych*. Zazwyczaj problemy te dotyczą pewnego zbioru *obiektów* (wariantów decyzyjnych, akcji, alternatyw, przypadków, obserwacji, opcji, kandydatów) opisanych za pomocą szeregu *atrybutów* (cech, charakterystyk, kryteriów, zmiennych). Atrybuty te mogą mieć charakter nominalny, porządkowy, lub numeryczny, a także posiadać porządek w zbiorze wartości wynikający np. z preferencji użytkownika. Zasadnicze typy problemów decyzyjnych to *klasyfikacja* obiektów do z góry zdefiniowanych klas decyzyjnych (nazywanych również kategoriami, lub krótko klasami), *ranking* (porządkowanie) obiektów od najlepszego do najgorszego, oraz *wybór* podzbioru najlepszych obiektów. Przykładowe problemy decyzyjne to: ocena ryzyka bankructwa firm, ranking uczelni wyższych oraz wybór najlepszego oferenta w przetargu.

Ze względu na złożoność wielu rzeczywistych problemów decyzyjnych, opracowano szereg naukowych metod wspomagania decyzji, m.in. w ramach dziedziny *Wielokryterialnego Wspomagania Decyzji* (WWD). Metody WWD dotyczą głównie problemów decyzyjnych ze skończonym zbiorem obiektów, w których dziedziny poszczególnych atrybutów są uporządkowane z punktu widzenia preferencji *decydenta* (użytkownika, eksperta). W szczególności, metody rozwijane dla problemu klasyfikacji dotyczą sytuacji, w których również klasy są *uporządkowane* względem preferencji decydenta. Problemy te nazywane są problemami *wielokryterialnego sortowania* lub problemami *klasyfikacji porządkowej z ograniczeniami monotonicznymi*.

Problemy decyzyjne adresowane są również przez dziedzinę *Uczenia Maszynowego* (UM). Tutaj jednak zazwyczaj nie rozważa się apriorycznych porządków na poszczególnych atrybutach. Dla problemu klasyfikacji, poza standardową wersją problemu z nieuporządkowanymi klasami, rozważa się również dwie inne wersje. Pierwsza z nich dotyczy *klasyfikacji rozmytej* (ang. soft label classification), czyli takiej, w której każdy obiekt ze zbioru treningowego może należeć w pewnym stopniu do każdej z klas. Przykładem tego typu problemu jest rozpoznawanie emocji w nagranych wcześniej wypowiedziach. W tej sytuacji naturalne jest, że poszczególne wypowiedzi mogą uzewnętrzniać różne emocje i to w różnym stopniu. Drugą wersją problemu klasyfikacji jest tzw. *klasyfikacja porządkowa*, czyli taka, która uwzględnia porządek klas decyzyjnych. Przykładem tej wersji problemu klasyfikacji jest przydział każdego z pacjentów, opisanych szeregiem testów medycznych,

do jednej z klas określających prawdopodobieństwo powikłań – słabe, średnie, lub wysokie.

Zanim temat niniejszej rozprawy doktorskiej został sprecyzowany, badania autora, prowadzone wspólnie z prof. Romanem Słowińskim i dr Jerzym Błaszczyńskim, dotyczyły zastosowania *teorii zbiorów przybliżonych opartej na dominacji* (ang. Dominance-based Rough Set Approach, DRSA), a także jej rozszerzeń ze zmienną spójnością, do klasyfikacji porządkowej z ograniczeniami monotonicznymi. Badania te pozwoliły zaobserwować przewagę DRSA nad innymi podejściami z WWD i UM. Po pierwsze, DRSA umożliwia uwzględnienie niespójności w przykładach decyzji (zbiorze treningowym) wprowadzając pojęcia dolnego i górnego przybliżenia unii uporządkowanych klas decyzyjnych; te przybliżenia są bazą do indukcji reguł decyzyjnych. Po drugie, DRSA umożliwia analizę danych heterogenicznych, zawierających (jednocześnie) atrybuty nominalne, porządkowe i ilościowe (numeryczne) – nie jest wymagana wstępna dyskretyzacja atrybutów numerycznych ani konwersja atrybutów nominalnych i porządkowych na ilościowe. Po trzecie, podstawowa idea DRSA jest zgodna z ważnym aktualnym trendem w WWD – indukcją *modelu preferencji* decydenta na podstawie *przykładów decyzji*. Po czwarte, zastosowanie DRSA pozwala na indukcję z przykładów decyzji zrozumiałego modelu preferencji decydenta w postaci zbioru *monotonicznych reguł decyzyjnych*.

Korzyści płynące z zastosowania DRSA do klasyfikacji porządkowej z ograniczeniami monotonicznymi stały się motywacją do podjęcia przez autora badań nad zastosowaniem DRSA również do innych problemów decyzyjnych. Niniejsza rozprawa jest wynikiem tych badań. Dotyczy ona dwóch problemów decyzyjnych. Pierwszym z nich jest *problem wielokryterialnego rankingu* obiektów rozważany w WWD, w którym dziedziny poszczególnych atrybutów są uporządkowane zgodnie z preferencjami decydenta. Drugim problemem jest *klasyfikacja na podstawie podobieństwa do znanych przypadków* (ang. case-based reasoning). W rozprawie zaprezentowano uzyskane dla obu problemów wyniki teoretyczne i eksperymentalne. Wyniki te zostały w większości opublikowane przez autora, wzbogacając w ten sposób obecne instrumentarium specjalności *Inteligentne Systemy Wspomagania Decyzji*.

Podziękowania

Niniejsza rozprawa została przygotowana w Zakładzie Inteligentnych Systemów Wspomagania Decyzji Instytutu Informatyki na Wydziale Informatyki Politechniki Poznańskiej. Do jej powstania przyczyniło się liczne grono osób, którym chciałbym w tym miejscu podziękować.

Serdecznie dziękuję mojemu promotorowi prof. dr hab. inż. Romanowi Słowińskiemu za inspirację, ciągłe wsparcie i zaangażowanie w proces przygotowania rozprawy. Dziękuję również moim kolegom i znajomym, od których wiele się nauczyłem, w tym w szczególności Salvatore Greco, Jurkowi Błaszczyńskiemu, Krzysztofowi Dembczyńskiemu, Wojciechowi

Kotłowskiemu i Miłoszowi Kadzińskiemu.

Pragnę również gorąco podziękować mojej ukochanej żonie Karolinie oraz moim rodzicom za ich nieustanne wsparcie, zachętę do pracy oraz nadzwyczajną cierpliwość okazaną podczas przygotowywania rozprawy.

Spis treści

1	Wprowadzenie	1
1.1	Opis Problemu	1
1.1.1	Problem Wielokryterialnego Rankingu	3
1.1.2	Problem Klasyfikacji na Podstawie Podobieństwa	7
1.2	Przegląd Istniejących Podejść do Problemu Wielokryterialnego Rankingu	8
1.2.1	Podejścia WWD do Problemu Wielokryterialnego Rankingu	8
1.2.2	Podejścia UP do Problemu Wielokryterialnego Rankingu	9
1.3	Przegląd Istniejących Podejść do Klasyfikacji na Podstawie Podobieństwa	10
1.3.1	Metody Oparte na Paradygmacie “Leniwego Uczenia się”	10
1.3.2	Metody Oparte na Paradygmacie “Aktywnego Uczenia się”	11
1.3.3	Pomiar Podobieństwa	12
1.4	Motywacja Zastosowania Teorii Zbiorów Przybliżonych Opartej na Dominacji do Wielokryterialnego Rankingu i Klasyfikacji na Podstawie Podobieństwa	12
1.5	Cel i Zakres Rozprawy	13
2	Podsumowanie	15
2.1	Wkład Rozprawy do Specjalności Inteligentnych Systemów Wspomagania Decyzji	15
2.1.1	Metoda dla Problemu Wielokryterialnego Rankingu Wykorzystująca VC-DRSA.	15
2.1.2	Analiza Procedur Eksploatacji Struktury Preferencji	18
2.1.3	Metoda Klasyfikacji na Podstawie Podobieństwa Wykorzystująca DRSA	19
2.1.4	Eksperymentalna Weryfikacja Metody dla Problemu Wielokryterialnego Rankingu	21
2.2	Kierunki Dalszych Badań	22
	Literatura	23

Rozdział 1

Wprowadzenie

1.1 Opis Problemu

Wielokryterialne Wspomaganie Decyzji (WWD). Dziedzina Wielokryterialnego Wspomagania Decyzji (WWD) [5, 25, 88] adresuje różne typy *problemów decyzyjnych* dotyczących skończonego zbioru *obiektów* (nazywanych również wariantami decyzyjnymi, akcjami, alternatywami, rozwiązaniami, opcjami, kandydatami, itp.) ocenionych z punktu widzenia skończonego zbioru kryteriów, tj. atrybutów, których dziedziny uporządkowane są względem preferencji decydenta. Przykładowo, w problemie wyboru notebooka, cena i pojemność dyskowa urządzenia są kryteriami, ponieważ niższa cena jest preferowana od wyższej, a wyższa pojemność dyskowa od niższej pojemności. Typowym założeniem przyjmowanym w WWD jest obecność *spójnego zbioru kryteriów* (spójnej rodziny kryteriów) [96], spełniającego następujące własności:

- *kompletność* (w zbiorze kryteriów uwzględniono wszystkie kryteria istotne z punktu widzenia decydenta),
- *monotoniczność* (poprawa ocen obiektu na kryteriach idzie w parze ze wzrostem stopnia w jakim jest on preferowany nad innymi obiektami),
- *nienadmiarowość* (brak kryterium, którego usunięcie nie naruszyłoby własności kompletności lub monotoniczności).

Zgodnie z pracami Roy [93], głównymi typami problemów decyzyjnych są: *opis*, *ranking* (porządkowanie), *wyбір* oraz *klasyfikacja porządkowa*.

WWD powstało w odpowiedzi na zapotrzebowanie na naukowe metody wspomaganie podejmowania decyzji w wielu rzeczywistych problemach decyzyjnych, w których rozważane kryteria reprezentują sprzeczne punkty widzenia, uniemożliwiając znalezienie obiektywnie optymalnego rozwiązania rozważanego problemu. WWD oferuje szerokie spektrum metod wspomagających *decydentów* w podejmowaniu decyzji poprzez strukturalizację problemu decyzyjnego, tworzenie *modelu preferencji* decydenta w oparciu o pozyskaną

od niego *informację preferencyjną* oraz eksploatację tego modelu w celu wypracowania ostatecznej *rekomendacji* [97].

Uczenie Maszynowe (UM). Dziedzina badań Uczenie Maszynowe (UM) (ang. Machine Learning) [7, 35, 63, 84, 85] adresuje różne typy problemów decyzyjnych. Jednym z najczęściej rozważanych jest *problem klasyfikacji*. Problemy klasyfikacji obiektów rozważane są w obszarze *uczenia nadzorowanego*, gdzie celem uczenia się jest stworzenie trafnego *klasyfikatora*, w oparciu o dostępne *obiekty treningowe*, dla których poprawna klasa jest znana. Trafność klasyfikatora rozumiana jest tutaj jako zdolność do możliwie dokładnego przewidywania klasy decyzyjnej dla obiektów z tzw. *zbioru testowego*, czyli zbioru obiektów, które były niedostępne dla klasyfikatora na etapie uczenia się.

Uczenie się Preferencji (UP). W ostatnich latach można zaobserwować gwałtowny rozwój dziedziny Uczenia się Preferencji (UP) (ang. Preference Learning) [36], będącej istotną poddziedziną UM. Większość badań w ramach UP dotyczy problemów “uczenia się rankingu” (ang. “learning to rank”) [19, 21, 80, 91]. W proponowanych metodach uczenia się częstą techniką jest sprowadzenie problemu rankingu do serii problemów klasyfikacji porządkowej. Takie podejście przedstawiono np. w pracy [3]. W obszarze zainteresowań UP znajdują się aplikacje związane z Internetem, w szczególności systemy rekomendacyjne oraz systemy wyszukiwania informacji (ang. information retrieval systems). W pierwszych systemach, zadanie polega na zarekomendowaniu użytkownikowi nowego przedmiotu (np. filmu, książki), który pasuje do preferencji tego użytkownika. Rekomendacja ustalana jest na podstawie informacji o wcześniejszych zachowaniach (wyborach) użytkownika. W drugich systemach, zadanie polega na uporządkowaniu, zgodnie z preferencjami użytkownika, dokumentów wyszukanych w odpowiedzi na zapytanie wydane przez tego użytkownika. Istnieje kilka algorytmów dostosowanych do tego typu problemów. Uczenie się jest najczęściej realizowane poprzez minimalizację empirycznej estymaty przyjętej *funkcji straty* [22]. Najbardziej popularne podejścia bazują na minimalizacji tzw. *błędu rangowego*. Wśród nich wyróżnić można warianty maszyny wektorów nośnych (ang. support vector machine) [64, 74] oraz podejścia typu “boosting” [21, 34].

W UM, a szczególnie w UP, zakłada się istnienie pewnego stochastycznego procesu generującego obserwowane dane. Stąd też, dane treningowe traktowane są jako próba losowa z pewnego nieznanego wielowymiarowego rozkładu prawdopodobieństwa. Co więcej, zakłada się, że dane treningowe są zniekształcone ze względu na obecność szumu losowego. Przy tych założeniach, celem uczenia się jest indukcja modelu preferencji, który generalizuje dane treningowe (tj., odnosi się do całej populacji obiektów) i jest w stanie dokonywać, *średnio*, dobrych predykcji dla obiektów niedostępnych na etapie uczenia. W UP kładzie się nacisk, z jednej strony, na trafność predykcji, a z drugiej strony, na skalowalność proponowanych algorytmów (co jest szczególnie istotne przy analizie dużych wolumenów danych).

Jak podkreślono w pracach [20, 21], WWD i UP współdzielą pewne cele, pojęcia i kwestie metodologiczne. Główna różnica pomiędzy nimi polega na sposobie tworzenia modelu preferencji decydenta. W UP, model preferencji jest wynikiem statystycznej analizy danych (obiektów treningowych). W WWD natomiast, model preferencji tworzony jest na podstawie informacji preferencyjnej pozyskanej od decydenta, często w sposób interaktywny.

1.1.1 Problem Wielokryterialnego Rankingu

Wielokryterialny ranking obiektów. Spośród różnych typów wielokryterialnych problemów decyzyjnych, w rozprawie rozważony został problem rankingu, w którym celem jest uporządkowanie zbioru obiektów opisanych za pomocą szeregu kryteriów. Uzyskany ranking może mieć postać *preporządku zupełnego* (ang. total preorder, weak order) lub *preporządku częściowego* (ang. partial preorder) (w którym pewne obiekty mogą być wzajemnie nieporównywalne). Problemy wielokryterialnego rankingu obiektów spotykane są często w takich dziedzinach jak finanse, ekonomia, zarządzanie i inżynieria [111, 114]. Typowe przykłady to ranking uczelni wyższych, szpitali, miast, państw, smartphonów, notebooków, linii lotniczych.

Problem rankingu nie jest, co oczywiste, problemem nowym. Był on już analizowany w wielu dziedzinach takich jak teoria decyzji, nauki społeczne, wyszukiwanie informacji, ekonomia, WWD i UP. W rozprawie przyjęto perspektywę WWD, biorąc również pod uwagę wybrane podejścia zaproponowane w dziedzinie UP.

W WWD, konstrukcja kryteriów z jawnymi monotonicznymi skalami preferencji jest ważnym krokiem w procesie wspomaganie decyzji. Kryteria są funkcjami o *skali porządkowej* lub *ilościowej* (interwałowej lub ilorazowej), zbudowanymi na elementarnych cechach obiektów, umożliwiającymi sensowne scharakteryzowanie poszczególnych obiektów, w ten sposób, że obiekty nierozróżnialne w sensie wartości kryteriów traktowane są jako równoważne. Z kolei w UP, zależności pomiędzy zbiorami wartości atrybutów a preferencjami decydenta odkrywane są z danych w trakcie uczenia. Oznacza to, że w UP monotoniczne skale preferencji konwertujące elementarne cechy obiektów na kryteria nie są używane w sposób jawny.

Biorąc pod uwagę taksonomię problemów “uczenia się rankingu” rozważaną w [37], problem rankingu rozważany w niniejszej rozprawie należy do kategorii “rankingu obiektów” (ang. object ranking) [19], gdzie zadanie polega na nauczeniu się z danych treningowych “dobrego” tzw. *rankera* – modelu przewidującego ranking obiektów – w oparciu o informację preferencyjną w postaci porównań obiektów parami.

Relacja dominacji. Mając dany skończony zbiór obiektów A oraz skończony zbiór $G = \{g_1, \dots, g_n\}$ zawierający kryteria przypisujące oceny $g_i(a)$ dla każdego obiektu $a \in A$, $i = 1, \dots, n$, *relacja dominacji* D na zbiorze A zdefiniowana jest w sposób następujący.

Mając dane obiekty $a, b \in A$, obiekt a *dominuje* obiekt b , co oznacza się aDb , wtedy i tylko wtedy, gdy ocena $g_i(a)$ jest co najmniej tak dobra jak ocena $g_i(b)$ dla każdego kryterium $g_i \in G$. Relacja dominacji jest relacją częściowego porządku na zbiorze A (jest zwrotna i przechodnia).

Informacja preferencyjna i model preferencji. W WWD, wnioski wynikające jedynie z analizy ocen obiektów na poszczególnych kryteriach, są zazwyczaj zbyt słabe aby móc przedstawić decydentowi użyteczną dla niego rekomendację. Wynika to z faktu, że często wiele obiektów jest ze sobą nieporównywalnych w sensie relacji dominacji (nie dominują się one wzajemnie).

W celu zwiększenia porównywalności obiektów, konieczne jest pozyskanie od decydenta dodatkowej informacji dotyczącej rozważanych obiektów. Informacja ta nazywana jest *informacją preferencyjną*. Poszczególne metody wspomaganie decyzji dla problemu wielokryterialnego rankingu różnią się między sobą przyjętym typem informacji preferencyjnej oraz typem *modelu preferencji* stworzonego z użyciem tej informacji. Model ten ma za zadanie agregację wektorów ocen obiektów w sposób zgodny z systemem wartości decydenta. Model preferencji może mieć postać funkcji (np. addytywnej funkcji użyteczności), jak przyjmuje się w wieloatrybutowej teorii użyteczności (ang. Multiple Attribute Utility Theory) [77], relacji binarnej (np. relacji przewyższania) [94], lub zbioru reguł decyzyjnych [48, 49, 51, 105]. Model preferencji indukuje *strukturę preferencji* w zbiorze obiektów. Właściwa *eksploatacja* tej struktury pozwala na uzyskanie rankingu obiektów, który jest następnie prezentowany decydentowi.

W UP, model preferencji jest wynikiem statystycznej analizy treningowego zbioru obiektów. Tak więc, dane treningowe są ekwiwalentem informacji preferencyjnej rozważanej w WWD.

Przykłady decyzji. Istotnym aktualnym trendem w WWD jest użycie *pośredniej informacji preferencyjnej*. Jednym z najbardziej znaczących przejawów tego trendu jest postulat indukcyjnej konstrukcji modelu preferencji na podstawie podanych przez decydenta *przykładów decyzji*. Przykłady te mogą dotyczyć zarówno rzeczywistych jak i fikcyjnych obiektów; mogą one także wynikać z obserwacji decyzji podjętych przez decydenta w przeszłości. Wyrażanie preferencji poprzez podejmowanie przykładowych decyzji nie wymaga od decydenta znajomości szczegółów zastosowanej metody wspomaganie decyzji, czy też specyfikacji wartości parametrów przyjętego modelu preferencji (takich jak progi, wagi, współczynniki substytucji, itp.).

Indukcja modelu preferencji na podstawie przykładów decyzji jest zgodna z paradygmatem uczenia się przez indukcję znanego ze sztucznej inteligencji [82]. Jest również zgodna z paradygmatem dezagregacji-agregacji [72] oraz z zasadą racjonalności a posteriori postulowaną przez March'a [81]. Uczenie się przez indukcję z przykładów decyzji wykorzystano już do konstrukcji różnych modeli preferencji, takich jak ogólna addytywna

funkcja użyteczności [28, 59], relacja przewyższania [60, 86], monotoniczne drzewo decyzyjne [40] i zbiór reguł decyzyjnych [51].

W rozważanym w rozprawie problemie wielokryterialnego rankingu obiektów, przykłady decyzji najczęściej mają postać *porównań parami* wybranych obiektów, tzw. *obiektów referencyjnych*. Jest to całkiem naturalne zważywszy, że pozycja obiektu w rankingu zależy od jego relacji z innymi obiektami znajdującymi się w tym samym rankingu. Dla dwóch obiektów a , b , najprostszy wynik porównania a z b to stwierdzenie istnienia lub braku słabej preferencji a nad b . W bardziej skomplikowanym przypadku, można również rozważać intensywność preferencji a nad b .

Czytelność modelu preferencji. Drugim ważnym postulatem WWD jest *czytelność* (ang. readability) tworzonego modelu preferencji. Postulat ten nawiązuje do koncepcji *przejrzystej skrzynki* (ang. glass box) [55], która jest przeciwieństwem “czarnej skrzynki” (ang. black box). Oznacza to, że model preferencji powinien być zrozumiały dla decydenta.

Regułowy model preferencji. Jednym z modeli spełniających oba powyższe postulaty WWD jest model logiczny w postaci zbioru reguł decyzyjnych typu “jeżeli warunki, to decyzja”, wyindukowanych z przykładów decyzji pozyskanych od decydenta. Model ten został wprowadzony do analizy decyzji przez Greco, Matarazzo i Słowińskiego [47, 49, 106]. Popularne powiedzenie, przypisywane Slovicowi [103] głosi, że “ludzie najpierw podejmują decyzje, a następnie poszukują reguł, które uzasadniają ich wybory”. Istotnie, reguły decyzyjne mają zdolność wyjaśniania przyczyn podjętych przez decydenta decyzji. Tak więc, wykorzystanie modelu preferencji w postaci zbioru reguł decyzyjnych wyindukowanych z przykładów decyzji umożliwia zarówno wyjaśnianie decyzji już podjętych, jak i rekomendację decyzji przyszłych.

Zaletą reprezentowania preferencji decydenta za pomocą reguł decyzyjnych jest możliwość wzięcia pod uwagę, jednocześnie, atrybutów różnej natury – nominalnych, porządkowych i ilościowych (numerycznych). Ponadto, nie ma konieczności dyskretyzacji atrybutów ilościowych. Regułowy model preferencji jest również atrakcyjny ze względu na wyniki analizy aksjomatycznej wszystkich trzech typów modelu preferencji rozważanych w WWD (tj., funkcji użyteczności, relacji przewyższania oraz zbioru reguł decyzyjnych) [50, 52]. Zgodnie z tą analizą, zbiór reguł decyzyjnych jest jedynym modelem preferencji, który rozważa najbardziej złożone interakcje pomiędzy kryteriami, jest niekompensacyjny, akceptuje porządkowe skale preferencji kryteriów oraz nie wymaga konwersji ocen porządkowych na ilościowe. Ponadto, reguły decyzyjne są w łatwy sposób interpretowalne przez decydentów, którzy mają w ten sposób większe zaufanie do rekomendacji proponowanych w oparciu o te reguły [51].

Wiedza dziedzinowa. Ważnym aspektem analizy problemów decyzyjnych jest *wiedza dziedzinowa*. Eksploatacja tej wiedzy umożliwia zwiększenie jakości tworzonego mo-

delu preferencji oraz zapewnienie jego kompatybilności z systemem wartości decydenta. W przypadku rozważanego problemu rankingu, wiedza dziedzinowa dotyczy zbiorów wartości kryteriów, skal preferencji tych kryteriów oraz następującej zależności monotonicznej: poprawa oceny obiektu na jednym lub więcej kryteriach nie powinna skutkować pogorszeniem się pozycji tego obiektu w relacji do pozostałych obiektów. Zgodnie z tą zależnością, decydent oczekuje od metody wspomaganie decyzji dla problemu rankingu, że końcowy ranking obiektów uzyskany tą metodą będzie zachowywał relację dominacji na zbiorze porządkowanych obiektów.

Spójność przykładów decyzji. W praktyce, przykłady decyzji podane przez decydenta są często *niespójne* ze względu na wahanie się decydenta, niestały charakter jego preferencji, lub nieobecność istotnych kryteriów [95]. The niespójności nie powinny być traktowane jako zwykłe błędy lub szum. Niosą one bowiem istotną informację, która powinna zostać wzięta pod uwagę podczas konstrukcji modelu preferencji decydenta. W przypadku informacji preferencyjnej w postaci porównań obiektów referencyjnych parami, typowej dla problemu wielokryterialnego rankingu, niespójność powstaje w efekcie naruszenia następującej ogólnej zależności monotonicznej: “jeżeli na każdym kryterium obiekt a jest preferowany nad obiektem b co najmniej tak silnie jak obiekt c jest preferowany nad obiektem d , to globalna preferencja a nad b jest nie słabsza niż globalna preferencja c nad d ”.

Teoria zbiorów przybliżonych oparta na dominacji (DRSA). W celu identyfikacji niespójności przykładów decyzji ze względu na relację dominacji, Greco, Matarazzo i Słowiński zaproponowali teorię zbiorów przybliżonych opartą na dominacji (ang. Dominance-based Rough Set Approach, DRSA) [43, 44, 47–49, 105], która została z sukcesem zastosowana do problemów klasyfikacji wielokryterialnej, wyboru i rankingu. W podejściu DRSA, niespójne przykłady decyzyjne nie są poprawiana ani ignorowane. Zamiast tego, stosuje się *strukturalizację* zbioru przykładów decyzji poprzez wyznaczenie *dolnych* i *górných* przybliżeń rozważanych zbiorów obiektów (tj., *unii klas decyzyjnych* w przypadku problemu klasyfikacji lub *relacji preferencji* w przypadku problemu rankingu), co pozwala na rozróżnienie, odpowiednio, wiedzy pewnej i możliwej. W ten sposób, możliwe jest zawężenie a priori zbioru przykładów decyzji służących do indukcji modelu preferencji do podzbioru przykładów spójnych, dzięki czemu wyindukowany model preferencji będzie cechował się większą wiarygodnością.

DRSA ze zmienną spójnością (VC-DRSA). Gdy liczba niespójnych przykładów decyzji jest relatywnie duża, dolne przybliżenia wyznaczone zgodnie z podejściem DRSA są często nieliczne, lub wręcz puste. Sytuacja taka utrudnia dalszą analizę danych. Z tego powodu, w literaturze zaproponowano liczne rozszerzenia DRSA, takie jak: różne wersje DRSA ze zmienną spójnością (VC-DRSA) (ang. Variable Consistency Dominance-based Rough Set Approach) [9–11, 61] oraz DRSA ze zmienną precyzją (VP-DRSA) (ang.

Variable Precision Dominance-based Rough Set Approach) [70]. Rozszerzenia te celowo osłabiają definicję dolnego przybliżenia unii klas decyzyjnych, co prowadzi do włączania do dolnych przybliżeń obiektów “wystarczająco spójnych”. Spójność obiektów mierzy się za pomocą przyjętej *miary spójności*, którą może być, np. przybliżona przynależność (ang. rough membership) [61, 90, 115, 116], współczynnik Bayesa (ang. Bayes factor) [102], wybrana miara confirmacji [31], lub miara ϵ [10] (systematyczny przegląd miar spójności dostępny jest w pracy [11]). W [11], autorzy porównali różne generalizacje DRSA pod kątem czterech pożądaných własności monotoniczności, oznaczonych przez (m1)–(m4). To porównanie oraz wyniki przeprowadzonych eksperymentów obliczeniowych [8, 12, 13], dotyczących problemów wielokryterialnej klasyfikacji obiektów, wykazały przewagę podejścia VC-DRSA z miarą spójności ϵ [8, 10, 11]. W dalszej części rozprawy podejście to oznaczane jest jako ϵ -VC-DRSA.

1.1.2 Problem Klasyfikacji na Podstawie Podobieństwa

Ludzie mają tendencję do rozwiązywania nowych problemów z wykorzystaniem rozwiązań *podobnych* problemów napotkanych w przeszłości. Ten proces często określa się mianem *wnioskowania na podstawie podobieństwa do znanych przypadków* (ang. case-based reasoning). Jak zaobserwowali Gilboa i Schmeidler [39], podstawowa idea takiego wnioskowania zawiera się w następującym zdaniu Hume’a [67]: “Po podobnych przyczynach spodziewamy się podobnych skutków”. Zdanie to można sparafrazować mówiąc: “Im bardziej podobne są przyczyny, tym bardziej podobnych spodziewamy się skutków”.

W rozprawie rozważono klasyfikację obiektów dokonywaną poprzez wnioskowanie na podstawie podobieństwa do znanych przypadków, określaną mianem *klasyfikacji na podstawie podobieństwa*. Konkretnie, przyjęto następujące sformułowanie problemu klasyfikacji na podstawie podobieństwa. Dany jest skończony zbiór obiektów U (*baza przypadków*) oraz skończona rodzina klas decyzyjnych \mathcal{D} . Obiekt $y \in U$ (“przypadek”) opisany jest za pomocą wartości cech (atrybutów) $f_1, \dots, f_n \in F$. Dla każdej cechy $f_i \in F$, dana jest *funkcja podobieństwa* (ang. marginal similarity function) $\sigma_{f_i} : U \times U \rightarrow [0, 1]$, taka, że wartość $\sigma_{f_i}(y, x)$ wyraża podobieństwo obiektu $y \in U$ do obiektu $x \in U$ ze względu na cechę f_i ; minimalne wymaganie stawiane przed funkcją σ_{f_i} jest następujące: dla każdego $x, y \in U$, $\sigma_{f_i}(y, x) = 1 \Leftrightarrow y$ oraz x mają tę samą wartość cechy f_i . Ponadto, dla każdego obiektu $y \in U$ dana jest informacja dotycząca (znormalizowanej) wiarygodności przynależności y do każdej z klas decyzyjnych z rodziny \mathcal{D} . W celu wyrażenia stopni przynależności do poszczególnych klas decyzyjnych, każda klasa $X \in \mathcal{D}$ modelowana jest jako zbiór rozmyty w U [117], scharakteryzowany przez funkcję przynależności $\mu_X : U \rightarrow [0, 1]$. Tak więc, każdy obiekt $y \in U$ może należeć z różnym stopniem przynależności do różnych klas decyzyjnych. Celem wspomaganie decyzji jest zaprezentowanie decydentowi rekomendacji dotyczącej nowego obiektu $z \notin U$, w postaci jego stopnia przynależności do poszczególnych klas decyzyjnych.

W UM, problemy klasyfikacji dopuszczające częściową przynależność do klas decyzyjnych często nazywane są problemami *klasyfikacji rozmytej* (ang. soft label classification) (np. [113]). Oczywiście, gdy $\mu_X \in \{0, 1\}$ dla każdej klasy $X \in \mathcal{D}$, problem klasyfikacji rozmytej sprowadza się do “standardowego” problemu klasyfikacji.

Przykładowe problemy klasyfikacji rozmytej opisane są w [87] i [109]. W pierwszej pracy, stopnie przynależności do dwóch rozważanych klas decyzyjnych podane zostały przez lekarzy, oceniających na podstawie rekordu pacjenta ryzyko powstania małopłytkowości na skutek podawania heparyny. Drugi artykuł dotyczy rozpoznawania emocji w nagranych wypowiedziach. W tej sytuacji naturalne jest, że poszczególne wypowiedzi mogą jednocześnie uzewnętrzniać różne emocje w różnym stopniu.

Klasyfikacja rozmyta nawiązuje do problemu analizy składu (ang. compositional data analysis) [112]. Przykładem może być tutaj analiza składu mleka uzyskanego od karmiących krów poddanych różnym dietom. Należy jednak zwrócić uwagę, że w problemie analizy składu, wyjścia (odpowiadające frakcjom komponentów w mieszance) są od siebie zależne, a ich suma wynosi jeden. Z kolei w praktycznych problemach klasyfikacji rozmytej, wyjścia (klasy) mogą być od siebie niezależne, a podane przez decydenta wiarygodności przynależności do poszczególnych klas nie muszą sumować się do jedynki. Przykładem takiej sytuacji jest problem diagnostyczny, polegający na określeniu wiarygodności występowania choroby X , Y oraz Z . Ponieważ pacjent $y \in U$ może cierpieć równocześnie na różne choroby, suma wiarygodności przynależności y do klas X , Y i Z może przekraczać jedynkę.

1.2 Przegląd Istniejących Podejść do Problemu Wielokryterialnego Rankingu

Problem rankingu jest przedmiotem rozważań zarówno w WWD jak i UP. Poniżej przedstawiono pokrótce różne podejścia do tego problemu zaproponowane w tych dwóch dziedzinach. Rozważane metody różnią się głównie typem wejściowej informacji preferencyjnej oraz postacią modelu preferencji tworzonego na podstawie tej informacji. Ponieważ rozważany w rozprawie problem rankingu ma naturę wielokryterialną, większą uwagę poświęcono metodom WWD, które wprost biorą pod uwagę wiedzę dziedzinową dotyczącą porządku ocen na poszczególnych kryteriach. Dzięki temu, modele preferencji tworzone w metodach WWD są zgodne z systemem wartości decydenta.

1.2.1 Podejścia WWD do Problemu Wielokryterialnego Rankingu

W WWD, problem rankingu jest adresowany przez wiele metod wielokryterialnego wspomaganie decyzji. Niektóre z tych metod wymagają od decydenta bezpośredniej spe-

cyfikacji parametrów modelu preferencji, podczas gdy inne bazują na indukcji modelu preferencji z przykładów decyzji. Spotyka się również metody “hybrydowe”. Zasadniczo, rozróżnia się trzy grupy metod, zależnie od przyjętej postaci modelu preferencji decydenta.

Pierwszą grupę tworzą metody wpisujące się w *wieloatrybutową teorię użyteczności* (ang. Multiple Attribute Utility Theory) [77], wykorzystujące model preferencji w postaci funkcji użyteczności. Wśród tych metod warto wyróżnić takie metody jak: ASSESS [24, 77], UTA [71, 101], UTA^{GMS} [59], GRIP [28], AHP [98, 99] oraz PAPRIKA [62].

Drugą grupę tworzą metody wykorzystujące model preferencji w postaci *relacji przewyższania* (ang. outranking methods). Metody należące do tej grupy dzielą się zasadniczo na dwie podgrupy: metody z rodziny ELECTRE [26, 27, 29] oraz metody z rodziny PROMETHEE [4, 14, 15]. W pierwszej podgrupie znajdują się metody wielokryterialnego rankingu takie jak: ELECTRE III [92], ELECTRE IV [66] oraz ELECTRE^{GKMS} [41, 75]. Z kolei do drugiej podgrupy zalicza się metody wielokryterialnego rankingu takie jak: PROMETHEE I, PROMETHEE II oraz PROMETHEE^{GKS} [75].

Trzecią grupę metod tworzą metody wykorzystujące model preferencji w postaci *zbioru reguł decyzyjnych*. Metody te wykorzystują przykłady decyzji podane przez decydenta w postaci porównań parami obiektów referencyjnych. Przykłady te tworzą tzw. tablicę porównań parami (ang. pairwise comparison table) [42, 45] i są analizowane z wykorzystaniem DRSA lub VC-DRSA celem identyfikacji przykładów (dostatecznie) spójnych, a następnie celem indukcji reguł decyzyjnych reprezentujących preferencje decydenta. Wygenerowane reguły decyzyjne są z kolei stosowane na zbiorze obiektów A podlegających porządkowaniu, a wynikowa struktura preferencji jest eksploatowana z wykorzystaniem wybranej procedury rankingowej celem wypracowania rekomendowanego rankingu obiektów w postaci preporządku zupełnego. Zasadniczo, można wyróżnić cztery metody WWD dla problemu rankingu wykorzystujące regułowy model preferencji: metoda α zaproponowana w [43, 44, 46, 47, 58] (oraz przypomniana w [48, 105]), metoda β scharakteryzowana w [48, 49, 104–106], metoda γ zaprezentowana w [51, 104] oraz metoda δ przedstawiona w [33]. Te cztery metody były punktem wyjścia do opracowania podejścia do problemu rankingu zaproponowanego w rozprawie.

Istnieją również metody WWD dla problemu rankingu, które nie kwalifikują się do żadnej z powyższych grup. Wśród nich można wymienić np. metodę sum ważonych (ang. weighted sum model) [30, 110]), metodę ważonych iloczynów (ang. weighted product model) [16, 83, 110] oraz metodę TOPSIS [68, 69].

1.2.2 Podejścia UP do Problemu Wielokryterialnego Rankingu

W UP, model preferencji decydenta tworzony jest przez indukcję, na podstawie danych treningowych. Model ten ma postać *funkcji użyteczności* lub *funkcji preferencji*, czyli funkcji zdefiniowanej dla par obiektów indukującej relację słabej preferencji w zbiorze porządkowanych obiektów A . W pierwszym przypadku, funkcja użyteczności wprost określa

preporządek zupełny w zbiorze A . W drugim przypadku, w celu uzyskania rekomendowanego rankingu obiektów, konieczna jest eksploatacja powstałej struktury preferencji z użyciem procedury rankingowej. Wśród metod UP wykorzystujących dane treningowe w postaci porównań obiektów referencyjnych parami warto wyróżnić następujące znane z literatury metody: SVM^{rank} [65, 73], RankBoost [34] oraz złożony klasyfikator regułowy (ang. ensembles of decision rules) [21]. Pierwsza z tych metod została użyta w przeprowadzonym porównawczym eksperymencie obliczeniowym.

Powyższe metody, chociaż efektywne obliczeniowo, mają szereg wad z punktu widzenia WWD. Po pierwsze, mają one tendencję do budowania modeli trudno interpretowalnych przez decydenta (jak np. SVM^{rank}). Po drugie, nie radzą sobie dobrze z kryteriami porządkowymi (dokonując arbitralnej konwersji skali porządkowej na ilościową). Po trzecie, nie biorą wprost pod uwagę wiedzy dziedzinowej odnośnie porządkowej natury kryteriów, co może prowadzić do tworzenia modeli preferencji niezgodnych z systemem wartości decydenta.

1.3 Przegląd Istniejących Podejść do Klasyfikacji na Podstawie Podobieństwa

Poniżej omówiono znane z literatury podejścia do problemu klasyfikacji wykorzystujące wnioskowanie na podstawie podobieństwa. Metody te można zasadniczo podzielić na dwie grupy. Pierwszą grupę tworzą metody oparte na paradygmacie “leniwego uczenia się” (ang. lazy learning). Drugą grupę tworzą metody oparte na paradygmacie “aktywnego uczenia się” (ang. active learning). Metody z pierwszej grupy, wykorzystują wnioskowanie na podstawie podobieństwa jedynie w celu zaklasyfikowania nowego (testowego) obiektu. Tymczasem, metody z drugiej grupy wykorzystują informację o wzajemnym podobieństwie obiektów (na poszczególnych cechach i/lub globalnym) w celu indukcji klasyfikatora.

1.3.1 Metody Oparte na Paradygmacie “Leniwego Uczenia się”

Metody oparte na paradygmacie “leniwego uczenia się” zasadniczo nie przetwarzają danych treningowych do czasu pojawienia się obiektu do zaklasyfikowania. Wtedy, nowy obiekt porównywany jest z obiektami treningowymi, tworzącymi tzw. bazę przypadków (ang. case base). Ostateczna klasyfikacja nowego obiektu wypracowywana jest wówczas na podstawie klasyfikacji obiektów najbardziej do niego podobnych ze względu na rozważane cechy. Wśród tego typu metod można wyróżnić następujące metody:

- metoda *k-najbliższych sąsiadów* (ang. *k-nearest-neighbors*) [32, 84];
- metody IB1, IB2, and IB3 [2];

- systemowe podejścia do wnioskowania na podstawie podobieństwa do znanych przypadków [1, 78], gdzie wnioskowanie rozumiane jest jako (cykliczny) cztero-etapowy proces złożony z faz *wyszukiwania podobnych przypadków* (ang. retrieve), ich *ponownego użycia* (ang. reuse), *rewizji* (ang. revise) oraz *zachowania* (ang. retain) do przyszłego użycia; podejścia te mają swoje korzenie w pracach Roberta Schanka dotyczących modelu pamięci dynamicznej [100];
- metoda decyzji opartych na wcześniejszych przypadkach (ang. case-based decisions) zaproponowana przez Gilboę i Schmeidlera [38, 39];
- metoda wykorzystująca *modelowanie rozmyte* (ang. fuzzy set modeling), zaproponowana w [23].

Należy zaznaczyć, że powyższe metody, ze względu na brak generalizacji danych treningowych, mogą być podatne na szum w tych danych, zarówno jeżeli idzie o nadmiarowe cechy, jak i o obserwacje odstające. Ponadto, ze względu na potencjalnie duży rozmiar bazy przypadków, klasyfikacja nowego obiektu może trwać relatywnie długo jeżeli obiekt ten jest porównywany ze wszystkim obiektami z bazy przypadków. Ta kwestia motywuje badania dotyczące technik redukcji bazy przypadków, opisane np. w [79]. Zaletą powyższych metod jest możliwość przyrostowego uczenia się poprzez włączanie nowych obiektów do bazy przypadków.

1.3.2 Metody Oparte na Paradygmacie “Aktywnego Uczenia się”

Metody oparte na paradygmacie “aktywnego uczenia się” dokonują jawnego uogólnienia zbioru treningowego w celu utworzenia klasyfikatora zdolnego do predykcji klasy decyzyjnej dla nowych obiektów. Wśród tego typu metod można wyróżnić następujące metody:

- *sztuczne sieci neuronowe z funkcjami o symetrii kołowej* (ang. radial-basis function artificial neural networks) [17, 89];
- podejścia stosujące DRSA do wnioskowania na podstawie podobieństwa do znanych przypadków (ang. dominance-based rough set approach to case-based reasoning) [53, 54, 56];
- podejścia operujące na macierzy globalnego podobieństwa obiektów [18];
- podejścia wykorzystujące algorytm Similarity Based Classification (SBC) [6].

W drugiej z powyższych metod, obiekty treningowe porównywane są parami a funkcje podobieństwa na poszczególnych cechach są agregowane z wykorzystaniem reguł decyzyjnych opartych na zależności monotonicznej “im bardziej obiekt y jest podobny do obiektu x

na rozważanych cechach, tym bardziej y należy do danej klasy X ". Metoda ta jest inspiracją dla podejścia do problemu klasyfikacji na podstawie podobieństwa zaproponowanego w rozprawie.

1.3.3 Pomiar Podobieństwa

Pomiar podobieństwa jest kluczową kwestią wszystkich podejść do wnioskowania na podstawie podobieństwa. Pytania związane z pomiarem podobieństwa nasuwają się na dwóch poziomach:

- na poziomie pojedynczych cech: jak zdefiniować właściwą miarę podobieństwa ze względu na daną cechę?
- na poziomie globalnym (wszystkich cech): jak właściwie agregować podobieństwa na pojedynczych cechach w globalne podobieństwo obiektów?

Oczywiście, właściwy dobór funkcji podobieństwa dla poszczególnych cech jest kwestią istotną. Tym niemniej, problem ten dotyka wszystkich metod klasyfikacji w oparciu o podobieństwo i dlatego jest pomijany w rozprawie (zgodnie z opisem problemu przedstawionym w Pkt. 1.1.2, zakłada się, że funkcje te są dane). Natomiast w rozprawie skoncentrowano się na drugim z ww. zagadnień, tj. na sposobie agregacji podobieństw na poszczególnych cechach w podobieństwo globalne. Agregacja ta dokonywana jest z wykorzystaniem reguł decyzyjnych, dzięki czemu unika się stosowania arbitralnych operatorów sprowadzających wektor podobieństw na poszczególnych cechach do pojedynczej liczby rzeczywistej (takich jak operator minimum, ważona norma L_p , itp.).

1.4 Motywacja Zastosowania Teorii Zbiorów Przybliżonych Opartej na Dominacji do Wielokryterialnego Rankingu i Klasyfikacji na Podstawie Podobieństwa

Pomimo, iż problemy rozważane w rozprawie, tj. wielokryterialny ranking oraz klasyfikacja na podstawie podobieństwa do znanych przypadków, są problemami o dużym praktycznym znaczeniu, wiele metod stosowanych dotychczas do tych problemów jest ciężkich w użyciu (gdyż np. wymagają podania przez decydenta zbyt obszernej informacji preferencyjnej), nie zawsze są one adekwatne do rozwiązywanego problemu (np. nie uwzględniają wprost kryteriów nominalnych/porządkowych) i/lub tworzą modele preferencji, które są niezrozumiałe dla decydenta.

Zalety zastosowania DRSA. W rozprawie zaproponowano zastosowanie do obu problemów DRSA, dzięki czemu stało się możliwe uchwycenie i modelowanie zależności monotonicznych typowych dla tych problemów. DRSA jest podejściem atrakcyjnym z punktu widzenia decydenta, ponieważ wymaga bardzo słabych założeń odnośnie dostępnych danych. DRSA może być stosowane dla danych heterogenicznych, zawierających (nawet równocześnie) atrybuty nominalne, porządkowe i ilościowe (numeryczne) – nie jest wymagana dyskretyzacja atrybutów numerycznych ani konwersja atrybutów porządkowych na ilościowe. Zastosowanie DRSA pozwala na wykorzystanie wiedzy dziedzinowej odnośnie zależności monotonicznych występujących w danych. DRSA umożliwia identyfikację niespójności z zasadą dominacji występujących w przykładach decyzji (danych uczących) poprzez wyznaczenie dolnych i górnych przybliżeń rozważanych zbiorów obiektów. Przybliżenia te są bazą do indukcji czytelnego modelu preferencji, względnie modelu klasyfikacyjnego, w postaci zbioru monotonicznych reguł decyzyjnych. Reguły te w sposób przejrzysty wskazują zależności logiczne występujące w przykładach decyzji.

1.5 Cel i Zakres Rozprawy

Ogólnym celem rozprawy jest opracowanie zastosowań DRSA do problemu wielokryterialnego rankingu oraz do problemu klasyfikacji na podstawie podobieństwa do znanych przypadków (ang. case-based reasoning). Cel ten został podzielony na cztery zadania:

- (o1) opracowanie metodyki dla problemu wielokryterialnego rankingu wykorzystującej VC-DRSA;
- (o2) analiza szeregu alternatywnych procedur eksploatacji struktury preferencji otrzymywanej w problemie wielokryterialnego rankingu w wyniku zastosowania zbioru wyindukowanych reguł decyzyjnych na zbiorze obiektów podlegających uporządkowaniu; w szczególności, analiza i porównanie własności szeregu znanych z literatury *metod rankingowych*, z których każda może zostać zastosowana w najbardziej obiecującej z rozważanych procedur eksploatacji struktury preferencji;
- (o3) opracowanie metodyki dla problemu klasyfikacji na podstawie podobieństwa do znanych przypadków wykorzystującej DRSA;
- (o4) eksperymentalna weryfikacja zaproponowanej metodyki dla problemu rankingu.

W rozprawie, główny nacisk położony został na proponowaną metodę wspomagania decyzji dla problemu wielokryterialnego rankingu. W celu weryfikacji tej metody, zaprojektowany i przeprowadzony został eksperyment obliczeniowy, w którym sześć wariantów tej metody porównano z metodą SVM^{rank} , uznawaną przez specjalistów od uczenia maszynowego za jedną z najlepszych metod uczenia się dla problemu rankingu.

Proponowana metoda klasyfikacji na podstawie podobieństwa, wykorzystująca DRSA, zilustrowana została przykładem pokazującym działanie tej nowej metody. Jej porównanie eksperymentalne z innymi metodami dla problemu klasyfikacji na podstawie podobieństwa byłoby bezzasadne, ponieważ inne metody nie uwzględniają tej samej wejściowej informacji klasyfikacyjnej (tj. wiarygodności przynależności obiektów do poszczególnych klas decyzyjnych). Ponadto, większość innych metod nie udostępnia również tej samej wyjściowej informacji klasyfikacyjnej (tj. wiarygodności przynależności obiektów do poszczególnych klas decyzyjnych, wynikającej z zastosowania wyindukowanych reguł decyzyjnych).

Rozdział 2

Podsumowanie

2.1 Wkład Rozprawy do Specjalności Inteligentnych Systemów Wspomagania Decyzji

Celem rozprawy było twórcze zastosowanie teorii zbiorów przybliżonych opartej na dominacji (DRSA) do problemu wielokryterialnego rankingu obiektów oraz do problemu klasyfikacji na podstawie podobieństwa do znanych przypadków (ang. case-based reasoning). Cel ten został zdekomponowany na cztery zadania (o1)-(o4) przedstawione w Pkt. 1.5. W ocenie autora, wszystkie te zadania zostały zrealizowane. Poniżej podsumowano najważniejsze wyniki badań przedstawionych w rozprawie.

2.1.1 Metoda dla Problemu Wielokryterialnego Rankingu Wykorzystująca VC-DRSA.

W rozprawie zaproponowano metodę porządkowania (tworzenia rankingu) obiektów VC-DRSA^{rank}. Główne cechy tej metody są następujące:

- wykorzystanie zrozumiałego dla decydenta modelu preferencji w postaci zbioru monotonicznych reguł decyzyjnych typu “jeżeli zachodzi koniunkcja warunków elementarnych na wybranych kryteriach, to sugerowane jest podjęcie określonej decyzji”; przykładowa reguła decyzyjna ma postać: “jeżeli różnica cen dysków a i b nie przekracza 100 zł oraz różnica pojemności tych dysków jest nie mniejsza niż 500 GB, to dysk a jest co najmniej tak dobry jak dysk b ”;
- zgodność z aktualnym trendem WWD – model preferencji indukowany jest na podstawie przykładów decyzji;
- wykorzystanie prostych przykładów decyzji w postaci porównań parami obiektów referencyjnych typu “obiekt a jest co najmniej tak dobry jak obiekt b ” oraz “obiekt a nie jest co najmniej tak dobry jak obiekt b ”; tego typu holistyczna informacja

preferencyjna, odpowiadająca relacji przewyższania S (ang. outranking) i relacji braku przewyższania S^c (ang. non-outranking), jest względnie prosta do pozyskania od decydenta;

- możliwość jednoczesnego uwzględnienia kryteriów ilościowych i porządkowych, bez konieczności dyskretyzacji kryteriów ilościowych oraz bez konieczności konwersji kryteriów porządkowych na numeryczne;
- metoda radzi sobie (wykorzystując pojęcie zbioru przybliżonego opartego na dominacji) z niespójnością przykładów decyzji, tj. z naruszeniem zależności monotonicznej “jeżeli na każdym kryterium obiekt a jest preferowany nad obiekt b co najmniej tak silnie jak obiekt c jest preferowany nad obiekt d , to globalna preferencja a nad b jest nie słabsza niż globalna preferencja c nad d ”; dzięki temu, metoda nie wymaga aby rozważany zbiór kryteriów był zbiorem spójnym w sensie WWD.

Ponadto, zaproponowana w rozprawie metoda porządkowania obiektów VC-DRSA^{rank} rozszerza i ulepsza wcześniejsze regułowe podejścia do problemu rankingu stosowane w WWD, wymienione w Pkt. 1.2.1. Najważniejsze rozszerzenia i ulepszenia omówiono poniżej.

- VC-DRSA^{rank} nie wymaga od decydenta definiowania gradualnych relacji preferencji dla poszczególnych kryteriów ilościowych, w przeciwieństwie do poprzednich regułowych metod dla problemu rankingu zaproponowanych w ramach WWD. Zamiast tego, wykorzystuje różnicę ocen na kryterium jako prostą miarę siły preferencji.
- VC-DRSA^{rank} jest lepiej przystosowana do rozwiązywania rzeczywistych problemów wielokryterialnego rankingu (często charakteryzujących się dużą liczbą niespójnych przykładów decyzji) ze względu na wykorzystanie DRSA ze zmienną spójnością (VC-DRSA). Dzięki takiemu podejściu, dolne przybliżenia relacji przewyższania S i braku przewyższania S^c mogą zawierać, oprócz spójnych par obiektów, również pary obiektów “dostatecznie spójne”. Spójność ta definiowana jest z wykorzystaniem miar spójności $\epsilon_S, \epsilon_{S^c}$. [10, 11]. Na bazie tak zdefiniowanych dolnych przybliżeń, VC-DRSA^{rank} indukuje probabilistyczne reguły decyzyjne, o kontrolowanej spójności.
- Rozważono dwa punkty widzenia dotyczące natury zbioru kryteriów G . Pierwszy z nich, typowy dla WWD, zakłada, że zbiór G jest spójnym zbiorem kryteriów (zgodnie z definicją przedstawioną w Pkt. 1.1). Drugi punkt widzenia, charakterystyczny dla UP, nie narzuca żadnych założeń odnośnie zbioru G (zbiór ten traktuje się jako niekoniecznie spójny zbiór kryteriów, a więc może on nie zawierać pewnych kryteriów istotnych z punktu widzenia decydenta i/lub może zawierać kryteria nadmiarowe). Dla obu punktów widzenia zaproponowano odpowiedni sposób konstrukcji tablicy porównań parami na podstawie podanych przez decydenta przykładów

decyzji. Ponadto, wskazano jak każdy z tych punktów widzenia powinien zostać wzięty pod uwagę na etapie zastosowania reguł decyzyjnych na zbiorze obiektów podlegających porządkowaniu.

- W celu indukcji probabilistycznych reguł decyzyjnych z dolnych przybliżeń relacji S i S^c , zaadoptowano nowoczesny algorytm VC-DomLEM [13]. W poprzednich regułowych podejściach WWD do problemu rankingu do indukcji reguł stosowany był algorytm DomLEM zaproponowany w 2001 roku [57, 107].
- Rozważono łącznie sześć wariantów metody VC-DRSA^{rank}, które odpowiadają sześciu sposobom definiowania struktury preferencji złożonej z relacji \mathbb{S} i \mathbb{S}^c wynikających z zastosowania wyindukowanych reguł decyzyjnych na zbiorze A obiektów podlegających porządkowaniu. Warianty te wynikają, po pierwsze, z różnego spojrzenia na zbiór kryteriów (spójny zbiór kryteriów; niekoniecznie spójny zbiór kryteriów), po drugie, z różnej interpretacji pokrycia par obiektów wyindukowanymi regułami (interpretacja jakościowa – czy para obiektów pokrywana jest przez co najmniej jedną regułę; interpretacja ilościowa – jaka jest siła najsilniejszej reguły pokrywającej parę obiektów), a po trzecie, z różnej definicji siły reguły w przypadku ilościowej interpretacji pokrycia regułami. W szczególności, rozważono:
 - wariant VC-DRSA^{rank}_{c0|1}, zakładający spójność zbioru kryteriów G (litera c pochodzi od ang. *consistent*), w którym relacje \mathbb{S} i \mathbb{S}^c definiowane są jako relacje ostre (tzn., para obiektów $(a, b) \in A \times A$ może należeć lub nie należeć do każdej z tych relacji);
 - wariant VC-DRSA^{rank}_{nc0|1}, nie zakładający spójności zbioru kryteriów G (litery *nc* pochodzą od ang. *not necessarily consistent*), w którym relacje \mathbb{S} i \mathbb{S}^c definiowane są jako relacje ostre;
 - wariant VC-DRSA^{rank}_{c0-1_{cr}}, zakładający spójność zbioru kryteriów G , w którym relacje \mathbb{S} i \mathbb{S}^c definiowane są jako relacje wartościowane (tzn., para obiektów $(a, b) \in A \times A$ może należeć do każdej z tych relacji w stopniu od 0 do 1), których wartość dla pary obiektów $(a, b) \in A \times A$ zależy od maksymalnej wiarygodności (ang. *credibility*) reguł pokrywających tę parę obiektów;
 - wariant VC-DRSA^{rank}_{nc0-1_{cr}}, nie zakładający spójności zbioru kryteriów G , w którym relacje \mathbb{S} i \mathbb{S}^c definiowane są jako relacje wartościowane, których wartość dla pary obiektów $(a, b) \in A \times A$ zależy od maksymalnej wiarygodności reguł pokrywających tę parę obiektów;
 - wariant VC-DRSA^{rank}_{c0-1_x}, zakładający spójność zbioru kryteriów G , w którym relacje \mathbb{S} i \mathbb{S}^c definiowane są jako relacje wartościowane, których wartość dla pary obiektów $(a, b) \in A \times A$ zależy od maksymalnej wartości iloczynu wiarygodności i współczynnika pokrycia (ang. *coverage factor*) reguł pokrywających tę parę obiektów;

- wariant VC-DRSA $_{nc0-1 \times}^{rank}$, nie zakładający spójności zbioru kryteriów G , w którym relacje \mathbb{S} i \mathbb{S}^c definiowane są jako relacje wartościowane, których wartość dla pary obiektów $(a, b) \in A \times A$ zależy od maksymalnej wartości iloczynu wiarygodności i współczynnika pokrycia reguł pokrywających tę parę obiektów.
- Zdefiniowano w sposób ogólny dwufazową procedurę eksploatacji struktury preferencji w zbiorze A , złożonej z ostrych/wartościowanych relacji \mathbb{S} i \mathbb{S}^c . Za pierwszą fazę przyjęto *transformację* tej struktury preferencji do pojedynczej relacji wartościowanej. Drugą fazę stanowi eksploatacja uzyskanej relacji za pomocą *metody rankingowej*, prowadząca do uzyskania preporządku zupełnego lub częściowego w zbiorze A . Zaletą tego typu procedury jest fakt, że w jej drugiej fazie można użyć dowolnej znanej z literatury metody rankingowej.
- Zdefiniowano nową miarę zgodności τ' uogólniającą współczynnik korelacji rangowej τ Kendalla. Miara ta ma zastosowanie w najbardziej ogólnym przypadku, w którym mierzy się zgodność pomiędzy rankingiem będącym preporządkiem częściowym w zbiorze A i porównaniami parami (wyrażonymi poprzez relacje S i S^c) obiektów z podzbioru zbioru A .

2.1.2 Analiza Procedur Eksploatacji Struktury Preferencji

Wkład rozprawy w zakresie eksploatacji struktury preferencji jest następujący.

- Rozważono cztery alternatywne procedury eksploatacji struktury preferencji powstającej w wyniku zastosowania wyindukowanych reguł decyzyjnych na zbiorze A obiektów podlegających porządkowaniu. Za najbardziej obiecującą procedurę uznaną opisaną powyżej procedurę dwufazową.
- Zaproponowano konkretny sposób realizacji pierwszej fazy dwufazowej procedury eksploatacji struktury preferencji.
- W drugiej fazie dwufazowej procedury eksploatacji struktury preferencji, rozważono zastosowanie jednej z pięciu metod rankingowych znanych z literatury: *Net Flow Rule*, *Iterative Net Flow Rule*, *Min in Favor*, *Iterative Min in Favor*, *Leaving and Entering Flows*.
- Ustalono, że procedura dwufazowa uogólnia stosowaną dotychczas w podejściach regułowych jednofazową procedurę eksploatacji struktury preferencji tworzącą ranking obiektów na podstawie ich bilansu przepływów (ang. net flow score).
- W poszukiwaniu “najlepszej” metody rankingowej, rozważono 11 pożądaných własności metod rankingowych. Część z tych własności było już analizowanych w literaturze, inne zostały zdefiniowane w rozprawie.

- W wyniku przeprowadzenia analizy formalnej pięciu wyżej wymienionych metod rankingowych pod kątem ich pożądaných własności, ustalono, które z własności są spełniane przez poszczególne metody. W efekcie stwierdzono, że najlepsze własności ma metoda *Net Flow Rule*.

2.1.3 Motoda Klasyfikacji na Podstawie Podobieństwa Wykorzystująca DRSA

W rozprawie zaproponowano regułową metodę klasyfikacji dla problemu klasyfikacji na podstawie podobieństwa opisanego w Pkt. 1.1.2. Główne cechy tej metody są następujące:

- stosuje DRSA do wnioskowania na podstawie podobieństwa do znanych przypadków (ang. case-based reasoning);
- unika agregacji podobieństwa obiektów na poszczególnych cechach w podobieństwo globalne za pomocą arbitralnych operatorów sprowadzających wektor podobieństw na poszczególnych cechach do pojedynczej liczby rzeczywistej (takich jak operator minimum, ważona norma L_p , itp.); zamiast tego, podobieństwo globalne obiektów modelowane jest przez reguły decyzyjne wykorzystujące relację dominacji, wyindukowane na podstawie przykładów klasyfikacji; regułowy model podobieństwa jest najmniej obciążony arbitralnymi założeniami odnośnie do agregacji podobieństw na poszczególnych cechach;
- jest metodą “aktywnego uczenia się”, uczącą się reprezentacji podobieństwa za pomocą zbioru reguł decyzyjnych; w porównaniu do metod wnioskowania na podstawie podobieństwa opartych na paradygmacie “leniwego uczenia się”, jest mniej podatna na szum w danych treningowych, zarówno jeżeli idzie o nadmiarowe cechy, jak i o obserwacje odstające;
- ponieważ indukowane reguły decyzyjne wykorzystują jedynie porządkowe własności funkcji podobieństwa zdefiniowanych dla poszczególnych cech, proponowana metoda jest niewrażliwa na dobór tych funkcji, pod warunkiem, iż są one porządkowo równoważne;
- zakłada, że każda klasa decyzyjna X jest zbiorem rozmytym w zbiorze obiektów treningowych U , scharakteryzowanym przez funkcję przynależności $\mu_X : U \rightarrow [0, 1]$;
- korzysta jedynie z porządkowych własności funkcji przynależności do poszczególnych klas decyzyjnych.

Ponadto, zaproponowana w rozprawie metoda klasyfikacji na podstawie podobieństwa rozszerza i ulepsza wcześniejsze regułowe podejścia do tego problemu wykorzystujące DRSA, wymienione w Pkt. 1.3.2. Najważniejsze rozszerzenia i ulepszenia omówiono poniżej.

- Zaproponowano *nową zależność monotoniczną*: “im bardziej obiekt y jest podobny do obiektu x na rozważanych cechach, tym bardziej podobna jest przynależność y do klasy X do przynależności x do klasy X ”. Zasada ta wiernie odzwierciedla monotoniczność charakterystyczną dla wnioskowania przez podobieństwo, tj. zależność pomiędzy globalnym podobieństwem obiektów a ich podobieństwem na poszczególnych cechach. Odstępstwa od tej zasady powodują niespójności w zbiorze obiektów, analizowane z wykorzystaniem DRSA.
- Dokonano rewizji definicji globalnych relacji bliskości obiektów, zaprezentowanych we wcześniejszej wersji proponowanej metody [108], poprzez dopuszczenie w jednej z tych definicji, że parametr α może być mniejszy niż zero, a parametr β może być większy niż zero.
- Wprowadzono pojęcia *tablicy podobieństw*, relacji *x -dominacji* oraz *x -pozytywnego* i *x -negatywnego stożka dominacji* w przestrzeni podobieństwa, gdzie x jest obiektem referencyjnym.
- Zaproponowano sposób indukcji reguł decyzyjnych. Reguły te są indukowane niezależnie dla każdej klasy decyzyjnej X i każdego obiektu referencyjnego x , dla dolnych lub górnych przybliżeń zbiorów takich obiektów, których przynależność do klasy X zawiera się w przedziale $[\alpha, \beta]$ oraz poza przedziałem (α, β) , gdzie α i β są wartościami funkcji przynależności do klasy X takimi, że przedział $[\alpha, \beta]$ zawiera przynależność referenta x . Przykładowa reguła decyzyjna ma postać: “jeżeli podobieństwo pacjenta y do pacjenta referencyjnego x na cesze *temperatura* jest ≥ 0.8 oraz podobieństwo y do x na cesze *ból mięśni* jest ≥ 1.0 , to przynależność pacjenta y do klasy decyzyjnej *grypa* zawiera się w przedziale $[0.8, 1.0]$ ”, gdzie obiekt referencyjny x posiada następujące wartości cech: *temperatura*=39, *ból mięśni*=tak, *ból głowy*=tak.
- Zaproponowano schemat zastosowania wyindukowanych pewnych/możliwych reguł decyzyjnych do nowego obiektu (nowego przypadku) z , w celu predykcji jego przynależności do danej klasy decyzyjnej X . Sposób ten polega na wyznaczeniu punktów (ang. score) dla każdego stopnia przynależności pokrytego przez część decyzyjną przynajmniej jednej reguły dopasowanej do obiektu z . Następnie, przewiduje się dla z ten stopień przynależności, który uzyskał najwięcej punktów.

Warto podkreślić, że zaproponowana w rozprawie metoda dla problemu klasyfikacji na podstawie podobieństwa do znanych przypadków dostarcza decydentowi szeregu użytecznych informacji, których inne metody stosowane dla tego problemu nie udostępniają. W szczególności, w związku z określeniem stopnia przynależności nowego obiektu z do klasy X , decydent może uzyskać informację o regułach pokrywających ten obiekt, a także, o obiektach uczących, które wspierają te reguły. Ponadto, reguły są czytelnym interpretatorem warunków podobieństwa.

2.1.4 Eksperymentalna Weryfikacja Metody dla Problemu Wielokryterialnego Rankingu

Przeprowadzony eksperyment miał na celu porównania sześciu wariantów metody VC-DRSA^{rank} z metodą SVM^{rank}, uznawaną przez specjalistów od uczenia maszynowego za jedną z najlepszych metod uczenia się dla problemu rankingu.

Eksperyment wykazał, że zaproponowane w rozprawie podejście do problemu wielokryterialnego rankingu jest konkurencyjne w stosunku do metody SVM^{rank}. Biorąc pod uwagę ponadto szerszy zakres stosowalności proponowanej metody (którą można stosować wprost również do problemów z kryteriami porządkowymi), a także dobrą interpretowalność regułowego modelu preferencji, metoda VC-DRSA^{rank} jawi się jako bardziej atrakcyjna z punktu widzenia decydenta.

W celu porównania siedmiu rozważanych w eksperymencie metod (6 wariantów metody VC-DRSA^{rank} oraz metody SVM^{rank}), użyto współczynnika korelacji rangowej τ Kendalla oraz jego wariantu τ^{-I} . Ten drugi współczynnik nie uwzględnia par obiektów $(a, b) \in A \times A$ takich, że zgodnie z prawdziwym rankingiem na zbiorze A , obiekt a ma tę samą rangę co obiekt b . Zgodnie z miarą τ , warianty VC-DRSA^{rank}_{c0|1} i VC-DRSA^{rank}_{nc0|1} uzyskały najlepszy wynik dla analizowanych 11 zbiorów danych. Z drugiej strony, wersja VC-DRSA^{rank}_{nc0-1 \times} uzyskała najlepszy wynik ze względu na miarę τ^{-I} .

Eksperyment wykazał również, że zastosowanie VC-DRSA z miarą spójności ϵ daje w problemie rankingu lepsze wyniki, niż zastosowanie DRSA.

Przeprowadzono również analizę statystyczną średnich wartości miar τ i τ^{-I} uzyskanych w eksperymencie. Celem tej analizy było ustalenie, czy porównywane metody są statystycznie istotnie różne, a jeżeli tak, to które konkretnie metody różnią się trafnością przewidywania rankingu. Przyjęto poziom istotności statystycznej $\alpha = 0.01$. Używając testu Friedmana, hipotezę zerową, mówiącą, iż wszystkie 7 metod działa równie dobrze, udało się odrzucić jedynie dla miary τ . Jednak zastosowany następnie test post-hoc Nemenyi nie wykazał znaczącej różnicy pomiędzy żadnymi dwoma metodami. Wykonano również serię sparowanych testów rangowych Wilcoxon w celu porównania poszczególnych metod w parach. Testy te pozwoliły na wyciągnięcie następujących konkluzji:

- VC-DRSA^{rank}_{c0|1} działa ze względu na miarę τ lepiej niż VC-DRSA^{rank}_{c0-1 ϵ} oraz lepiej niż VC-DRSA^{rank}_{nc0-1 ϵ} ,
- VC-DRSA^{rank}_{nc0|1} działa ze względu na miarę τ lepiej niż VC-DRSA^{rank}_{c0-1 ϵ} oraz lepiej niż VC-DRSA^{rank}_{nc0-1 ϵ} ,
- VC-DRSA^{rank}_{nc0-1 \times} działa ze względu na miary τ i τ^{-I} lepiej niż VC-DRSA^{rank}_{c0-1 \times} .

2.2 Kierunki Dalszych Badań

Poniżej zaprezentowano kilka możliwych kierunków dalszych badań.

- Rozszerzenie eksperymentu obliczeniowego dotyczącego metody VC-DRSA^{rank}, po pierwsze poprzez porównanie także z innymi metodami niż SVM^{rank}, a po drugie poprzez rozważenie większej liczby zbiorów danych, włączając w to rzeczywiste problemy decyzyjne.
- Rozwój innych obiecujących zastosowań DRSA, np. dla problemu porządkowania z hierarchiczną strukturą kryteriów oraz dla problemu porządkowania, w którym informacja preferencyjna dana jest w postaci przypisań obiektów do uporządkowanych klas decyzyjnych [76].
- Dalszy rozwój zaproponowanej w rozprawie metody dla problemu klasyfikacji na podstawie podobieństwa, w szczególności poprzez opracowanie algorytmów selekcji obiektów referencyjnych oraz poprzez zastosowanie modelu ze zmienną spójnością (VC-DRSA).
- Eksperymentalne porównanie zaproponowanej w rozprawie metody klasyfikacji na podstawie podobieństwa z innymi metodami, godząc się na pewne uproszczenia w sformułowaniu problemu, takie jak:
 - przyjęcie, że klasy decyzyjne są zbiorami ostrymi (a nie rozmytymi) – w takim przypadku staje się możliwe porównanie proponowanej metody z podejściami do problemu klasyfikacji na podstawie podobieństwa rozważanymi w dziedzinie Uczenia Maszynowego, np. z metodą k-NN,
 - przyjęcie, że klasyfikacja nie musi być dokonywana na podstawie podobieństwa – w takim przypadku staje się możliwe porównanie proponowanej metody z podejściami do problemu klasyfikacji rozmytej rozważanymi w dziedzinie Uczenia Maszynowego, np. z metodą fuzzy-input fuzzy-output SVM [109].

Ponadto, interesujące byłoby również zastosowanie zaproponowanej w rozprawie metody klasyfikacji na podstawie podobieństwa do problemów klasyfikacji porządkowej. Takie zastosowanie pozwoliłoby porównać tę metodę z klasycznym podejściem do klasyfikacji porządkowej wykorzystującym DRSA.

Literatura

- [1] Aamodt Agnar and Enric Plaza. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *Artificial Intelligence Communications*, 7:39–52, 1994.
- [2] David W. Aha, Kibler Dennis, and Albert Marc K. Instance-based learning algorithms. *Machine Learning*, 6:37–66, 1991.
- [3] Nir Ailon and Mehryar Mohri. An efficient reduction of ranking to classification. In Rocco Servedio and Tong Zhang, editors, *Proceedings of the 21st Annual Conference on Learning Theory, Helsinki, Finland, 9-12 July 2008*, pages 87–97, 2008.
- [4] Majid Behzadian, R.B. Kazemzadeh, A. Albadvi, and M. Aghdasi. PROMETHEE: A comprehensive literature review on methodologies and applications. *European Journal of Operational Research*, 200(1):198–215, 2010.
- [5] Valerie Belton and Theodor Stewart. *Multiple Criteria Decision Analysis: An Integrated Approach*. Kluwer, Dordrecht, 2002.
- [6] Axel E. Bernal, Karen Hospevian, Tayfun Karadeniz, and Jean-Louis Lassez. Similarity based classification. In M.R. Berthold et al., editors, *IDA 2003*, volume 2810 of *Lecture Notes in Computer Science*, pages 187–197, Berlin, 2003. Springer.
- [7] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [8] Jerzy Błaszczyński. *Rule Models for Ordinal Classification in Variable Consistency Rough Set Approaches*. PhD thesis, Poznań University of Technology, 2010.
- [9] Jerzy Błaszczyński, Salvatore Greco, Roman Słowiński, and Marcin Szelaĝ. On variable consistency dominance-based rough set approaches. In S. Greco, Y. Hata, S. Hirano, M. Inuiguchi, S. Miyamoto, H. S. Nguyen, and R. Słowiński, editors, *Rough Sets and Current Trends in Computing 2006*, volume 4259 of *Lecture Notes in Artificial Intelligence*, pages 191–202, Berlin Heidelberg, 2006. Springer.
- [10] Jerzy Błaszczyński, Salvatore Greco, Roman Słowiński, and Marcin Szelaĝ. Monotonic variable consistency rough set approaches. In J. Yao, P. Lingras, W. Wu, M. Szczuka, N. J. Cercone, and Dominik Ślęzak, editors, *Rough Sets and Knowledge*

- Technology 2007*, volume 4481 of *Lecture Notes in Artificial Intelligence*, pages 126–133, Berlin Heidelberg, 2007. Springer.
- [11] Jerzy Błaszczyński, Salvatore Greco, Roman Słowiński, and Marcin Szelaĝ. Monotonic variable consistency rough set approaches. *International Journal of Approximate Reasoning*, 50(7):979–999, 2009.
- [12] Jerzy Błaszczyński, Roman Słowiński, and Marcin Szelaĝ. Probabilistic rough set approaches to ordinal classification with monotonicity constraints. In Eyke Hüllermeier, Rudolf Kruse, and F. Hoffmann, editors, *IPMU 2010*, volume 6178 of *Lecture Notes in Artificial Intelligence*, pages 99–108, Berlin Heidelberg, 2010. Springer.
- [13] Jerzy Błaszczyński, Roman Słowiński, and Marcin Szelaĝ. Sequential covering rule induction algorithm for variable consistency rough set approaches. *Information Sciences*, 181:987–1002, 2011.
- [14] Jean Pierre Brans and Bertrand Mareschal. PROMETHEE methods. In Jose Figueira, Salvatore Greco, and Matthias Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, Operations Research & Management Science, pages 163–196. Springer, New York, 2005.
- [15] Jean Pierre Brans, Bertrand Mareschal, and Philippe Vincke. Promethee: a new family of outranking methods in multicriteria analysis. In J.P. Brans, editor, *Operational Research, IFORS 84*, pages 477–490. North Holland, Amsterdam, 1984.
- [16] Percy Williams Bridgman. *Dimensional Analysis*. Yale University Press, 1922.
- [17] David S. Broomhead and David Lowe. Radial basis functions, multi-variable functional interpolation and adaptive networks. Technical Report 4148, Royal Signals and Radar Establishment, 1988.
- [18] Yihua Chen, Eric K. Garcia, and Maya R. Gupta. Similarity-based classification: Concepts and algorithms. *Journal of Machine Learning Research*, 10:747–776, 2009.
- [19] William W. Cohen, Robert E. Schapire, and Yoram Singer. Learning to order things. *Journal of Artificial Intelligence Research*, 10:243–270, 1999.
- [20] Salvatore Corrente, Salvatore Greco, Miłosz Kadziński, and Roman Słowiński. Robust ordinal regression in preference learning and ranking. *Machine Learning*, 93:381–422, 2013.
- [21] Krzysztof Dembczyński, Wojciech Kotłowski, Roman Słowiński, and Marcin Szelaĝ. Learning of rule ensembles for multiple attribute ranking problems. In Johannes Fürnkranz and Eyke Hüllermeier, editors, *Preference Learning*, pages 217–247. Springer, Berlin, 2010.

- [22] Michael Doumpos and Constantin Zopounidis. Preference disaggregation and statistical learning for multicriteria decision support: a review. *European Journal of Operational Research*, 209(3):203–214, 2012.
- [23] Didier Dubois, Henri Prade, Francesca Esteva, Pere Garcia, Lluís Godo, and Ramon Lopez de Mantaras. Fuzzy set modelling in case-based reasoning. *International Journal of Intelligent Systems*, 13:345–373, 1998.
- [24] Peter H. Farquhar. State of the art – utility assessment methods. *Management Science*, 30(11):1283–1300, 1984.
- [25] José Figueira, Salvatore Greco, and Matthias Ehrgott, editors. *Multiple Criteria Decision Analysis: State of the Art Surveys*. Operations Research & Management Science. Springer, New York, 2005.
- [26] José Figueira, Salvatore Greco, Bernard Roy, and Roman Słowiński. ELECTRE methods: Main features and recent developments. In P. Pardalos, D. Hearn, and C. Zopounidis, editors, *Handbook of Multicriteria Analysis*, volume 103 of *Applied Optimization*, pages 51–89. Springer, 2010.
- [27] José Figueira, Salvatore Greco, Bernard Roy, and Roman Słowiński. An overview of ELECTRE methods and their recent extensions. *Journal of Multi-Criteria Decision Analysis*, 20(1-2):61–85, 2013.
- [28] José Figueira, Salvatore Greco, and Roman Słowiński. Building a set of additive value functions representing a reference preorder and intensities of preference: Grip method. *European Journal of Operational Research*, 195(2):460–486, 2009.
- [29] José Figueira, Vincent Mousseau, and Bernard Roy. ELECTRE methods. In José Figueira, Salvatore Greco, and Matthias Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, Operations Research & Management Science, pages 133–162. Springer, New York, 2005.
- [30] Peter C. Fishburn. *Additive Utilities with Incomplete Product Set: Applications to Priorities and Assignments*. Operations Research Society of America (ORSA), 1967.
- [31] Branden Fitelson. *Studies in Bayesian Confirmation Theory*. PhD thesis, University of Wisconsin-Madison, 2001.
- [32] Evelyn Fix and J. L. Hodges. Discriminatory analysis, nonparametric discrimination: Consistency properties. Technical report 4, USAF School of Aviation Medicine, Randolph Field, Texas, 1951.
- [33] Philippe Fortemps, Salvatore Greco, and Roman Słowiński. Multicriteria decision support using rules that represent rough-graded preference relations. *European Journal of Operational Research*, 188(1):206–223, 2008.

- [34] Yoav Freund, Raj Iyer, Robert E. Schapire, and Yoram Singer. An efficient boosting algorithm for combining preferences. *Journal of Machine Learning Research*, 6(4):933–969, 2003.
- [35] Jerome Friedman. Recent advances in predictive (machine) learning. *Journal of Classification*, 23:175–197, 2006.
- [36] Johannes Fürnkranz and Eyke Hüllermeier, editors. *Preference Learning*. Springer, Berlin, 2010.
- [37] Johannes Fürnkranz and Eyke Hüllermeier, editors. *Preference Learning: An Introduction*, pages 3–8. Springer, 2010.
- [38] Itzhak Gilboa and David Schmeidler. Case-based decision theory. *The Quarterly Journal of Economics*, 110(3):605–639, August 1995.
- [39] Itzhak Gilboa and David Schmeidler. *A Theory of Case-Based Decisions*. Number 9780521003117 in Cambridge Books. Cambridge University Press, February 2001.
- [40] Silvio Giove, Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. Variable consistency monotonic decision trees. In James J. Alpigini, James F. Peters, Andrzej Skowron, and Ning Zhong, editors, *Rough Sets and Current Trends in Computing (RSCTC 2002)*, volume 2475 of *Lecture Notes in Artificial Intelligence*, pages 247–254. Springer, Malvern, USA, 2002.
- [41] Salvatore Greco, Miłosz Kadziński, Vincent Mousseau, and Roman Słowiński. Electre^{GKMS}: Robust ordinal regression for outranking methods. *European Journal of Operational Research*, 214(10):118–135, 2011.
- [42] Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. Rough set approach to multi-attribute choice and ranking problems. Technical Report 38/95, ICS, Warsaw University of Technology, Warsaw, Poland, 1995.
- [43] Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. Rough approximation of a preference relation by dominance relations. In S. Tsumoto, S. Kobayashi, T. Yokomori, H. Tanaka, and A. Nakamura, editors, *Proc. of the Fourth Internat. Workshop on Rough Sets, Fuzzy Sets and Machine Discovery (RSFD'96)*, pages 125–130. Tokyo University Press, 1996.
- [44] Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. Rough approximation of preference relation by dominance relations. ICS Research Report 16/96, Warsaw University of Technology, Warsaw, 1996.

- [45] Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. Rough set approach to multi-attribute choice and ranking problems. In G. Fandel and T. Gal, editors, *Multiple Criteria Decision Making, Proceedings of the Twelfth International Conference, Hagen, Germany*, pages 318–329, Berlin, 1997. Springer.
- [46] Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. Rough approximation of a preference relation in a pairwise comparison table. In Polkowski Lech et al., editors, *Rough Sets in Knowledge Discovery*, volume 2, chapter 2, pages 13–36. Springer, Berlin Heidelberg, 1998.
- [47] Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. Rough approximation of a preference relation by dominance relations. *European Journal of Operational Research*, 117:63–83, 1999.
- [48] Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. The use of rough sets and fuzzy sets in MCDM. In Tomas Gal, Theodor J. Stewart, and Thomas Hanne, editors, *Multicriteria Decision Making: Advances in MCDM models, Algorithms, Theory, and Applications*, volume 21 of *International Series in Operations Research & Management Science*, chapter 14, pages 14.1–14.59. Kluwer, Dordrecht, 1999.
- [49] Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. Rough sets theory for multicriteria decision analysis. *European Journal of Operational Research*, 129(1):1–47, 2001.
- [50] Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. Axiomatic characterization of a general utility function and its particular cases in terms of conjoint measurement and rough-set decision rules. *European Journal of Operational Research*, 158(2):271–292, 2004.
- [51] Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. Decision rule approach. In Jose Figueira, Salvatore Greco, and Matthias Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, Operations Research & Management Science, chapter 13, pages 507–562. Springer, New York, 2005.
- [52] Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. Preference representation by means of conjoint measurement and decision rule model. In Denis Bouyssou, Eric Jacquet-Lagrèze, Patrice Perny, Roman Słowiński, Daniel Vanderpooten, and Philippe Vincke, editors, *Aiding Decisions with Multiple Criteria - Essays in Honor of Bernard Roy*, pages 263–313. Kluwer, Boston, 2005.
- [53] Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. Dominance-based rough set approach to case-based reasoning. In V. Torra, Y. Narukawa, A. Valls, and J. Domingo-Ferrer, editors, *Modelling Decisions for Artificial Intelligence*, volume

- 3885 of *Lecture Notes in Artificial Intelligence*, pages 7–18, Berlin Heidelberg, 2006. Springer.
- [54] Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. Case-based reasoning using gradual rules induced from dominance-based rough approximations. In G. Wang, T. Li, J. W. Grzymała-Busse, D. Miao, Andrzej Skowron, and Yiyu Y. Yao, editors, *Rough Sets and Knowledge Technology (RSKT 2008)*, volume 5009 of *Lecture Notes in Artificial Intelligence*, pages 268–275, Berlin, 2008. Springer.
- [55] Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. Dominance-based rough set approach to interactive multiobjective optimization. In Juergen Branke, Kalyanmoy Deb, Kaisa Miettinen, and Roman Słowiński, editors, *Multiobjective Optimization: Interactive and Evolutionary Approaches*, volume 5252 of *Lecture Notes in Computer Science*, chapter 5, pages 121–156. Springer, Berlin, 2008.
- [56] Salvatore Greco, Benedetto Matarazzo, and Roman Słowiński. Granular computing for reasoning about ordered data: the dominance-based rough set approach. In Witold Pedrycz, Andrzej Skowron, and Vladik Kreinovich, editors, *Handbook of Granular Computing*, chapter 15, pages 347–373. John Wiley and Sons, Chichester, 2008.
- [57] Salvatore Greco, Benedetto Matarazzo, Roman Słowiński, and Jerzy Stefanowski. An algorithm for induction of decision rules consistent with the dominance principle. In Wojciech Ziarko and Yiyu Y. Yao, editors, *Rough Sets and Current Trends in Computing 2001*, volume 2005 of *Lecture Notes in Artificial Intelligence*, pages 304–313, Berlin, 2001. Springer.
- [58] Salvatore Greco, Benedetto Matarazzo, Roman Słowiński, and Alexis Tsoukiàs. Exploitation of a rough approximation of the outranking relation in multicriteria choice and ranking. In Theodor J. Stewart and Robin C. van den Honert, editors, *Trends in Multicriteria Decision Making*, volume 465 of *Lecture Notes in Economics and Mathematical Systems*, pages 45–60, Berlin, 1998. Springer.
- [59] Salvatore Greco, Vincent Mousseau, and Roman Słowiński. Ordinal regression revisited: multiple criteria ranking using a set of additive value functions. *European Journal of Operational Research*, 191(2):415–435, 2008.
- [60] Salvatore Greco, Roman Słowiński, Jose Figueira, and Vincent Mousseau. Robust ordinal regression. In Matthias Ehrgott, Jose Figueira, and Salvatore Greco, editors, *Trends in Multiple Criteria Decision Analysis*, International Series in Operations Research & Management Science, chapter 9, pages 241–283. Springer, New York, 2010.

- [61] Salvatore Greco, Roman Słowiński, Benedetto Matarazzo, and Jerzy Stefanowski. Variable consistency model of dominance-based rough sets approach. In Wojciech Ziarko and Yiyu Y. Yao, editors, *Rough Sets and Current Trends in Computing 2001*, volume 2005 of *Lecture Notes in Artificial Intelligence*, pages 170–181, Berlin, 2001. Springer.
- [62] Paul Hansen and Franz Ombler. A new method for scoring additive multi-attribute value models using pairwise rankings of alternatives. *Journal of Multi-Criteria Decision Analysis*, 15:87–107, 2009.
- [63] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer, 2001.
- [64] Ralf Herbrich, Thore Graepel, and Klaus Obermayer. Regression models for ordinal data: A machine learning approach. Technical report TR-99/03, TU Berlin, 1999.
- [65] Ralf Herbrich, Thore Graepel, and Klaus Obermayer. *Large Margin Rank Boundaries for Ordinal Regression*, chapter 7, pages 115–132. MIT Press, January 2000.
- [66] Jean-Christophe Hugonnard and Bernard Roy. Le plan d’extension du métro en banlieue parisienne, un cas type d’application de l’analyse multicritère. *Les Cahiers Scientifiques de la Revue Transports*, 6(6):77–108, 1982. 1er trimestre.
- [67] David Hume. *An Enquiry Concerning Human Understanding*. Clarendon Press, Oxford, 1748.
- [68] Ching-Lai Hwang, Young-Jou Lai, and Ting-Yun Liu. A new approach for multiple objective decision making. *Computers and Operational Research*, 20:889–899, 1993.
- [69] Ching-Lai Hwang and Kwangsun Yoon. *Multiple Attribute Decision Making: Methods and Applications*. Springer, New York, 1981.
- [70] Masahiro Inuiguchi and Yukihiro Yoshioka. Variable-precision dominance-based rough set approach. In S. Greco, Y. Hata, S. Hirano, M. Inuiguchi, S. Miyamoto, H. S. Nguyen, and R. Słowiński, editors, *Rough Sets and Current Trends in Computing 2006*, volume 4259 of *Lecture Notes in Artificial Intelligence*, pages 203–212, Berlin Heidelberg, 2006. Springer.
- [71] Eric Jacquet-Lagrèze and Yannis Siskos. Assessing a set of additive utility functions for multicriteria decision making: The UTA method. *European Journal of Operational Research*, 10:151–164, 1982.
- [72] Eric Jacquet-Lagrèze and Yannis Siskos. Preference disaggregation: 20 years of mcda experience. *European Journal of Operational Research*, 130(2):233–245, 2001.

- [73] Thorsten Joachims. Optimizing search engines using clickthrough data. In *Proceedings of the International Conference on Knowledge Discovery and Data Mining (ACM SIGKDD 2002)*, pages 133–142, Alberta, Canada, 2002. ACM.
- [74] Thorsten Joachims. Training linear SVMs in linear time. In *Proceedings of the ACM Conference on Knowledge Discovery and Data Mining (KDD)*, 2006.
- [75] Miłosz Kadziński. *New directions in Robustness Analysis and Preference Modeling in Multiple Criteria Decision Aiding*. NAKOM, 2012.
- [76] Miłosz Kadziński, Roman Słowiński, and Marcin Szeląg. Dominance-based rough set approach to multiple criteria ranking with sorting-specific preference information. In Stan Matwin and Jan Mielniczuk, editors, *Challenges in Computational Statistics and Data Mining*, volume 605 of *Studies in Computational Intelligence*. Springer, to appear.
- [77] Ralph L. Keeney and Howard Raiffa. *Decisions with multiple objectives: Preferences and value tradeoffs*. Wiley, 1976.
- [78] Janet Kolodner. *Case-Based Reasoning*. Morgan Kaufmann, San Mateo, 1993.
- [79] Yan Li, Simon Chi Keung Shiu, Sankar Kumar Pal, and James N.-K. Liu. A rough set-based case-based reasoner for text categorization. *International Journal of Approximate Reasoning*, 41:229–255, 2006.
- [80] Tie-Yan Liu. *Learning to Rank for Information Retrieval*. Springer, Berlin, 2011.
- [81] James G. March. Bounded rationality, ambiguity, and the engineering of choice. In David E. Bell, Howard Raiffa, and Amos Tversky, editors, *Decision Making, Descriptive, Normative and Prescriptive Interactions*, pages 33–58. Cambridge University Press, New York, 1988.
- [82] Ryszard S. Michalski. A theory and methodology of inductive learning. In Ryszard S. Michalski, Jaime G. Carbonell, and Tom M. Mitchell, editors, *Machine Learning: An Artificial Intelligence Approach*, pages 83–129. Tioga Publishing, Palo Alto, 1983.
- [83] Dawid W. Miller and Martin K. Starr. *Executive Decisions and Operations Research*. Prentice-Hall, Inc, 1969.
- [84] Tom Mitchell. *Machine Learning*. McGraw Hill, 1997.
- [85] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of Machine Learning*. Adaptive Computation and Machine Learning. MIT Press, 2012.

- [86] Vincent Mousseau and Roman Słowiński. Inferring an ELECTRE TRI model from assignment examples. *Journal of Global Optimization*, 12(2):157–174, 1998.
- [87] Quang Nguyen, Hamed Valizadegan, and Milos Hauskrecht. Learning classification models with soft-label information. *Journal of the American Medical Informatics Association*, 21:501–508, 2014.
- [88] Włodzimierz Ogryczak and Rudolf Vetschera, editors. *Methodological foundations of multi-criteria decision making*, volume 158(2) of *European Journal of Operational Research*, pages 267–532. Springer, 2004.
- [89] Yen-Jen Oyang, Shien-Ching Hwang, Yu-Yen Ou, Chien-Yu Chen, and Zhi-Wei Chen. Data classification with radial basis function networks based on a novel kernel density estimation algorithm. *IEEE Transactions on Neural Networks*, 16(1):225–236, 2005.
- [90] Zdzisław Pawlak and Andrzej Skowron. Rough membership functions. In Ronald R. Yager, Mario Fedrizzi, and Janusz Kacprzyk, editors, *Advances in the Dempster-Shafer Theory of Evidence*, pages 251–271. John Wiley & Sons, Inc., New York, 1994.
- [91] Filip Radlinski and Thorsten Joachims. Learning to rank from implicit feedback. In *Proceedings of the ACM Conference on Knowledge Discovery and Data Mining (KDD-05)*, pages 239–248, 2005.
- [92] Bernard Roy. ELECTRE III : Un algorithme de classements fondé sur une représentation floue des préférences en présence de critères multiples. *Cahiers du CERO*, 20(1):3–24, 1978.
- [93] Bernard Roy. *Méthodologie Multicritère d’Aide à la Décision*. Economica, Paris, 1985.
- [94] Bernard Roy. The outranking approach and the foundation of ELECTRE methods. *Theory and Decision*, 31:49–73, 1991.
- [95] Bernard Roy. *Multicriteria Methodology for Decision Aiding*. Kluwer, Dordrecht, 1996.
- [96] Bernard Roy and Denis Bouyssou. *Aide Multicritère à la Décision: Méthodes et Cas*. Economica, Paris, 1993.
- [97] Bernard Roy and Roman Słowiński. Questions guiding the choice of a multicriteria decision aiding method. *EURO Journal on Decision Processes*, 1(1):69–97, 2013.
- [98] Thomas Saaty. *The Analytic Hierarchy Process*. McGraw Hill, New York, 1980.

- [99] Thomas L. Saaty. *Fundamentals of Decision Making and Priority Theory*. RWS Publications, Pittsburgh, Pennsylvania, 2001.
- [100] Roger Schank. *Dynamic Memory: A Theory of Learning in Computers and People*. Cambridge University Press, 1982.
- [101] Yannis Siskos, Evangelos Grigoroudis, and Nikolaos F. Matsatsinis. UTA methods. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 297–344. Springer, Boston, 2005.
- [102] Dominik Ślęzak. Rough Sets and Bayes Factor. In James F. Peters and Andrzej Skowron, editors, *Transactions on Rough Sets III*, volume 3400 of *Lecture Notes in Computer Science*, pages 202–229. Springer, Berlin, 2005.
- [103] Paul Slovic. Choice between equally-valued alternatives. *Journal of Experimental Psychology: Human Perception Performance*, 1(3):280–287, 1975.
- [104] Roman Słowiński, Salvatore Greco, and Benedetto Matarazzo. Mining decision-rule preference model from rough approximation of preference relation. In *Proc. 26th IEEE Annual Int. Conference on Computer Software and Applications (COMPSAC 2002)*, pages 1129–1134, Los Alamitos, CA, 2002. IEEE Computer Society Press.
- [105] Roman Słowiński, Salvatore Greco, and Benedetto Matarazzo. Rough sets in decision making. In R.A. Meyers, editor, *Encyclopedia of Complexity and Systems Science*, pages 7753–7786. Springer, New York, 2009.
- [106] Roman Słowiński, Salvatore Greco, and Benedetto Matarazzo. Rough set based decision support. In Edmund K. Burke and Graham Kendall, editors, *Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques*, chapter 19, pages 557–609. Springer, New York, 2nd edition, 2014.
- [107] Jerzy Stefanowski. *Algorytmy Indukcji Reguł Decyzyjnych w Odkrywaniu Wiedzy*, volume 361 of *Rozprawy*. Wydawnictwo Politechniki Poznańskiej, 2001.
- [108] Marcin Szelaǵ, Salvatore Greco, Jerzy Błaszczyński, and Roman Słowiński. Case-based reasoning using dominance-based decision rules. In J.T. Yao et al., editors, *Rough Sets and Knowledge Technology 2011*, volume 6954 of *Lecture Notes in Computer Science*, pages 404–413, Heidelberg, 2011. Springer.
- [109] Christian Thiel, Stefan Scherer, and Friedhelm Schwenker. Fuzzy-input fuzzy-output one-against-all support vector machines. In B. Apolloni et al., editor, *KES 2007/ WIRN 2007, Part III*, volume 4694 of *Lecture Notes in Artificial Intelligence*, pages 156–165, Berlin, 2007. Springer.

- [110] Evangelos Triantaphyllou. *Multi-Criteria Decision Making: A Comparative Study*. Kluwer Academic Publishers, Dordrecht, Netherlands, 2000.
- [111] Tadeusz Trzaskalik. *Multiple criteria methods in the Polish financial market (in Polish)*. Polskie Wydawnictwa Ekonomiczne S.A., 2006.
- [112] Jan Verwaeren. *Mathematical optimization methods for the analysis of compositional data: subset selection, unmixing and prediction*. PhD thesis, Ghent University, 2014.
- [113] Willem Waegeman, Jan Verwaeren, Bram Slabbinck, and Bernard De Baets. Supervised learning algorithms for multi-class classification problems with partial class memberships. *Fuzzy S*, 184:106–125, 2011.
- [114] Margaret M. Wiecek, Matthias Ehrgott, Georges Fadel, and Jose Rui Figueira. Multiple criteria decision making for engineering. *Omega*, 36(3):337–339, 2008.
- [115] S. K. Michael Wong and Wojciech Ziarko. Comparison of the probabilistic approximate classification and the fuzzy set model. *Fuzzy Sets and Systems*, 21:357–362, 1987.
- [116] Yiyu Y. Yao. Probabilistic rough set approximations. *International Journal of Approximate Reasoning*, 49(2):255–271, 2008.
- [117] Lotfi Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965.