



Zachodniopomorski
Uniwersytet Technologiczny
w Szczecinie

WYDZIAŁ INFORMATYKI
Katedra Inżynierii Systemów Informatycznych

Szczecin, dnia 27 czerwca 2024 r.

Dr hab. inż. Piotr Sulikowski, prof. ZUT
Katedra Inżynierii Systemów Informatycznych
Wydział Informatyki
Zachodniopomorski Uniwersytet Technologiczny
w Szczecinie

Rada Naukowa
Dyscyplin Matematyka i Informatyka
Uniwersytet im. Adama Mickiewicza
w Poznaniu

Recenzja rozprawy doktorskiej

mgra Łukasza Szymali

pt. „Zastosowanie teorii zbiorów rozmytych w klasyfikacji dominującej
dyscypliny naukowej autorów w naukometrii”

opracowanej pod kierunkiem prof. UAM dra hab. Krzysztofa Dyczkowskiego
i prof. dra hab. Marka Kwieka

1. Cel, zakres i układ rozprawy

Jednym z wyzwań naukometrii jest właściwa klasyfikacja dyscyplin naukowych autorów, która ma istotne znaczenie dla oceny ich dorobku naukowego. Tradycyjne podejście, nazywane w rozprawie bazowym, oparte na wartości modalnej z dyscyplin czasopism, korzysta z prostych algorytmów i ograniczonej interpretacji bogatych danych. W praktyce podejście to często skutkuje niejednoznacznościami w klasyfikacji.

Stąd główny cel recenzowanej rozprawy, jakim było zbadanie możliwości wykorzystania teorii zbiorów rozmytych do ulepszenia algorytmu klasyfikacji dominującej dyscypliny naukowej autorów przy pomocy wartości modalnej, jest jak najbardziej użyteczny i interesujący.

Rozprawa datowana jest na 2024 rok, liczy 122 strony (w tym 114 numerowanych i 8 nienumerowanych) i oprócz treści głównej obejmuje spis treści, streszczenia w języku polskim i angielskim, bibliografię oraz trzy załączniki. Autor podzielił pracę na 7 rozdziałów, w tym wprowadzenie, 5 rozdziałów właściwych i podsumowanie. Rozdziały właściwe zostały podzielone na 3 do 5 podrozdziałów, zawierających nie więcej niż 8 podpodrozdziałów.

Rozdział pierwszy stanowi krótki wstęp do tematyki i zakresu poszczególnych części pracy. W rozdziale drugim omówiono ogólne pojęcia naukometrii, wybrane bazy bibliometryczne i wskaźniki naukometryczne, a także zaprezentowano podejście bazowe ze szczególnym uwzględnieniem jego niedoskonałości. W rozdziale trzecim przedstawiono podstawowe pojęcia teorii zbiorów rozmytych, zajmującej się reprezentacją, przetwarzaniem oraz sterowaniem na podstawie nieprecyzyjnych informacji, w tym trzy elementy, na których oparte są proponowane przez autora rozwiązania, tj. zmienne lingwistyczne, sterowniki rozmyte i operatory uporządkowanej średniej ważonej (OWA – *Ordered Weighted Averaging*). Kolejne trzy rozdziały poświęcone są proponowanym algorytmom klasyfikacji dominującej dyscypliny naukowej w podejściu rozmytym. Rozdział siódmy, ostatni, zawiera krótkie podsumowanie rozprawy i kierunki rozwoju przedstawionego w niej podejścia.

2. Ocena rozprawy

Oceniając recenzowaną rozprawę doktorską, należy odnieść się do przepisów Ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2022 r., poz. 574, z późn. zm.), w oparciu o które toczy się przedmiotowe postępowanie w sprawie nadania stopnia doktora. Artykuł 187 tej ustawy definiuje następujące wymogi względem rozpraw doktorskich:

- 1. Rozprawa doktorska prezentuje ogólną wiedzę teoretyczną kandydata w dyscyplinie albo dyscyplinach oraz umiejętność samodzielnego prowadzenia pracy naukowej lub artystycznej.*
- 2. Przedmiotem rozprawy doktorskiej jest oryginalne rozwiązanie problemu naukowego, oryginalne rozwiązanie w zakresie zastosowania wyników własnych badań naukowych w sferze gospodarczej lub społecznej albo oryginalne dokonanie artystyczne.*
- 3. Rozprawę doktorską może stanowić praca pisemna, w tym monografia naukowa, zbiór opublikowanych i powiązanych tematycznie artykułów naukowych, praca projektowa, konstrukcyjna, technologiczna, wdrożeniowa lub artystyczna, a także samodzielna i wyodrębniona część pracy zbiorowej.*
- 4. Do rozprawy doktorskiej dołącza się streszczenie w języku angielskim, a do rozprawy doktorskiej przygotowanej w języku obcym również streszczenie w języku*

polskim. W przypadku gdy rozprawa doktorska nie jest pracą pisemną, dołącza się opis w językach polskim i angielskim.

Rozprawa mgra Łukasza Szymuli ma charakter teoretyczno-praktyczny. Zaproponowane przez autora rozmyte podejście do klasyfikacji zostało zastosowane w obszarze naukometrii i wpisuje się w dyscyplinę informatyka, w szczególności obszar metodologii obliczeniowych [ang. *Computing methodologies*], podobszar sztucznej inteligencji [ang. *Artificial intelligence*], reprezentacji wiedzy i wnioskowania [ang. *Knowledge representation and reasoning*], nieprecyzyjności i logiki rozmytej [ang. *Vagueness and fuzzy logic*] według ostatniej klasyfikacji *Association for Computing Machinery – ACM Computing Classification System*.

Poziom szczegółowości opisów wybranych elementów teorii zbiorów rozmytych, a także przeprowadzone badania o charakterze aplikacyjnym wskazują na znaczącą wiedzę teoretyczną autora i umiejętność jej zastosowania w pracy naukowej. Kandydat do stopnia doktora zapoznał się ze znaczną liczbą prac dotyczących zbiorów rozmytych i naukometrii. Bibliografia obejmuje 96 pozycji, stanowiących w większości recenzowane artykuły naukowe w języku angielskim. Opisy bibliograficzne przygotowane zostały w oparciu o styl APA (*American Psychological Association*).

Autor odnosi się do źródeł w czytelny sposób w treści pracy, chociaż nie zawsze konsekwentnie stosuje wybrany styl cytowania. Struktura pracy jest przemyślana i kompletna, brakuje jednak takich elementów jak spis rysunków, spis tabel czy wykaz zastosowanych skrótów. Praca została napisana językiem poprawnym, odpowiednim dla tego typu tekstów. Można natknąć się na nieliczne literówki, błędy interpunkcyjne itp., czy braki w numeracji wzorów. Znaczenie niektórych zmiennych lingwistycznych nie zostało podane wprost podczas opisu ich konstrukcji w podrozdziale 4.2.1, ale z kontekstu można było wywnioskować, że np. zmienna „Zespół” odnosi się do metryki „Liczba autorów”, zmienna „Percentyl” do metryki „Percentyl czasopisma publikacji”, a zmienna „Rok” oznacza rok wydania publikacji. Z drugiej strony autor zadbał o wyróżnienie nazw zmiennych, co zwiększyło czytelność pracy. W spisie treści nie uwzględniono załączników zawierających szczegółowe wyniki metryk ewaluacyjnych. Wskazane niedoskonałości mają w większości charakter redakcyjny i nie mają istotnego wpływu na pozytywny odbiór pracy.

Mając na uwadze powyższe, uważam, że recenzowana rozprawa prezentuje ugruntowaną wiedzę teoretyczną kandydata w dyscyplinie informatyka i potwierdza umiejętność samodzielnego prowadzenia pracy naukowej, **spełniając wymogi art. 187 pkt 1 ustawy.**

Tematem recenzowanej rozprawy jest „Zastosowanie teorii zbiorów rozmytych w klasyfikacji dominującej dyscypliny naukowej autorów w naukometrii”. Na str. 17-18 pracy autor wskazuje hipotezy badawcze, które w istocie stanowią pytania badawcze:

- Czy zastosowanie metod z aparatu pojęciowego teorii zbiorów rozmytych w algorytmie klasyfikacji dominującej dyscypliny może jednoznacznie zaklasyfikować większą ilość obserwacji niż podejście bazowe?
- Które rozwiązania znane z teorii zbiorów rozmytych pozwalają na poprawę liczby jednoznacznie zaklasyfikowanych autorów w podejściu rozmytym?
- Które zmienne lingwistyczne, termy i funkcje wagowe są kluczowe w poprawie jednoznacznej klasyfikacji?
- Dla jakich parametrów wejściowych sterownika rozmytego i funkcji wagowych możliwe jest uzyskanie poprawy jednoznacznej klasyfikacji dyscyplin?
- Jak dobór wag w operatorze OWA wpływa na poprawę jednoznacznie zaklasyfikowanych autorów do dyscyplin?
- Czy klasyfikacja wykonana podejściem rozmytym przypisuje takie same klasy jak w przypadku podejścia bazowego?

Wnioskując z całości wyводу autora rozprawy, za główną hipotezę badawczą należałoby przyjąć przypuszczenie, że możliwe jest uzyskanie zwiększenia jednoznacznej klasyfikacji dominującej dyscypliny naukowej w porównaniu z podejściem bazowym, dzięki zastosowaniu metod z obszaru teorii zbiorów rozmytych, takich jak (i) zmienne lingwistyczne, (ii) sterowniki rozmyte, (iii) agregacje wartości rozmytych z wykorzystaniem operatorów OWA. Hipoteza taka wydaje się być hipotezą prawdopodobną o wysokich szansach praktycznej weryfikacji.

W naukometrii można operować pojęciami o charakterze nieprecyzyjnym. Zaproponowana przez autora modyfikacja podejścia bazowego zakłada zastosowanie ww. metod na etapie wyznaczania częstości występowania danej dyscypliny. Przeprowadzone eksperymenty, opisane w kluczowych według mnie rozdziałach 4, 5 i 6, dały bardzo obiecujące efekty, pozwalając na wyraźne zmniejszenie niejednoznaczności klasyfikacji dominującej dyscypliny naukowej, i wskazując na przewagę rozmytego podejścia obliczeniowego nad podejściem bazowym. O ile w algorytmie standardowym jednoznaczną klasyfikację osiągnięto dla ok. 69% naukowców z bazy, to w przypadku zmiennych lingwistycznych udało się tak zaklasyfikować nawet ok. 92% naukowców z bazy (najlepsze efekty dla zmiennej wyrażającej rok wydania publikacji). W podejściu opartym na sterownikach rozmytych zaklasyfikowano ok. 95% obserwacji. Podobny wynik uzyskano przy zastosowaniu operatorów OWA. Autor dodatkowo zweryfikował podobieństwo klasyfikacji uzyskanej przez zaproponowane modyfikacje w stosunku do podejścia standardowego, nie uwzględniającego nieprecyzyjności informacji. Zabrakło mi szerszego, agregującego podsumowania wszystkich wyników, uzyskanych z wykorzystaniem różnych metod, popartego głębszą dyskusją.

Na uwagę zasługuje fakt, że eksperymenty prowadzone były na realnych danych, z wykorzystaniem platformy ICSR Lab na pełnej bibliometrycznej bazie Scopus (60 mln rekordów), do której dostęp otrzymano w ramach umowy zawartej pomiędzy Centrum Studiów nad Polityką Publiczną Uniwersytetu im. Adama Mickiewicza w Poznaniu

a International Center for the Study of Research (ICSR) wydawnictwa Elsevier. W związku z ograniczeniami ICSR Lab, skrypty pozyskujące wyniki wymagały od autora rozprawy własnej implementacji rozwiązań teorii zbiorów rozmytych.

Zaprezentowane w rozprawie opisy prac i wyniki eksperymentów wskazują na oryginalność rozwiązania postawionego problemu, a także wykorzystanie wiedzy teoretycznej w badaniach prowadzonych pod kątem ściśle określonych potrzeb praktyki. Tym samym oceniam, że przedstawiona rozprawa **spełnia wymogi art. 187 pkt 2 ustawy**.

Recenzowana rozprawa stanowi pracę pisemną, stanowi spójną całość, opracowaną w sposób, jakiego oczekuje się od prac naukowych. Rozprawa została napisana w języku polskim, a dodatkowo uwzględniono streszczenie w języku angielskim. Należy zatem uznać, że **wymogi art. 187 pkt 3 i 4 ustawy również zostały spełnione**.

3. Wnioski końcowe

Podsumowując, w mojej ocenie rozprawa doktorska mgra Łukasza Szymuli pt. „Zastosowanie teorii zbiorów rozmytych w klasyfikacji dominującej dyscypliny naukowej autorów w naukometrii” **spełnia wszystkie wymogi formalnie stawiane rozprawom doktorskim**, określone w art. 187 Ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2022 r., poz. 574, z późn. zm.). W związku z tym wnoszę o dopuszczenie kandydata do publicznej obrony i dalszego procedowania w postępowaniu o nadanie stopnia naukowego doktora w dyscyplinie informatyka.



dr hab. inż. Piotr Sulikowski, prof. ZUT