

Zastosowanie teorii zbiorów przybliżonych do identyfikacji zależności pomiędzy wizerunkiem Szczecina a planami studentów w zakresie miejsca realizacji kariery zawodowej

Karol Kuczera¹

1. Wstęp

Rozwój technologii komunikacyjnych i informacyjnych (ICT), w tym technik obliczeniowych, pozwala na operowanie coraz większymi zbiorami danych oraz poszukiwanie w nich coraz bardziej złożonych zależności. Wydaje się jednak, iż ważniejsza od zmian ilościowych jest ewolucja jakościowa. Początkowo technologia ta stawiała wysokie wymagania odnośnie formalizacji i strukturalizacji danych. Pociągało to za sobą wiele problemów i ograniczeń oraz skutkowało często zniekształcaniem prawdziwego (obserwowanego) obrazu jakościowej rzeczywistości kwantyfikowanego i dostosowywanego do wymogów technologii obliczeniowej.

Jednak szybki rozwój ICT i duże inwestycje w rozwój infrastruktury zlikwidowały większość barier i przeszkód w tym względzie [1]. Obecnie obserwuje się złagodzenie wymogów stawianych zasileniom systemów informatycznych, implementacja coraz bardziej wyrafinowanych metod obliczeniowych pozwala operować danymi o różnych postaciach, a także zwiększa tolerancję dla zasileń niekompletnych, niespójnych czy wręcz sprzecznych. Pomimo, że komputery są *per se* maszynami matematycznymi „wyposażane” są w funkcje pozwalające na odchylenie się od bezwzględnej ścisłości matematycznej.

Tendencja ta zbieżna jest z sugestiami pojawiającymi się wraz z rozwojem metod wspomagania decyzji. Jak podkreśla T. Saaty podejmowanie decyzji jest najczęstszą czynnością, którą realizują ludzie, wszyscy i przez cały czas. Zagadnienia decyzyjne są zaś w większości opisowe i predykcyjne co pomaga w inteligentnym radzeniu sobie ze złożonym światem. Decyzje powinny być

¹ Uniwersytet Szczeciński, Wydział Nauk Ekonomicznych i Zarządzania, karol@kuczera.szczecin.pl

podejmowane zgodnie z własnym rozumieniem, dając poczucie wyboru całkowicie uzasadnionego indywidualnymi bądź grupowymi wartościami, wierzeniami czy przekonaniem. Prowadzi to do sugestii, że techniki decyzyjne powinny być możliwie zbliżone do aspektów biologicznych i psychologicznych ludzi. Metody wspierające podejmowanie decyzji powinny być bardziej opisowe i przejrzyste, zdolne do uchwycenia norm oceny przechowywanych w pamięci decydenta i do opisu decyzji w sposób normatywny (dobre, złe) [15]. Pogląd ten nie wydaje się odosobniony [zob. np.: 12]. Szczególnie często kryteria i wartości jakościowe stanowiące przejaw wiedzy, doświadczenia, opinii a także emocji decydentów ujawniają się w obszarze dylematów ekonomicznych oraz w zarządzaniu, wyraźnie podatnych ostatnio na absorpcję zależności socjologicznych i psychologicznych [4]. Podejście takie zaobserwować można m.in. w pracy [7], gdzie w (opartym na metodzie AHP) procesie pozyskiwania informacji ze sformalizowanych dokumentów strategii przedsiębiorstwa ostatecznie najistotniejsza okazuje się wiedza i doświadczenie eksperta strategii wykorzystywana bezpośrednio w procesie podejmowania decyzji.

Znakomitym przykładem zasygnalizowanych mechanizmów wydaje się być teoria zbiorów przybliżonych Pawlaka ogłoszona w roku 1982². Ze względu na stosunkowo duże wymagania kalkulacyjne początkowo funkcjonowała raczej w warstwie teoretycznej, a jej popularność zaczęła rosnąć wraz rozwojem i upowszechnianiem technologii komputerowych otwierających szerokie możliwości implementacji i praktycznego wykorzystania metod na niej opartych. Obecnie uważana jest powszechnie za jedną z najszybciej rozwijających się metod sztucznej inteligencji. Synergia rosnących mocy obliczeniowych oferowanych przez ICT oraz „tolerancyjność” zbiorów przybliżonych w stosunku do zasileń informacyjnych opisujących analizowaną rzeczywistość w warunkach rosnącej złożoności świata oraz konieczności operowania lawinowo powiększającymi się zbiorami danych empirycznych, z natury nieprecyzyjnymi i niepełnymi przenosi jedno z kluczowych zagadnień zarządzania – podejmowanie decyzji – na zupełnie nowy poziom jakościowy.

Celem niniejszego opracowania jest próba odkrycia wiedzy ukazującej powiązania pomiędzy elementami wizerunku Szczecina w oczach studentów, a deklaracjami o pozostaniu bądź opuszczeniu miasta po zakończeniu studiów, inaczej mówiąc, jaki wizerunek miasta wpływać może na wiązanie swojej przyszłości przez absolwentów ze Szczecinem. Drugim przyjętym celem jest wskazanie możliwości wykorzystania metody zbiorów przybliżonych do indukcji reguł decyzyjnych w podjętym obszarze. Zakładane wnioski z jednej strony przyczynić mogą się do weryfikacji i popularyzacji wybranej metody i jej implementacji w środowisku technologii obliczeniowej, z drugiej, utylitarnej, stanowić mogą sugestie dla władarzy odnośnie potencjalnych kierunków działań

² Przyjmuje się, iż została zaprezentowana po raz pierwszy w publikacji [8] choć już wcześniejsze publikacje odnosiły się do tego zagadnienia, np. [6]; por. [17]

związanych z zachęcaniem absolwentów do związania swojej kariery zawodowej ze Szczecinem i osłabienia niekorzystnej tendencji utraty wykształconych kadr.

2. Badania

Niniejszy artykuł jest kontynuacją rozważań podjętych we wcześniejszej publikacji autora [5], której celem była prezentacja i interpretacja wyników badań dotyczących wizerunku Szczecina w opinii studentów. Tło prowadzonej dyskusji stanowiło m.in. zagadnienie kształtowania marki miejsc i roli jaką zajmuje w zarządzaniu miastami, które w coraz większym stopniu odpowiada regułom zarządzania firmami operującymi na rynku konkurencyjnym. Podkreślono, że narastająca konkurencja pomiędzy miastami zarówno o turystów jak i inwestorów, przedsiębiorców i wykształconą kadrę wymaga, aby miasta stawały się atrakcyjnym „produktem”, a zarządzanie nimi oparte było o strategię kształtowania i rozwoju wraz ze zintegrowaną strategią komunikacyjną oraz budową marki miasta. Pojęcie marki rozumiane jest dualnie, w kategoriach tożsamości, czyli sposobu w jaki zarządzający marką chcieliby, aby była postrzegana oraz wizerunku – jej faktycznego odbioru czy też stosunku, często emocjonalnego, między odbiorcą a marką. Wizerunek miasta staje się dynamicznym konglomeratem skojarzeń na jego temat powstającym w umysłach odbiorców działań marketingowych, także pod wpływem czynników pośrednich, jak np. opinii innych [3]. Szczególną uwagę skierowano na wizerunek marki Szczecin. Prezentacja części pozyskanego materiału empirycznego oparta została o dyferencjał semantyczny oraz statystykę opisową, w tym głównie pozycyjne miary rozkładu ocen. Do udziału w badaniach ankietowych zaproszono studentów kierunku zarządzanie wszystkich lat, zarówno pierwszego jak i drugiego stopnia, studiów stacjonarnych Wydziału Nauk Ekonomicznych i Zarządzania Uniwersytetu Szczecińskiego. O arbitralnym wyborze próby badawczej zdecydował fakt, iż studenci zarządzania stanowią potencjalne lokalne kadry zarządzające, a ich decyzje o pozostaniu w mieście bądź realizacji kariery poza Szczecinem wpływać będą na przyszły rozwój miasta, jego siłę i pozycję konkurencyjną, a tym samym i wizerunek. Na wybór próby wpływ miało także miejsce pracy autora. W badaniu pozyskano 287 kwestionariuszy ankietowych (przy 100% zwrotności) na 293 osoby będące studentami kierunku zarządzanie, wykorzystano natomiast 253 kompletnych i bezbłędnie wypełnionych kwestionariuszy ankietowych.

3. Teoria zbiorów przybliżonych

W niniejszej publikacji podjęto próbę wykorzystania do analizy uzyskanych danych teorii zbiorów przybliżonych Pawlaka będącą uzupełnieniem teorii mnogości i stanowiącą odpowiedź na problem „niedoskonałej” wiedzy, który

znajdował się od długiego czasu w kręgu zainteresowania zarówno filozofów, jak i logików i matematyków, a w związku z rozwojem technik komputerowych, w tym obszarze sztucznej inteligencji, wzbudzał zainteresowanie również wśród informatyków. Z logicznego punktu widzenia stanowi ona nowe podejście matematyczne do pojęć nieostrych, z punktu widzenia praktycznego, jest nową metodą analizy danych pozwalającą między innymi poszukiwać zależności między danymi, redukować ich nadmiarowość, określać wagi poszczególnych składników danych oraz generować z danych reguły decyzyjne. Z. Pawlak analizę danych z wykorzystaniem swojej metody traktuje jako szczególny przypadek wnioskowań indukcyjnych. Wychodząc od pewnych, częściowych faktów dotyczących badanej rzeczywistości, w drodze uogólnień, dąży się tu do tworzenia wiedzy o szerszym świecie niż ten, który stanowił punkt wyjścia wnioskowań. Działania takie zmierzają do wyciągania wniosków prawdopodobnych (możliwych), a weryfikacja hipotez odbywa się w drodze eksperymentów³ [11].

Językiem teorii zbiorów opisywane są „przybliżone operacje na zbiorach, jak również „dokładne” operacje na zbiorach przybliżonych. Jeżeli (...) Proponowane podejście opiera się na niedostrzegalnej relacji, która „skleja” razem wszystkie obiekty, których nie moglibyśmy rozróżnić za pomocą dostępnych środków obserwacji (pomiar) lub wyrazu (język).” [10] Odmienne niż ma to miejsce w przypadku (klasycznej) teorii mnogości, gdzie definiowanie zbioru odbywa się poprzez jego elementy bez uwzględniania dodatkowej wiedzy o elementach uniwersum, z których powstają zbiory, w przypadku zbiorów przybliżonych zakłada się, iż pewne dane o elementach uniwersum są znane oraz wykorzystywane podczas tworzenia zbiorów [11].

„Istota metody polega na wprowadzeniu dwóch funkcji przynależności $\underline{\epsilon}$ i $\bar{\epsilon}$, które oznaczają odpowiednio „z pewnością należy” i „możliwe, że należy” oraz pojęć górnego i dolnego przybliżenia zbioru A w strukturze relacyjnej $A = (U, R)$, gdzie R jest relacją równoważności nazywaną tutaj obszarem przybliżenia (ang. *approximation space*)” [9].

Stale rozszerzana i uzupełniana, teoria zbiorów przybliżonych znalazła szerokie zastosowanie w wielu dziedzinach o charakterze praktycznym, w tym w zarządzaniu, a w szczególności w podejmowaniu decyzji. Ze względu na ograniczenia redakcyjne pominięto szczegółowy opis wykorzystanej teorii, który znaleźć można w wielu innych publikacjach [zob. m.in.: 18; 19].

³ Drugi rodzaj wnioskowania stanowi dedukcja wyprowadzająca twierdzenia prawdziwe z innych twierdzeń prawdziwych, posiada dobrze ugruntowane podstawy teoretyczne, a hipotezy weryfikuje w drodze dowodów, co jednak wymaga odnotowania, nie znalazła szerszego zastosowania poza rozumowaniami matematycznymi.

4. Wyniki badań

Na potrzeby niniejszego opracowania wykorzystano program Rose2 (Rough Sets Data Explorer) opracowany i udostępniany przez Laboratorium Inteligentnych Systemów Wspomagania Decyzji Politechniki Poznańskiej⁴.

Jedno z pytań ankiety przyjęło postać dyferencjału semantycznego zawierającego 13 par antonimów dotyczących odbioru Szczecina, z którego zbudowano spolaryzowany profil ocen (zob. Kuczera 2014). Pary te to: (1) Przyjacielski-Chłodny, (2) Romantyczny-Przyziemny, (3) Bezpieczny-Niebezpieczny, (4) Czysty-Brudny, (5) Interesujący-Nudny, (6) Tętniący życiem-Śpiący, (7) Ładny-Brzydki, (8) Wyrafinowany-Kiczowaty, (9) Naturalny-Sztuczny, (10) Otwarty-Zamknięty, (11) Pełen harmonii-Pełen nieładu, (12) Nowoczesny-Zacofany, (13) Wielkomijski-Zaściankowy. Do dyspozycji ankietowanych była skala: 3, 2, 1, 0, 1, 2, 3, gdzie 0 wyrażało wartość neutralną, pozostałe zaś wartości pozwalały na wskazanie ukierunkowania i intensywności oceny w każdej parze. Ankieta zawierała także pytanie dotyczące planów po zakończeniu nauki, w którym jeden z wariantów odpowiedzi brzmiał: „planuję związać swoją karierę ze Szczecinem”, respondenci ustosunkowywali się do tego stwierdzenia wybierając jedno z możliwych wskazań: (1) całkowicie się zgadzam, (2) raczej się zgadzam, (3) ani się zgadzam, ani się nie zgadzam, (4) raczej się nie zgadzam, (5) całkowicie się nie zgadzam.

O ile w poprzednim opracowaniu zaprezentowano odpowiedzi na poszczególne pary antonimów w różnych ujęciach, tym razem podjęto próbę odkrycia zależności pomiędzy wskazanymi pytaniami. Próbę skonstruowania bazy w postaci reguł decyzyjnych ukazujących jakie oceny dotyczące wizerunku Szczecina związane są (wpływają na)z planami o pozostaniu w Szczecinie, wyjazdem z miasta bądź niewiadomą w tym obszarze.

Z posiadanych danych zbudowano tablicę decyzyjną⁵ zawierającą argumenty warunkowe – wartości dla 13 skal dyferencjału semantycznego oraz jako argument decyzyjny informację o planach po zakończeniu studiów. Analizy tej tablicy decyzyjnej nie prowadziły do satysfakcjonujących rezultatów co skłoniło autora do przekształcenia pierwotnej tablicy decyzyjnej na jej postać wtórną. Zmiana dotyczyła zakresu wartości cech zmiennych warunkowych. Skalę 7 punktową zamieniono na 3 punktową, gdzie oceny preferujące element znajdujący się po lewej stronie pary kodowane były pod cyfrą 1, neutralne 2, natomiast te, które preferowały element znajdujący się po prawej stronie otrzymywały wartość 3. Nie dokonano natomiast przekształceń zmiennej decyzyjnej pozostawiając zakres pięciu odpowiedzi kodowanych kolejnymi liczbami od 1 do 5 (rys. 1).

⁴ <http://idss.cs.put.poznan.pl/site/rose.html>; opis programu dostępny jest w [13; 14]

⁵ Tablica decyzyjna stanowi specjalny rodzaj tablicy informacyjnej, wyróżniono w niej argumenty warunkowe oraz decyzyjne

	1	2	3	4	5	6	7	8	9	10	11	12	13	konkluzja (D)
1	1	3	1	2	1	1	1	1	1	1	2	3	2	1
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	2	1	1	1	1	1	1	1	1	2	1	1	1
5	1	1	1	1	1	1	1	2	1	1	1	1	1	1
6	1	1	1	3	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	1	2	2	3	3	3	3	2	2	2	3	1	1	1
9	1	1	1	3	3	3	1	3	1	1	1	1	1	1
10	1	1	1	2	3	3	1	1	3	3	3	3	3	1
11	3	2	1	3	2	3	2	1	1	1	1	2	3	1
12	1	1	1	1	1	1	1	1	1	1	2	2	1	1

Rys. 1. Fragment wtórnej tablicy decyzyjnej

Źródło: opracowanie własne

253 przykłady zawarte w tablicy decyzyjnej pozwoliły na wyznaczenie 209 zbiorów elementarnych zawierających takie same wartości atrybutów warunkowych. Zebrany materiał cechował się jakością klasyfikacji na poziomie 0,7984, mówiącej o dość dużym stopniu zależności atrybutów decyzyjnych od atrybutów warunkowych. Oznacza, iż na podstawie prawie 80% przykładów z tablicy możliwe jest wyznaczenie reguł pewnych. Wskaźnik ten obliczany jest jako relacja liczebności przykładów zawartych w pozytywnym obszarze zbioru (w jego dolnym przybliżeniu) i liczebności w całym uniwersum przykładów.

Zwrócić należy uwagę, iż jakość klasyfikacji =1 wskazywałaby, iż zbiór atrybutów decyzyjnych zależy jednoznacznie od zbioru atrybutów warunkowych (zbiór dokładny) świadcząc o dobrym (deterministycznym, spójnym) określeniu tablicy decyzyjnej – wszystkie reguły decyzyjne otrzymane z przykładów zawartych w tablicy mają charakter deterministyczny. Deterministyczne reguły tablicy decyzyjnej to taki ich zbiór, w którym każda para reguł o identycznych wartościach wszystkich atrybutów warunkowych wskazuje na identyczne wartości atrybutów decyzyjnych.

Jakość klasyfikacji <1 świadczy o występowaniu reguł odchylających się od tej zależności, czyli reguł niedeterministycznych⁶. Oznacza to, że w oparciu o zbiór atrybutów warunkowych nie można wnioskować o jednoznacznej przynależności przykładów do poszczególnych klas abstrakcji – zbiorów decyzyjnych wyznaczonych na podstawie wartości atrybutu decyzyjnego. W tym przypadku tablica decyzyjna traktowana jest jako niespójna (źle określona) jednakże klasyfikacja w dalszym ciągu może zostać opisana – w sposób przybliżony. Nadmienić należy, że analiza zbioru reguł niedeterministycznych ma istotne znaczenie poznawcze i praktyczne, stanowić może podstawę do weryfikacji procesu akwizycji danych, co przełożyć się może na poprawę opisu badanych zjawisk i procesów.

⁶ W literaturze znaleźć można także określenia reguł pewnych i możliwych czy pewnych i sprzecznych, jak również reguł częściowo dyskryminujących, więcej na ten temat znaleźć można w [16]

Oprócz jakości klasyfikacji obliczono także dokładność przybliżenia poszczególnych klas abstrakcji jako relację dolnego i górnego przybliżenia przykładów należących do danej klasy (rys. 2). Najlepsza wartość przybliżenia ujawniła się dla klasy 5 – „całkowicie się nie zgadzam” (, iż planuję związać swoją karierę ze Szczecinem).

The screenshot shows a window titled 'Approximation Viewer -C:\Users\Karol Kuczera\...'. It displays the 'Quality of classification' as 0.7984. Below this is a table of approximations for five classes. The table has five columns: Class, # of Objects, Lower Approx., Upper Approx., and Accuracy. The data is as follows:

Class	# of Objects	Lower Approx...	Upper Approx...	Accuracy
1	28	17	54	0.3148
2	82	62	109	0.9688
3	61	50	99	0.9051
4	52	47	79	0.5949
5	30	26	35	0.7429

At the bottom of the window, it states 'Number of atoms: 209'.

Rys. 2 Parametry aproksymacji tabeli decyzyjnej

Źródło: opracowanie własne, zrzut ekranu programu Rose2

Teoria zbiorów przybliżonych pozwala na identyfikowanie redundancji danych w tablicach informacyjnych oraz jej eliminację – tworzenie reduktów. W stosunku do tablic decyzyjnych zastosowanie ma pojęcie reduktu względnego – reduktu ze względu na atrybuty decyzyjne. Redukt warunkuje zbiór decyzyjny w takim samym stopniu jak pełen zbiór argumentów warunkowych pomimo operowania mniejszą liczbą cech charakteryzujących obiekty należące do uniwersum przykładów. Zwrócić należy uwagę, iż w tablicy występować może więcej niż jeden redukt, zaś ich część wspólna nosi nazwę rdzenia bądź jądra i zawiera te atrybuty, których nie można usunąć bez utraty jakości przybliżenia klasyfikacji.

W badanym przykładzie otrzymano 2 redukty: {2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13}, {2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13}, oba o długości 11 atrybutów, tym samym nadmiarowość zgromadzonego materiału nie była znaczna i większość uwzględnionych w ankiecie atrybutów okazała się istotna. Rdzeń przyjął postać: {2, 3, 4, 6, 7, 8, 9, 11, 12, 13}.

Wykorzystując dostępne w systemie Rose2 algorytmy indukcji reguł decyzyjnych przeprowadzono eksperymenty, z których interesujące efekty ujawniły się przy zastosowaniu metody rozszerzonego minimalnego pokrycia (ang. *extended minimal covering*) oraz wykorzystaniu przy budowie reguł punktów odcięcia minimalizujących miarę Laplace. Metoda ta generuje reguły o rozszerzonej składni warunków w postaci atrybut \leq wartość lub atrybut \geq wartość, pozwala także na operowanie danymi ciągłymi, bez potrzeby ich wcześniejszej dyskretyzacji, co w analizowanym przypadku nie miało akurat znaczenia.

Otrzymano 66 reguł, w tym 58 pewnych i 8 przybliżonych. Wiele reguł posiadało jednak niewielkie wsparcie, czyli liczbę przypadków popierających regułę. Skłoniło to do ponownej próby indukcji reguł, tym razem metodą

pozwalająca na określenie parametrów satysfakcji (ang. *satisfactory description*). Wskazano, iż reguły powinny posiadać co najmniej 10% wsparcie (10% przykładów należących do danej klasy abstrakcji popiera-jest zgodnych z daną regułą, lub mówiąc inaczej, reguła pokrywa co najmniej 10% przykładów z danej klasy) oraz 100% dyskryminację (tylko reguły pewne, gdzie wszystkie przykłady spełniające część warunkową są jednoznacznie przypisywane do jednej klasy). Parametr określający maksymalną długość reguł (liczbę argumentów warunkowych) pozostawiono na domyślnej wartości równej długości użytego reduktu (11) – otrzymane reguły zawierały nie więcej niż 5 warunków, co potwierdziło przypuszczenia, iż parametr ten nie odegra istotnej roli.

Otrzymano zestaw 10 reguł decyzyjnych:

Reguła 1. Jeżeli $(2 = 1)$ i $(4 = 1)$ i $(8 = 2)$ i $(12 = 2)$ to (konkluzja = 1); [3, 3, 10.71%, 100.00%] [3, 0, 0, 0, 0] [{21, 27, 28}, {}, {}, {}, {}]

Reguła 2. Jeżeli $(4 = 3)$ i $(10 = 2)$ i $(13 = 3)$ to (konkluzja = 5); [3, 3, 10.00%, 100.00%] [0, 0, 0, 0, 3] [{}, {}, {}, {}, {228, 234, 235}]

Reguła 3. Jeżeli $(10 = 2)$ i $(11 = 3)$ i $(13 = 3)$ to (konkluzja = 5); [3, 3, 10.00%, 100.00%] [0, 0, 0, 0, 3] [{}, {}, {}, {}, {228, 235, 239}]

Reguła 4. Jeżeli $(2 = 3)$ i $(3 = 2)$ i $(4 = 3)$ i $(7 = 3)$ to (konkluzja = 5); [4, 4, 13.33%, 100.00%] [0, 0, 0, 0, 4] [{}, {}, {}, {}, {227, 230, 233, 249}]

Reguła 5. Jeżeli $(2 = 3)$ i $(3 = 2)$ i $(4 = 3)$ i $(8 = 3)$ to (konkluzja = 5); [3, 3, 10.00%, 100.00%] [0, 0, 0, 0, 3] [{}, {}, {}, {}, {227, 230, 249}]

Reguła 6. Jeżeli $(2 = 2)$ i $(3 = 3)$ i $(4 = 3)$ i $(12 = 3)$ to (konkluzja = 5); [3, 3, 10.00%, 100.00%] [0, 0, 0, 0, 3] [{}, {}, {}, {}, {234, 235, 250}]

Reguła 7. Jeżeli $(2 = 2)$ i $(4 = 3)$ i $(12 = 3)$ i $(13 = 3)$ to (konkluzja = 5); [3, 3, 10.00%, 100.00%] [0, 0, 0, 0, 3] [{}, {}, {}, {}, {234, 235, 250}]

Reguła 8. Jeżeli $(2 = 2)$ i $(7 = 3)$ i $(8 = 3)$ i $(12 = 3)$ to (konkluzja = 5); [3, 3, 10.00%, 100.00%] [0, 0, 0, 0, 3] [{}, {}, {}, {}, {234, 236, 250}]

Reguła 9. Jeżeli $(2 = 2)$ i $(7 = 3)$ i $(8 = 3)$ i $(13 = 3)$ to (konkluzja = 5); [3, 3, 10.00%, 100.00%] [0, 0, 0, 0, 3] [{}, {}, {}, {}, {234, 236, 250}]

Reguła 10. Jeżeli $(7 = 1)$ i $(8 = 1)$ i $(10 = 1)$ i $(11 = 2)$ i $(13 = 3)$ to (konkluzja = 5); [3, 3, 10.00%, 100.00%] [0, 0, 0, 0, 3] [{}, {}, {}, {}, {226, 242, 253}]

Każda reguła zawiera część warunkową i konkluzję, a ponadto opisana jest dodatkowymi charakterystykami, które w kolejności mówią o:

- liczbie pokrytych przez regułę przykładów – ile z nich ma identyczne wartości warunkowe,
- liczbie pokrytych przykładów pozytywnych 7 – ile z nich ma identyczne wartości warunkowe oraz konkluzję,
- wsparciu reguły – stosunku liczby pozytywnych pokrytych przykładów do liczby przykładów zgodnych z konkluzją,

⁷ Ze względu na ustawiony 100% poziom dyskryminacji, wszystkie reguły są regułami pewnymi i pokrywają jedynie przykłady pozytywne.

- zaufaniu reguły - stosunku liczby pozytywnych pokrytych przykładów do liczby pokrytych przykładów,
- rozkładzie przynależności do klas decyzyjnych obiektów pokrywanych przez regułę,
- numerach przykładów w odpowiedniej klasie abstrakcji.

Reguły ujawniły się jedynie dla skrajnych wartości konkluzji, co sugerować może, iż osoby deklarujące silne postawy mają podobny odbiór miasta, szczególnie te, które zamierzają je po studiach opuścić – 9 z 10 reguł z konkluzją=5 (całkowicie się nie zgadzam, że po zakończeniu nauki planuję związać swoją karierę ze Szczecinem). W regułach powtarzają się z większą intensywnością wybrane argumenty warunkowe: warunek 2 – 7 razy, 4 – 6 razy, 8 i 13 po 5 razy, co interpretować można jako ich ponadprzeciętne znaczenie do ujawniania się silnych konkluzji – zdecydowanie zostaję lub zdecydowanie wyjeżdżam.

Z reguły 1 odczytać można, iż 10% osób deklarujących chęć wiązania swojej kariery ze Szczecinem postrzega miasto jako romantyczne, czyste oraz ma neutralny stosunek do relacji wyrafinowania a kiczowatości i nowoczesności a zacofania.

Wśród osób zdeterminowanych do wyjazdu powtarzająca się ocena dotyczy warunku 2 – romantyczny a przyziemny, i osiąga wartości neutralne (2) lub negatywne (3). 6 reguł operuje argumentem 4, osoby pozostające widzą miast jako czyste (reguła 1), wszyscy wyjeżdżający jako brudne (reguły: 2, 4, 5, 6, 7). Warto odnotować, iż ta para antonimów jako jedyna uzyskała negatywną wartość średnią ocen [por.: 5].

Argument warunkowy 13 (wielkowiejski-zaściankowy) ujawnił się w 5 regułach o konkluzji zdecydowanie wyjeżdżam i we wszystkich przypadkach osiągnął wartość negatywną (zaściankowy). Mniej jednoznacznie rysuje się odbiór antonimów wyrafinowany-kiczowaty – warunek 8. W regule z konkluzją „zostaję” przyjął wartość neutralną, w pozostałych regułach zarówno negatywną – trzykrotnie, jak i pozytywną – 1 raz.

Podjęto kolejną próbę indukcji reguł nakierowaną na reguły o konkluzjach mniej stanowczych, ponownie wykorzystano metodę z parametrami satysfakcji jednakże obniżając dyskryminację do poziomu 80%, co skutkowało pojawieniem się kolejnych 6 reguł oraz zastąpieniem reguły 4 przez regułę bardziej ogólną o zredukowanej liczbie argumentów warunkowych (nowa reguła 7). Nowe reguły są regułami możliwymi (częściowo dyskryminującymi):

Reguła 1. Jeżeli $(3 = 1)$ i $(10 = 1)$ i $(12 = 3)$ to $(\text{konkluzja} = 2)$; [15, 12, 14.63%, 80.00%] [1, 12, 1, 0, 1] [{}], {38, 40, 41, 42, 44, 60, 67, 71, 86, 88, 108, 110}, {133}, {}, {242}]

Reguła 2. Jeżeli $(3 = 1)$ i $(11 = 3)$ i $(12 = 3)$ to $(\text{konkluzja} = 2)$; [11, 9, 10.98%, 81.82%] [1, 9, 0, 0, 1] [{}], {38, 44, 60, 67, 71, 88, 100, 108, 110}, {}, {}, {239}]

Reguła 3. Jeżeli $(3 = 3)$ i $(7 = 1)$ i $(10 = 1)$ i $(11 = 1)$ to (konkluzja = 3); [10, 8, 13.11%, 80.00%] [0, 0, 8, 1, 1] [{}, {}, {114, 116, 119, 144, 146, 151, 156, 171}, {210}, {241}]

Reguła 4. Jeżeli $(3 = 2)$ i $(8 = 2)$ i $(11 = 2)$ to (konkluzja = 4); [7, 6, 11.54%, 85.71%] [0, 0, 1, 6, 0] [{}, {}, {157}, {181, 192, 206, 207, 208, 223}, {}]

Reguła 5. Jeżeli $(7 = 3)$ i $(10 = 1)$ i $(12 = 2)$ to (konkluzja = 4); [7, 6, 11.54%, 85.71%] [0, 1, 0, 6, 0] [{}, {58}, {}, {178, 190, 192, 195, 197, 205}, {}]

Reguła 6. Jeżeli $(3 = 2)$ i $(4 = 3)$ i $(10 = 1)$ i $(12 = 1)$ to (konkluzja = 4); [8, 7, 13.46%, 87.50%] [0, 1, 0, 7, 0] [{}, {35}, {}, {179, 184, 206, 207, 217, 220, 221}, {}]

Reguła 7. Jeżeli $(2 = 3)$ i $(3 = 2)$ i $(7 = 3)$ to (konkluzja = 5); [5, 4, 13.33%, 80.00%] [0, 1, 0, 0, 4] [{}, {65}, {}, {}, {227, 230, 233, 249}]

Przykładowa reguła 1 o konkluzji=2 (raczej pozostanę w Szczecinie) operuje argumentami warunkowymi 3 (bezpieczny-niebezpieczny), 10 (otwarty-zamknięty) oraz 12 (nowoczesny-zacofany) przyjmując w dwóch pierwszych przypadkach wartość pozytywną oraz negatywną w przypadku trzecim. Reguła pokrywa 15 przykładów, posiada wsparcie na poziomie prawie 15% (12 pozytywnie pokrytych przykładów do 82 przykładów z konkluzją=2) i 80% zaufanie – 12 z 15 przykładów, które pokrywa są przykładami pozytywnymi (konkluzja=2), pozostałe 3 przykłady posiadają w tablicy decyzyjnej konkluzje 1, 3 i 5.

Reguła 3 ujawnia, iż osoby, które nie zadeklarowały ani chęci wyjazdu z miasta, ani pozostania w nim postrzegają Szczecin jako ładny, otwarty i pełen harmonii, lecz nie czują się w tu bezpiecznie.

Dotychczasowe analizy opierały się na jednym z dwóch dostępnych reduktów, wykorzystanie drugiego z nich prowadziło do podobnych reguł, zwrócić należy uwagę, iż redukty różnią się pomiędzy sobą nieznacznie, oprócz rdzenia jeden operuje argumentem 5 pomijając 10, w drugim jest sytuacja odwrotna.

Jednym z pytań zawartych w kwestionariuszu ankietowym było pytanie o (14) ogólne zadowolenie z życia w Szczecinie, które następnie uściślone zostało poprzez obszary: (15) sport i rekreacja, (16) jakość oferty szkół wyższych, (17) kultura, (18) rozrywka, (19) rozwój swoich pasji, (20) samorealizacja zawodowa, (21) poznawanie ciekawych ludzi, (22) bezpieczeństwo, (23) ochrona zdrowia, (24) komunikacja publiczna, (25) warunki mieszkaniowe. Także odnośnie tego pytania podjęto próbę uzyskania reguł decyzyjnych, gdzie poszczególne aspekty potraktowane zostały jako argumenty warunkowe tablicy decyzyjnej. Przyjmowały one wartości od 1 do 5, gdzie 1 oznaczał najniższy poziom zadowolenia, 5 zaś poziom najwyższy. Analogicznie jak miało to miejsce w poprzednim przypadku, argumentem decyzyjnym była deklaracja odnośnie planów po zakończeniu studiów. Otrzymanych zostało 14 reduktów o długości od 8 do 10 argumentów i rdzeń trzelementowy o bardzo niskiej jakości klasyfikacji. Sytuacja to skłoniła do ręcznej konstrukcji reduktu, który dla 7 zmiennych ({14, 16, 18, 20, 22, 23, 24}) osiągnął wysoką jakości klasyfikacji równą 0,967.

Indukcja reguł z przynajmniej 10% wsparciem i 100% poziomem dyskryminacji zaowocowała poniższym ich zestawem, w którym ponownie ujawniły się najsilniejsze postawy oraz ich dysproporcja – 1 reguła z konkluzją „zdecydowanie zostaje” i 9 reguł „zdecydowanie wyjeżdżam”:

Reguła 1. Jeżeli (16 = 5) i (22 = 3) i (23 = 3) to (konkluzja = 1); [4, 4, 14.29%, 100.00%] [4, 0, 0, 0, 0] [{1, 7, 16, 25}, {}, {}, {}, {}]

Reguła 2. Jeżeli (14 = 1) i (22 = 1) to (konkluzja = 5); [3, 3, 10.34%, 100.00%] [0, 0, 0, 0, 3] [{} , {} , {} , {} , {220, 228, 239}]

Reguła 3. Jeżeli (14 = 1) i (23 = 1) to (konkluzja = 5); [3, 3, 10.34%, 100.00%] [0, 0, 0, 0, 3] [{} , {} , {} , {} , {220, 228, 239}]

Reguła 4. Jeżeli (14 = 1) i (24 = 1) to (konkluzja = 5); [3, 3, 10.34%, 100.00%] [0, 0, 0, 0, 3] [{} , {} , {} , {} , {220, 228, 239}]

Reguła 5. Jeżeli (16 = 1) i (20 = 1) to (konkluzja = 5); [3, 3, 10.34%, 100.00%] [0, 0, 0, 0, 3] [{} , {} , {} , {} , {220, 233, 236}]

Reguła 6. Jeżeli (14 = 2) i (16 = 3) i (22 = 3) to (konkluzja = 5); [3, 3, 10.34%, 100.00%] [0, 0, 0, 0, 3] [{} , {} , {} , {} , {224, 241, 245}]

Reguła 7. Jeżeli (14 = 2) i (18 = 4) i (22 = 2) i (24 = 3) to (konkluzja = 5); [3, 3, 10.34%, 100.00%] [0, 0, 0, 0, 3] [{} , {} , {} , {} , {222, 230, 232}]

Reguła 8. Jeżeli (18 = 1) i (20 = 1) i (23 = 1) i (24 = 1) to (konkluzja = 5); [3, 3, 10.34%, 100.00%] [0, 0, 0, 0, 3] [{} , {} , {} , {} , {220, 233, 243}]

Reguła 9. Jeżeli (18 = 2) i (20 = 1) i (22 = 1) i (24 = 1) to (konkluzja = 5); [3, 3, 10.34%, 100.00%] [0, 0, 0, 0, 3] [{} , {} , {} , {} , {228, 237, 239}]

Reguła 10. Jeżeli (20 = 1) i (22 = 1) i (23 = 1) i (24 = 1) to (konkluzja = 5); [5, 5, 17.24%, 100.00%] [0, 0, 0, 0, 5] [{} , {} , {} , {} , {220, 228, 233, 237, 239}]

W odniesieniu do reguły 1 wgląd do danych źródłowych pozwala stwierdzić, iż wśród osób, które w argumentum warunkowym 16 wskazały wartość najwyższą – 5 także ci, którzy w pozostałych dwóch argumentach reguły wskazali wartości „lepsze”, niż ukazane w zaprezentowanej regule, deklarują silną chęć pozostania w Szczecinie – wartość 1, z wyjątkiem jednego przykładu gdzie konkluzja wynosi 2. Tym samym dalsza analiza, w tym łączenie reguł podobnych i ich uogólnianie – (22=3 lub 4 lub 5), czyli zaspokojenie danej potrzeby na poziomie nie niższym niż 3 bądź (konkluzja =1 lub 2) – zwiększyłaby wsparcie reguły. Ze względu na ograniczenia wydawnicze dalszy kierunek działań został jedynie zasygnalizowany.

Podkreślenia wymaga fakt, aby nie odnieść mylnego wrażenia, iż ujawniające się głównie reguły z częścią decyzyjną wskazującą na chęć wyjazdu z miasta nie świadczą o dominacji takich deklaracji w badanej zbiorowości. Wskazują natomiast na bardziej powtarzalne zachowania decyzyjne osób pragnących realizować swoją karierę poza Szczecinem. Trudniejsze do identyfikacji, a tym samym do przewidzenia, są reguły jakimi kierują się osoby pragnące wiązać swoją karierę ze Szczecinem lub te, które nie podjęły jeszcze decyzji. Liczebności poszczególnych klas zostały ukazane na rys. 1 (# of objects), 110 osób deklaruje chęć pozostania w Szczecinie (klasa 1 i 2), 61 nie podjęło jeszcze decyzji,

natomiast o wyjeździe myślą 82 osoby. Nadmienić należy, iż ponad połowa respondentów badania nie pochodzi z Szczecina.

5. Podsumowanie

Zaprezentowany niewielki wycinek możliwości i kierunków zastosowania teorii zbiorów przybliżonych wykorzystującej współczesną technologię obliczeniową pozwala podkreślić znaczenie technologii w szeroko pojętym zarządzaniu, a w szczególności w procesach wspomaganego podejmowania decyzji. Teoria zbiorów przybliżonych bez wsparcia technologicznego pozostałaby najprawdopodobniej w znacznej mierze w obszarze rozważań teoretycznych o niewielkim znaczeniu utylitarnym. Wsparcie technologiczne czyni z niej wartościowe narzędzie wspierające zarządzanie, pozwalające, odmiennie niż tradycyjne metody ilościowe, uchwycić rzeczywiste zależności opisane w sposób jakościowy, niespójny i niepełny, przenosi to także technologię na wyższy poziom zastosowań.

Dzięki symbiozie metod i technik obliczeniowych stosunkowo prosty i relatywnie mało czasochłonny proces obróbki danych pozwolił na identyfikację i redukcję danych nadmiarowych oraz budowę regułowej bazy wiedzy ukazującej relacje pomiędzy poszczególnymi elementami wizerunku Szczecina a planami na przyszłość wśród osób stanowiących potencjalnie wysoko wykształcone kadry Szczecina. Reguły w obu podjętych obszarach – wizerunku i zaspokajania potrzeb – ujawniły dość silną koncentrację zachowań osób pragnących wyjechać po zakończeniu studiów. Reguły te, oraz sposób ich indukcji, stać się mogą przyczynkiem do dalszej analizy oraz nakierowania działań, które stratedzy miejsca mogliby podjąć celem osłabienia niekorzystnej tendencji związanej z wyjazdami młodych, dobrze wykształconych ludzi szukających miejsca do realizacji swoich ambicji zawodowych poza Szczecinem, co w warunkach nasilającej się konkurencji pomiędzy miastami wydaje się istotnym elementem zarządzania jednostką terytorialną [por.: 2].

Bibliografia

- [1] Czaplewski M., E-biznes jako kierunek doskonalenia usług ubezpieczeń gospodarczych, Polskie Towarzystwo Ekonomiczne, Szczecin, 2007, s. 7
- [2] Gębarowski M., Oferta miasta akademickiego a oczekiwania studentów (na przykładzie Rzeszowa oraz studentów Politechniki Rzeszowskiej), w: Rosa G., Smalec A. (red.), Marketing przyszłości . Trendy. Strategie. Instrumenty. Strategie marketingowe miast i regionów, Zeszyty Naukowe Uniwersytetu Szczecińskiego, nr 709, Szczecin 2012, s. 161, 164-165
- [3] Gębarowski M., Studenci – lokalni patrioci czy obywatele świata? Znaczenie poczucia przynależności studentów do miejsca studiowania w kontekście kształtowania wizerunku ośrodka akademickiego, w: Smalec A. (red.), Marketing

- przyszłości. Trendy. Strategie. Instrumenty. Orientacja marketingowa jednostek samorządu terytorialnego i instytucji publicznych, Zeszyty Naukowe Uniwersytetu Szczecińskiego, nr 775, Szczecin 2013, s. 466
- [4] Kuczera K., Metoda AHP w identyfikacji preferencji decydentów, *Ekonomika i Organizacja Przedsiębiorstw* 2015, nr 3, s. 44
 - [5] Kuczera K., Strategia rozwoju miasta Szczecina a jego wizerunek w świetle badań naukowych, *Przegląd organizacji* 2014, nr 11, s. 38
 - [6] Marek W., Pawlak Z., Rough sets and information systems, ICS PAS Reports 441/81, Institute of Computer Science Polish Academy of Sciences (ICS PAS), Warsaw, 1981, s. 1-15
 - [7] Mastalerz M. W., Procedura wyboru alternatywy informatyzacji przedsiębiorstwa oparta na metodzie AHP, w: Kiełtyka L., Niedbał R. (Red.), *Wybrane zastosowania technologii informacyjnych wspomagających zarządzanie w organizacjach*, Wydawnictwo Politechniki Częstochowskiej, Częstochowa 2015, s. 51-66
 - [8] Pawlak Z., Rough sets, *International Journal of Information and Computer Sciences* 1982, Volume 11, Issue 5, s. 341-356
 - [9] Pawlak Z., Rough sets. Basic notions, ICS PAS Reports 431/81, Institute of Computer Science Polish Academy of Sciences (ICS PAS), Warsaw, 1981, s. 5
 - [10] Pawlak Z., Some remarks about rough sets, ICS PAS Reports 456/82, Institute of Computer Science Polish Academy of Sciences (ICS PAS), Warsaw, 1982, s. 4-5
 - [11] Pawlak Z., Zbiory przybliżone. Nowa matematyczna metoda analizy danych, Osiągnięcia nauki i techniki. Kierunki rozwoju i metody. Konwersatorium Politechniki Warszawskiej, nr 5, Politechnika Warszawska, Warszawa 2004, s. 1-6, 11
 - [12] Pilipczuk O., Grafika kognitywna w podejmowaniu decyzji, *Wydawnictwo Naukowe Uniwersytetu Szczecińskiego*, Szczecin 2013, s. 50, 116
 - [13] Predki B., Słowinski R., Stefanowski J., Susmaga R., Wilk Sz.: Rose - Software Implementation of the Rough Set Theory, w: Polkowski L., Skowron A., (red.) *Rough Sets and Current Trends in Computing, Lecture Notes in Artificial Intelligence*, vol. 1424. Springer-Verlag, Berlin 1998, s. 605-608
 - [14] Predki B., Wilk Sz., Rough Set Based Data Exploration Using ROSE System, w: Ras Z.W., Skowron A. (red.), *Foundations of Intelligent Systems, Lecture Notes in Artificial Intelligence*, vol. 1609. Springer-Verlag, Berlin 1999, s. 172-180
 - [15] Saaty T., *The Analytic Hierarchy and Analytic Network Process for the Measurement of Intangibles and for Decision Making* w: Figueira J., Greco S., Ehrgott M., *Multiple Criteria Decision Analysis: State of the Art Surveys*, Springer Science + Business Media, Boston 2005, s. 346
 - [16] Stefanowski J., *Algorytmy indukcji reguł decyzyjnych w odkrywaniu wiedzy*, rozprawa habilitacyjna, Seria Rozprawy nr 361, Wydawnictwo Politechniki Poznańskiej, Poznań 2001, s. 25-26
 - [17] Suraj Z., *An Introduction to Rough Set Theory and Its Applications. A tutorial*, ICENCO'2004, Cairo, s. 2-3
 - [18] Wawrzyniak A., Wąsikowska B., *Badanie preferencji wyborczych mieszkańców województwa zachodniopomorskiego przy użyciu zbiorów przybliżonych*, *Polskie Stowarzyszenie Zarządzania Wiedzą, Studia i Materiały* nr 56, Bydgoszcz 2011

- [19] Wąsikowska B., Furajji F., Badanie preferencji zakupowych kobiet i mężczyzn metodą zbiorów przybliżonych. *Handel Wewnętrzny*, tom I, maj-czerwiec 2012