Studia i Materiały Informatyki Stosowanej, Tom 14, Nr 2, 2022, str. 25-32

ZASTOSOWANIE TEORII ZBIORÓW PRZYBLIŻONYCH DO ANALIZY PACJENTÓW

Olga Małolepsza

Uniwersytet Kazimierza Wielkiego
Instytut Informatyki
ul. Kopernika 1, 85-064 Bydgoszcz
e-mail: olgaaaem@student.ukw.edu.pl

Streszczenie: *Teoria zbiorów przybliżonych jest jedną z najczęściej używanych metod do analizy danych w formie tabeli. Teoria zbiorów przybliżonych znalazła zastosowanie w analizie, przetwarzaniu i redukcji zbiorów danych. Jest wykorzystywana w programach medycznych, farmakologicznych, inżynierskich, bankowych, w zagadnieniach sterowania, rozpoznawania obiektów oraz w wielu innych dziedzinach. Niniejszy artykuł przedstawia zastosowanie teorii zbiorów przybliżonych do analizy danych dotyczących pacjentów, których celem była ocena, czy pacjent ma grypę, czy nie ma.*

Słowa kluczowe: *Zbiory przybliżone, zbiór danych, RSES, grypa, tablica decyzyjna*

APPLICATION OF ROUGH SET THEORY TO PATIENT ANALYSIS

Abstarct: *Rough set theory is one of the most widely used methods for analyzing data in table form. Rough set theory has found applications in the analysis, processing and reduction of data sets. It is used in medical, pharmacological, engineering, banking programs, control issues, object recognition and many other fields. This article presents an application of rough sets theory to the analysis of patient data, the purpose of which was to access whether a patient has influenza or not.*

Keywords: *Rough Sets, dataset, RSES, influenza, decision-making board*

1. WSTĘP

Teoria zbiorów przybliżonych jest matematyczną metodą radzenia sobie z pojęciami rozmytymi. Stworzona została przez Polaka, prof. Zdzisława Pawlaka. W praktyce teoria ta oferuje nowe podejście do analizy danych. W teorii zbiorów przybliżonych zbiór jest definiowany przez jego elementy i nie jest wymagana żadna dodatkowa wiedza o elementach uniwersum tworzących zbiór. Z drugiej strony, teoria zbiorów przybliżonych zakłada, że mamy pewne dane o elementach uniwersum, które wykorzystujemy do tworzenia zbiorów. Elementy, które mają tę samą informację są podobne i tworzą tak zwane zbiory elementarne. Są one podstawą rozumowania w teorii zbiorów przybliżonych.

Zastosowany zbiór danych ma na celu określenie czy pacjent ma grypę na podstawie kilku atrybutów. Wynikiem analizy danych za pomocą teorii zbiorów przybliżonych jest zbiór reguł decyzyjnych.

2. METODY

Teoria zbiorów przybliżonych zajmuje się klasyfikacją danych zgromadzonych w tabelach. Dane te mogą pochodzić z pomiarów, testów lub od ekspertów. Główny cel analizy danych wejściowych to określenie aproksymacji idei (przybliżenia), w celu umożliwienia dokładnej analizy pytań, określenia relacji i powiązań między nimi oraz uzyskanie narzędzia do klasyfikacji nowych przypadków.

W przeciwieństwie do innych metod wnioskowania, teoria zbiorów przybliżonych pozwala na tzw. nieprecyzyjne dane, sprzeczne dane oraz niekompletne dane.

2.1. System informacyjny

System informacyjny jest to zorganizowany zestaw danych, w postaci tabeli, w której wiersze reprezentują indywidualne obiekty, np. pomierzone w eksperymencie, obserwowane, historyczne itd., a atrybuty obiektów zapisane są w osobnych kolumnach tej tabeli.
U – zbiór obiektów (od słowa „uniwersum"), np. U = {x_1, x_2, x_3, x_4}, gdzie x_i to obiekt i-ty,
A – zbiór odwzorowań, pomiarów, oznaczanych jako a, z obiektu na wartość jego cechy, tj. a: U ->V_a dla każdego a $\in$A. Obiektowi x_n z U w wyniku pomiaru a_n, przypisywana jest wartość mierzonej cechy V_a, dla całej puli A metod pomiaru a_n.
System informacyjny oznaczamy jako SI = (U, A).
Przykładem takiego systemu może być medyczny system informacyjny (tab. 1), w którym obiektami są pacjenci z urazem kręgosłupa.

	Age	*LEMS*
x_1	16-30	50
x_2	16-30	0
x_3	31-45	1-25
x_4	31-45	1-25
x_5	46-60	26-49
x_6	16-30	26-49
x_7	46-60	26-49

Tabela 1. Przykład systemu informacyjnego

2.2. System decyzyjny

W wyniku obserwacji, dedukcji, badań lub doświadczenia eksperta, można określić wynikową decyzję, dla każdego z obiektów, w skutek czego powstaje **system decyzyjny** określony wzorem SD = (U, A, {d}), gdzied to zbiór możliwych decyzji, adekwatnych dla obiektów danego uniwersum U. System decyzyjny dla powyższego przykładu przyjmuje postać (tab. 2). Porównując tabelę 1 z tabelą 2, w systemie decyzyjnym dodano trzecią kolumnę, która zawiera informacje o zdolności pacjenta do chodzenia.

	Age	*LEMS*	*Walk*
x_1	16-30	50	Yes
x_2	16-30	0	No
x_3	31-45	1-25	No
x_4	31-45	1-25	Yes
x_5	46-60	26-49	No
x_6	16-30	26-49	Yes
x_7	46-60	26-49	No

Tabela 2. Przykład systemu decyzyjnego

2.3. Reguły decyzyjne

Nawiązując do tabel 1 oraz 2 można z systemu decyzyjnego odczytać zbiór siedmiu reguł logicznych. Przykład dla pacjenta x_7:

JEŻELI Age="46-60" I LEMS="26-49" TO Walk="No"

Reguły takie wykorzystane mogą być do klasyfikacji przypadków nowych, dla których nie jest znana wartość decyzji i celem jest jej prognoza. Algorytm wnioskujący musi wówczas przeanalizować poprzednikireguł i znaleźć regułę o wartościach atrybutów odpowiadających nowemu przypadkowi i odczytać następnik reguły i zwrócić go jako wynik.

2.4. Tożsamość obiektów

W tabeli 3 zaznaczeni zostali pacjenci x_3 oraz x_4. Para ta jest nierozróżnialna (tożsama).

	Age	*LEMS*	*Walk*
x_1	16-30	50	Yes
x_2	16-30	0	No
x_3	31-45	1-25	No
x_4	31-45	1-25	Yes
x_5	46-60	26-49	No
x_6	16-30	26-49	Yes
x_7	46-60	26-49	No

Tabela 3. Przykład systemu decyzyjnego wraz z tożsamością

Dodanie nowego atrybutu, np. płci, wagi lub historii choroby pacjenta, może spowodować, że obiekty te staną

się rozróżnialne – tożsamość określa się względem wybranego podzbioru atrybutów B ⊆ A.

2.5. Aproksymacja zbioru

Relacja równoważności prowadzi do podziału uniwersum na tzw. zbiory elementarne. Każda kombinacja zbiorów elementarnych to zbiór definiowalny. Rodzina zbiorów oznaczana jest jako Def(B). Podziały uniwersum na zbiory elementarne za pomocą relacji równoważności służą do tworzenia podzbiorów uniwersum, które wykorzystuje się w zadaniachklasyfikacji i wnioskowania. Zwykle poszukiwane są podzbiory definiowalne charakteryzujące się takąsamą wartością atrybutu decyzyjnego.

	Age	*LEMS*	*Walk*
x_1	16-30	50	Yes
x_2	16-30	0	No
x_3	31-45	1-25	No
x_4	31-45	1-25	Yes
x_5	46-60	26-49	No
x_6	16-30	26-49	Yes
x_7	46-60	26-49	No

Tabela 4. Obiekty tożsame o sprzecznych decyzjach w systemie decyzyjnym

W tabeli 4 oznaczono obiekty, które są tożsame, czyli są w tym samym zbiorze elementarnym, jednakże posiadają różne atrybuty decyzyjne. Korzystając ze zbiorów elementarnych nie da się w tym przypadku stworzyć zbioru definiowalnego pacjentów o decyzji jednoznacznej Yes lub decyzji No. Poszukiwany zbiór pacjentów jest niedefiniowalny. Należy posłużyć się jego aproksymacją.

2.6. Redukty

Różne podzbiory atrybutów $B_1 \subseteq A$ i $B_2 \subseteq A$ mogą prowadzić do identycznych podziałów uniwersum, czyli takich, że $IND_{B1}(X)=IND_{B2}(X)$. Oznacza to, że zbiory elementarne uzyskane z tych podziałów są identyczne, a w konsekwencji także i aproksymacja jest równie precyzyjna.
Redukt B to taki podzbiór atrybutów, który ma minimalną ilość atrybutów, a ponadto $IND_B(X)=IND_A(X)$.
Wyznaczenie reduktu polega na pozostawieniu w podzbiorze tylko tych atrybutów, które zachowują relację równoważności, czyli nie zmieniają aproksymacji zbioru i na usunięciu pozostałych.

Zwykle dla danego systemu decyzyjnego istnieć może wiele reduktów. Czasami ze względów praktycznych nie korzysta się z reduktu najkrótszego (z najmniejszą liczbą atrybutów), tylko z takiego, którego implementacja jest np. najprostsza, a sposób pomiaru wartości atrybutu najmniej kosztowny, najszybszy itd.

2.7. Wykorzystanie reguł decyzyjnych w klasyfikacji

Po wyznaczeniu reduktów możliwe jest wygenerowanie z danych dostępnych w systemie decyzyjnym reguł logicznych o postaci JEŻELI … I … TO …. Przykładowa uproszczona funkcja logiczna dla zagadnienia rekrutacji do pracy posiada dwa redukty: f_A (d,e,f,r)=(d∧ e) ∨ (e∧ r).
Dla każdego z nich odczytujemy dane z systemu decyzyjnego występujące tam wartości atrybutów oraz odpowiadające im decyzje, dzięki czemu tworzy się zestaw reguł, np. JEŻELI Diploma=MBA I Experience=Medium TO Decision=Accept (tab. 5).

	Diploma	*Experience*	*French*	*Reference*	*Decision*
x_1	MBA	Medium	Yes	Excellent	Accept
x_4	MSc	High	Yes	Neutral	Accept
x_6	MSc	High	Yes	Excellent	Accept
x_7	MBA	High	No	Good	Accept
x_2	MBA	Low	Yes	Neutral	Reject
x_3	MCE	Low	Yes	Good	Reject
x_5	MSc	Medium	Yes	Neutral	Reject
x_8	MCE	Low	No	Excellent	Reject

Tabela 5. Macierz rozróżnialności dla decyzji

2.8. Dyskretyzacja parametrów

Obiekty w systemie decyzyjnym mogą w ogólności być opisane atrybutami liczbowymi o dowolnej dokładności z ciągłej dziedziny.
Na rysunku 1 można zauważyć, że u_3 i u_5(podobnie jak u_4 i u_7) stały się tożsame po dyskretyzacji, jednak mają tę samą decyzję d. W wyniku dyskretyzacji utracono możliwość rozróżniania obiektów między sobą, jednak bez obniżenia miary jakości aproksymacji zbiorów.

A	a	b	d
u_1	0.8	2	1
u_2	1	0.5	0
u_3	1.3	3	0
u_4	1.4	1	1
u_5	1.4	2	0
u_6	1.6	3	1
u_7	1.3	1	1

$\mathbf{A^P}$	a^P	b^P	d
u_1	0	2	1
u_2	1	0	0
u_3	1	2	0
u_4	1	1	1
u_5	1	2	0
u_6	2	2	1
u_7	1	1	1

Rysunek 1. System decyzyjny przed i po dyskretyzacji

2.9. Rough Set Exploration System

Rough Set Exploration System (RSES) jest narzędziem komputerowym umożliwiającym analizę danych w postaci tablicowej, z zastosowaniem teorii zbiorów przybliżonych, który został stworzony przez zespół badawczy z Uniwersytetu Warszawskiego. Udostępnia wszystkie najważniejsze operacje dotyczące teorii zbiorów przybliżonych i jej zastosowań. Praca z programem polega na budowaniu grafu przetwarzania danych.

Do analizy zbioru danych posłużył zbiór „grypa", który zawiera dane dotyczące pacjentów. Atrybutami warunkowymi są:

- bol_glowy – tak, nie
- bol_miesni - tak, nie
- bol_gardla - tak, nie
- kaszel - tak, nie
- temperatura – normalna, wysoka, bardzo wysoka

Atrybutem decyzyjnym jest:

- grypa - tak, nie.

Uniwersum złożone jest z 22 obiektów, które zostały przedstawione w modelu regułowym odnośnie grypy (rys. 2).

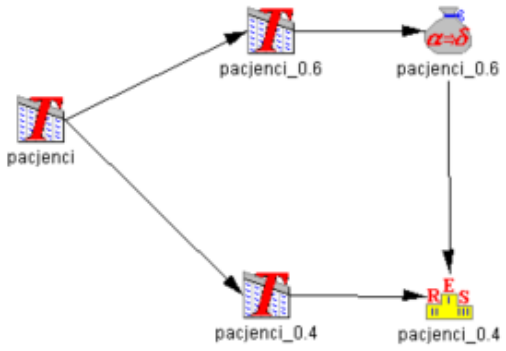
Table: pacjenci

22 / 7	bol_glowy	bol_miesni	bol_gardla	kaszel	katar	tempe...	grypa
O:1	'tak'	'tak'	'tak'	'tak'	'tak'	"bardzo wy...	'tak'
O:2	'tak'	'tak'	'tak'	'tak'	'tak'	'wysoka'	'tak'
O:3	'nie'	'nie'	'nie'	'nie'	'nie'	'normalna'	'nie'
O:4	'nie'	'nie'	'nie'	'nie'	'nie'	'wysoka'	'nie'
O:5	'nie'	'tak'	'nie'	'nie'	'nie'	"bardzo wy...	'tak'
O:6	'nie'	'tak'	'nie'	'nie'	'nie'	'wysoka'	'tak'
O:7	'tak'	'nie'	'nie'	'nie'	'nie'	"bardzo wy...	'tak'
O:8	'tak'	'nie'	'nie'	'nie'	'nie'	'wysoka'	'nie'
O:9	'nie'	'tak'	'nie'	'nie'	'nie'	'normalna'	'nie'
O:10	'nie'	'nie'	'tak'	'nie'	'nie'	'normalna'	'nie'
O:11	'nie'	'nie'	'nie'	'tak'	'nie'	'normalna'	'nie'
O:12	'nie'	'nie'	'nie'	'nie'	'tak'	'normalna'	'nie'
O:13	'tak'	'tak'	'nie'	'nie'	'nie'	'normalna'	'nie'
O:14	'tak'	'tak'	'nie'	'nie'	'nie'	'wysoka'	'tak'
O:15	'tak'	'tak'	'nie'	'nie'	'nie'	"bardzo wy...	'tak'
O:16	'nie'	'nie'	'tak'	'tak'	'nie'	'normalna'	'nie'
O:17	'tak'	'nie'	'nie'	'nie'	'nie'	'wysoka'	'tak'
O:18	'nie'	'tak'	'nie'	'nie'	'nie'	'wysoka'	'nie'
O:19	'nie'	'nie'	'tak'	'nie'	'nie'	'wysoka'	'nie'
O:20	'nie'	'nie'	'nie'	'tak'	'nie'	'wysoka'	'nie'
O:21	'nie'	'nie'	'nie'	'tak'	'tak'	'wysoka'	'nie'
O:22	'nie'	'nie'	'nie'	'nie'	'tak'	'wysoka'	'nie'

Rysunek 2. Model regułowy - grypa

3. REZULTATY

Zbiór danych podzielono na dwie nowe tablice:pacjenci_0.6 oraz pacjenci_0.4. Pierwsza tablica zostanie użyta do nauki, a druga do testowania. Tabele oznaczone są literą T. Kolejno liczymy reguły dla tablicy pacjenci_0.6, wskutek czego powstaje nowy obiekt (ikona worka) – zbiór reguł o nazwie takiej jak tabelka. Następnie testujemy wyliczone reguły na tablicypacjenci_0.4i powstaje nowy obiekt z wynikami, który oznaczony jest ikoną podium. Widoczne jest to na grafie (rys. 3).



Rysunek 3. Graf przetwarzania danych

Przykładowe reguły są widoczne na rysunku 4.

Rule set: pacjenci_0.6

(1-26)	Match	Decision rules
1	6	(bol_glowy='nie')&(bol_miesni='nie')=>(grypa={'nie'[6]})
2	4	(bol_miesni='nie')&(temperatura='wysoka')=>(grypa={'nie'[
3	3	(bol_glowy='tak')=>(grypa={'tak'[3]})
4	3	(temperatura="bardzo wysoka")=>(grypa={'tak'[3]})
5	3	(temperatura='normalna')=>(grypa={'nie'[3]})
6	3	(bol_glowy='nie')&(kaszel='tak')=>(grypa={'nie'[3]})
7	3	(bol_miesni='nie')&(kaszel='tak')=>(grypa={'nie'[3]})
8	2	(bol_glowy='nie')&(bol_miesni='tak')&(temperatura='wysok
9	2	(bol_miesni='tak')&(bol_gardla='nie')&(temperatura='wyso
10	2	(bol_miesni='tak')&(kaszel='nie')&(temperatura='wysoka')
11	2	(bol_miesni='tak')&(katar='nie')&(temperatura='wysoka')=
12	2	(bol_glowy='nie')&(katar='tak')=>(grypa={'nie'[2]})
13	2	(bol_miesni='nie')&(katar='tak')=>(grypa={'nie'[2]})
14	2	(bol_gardla='nie')&(katar='tak')=>(grypa={'nie'[2]})
15	2	(bol_glowy='nie')&(bol_gardla='tak')=>(grypa={'nie'[2]})
16	2	(bol_miesni='nie')&(bol_gardla='tak')=>(grypa={'nie'[2]})
17	2	(bol_gardla='tak')&(katar='nie')=>(grypa={'nie'[2]})
18	2	(kaszel='tak')&(katar='nie')=>(grypa={'nie'[2]})
19	2	(bol_gardla='nie')&(kaszel='tak')=>(grypa={'nie'[2]})
20	1	(bol_miesni='tak')&(bol_gardla='tak')=>(grypa={'tak'[1]})
21	1	(bol_miesni='tak')&(kaszel='tak')=>(grypa={'tak'[1]})
22	1	(bol_miesni='tak')&(katar='tak')=>(grypa={'tak'[1]})
23	1	(bol_gardla='tak')&(katar='tak')=>(grypa={'tak'[1]})
24	1	(bol_gardla='tak')&(kaszel='tak')&(temperatura='wysoka')
25	1	(kaszel='nie')&(katar='tak')=>(grypa={'nie'[1]})
26	1	(bol_gardla='tak')&(kaszel='nie')=>(grypa={'nie'[1]})

Rysunek 4. Treść przykładowych reguł

Match oznacza, ile obiektów ze zbioru treningowego (pacjenci_0.6) pasuje do danej reguły.

4. ANALIZA ZBIORU ZA POMOCĄ PROGRAMU RSES

Pierwszą metodą jest train&test, która pozwala sprawdzić działanie danego klasyfikatora. Wyniki klasyfikacji (rys. 5) zostały przedstawione w postaci tzw. macierzy pomyłek. Actual to wartości właściwe i oczekiwane, które były dostępne w tabeli testowej, Predicted to odpowiedź klasyfikatora. Dokładność dla opcji „tak" wynosi 1, a dla opcji „nie" 0,667.

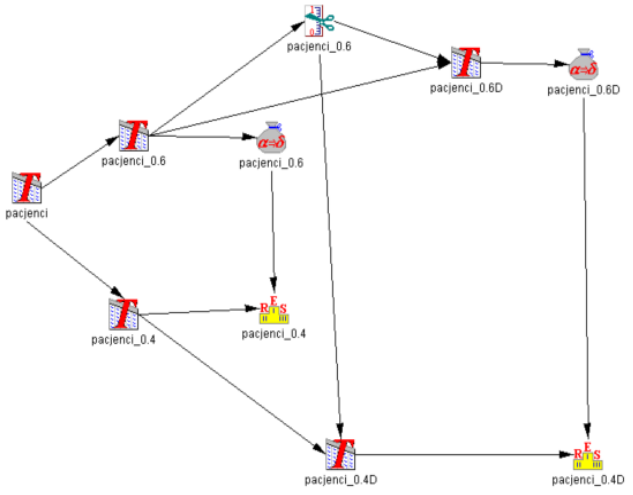
Results of experiments by train&test method: pacjenci_0.4

		Predicted				
		'tak'	'nie'	No. of obj.	Accuracy	Coverage
Actual	'tak'	3	0	3	1	1
	'nie'	2	4	6	0.667	1
	True positive rate	0.6	1			

Total number of tested objects: 9
Total accuracy: 0.778
Total coverage: 1

Rysunek 5. Wyniki klasyfikacji train&test

Następnie utworzono zbiór cięć (oznaczony ikoną z nożyczkami) o nazwie pacjenci_0.6. Kolejnym krokiem jest dyskretyzacja tablicy pacjenci_0.6 oraz pacjenci_0.4, wskutek czego powstają odpowiadające tablice pacjenci_0.6D i pacjenci_0.4D. I tak jak poprzednio wyliczamy reguły oraz je testujemy (rys. 6).



Rysunek 6. Wygląd grafu po dyskretyzacji.

Tabela z danymi po dyskretyzacji zostałaprzedstawiona na rysunku 7.

Table: pacjenci_0.6D

13 / 7	bol_gl...	bol_m...	bol_g...	kaszel	katar	tempe...	grypa
O:1	'tak'	'tak'	'tak'	'tak'	'tak'	'wysoka'	'tak'
O:2	'nie'	'nie'	'nie'	'nie'	'nie'	'wysoka'	'nie'
O:3	'nie'	'tak'	'nie'	'nie'	'nie'	"bardzo wy...	'tak'
O:4	'nie'	'tak'	'nie'	'nie'	'nie'	'wysoka'	'tak'
O:5	'tak'	'nie'	'nie'	'nie'	'nie'	"bardzo wy...	'tak'
O:6	'nie'	'tak'	'nie'	'nie'	'nie'	'normalna'	'nie'
O:7	'nie'	'nie'	'nie'	'nie'	'tak'	'normalna'	'nie'
O:8	'tak'	'tak'	'nie'	'nie'	'nie'	"bardzo wy...	'tak'
O:9	'nie'	'nie'	'tak'	'tak'	'nie'	'normalna'	'nie'
O:10	'nie'	'tak'	'nie'	'nie'	'nie'	'wysoka'	'nie'
O:11	'nie'	'nie'	'tak'	'nie'	'nie'	'wysoka'	'nie'
O:12	'nie'	'nie'	'nie'	'tak'	'nie'	'wysoka'	'nie'
O:13	'nie'	'nie'	'nie'	'tak'	'tak'	'wysoka'	'nie'

Rysunek 7. Tabela z danymi po dyskretyzacji

Natomiast wyniki reguł decyzyjnych po skalowaniu przedstawiono na rysunku 8. Są one takie same jak podczas klasyfikacji przedstawionej na rysunku 5.

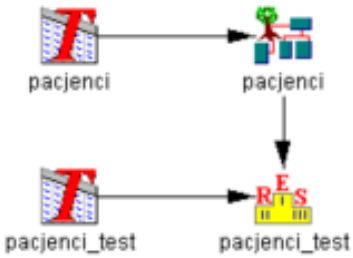
Results of experiments by train&test method: pacjenci_0.4D

		Predicted				
		'tak'	'nie'	No. of obj.	Accuracy	Coverage
Actual	'tak'	3	0	3	1	1
	'nie'	2	4	6	0.667	1
	True positive rate	0.6	1			

Total number of tested objects: 9
Total accuracy: 0.778
Total coverage: 1

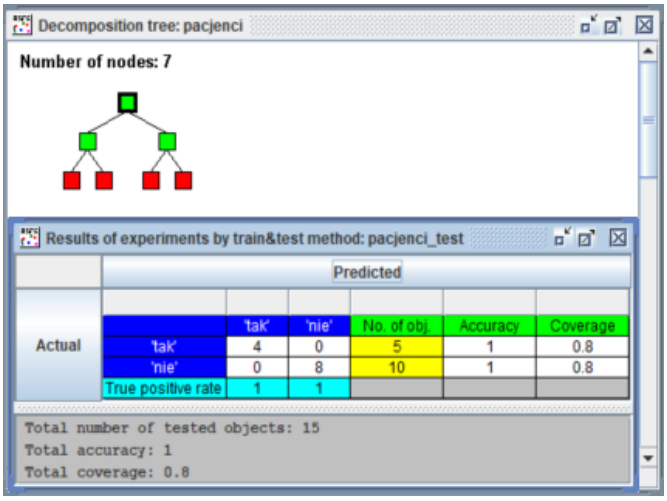
Rysunek 8. Reguły decyzyjne po skalowaniu

W przypadku większych zbiorów danych (powyżej 1000 obiektów),stosowanie tradycyjnych klasyfikatorów regułowych jest utrudnione, gdyż złożoność obliczeniowa algorytmów liczących reguły wymaga czasochłonnych obliczeń. Związku z powyższym w systemie RSES zaimplementowanometodę dekompozycji danych, która potrafi tworzyć klasyfikatory nawet dladanych liczących setki tysięcy obiektów. I choć w tym przypadku nie jest potrzebna ta metoda, postanowiono ją przetestować.
Utworzono drzewo dekompozycji dla tablicy pacjenci o tej samej nazwie z ikoną drzewa (rys. 9), a następnie przetestowano drzewo dekompozycji na tablicypacjenci_test.



Rysunek 9. Klasyfikator oparty na drzewie dekompozycji

Wyniki testu drzewa dekompozycji przedstawiono na rysunku 10, gdzie widoczne są wyniki klasyfikacji wraz z drzewem dekompozycji. Pomimo stosowania tej klasyfikacji przy dużych bazach danych, widać, że klasyfikator poradził sobie lepiej niż we wcześniejszych przypadkach, ponieważ dokładność jest rzędu 1 w obu przypadkach. Zmniejszyło się jednak pokrycie – wynosiło 1 w obu przypadkach, a teraz 0,8.

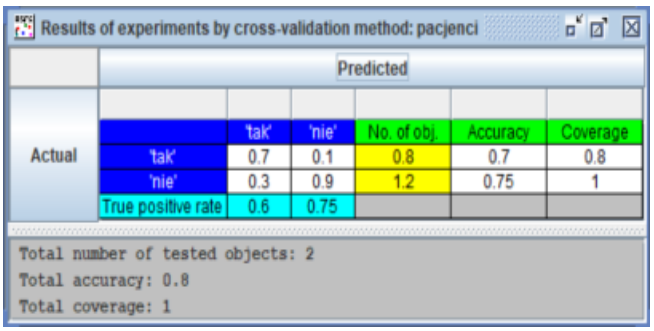


Rysunek 10. Wyniki klasyfikacji wraz z drzewem dekompozycji

Jako, że dla zbiorów danych liczących poniżej 1000 obiektów, najczęściej stosowanąmetodą testowania klasyfikatorów jest metodawalidacji krzyżowej (ang. cross-validation), na koniec została użyta (rys. 11). Metoda ta polega na tym, że wejściowy zbiór danych jest dzielony na kilka równych i rozłącznych części. Podczas pojedynczego testu jednąz wyznaczonych części używamy do testowania klasyfikatora, pozostałe zaśdo jego budowy. Wykonujemy tyle testów na ile części podzieliliśmy zbiórdanych. Końcowy wynik klasyfikacji jest średnią arytmetyczną wyników zewszystkich wykonanych testów.

Rysunek 11. Walidacja krzyżowa

Wyniki klasyfikacji przedstawiono na rysunku 12. Dokładność opcji „tak” wynosi 0.7, natomiat opcji „nie” 0.75, a pokrycie odpowiednio 0.8 i 1. W nawiązaniu do poprzednich metod, ta wypadła najgorzej.



Rysunek 12. Wyniki klasyfikacji walidacji krzyżowej

5. WNIOSKI

W niniejszym artykule przedstawiono wykorzystanie teorii zbiorów przybliżonych w oparciu o zbiór danych „grypa”, zawierający dane dotyczące pacjentów.
Przedstawiony zbiór w przypadku metody zbioru testowego oraz operacji dyskretyzacji i utworzenia zbioru reduktów okazał się być poprawnie sklasyfikowany. Najlepszy wynik uzyskano podczas klasyfikacji wraz z drzewem dekompozycji, a najgorszy podczasklasyfikacji walidacji krzyżowej, choć nadal nie wypadło aż tak źle. Podsumowując wyniki są one zadowalające.
Program sam w sobie okazał się potężnym narzędziem, dzięki któremu możemy ocenić wiele innych zbiorów danych.

Literatura

1. Pawlak Z., „Zbiory przybliżone. Nowa matematyczna metoda analizy danych”, Instytut Informatyki Teoretycznej i Stosowanej PAN

2. J. Komorowski, Z. Pawlak, L. Polkowski, A. Skowron, „Rough Sets: A Tutorial" In: S. K. Paland, A. Skowron (editors) Rough Fuzzy Hybridization, A New Trend in and Decision Making, pp 3-98. Springer-Verlag, Singapore, 1999.
3. https://multimed.org/student/sim/instrukcja_projekt.pdf - data dostępu 10.03.2022r.
4. RSES 2.1. Rough Set Exploration System PODRĘCZNIK UŻYTKOWNIKA, 2004.