

dr hab. Agnieszka Nowak – Brzezińska, Prof. UŚ
Uniwersytet Śląski
Wydział Nauk Ścisłych i Technicznych
Instytut Informatyki
e-mail: agnieszka.nowak-brzezinska@us.edu.pl

Katowice, 8 stycznia 2026 r.

RECENZJA ROZPRAWY DOKTORSKIEJ

mgr inż. **Dawida Kosiora**

pt. „Nowe metody uwzględniania niepewności i nieprecyzyjności danych
w klasyfikacji. Zastosowania w medycynie.”

przygotowanej pod kierunkiem

dr hab. Barbary Pękali, prof. UR

Uwagi wstępne

Recenzja opracowana została na podstawie pisma Zastępcy Dyrektora Instytutu Badań Systemowych Polskiej Akademii Nauk ds. Naukowych dra hab. inż. Jana W. Owsńskiego, z dnia 10 października 2025 r.

1. Przedmiot rozprawy, aktualność i ważność tematyki rozprawy

Rozprawa poświęcona jest problemowi niepewności i nieprecyzyjności danych w zadaniach klasyfikacji. Tematyka ta w ostatnim czasie jest podstawą wielu badań. W dobie gwałtownego wzrostu złożoności systemów informatycznych kluczowym wyzwaniem naukowym staje się adekwatna reprezentacja wiedzy, której nieodłącznym elementem jest niepewność oraz informacyjna nieprecyzyjność. Badania nad zbiorami rozmytymi są w tym kontekście fundamentalne, gdyż oferują dojrzały aparat matematyczny pozwalający na sformalizowanie i efektywne przetwarzanie danych, które ze względu na swój nieostry charakter wymykają się rygorom klasycznej logiki.

Współczesna nauka stoi przed wyzwaniem budowy systemów, które potrafią operować w środowisku naturalnie nasyconym niepewnością i nieprecyzyjnością, co jest kluczowe dla zbliżenia sposobu przetwarzania danych przez maszyny do ludzkich procesów poznawczych.

W informatyce pojęcia te, choć często stosowane zamiennie, opisują dwa różne wymiary niedoskonałości informacji, których sformalizowanie za pomocą zbiorów rozmytych jest niezbędne dla wiarygodności systemów decyzyjnych.

Klasyczna teoria zbiorów rozmytych pozwala opisywać pojęcia nieprecyzyjne (nieostre) poprzez wprowadzenie stopni przynależności, które wykraczają poza binarne ujęcie należenia elementu do zbioru i związane z tym binarne pojęcie prawdy i fałszu. Jednak w przypadku danych niepełnych lub niedoskonałych, gdzie dokładne określenie stopnia przynależności jest trudne lub niemożliwe, taka reprezentacja okazuje się niewystarczająca. Mamy wtedy do czynienia z nieprecyzyznością i związaną z nią niepewnością. W odpowiedzi na te ograniczenia, badacze opracowali szereg rozszerzeń klasycznej teorii zbiorów rozmytych, które umożliwiają lepsze modelowanie i przetwarzanie informacji obciążonej niepewnością i nieprecyzyznością. Proces wnioskowania rozmytego propaguje niepewność od poziomu danych wejściowych aż do decyzji. W wielu zastosowaniach sama funkcja przynależności także jest niepewna co prowadzi do rozwinięć teorii w postaci niepewnych zbiorów rozmytych czy zbiorów rozmytych typu 2. W takich modelach niepewność wiedzy jest reprezentowana na dwóch poziomach: rozmytość opisuje nieostrość pojęcia, a dodatkowe struktury (np. przedziały, funkcje wyższego rzędu) opisują niepewność co do samej funkcji przynależności. W rozprawie zaproponowano nowe podejścia do uwzględniania niedoskonałości danych, rozumianej jako niepewność lub nieprecyzyzność, w klasyfikacji danych, analizowanej zarówno w ujęciu teoretycznym, jak i praktycznym. Część teoretyczna obejmowała opracowanie nowych miar zawierania, podobieństwa oraz entropii dla przedziałowych zbiorów rozmytych, w kontekście danych niepewnych i/lub nieprecyzyjnych. W części aplikacyjnej zaproponowano metody wykorzystania tych miar na różnych etapach tworzenia klasyfikatora. Teoria zbiorów rozmytych oraz jej rozszerzenia stanowią skuteczne narzędzie do reprezentacji i przetwarzania takich danych. Badania koncentrowały się przede wszystkim na zastosowaniach praktycznych, ze szczególnym uwzględnieniem diagnostyki medycznej. W tym kontekście tematykę podjętą w rozprawie doktorskiej Pana magistra inżyniera Dawida Kosiora, dotyczącą rozwoju metod reprezentacji niepewności i nieprecyzyzności z użyciem przedziałowych zbiorów rozmytych, należy uznać za ważną, celową i aktualną. W ostatnich latach obserwuje się istotny wzrost liczby publikacji w tym obszarze, co potwierdza, że w wielu przypadkach przedziałowe zbiory rozmyte pozwalają uzyskać lepsze wyniki niż klasyczne zbiory rozmyte, gdyż umożliwiają one uwzględnienie nieprecyzyzności nie tylko na poziomie

samych danych, ale również na poziomie konstrukcji funkcji przynależności i rozbieżności ocen eksperckich.

Głównym celem rozprawy jest opracowanie nowych metod i algorytmów uwzględniających niepewność i nieprecyzyjność danych w celu poprawy jakości klasyfikatorów tworzonych za pomocą metody k -najbliższych sąsiadów, bardzo szeroko stosowanej i dającej zwykle bardzo dobre wyniki. Cel ten podzielony został na cele podrzędne tj.:

- Opracowanie miar i metod uwzględniających niepewność i nieprecyzyjność, umożliwiających efektywną klasyfikację;
- Opracowanie metod klasyfikacji uwzględniających dane niepewne, opartych o algorytm k -NN z wykorzystaniem nowych miar, a także wnioskowania rozmytego z nowymi agregacjami uwzględniającymi niepewność;
- Selekcja atrybutów dla poszczególnych klas obiektów - zaproponowanie efektywnej metody selekcji cech przy wykorzystaniu miar uwzględniających niepewność, w szczególności nowej miary podobieństwa;
- Oszacowanie wpływu brakujących wartości na jakość klasyfikacji - zaproponowanie efektywnej metody uzupełniania brakujących wartości z wykorzystaniem miar uwzględniających niepewność, w szczególności nowej miary entropii;
- Weryfikację przydatności proponowanych metod reprezentacji i manipulacji danymi z niepewnością, która została przeprowadzona poprzez sprawdzenie ich wpływu na jakość klasyfikacji.

Biorąc pod uwagę zawartość rozprawy uważam, że główna teza rozprawy, z godnie z którą uwzględnienie niepewności i nieprecyzyjności może znacząco poprawić jakość działania klasyfikatorów opartych na idei k -najbliższych sąsiadów, zarówno w warunkach danych niekompletnych jak i w procesie optymalizacji zbioru cech, została potwierdzona.

2. Zawartość rozprawy

Łączna objętość przedłożonej do recenzji pracy wynosi 137 stron. Doktorant zadbał o to by na początku uwzględnić wykaz skrótów stosowanych w rozprawie. Wprowadzenie do rozprawy pozwala już na samym początku poznać cele jakie przyjął Doktorant dla swojej pracy badawczej. W rozdziale drugim doktorant wprowadza podstawowe pojęcia takie jak zbiór rozmyty, rozszerzenie zbiorów rozmytych w postaci przedziałowych zbiorów rozmytych, relacje porządku oraz dla przedziałowych zbiorów rozmytych odpowiednio rozmyte agregacje,

negacje oraz relacje. W kolejnym rozdziale doktorant wyjaśnia miary zawierania się przedziałów i przedziałowych zbiorów rozmytych, miary podobieństwa dla przedziałowych zbiorów rozmytych oraz entropię tychże. Rozdział ten zatem kompletuje wiedzę z zakresu przedziałowych zbiorów rozmytych będącą podstawą badań doktoranta. Celem rozprawy jest wykazanie wsparcia jakie w zadaniu klasyfikacji danych daje logika rozmyta i jej rozszerzenie w postaci przedziałowych zbiorów rozmytych i z tego powodu w rozdziale czwartym doktorant wprowadza w tematykę zadania klasyfikacji, przedstawiając wybrane metody klasyfikacji i wybrane miary jakości klasyfikacji. Rozdział dość skrótowo traktuje tematykę klasyfikacji omawiając bardzo ogólnie jedynie wybrane metody oraz miary jakości klasyfikacji. W rozdziale piątym autor omawia metody wspierające klasyfikację w diagnostyce medycznej w oparciu o kryterium sąsiedztwa. Przedstawia najpierw ideę wnioskowania rozmytego z użyciem algorytmu k najbliższych sąsiadów dla przedziałowych zbiorów rozmytych, proponuje algorytm progowania dla wyboru decyzji, z analizą jego zastosowania dla medycznego zbioru danych, technikę uzupełniania braków w danych z uwzględnieniem przedziałowych zbiorów rozmytych a następnie przedstawia koncepcję selekcji atrybutów z użyciem zaproponowanego algorytmu *IVRelief*. Rozdział opisuje także konkretny przypadek systemu diagnostyki depresji, który od początku do końca bazuje na zastosowaniu koncepcji przedziałowych zbiorów rozmytych, wnioskowania rozmytego i zadania klasyfikacji. Rozdział szósty potwierdza adaptowalność przedziałowych zbiorów rozmytych w kontekście wnioskowania rozmytego w systemach regułowych na przykładzie systemu wspierającego diagnostykę chorób sercowo-naczyniowych. Rozdział ten zawiera także opis biblioteki IFIS będącej implementacją systemu wnioskowania rozmytego w środowisku Python. Biblioteka ta rozszerza klasyczne wnioskowanie rozmyte o obsługę wartości przedziałowych, co pozwala na lepsze odwzorowanie niepewności w modelach decyzyjnych. Ostatni rozdział rozprawy jest podsumowaniem osiągniętych wyników oraz opisem planowanych badań. Pracę kończą skorowidz, spis bibliografii oraz spisy rysunków i tabel.

Podsumowując, w mojej ocenie struktura przedłożonej rozprawy doktorskiej jest w pełni poprawna, a układ rozdziałów wyczerpuje podjętą problematykę badawczą, czyniąc pracę konstrukcją kompletną i spójną.

3. Oryginalne osiągnięcia

Oryginalny dorobek rozprawy zawarty jest w rozdziałach 2,3 oraz 5 rozprawy. Do dorobku tego należy zaliczyć:

- Zaproponowano uogólniony rozmyty klasyfikator wspomagający procesy decyzyjne, oparty na niepewnych danych z zastosowaniem przedziałowych danych. W efekcie stworzono narzędzie pozwalające na modelowanie i przetwarzanie danych wejściowych oraz prezentację wyników w sposób pozwalający na zachowanie istotnych informacji dotyczących poziomu niepewności na każdym etapie.
- Zaproponowano metodę uzupełniania brakujących wartości z użyciem przedziałowej miary entropii.
- Zaproponowano nowy algorytm wyznaczania istotności atrybutów w zbiorze danych - *IVRelief* (rozdział 5.3.1.). Algorytm wywodzi się od zaproponowanego wcześniej algorytmu Relief, dostosowanego do danych przedziałowych. Nowością jest wykorzystanie miary podobieństwa pomiędzy przedziałowymi zbiorami rozmytymi, uwzględniające niepewność informacji.
- Opracowano nowy test diagnostyczny *senDD* i system diagnostyczny System Prewencji Depresji (SPD) bazujący na teście *SenDD*. Zadanie to obejmowało wprowadzenie algorytmu wyznaczania wstępnej diagnozy wykorzystującego przedziałowe zbiory rozmyte oraz algorytmu *SENSDEPR* zwiększającego czułość diagnostyczną do weryfikacji decyzji ostatecznej.
- Zaproponowano metody porządkowania danych przedziałowych a także agregacji i miary entropii oraz podobieństwa (nową metodę konstrukcji przedziałowej miary podobieństwa zgodną z porządkiem częściowym lub liniowym, czyli w ujęciu klasycznym).
- Opracowano miary i metody uwzględniające niepewność i nieprecyzyjność umożliwiające efektywną diagnostykę medyczną oraz przeprowadzono weryfikację przydatności proponowanych metod poprzez sprawdzenie ich wpływu na jakość klasyfikacji.
- Zaproponowano uogólniony algorytm przedziałowego wnioskowania rozmytego (Algorytm 6, strona 104) dedykowany przypadkom w których eksperci mają problem z określeniem stopni przynależności elementów do rozważanych zbiorów rozmytych. Algorytm wykorzystuje m.in. klasy agregacji możliwości czy konieczności. Klasyczny schemat *modus ponens* został zastąpiony w modelu przedziałowego wnioskowania przez uogólniony *modus ponens*, gdzie przesłanki i wnioski są przedziałowymi zbiorami rozmytymi.

- Doktorant jest także współautorem biblioteki open-source do wnioskowania przybliżonego opartego na wartościach przedziałowych IFIS (rozdział 6.2.) dla języka Python.
- Pewnym globalnym miernikiem uznania istotności prac Autora mogą być indeksy bibliometryczne, tj. liczba punktów przyznanych publikacji, wskaźnik Impact Factor czy licznik cytowań. W dorobku Doktoranta znajduje się 16 publikacji o zasięgu międzynarodowym, lecz tylko w jednej z nich Doktorant jest pierwszym autorem. Oceniając całościowo dorobek rozprawy stwierdzam, że Autor w sposób systematyczny zrealizował zaplanowane badania oraz ocenił ich wyniki, wystarczająco dobrze je dokumentując. Wyniki przeprowadzonych prac badawczych zostały opublikowane w artykułach naukowych m.in. w przypisanych do dyscypliny Informatyka Techniczna i Telekomunikacja czasopismach *Applied Soft Computing* (200 pkt. MNiSW), *Journal of Artificial Intelligence and Soft Computing Research* (140 pkt. MNiSW), *International Journal of Intelligent Systems* (100 pkt. MNiSW) i *Entropy* (100 pkt. MNiSW). Dorobek ten uzupełniają artykuły opublikowane w materiałach konferencyjnych oraz rozdziały w monografiach naukowych. Dotychczasowe prace kandydata były cytowane łącznie 12 razy wg Web Of Science, 39 razy wg Scopus i 46 razy wg Scholar Google, większość baz publikacji wskazuje indeks Hirscha na poziomie 4. Warto podkreślić, że znacząca część tych cytowań dotyczy artykułów opublikowanych w czasopismach o wysokim współczynniku wpływu (Impact Factor). Potwierdza to zainteresowanie wynikami prowadzonych przez Doktoranta prac badawczych oraz samą dziedziną.

W związku z powyższymi osiągnięciami można stwierdzić, że wszystkie cele postawione w rozprawie zostały osiągnięte. Realizacja ww. celów stanowi potwierdzenie tezy postawionej w niniejszej rozprawie, która brzmi: uwzględnienie niepewności i nieprecyzyjności może znacząco polepszyć jakość działania klasyfikatorów opartych na idei k-najbliższych sąsiadów, nawet w warunkach danych niekompletnych lub nadmiarowych.

4. Analiza źródeł

Doktorant wykazał się dobrą znajomością dorobku literaturowego dotyczącego zagadnień, którym rozprawa jest poświęcona. Bibliografia liczy 139 prac, z czego około 40% stanowią pracę z ostatnich 10 lat. W bibliografii doktorant przywołuje 10 ze swoich 16

publikacji. Uwzględnił prace odnoszące się zarówno do najnowszych badań w tematyce rozprawy, jak i prace dokumentujące rozwój teorii zbiorów rozmytych i ogólnie metod klasyfikacji danych na przestrzeni ostatnich kilkudziesięciu lat. Taki dobór literatury potwierdza dostateczną wiedzę Doktoranta w obszarze badawczym rozprawy doktorskiej i należy ocenić go pozytywnie.

5. Wady i słabe strony rozprawy, uwagi dyskusyjne

Chociaż cele pracy zostały zrealizowane, a postawiona teza została udowodniona, rozprawa posiada również swoje słabe strony. Treść rozprawy, użyty język, sposób prezentacji wyników, a także sformułowane wnioski, które cechuje ceniona zwięzłość i przejrzystość, świadczą o wysokich umiejętnościach Autora do poprawnego i przekonującego przedstawienia uzyskanych wyników.

Wśród najważniejszych uwag dyskusyjnych wymienić należy:

- a) Doktorant jest współautorem biblioteki IFIS (rozdział 6.2.) dla języka Python. Niestety nie określono stopnia udziału doktoranta w tym procesie. Jeśli biblioteka ma być wliczona w dorobek doktoranta powinniśmy znać stopień jego wkładu w jej tworzenie. Źródło biblioteki wskazuje na 9 współautorów.
- b) Brakuje mi w tekście rozprawy wyraźnej korelacji między prezentowanymi wynikami badań a dorobkiem publikacyjnym Doktoranta, co pozwoliłoby na lepszą ocenę wkładu własnego Doktoranta w rozwój danej dziedziny na tle jego dotychczasowych osiągnięć.
- c) Poważnym uchybieniem metodologicznym pracy jest brak precyzyjnego zdefiniowania aparatu pojęciowego dotyczącego niedoskonałości informacji. W pracy autor od samego początku posługuje się pojęciem niepewności wiedzy, nieostrości, nieprecyzyjności, niedoskonałości. Robi to bardziej w sposób intuicyjny zakładając, że czytający pracę będzie rozumiał dokładnie tak samo wspomniane pojęcia. W dysertacji poświęconej logice rozmytej takie rozróżnienie jest fundamentalne, gdyż każde z tych pojęć wymaga innego aparatu matematycznego (odpowiednio: zbiorów rozmytych, miar możliwości lub rachunku prawdopodobieństwa). Brak tych definicji na wstępie pracy rzutuje nieco na spójność dalszych rozważań formalnych.
- d) Wykaz publikacji Doktoranta obejmuje niemal wyłącznie prace wieloautorskie, w których Doktorant figuruje na dalszych pozycjach w liście autorów. Tylko

w jednej pracy konferencyjnej z 2020 roku, Doktorant jest pierwszym autorem. Choć formalne wymogi ustawowe zostały spełnione (dorobek publikacyjny Doktoranta w zupełności wypełnia oczekiwania), brak czy znikoma liczba publikacji z pierwszym autorstwem utrudnia jednoznaczną ocenę samodzielności naukowej recenzowanej osoby. Proszę o wyjaśnienie podczas obrony, czy i w jakiej mierze koncepcje badawcze zawarte w tych pracach były autorskimi pomysłami Doktoranta, a na ile stanowiły realizację szerszych grantów pod kierunkiem osób trzecich.

- e) Doktorant odwołuje się do raportu z 2008 roku, umieszczającego algorytm kNN w gronie 10 najlepszych algorytmów uczenia maszynowego, co wydaje się uzasadnione w kontekście ówczesnych zestawień, ale wymaga aktualizacji w świetle nowszych analiz. Współczesne rankingi z lat 2024-2025 nadal wymieniają kNN wśród kluczowych metod (np. obok Random Forest i SVM), doceniając jego prostotę i brak fazy treningu (lazy learning), lecz podkreślają dominację bardziej skalowalnych rozwiązań jak XGBoost czy modele głębokiego uczenia ze względu na złożoność obliczeniową i wrażliwość na dużą wymiarowość danych.
- f) W przypadku metody k najbliższych sąsiadów doktorant przedstawia metodę zakładając, że klasyfikacja obejmuje cechę objaśnianą na skali kategoriowej, wybierając dla obiektu klasyfikowanego decyzję większościową. Jednakże należy pamiętać, że atrybut decyzyjny może mieć skalę ilościową i wówczas proces klasyfikacji przebiega w inny sposób.
- g) Jako, że praca dotyczy dyscypliny informatyka spodziewałam się nieco większego wkładu Autora w zagadnienia związane z analizą proponowanych algorytmów w zakresie np. złożoności obliczeniowej (jako kryterium pozwalającego ocenić wagę danego algorytmu). Ocena złożoności algorytmów operujących na przedziałowych zbiorach rozmytych jest istotnym elementem prac naukowych, ponieważ stanowią one znaczne obciążenie dla procesora w porównaniu do klasycznych zbiorów typu 1.
- h) Autor przyjął arbitralnie wartości parametru $k \in \{1, 3, 5\}$ bez przedstawienia uzasadnienia teoretycznego lub empirycznego. Taki wybór jest wysoce ryzykowny z punktu widzenia wiarygodności wyników bowiem przy $k=1$ model jest ekstremalnie czuły na szum i wartości odstające (outliers). Przy dużej liczbie danych, niskie wartości k mogą nie oddawać rzeczywistej struktury rozkładu klas.

Należy wyjaśnić, czy przeprowadzone zostały testy dla wyższych wartości k i dlaczego ostatecznie przestano na tak niskim zakresie.

- i) Doktorant przywołuje macierz pomyłek dla dwóch klas decyzyjnych podczas gdy dużo bardziej uniwersalnym byłoby zdefiniowanie macierzy pomyłek dla wielu klas.
- j) Liczne błędy edycyjne/redakcyjne:
 - Wskazane jest stosowanie skróconego formatu odsyłaczy bibliograficznych w przypadku cytowań ciągłych. Zapisy typu [1,2,3...9] lepiej zastąpić formatem [1 - 9], zgodnie z powszechnie przyjętymi standardami edytorskimi (np. IEEE). Przykładowo na stronie 8 zamiast [17,18,19,20,21,22,23,24] lepiej użyć [17-24].
 - W pracy zauważono liczne odwołania do obiektów graficznych za pomocą zwrotów niedookreślonych (np. „poniższa tabela”, „powyższy rysunek”). Obecna forma jest ryzykowna ze względu na potencjalne przesunięcia obiektów w składzie tekstu i utrudnia nawigację po pracy. Strona 18: “poniższy zapis”, “poniższe definicje” przed rozdziałem 2.1, czy na stronie 103, gdzie autor odwołuje się do rysunku 6.1. stosując zapis “Powyższy rysunek” podczas gdy rysunek mieści się na stronie 102. Podobnie jest w przypadku tekstu “*Tabele prezentują*” (pod tabelami 5.5 i 5.6) na stronie 86, bez odwołania do konkretnych tabel poprzez ich numery co sprawia, że nie wiemy do końca których tabel ostatecznie komentarz dotyczy.
 - W pracy występują obiekty graficzne (np. Rys. 6.2, strona 107), do których autor w ogóle nie odwołuje się w tekście głównym. Pozostawienie „niezakotwiczonych” w tekście ilustracji obniża spójność wyводу i sprawia wrażenie braku należytej staranności w redakcji ostatecznej wersji dysertacji. Podobnie jest z dla rysunków, np. 4.1, 4.3, 4.4 i 4.6. To samo dotyczy np. tabel 5.7. i 5.8.
 - Struktura pracy budzi zastrzeżenia pod względem kompozycyjnym. Doktorant stosuje bezpośrednio sąsiedztwo nagłówków rozdziałów i podrozdziałów. Brakuje tekstów wprowadzających na początku głównych rozdziałów, które nakreśliłyby ich cel i zapowiedziały strukturę wewnętrzną, co negatywnie wpływa na płynność wyводу naukowego. Na stronie 19 brak treści między 2.2. a 2.2.1, zaś na stronie 57 brakuje określenia co będzie w rozdziale czwartym podczas gdy dla poprzednich i kolejnych rozdziałów zasada ta była zachowana. To samo dotyczy sekcji 5.4 i 5.4.1. na stronie 92.
 - W sferze redakcyjnej przedłożonej rozprawy należy zwrócić uwagę na systematyczny błąd w formatowaniu obiektów graficznych. Doktorant umieszcza

tytuły tabel poniżej ich treści, co jest niezgodne z powszechnie przyjętymi standardami edytorskimi prac naukowych oraz zasadami składu tekstu. O ile podpisy rysunków i schematów słusznie znajdują się pod nimi, o tyle nagłówki tabel umieszcza się nad tabelą, aby zapewnić czytelnikowi właściwy kontekst przed analizą danych liczbowych. Przykład: strona 77, Tabela 5.1.

- W tabelach prezentujących wyniki (np. Tabela 6.2) autor stosuje kropkę jako separator dziesiętny. W rozprawie pisanej w języku polskim standardem jest użycie przecinka. W tym samym względzie zauważono niespójność, gdyż na stronie 73 widzimy zarówno wartość 0.5 jak i 0,5.
- Warto zadbać o większą dyscyplinę w posługiwaniu się skrótami i nazewnictwem metod. Doktorant wielokrotnie zmienia sposób określania metody k-najbliższych sąsiadów. Wskazane byłoby zdefiniowanie skrótu k-NN przy pierwszej wzmiance i trzymanie się tej nomenklatury w dalszych częściach pracy. W pracy na stronie 10 zauważamy 4 różne sposoby odwołań do metody tj. “k-najbliższych sąsiadów kNN”, “Metody k-NN”, “Metody k-najbliższych sąsiadów” oraz “Algorytm k-NN”.
- Brak spójności językowej w warstwie ilustracyjnej pracy. Część rysunków posiada opisy w języku polskim, podczas gdy inne zachowano w języku angielskim. Przykładowo rysunki 6.3 i 6.4 (strona 110) są przedstawione w języku angielskim podczas gdy inne w języku polskim i sama rozprawa jest napisana w języku polskim. Rozprawa doktorska pisana w języku polskim powinna mieć wszystkie elementy składowe (w tym legendy, opisy osi wykresów i teksty na schematach) w języku polskim.
- W tekście dysertacji występuje niespójność w zakresie formy narracji. Autor zamiennie stosuje formy bezosobowe, pierwszą osobę liczby pojedynczej oraz pierwszą osobę liczby mnogiej (nierzadko w obrębie jednego podrozdziału). Takie rozwiązanie jest niezgodne z zasadami redagowania tekstów naukowych, które wymagają zachowania jednolitego stylu w całej pracy. Zmienność ta wprowadza niejasności co do autorstwa poszczególnych etapów badań oraz sprawia wrażenie braku spójności redakcyjnej tekstu. Styl w pierwszej osobie l.p. na stronie 36 “*będę posługiwał się*” a na stronie 37 “*nazywamy, oznaczamy, zwracamy...*”.
- W pracy Doktorant posługuje się terminem “fazyfikacja” oraz “defazyfikacja”. Są to w mojej ocenie niefortunne kalki z języka angielskiego. W polskiej literaturze

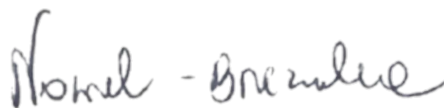
przedmiotu (prace prof. Kacprzyka) przyjęta jest terminologia oparta na pojęciach “rozmywanie” oraz “wyostrzanie”.

- Brak konsekwencji przy definiowaniu pojęć. Wzór 4.6 na stronie 67 określa czułość (jest zdefiniowany angielską nazwą sensitivity), i choć odpowiada skrótowi TPR z rysunku 4.6, nie posiada go, zaś drugi czynnik z rysunku 4.6., FPR, został użyty do zdefiniowania swoistości, i choć posiada skrót to nie posiada numeru przy wzorze. Ten brak konsekwencji jest zauważalny w wielu miejscach w tekście rozprawy. W tabelach 5.1 - 5.8 wyłączając tabelę 5.7. stosuje się pojęcie “precision” podczas gdy w tabeli 5.9. pojęcie “PREC”. Jeśli w obu przypadkach chodzi o tę samą miarę Doktorant powinien stosować jedną nazwę.
- Brak konsekwencji przy skrótach, np. na stronie 25 wyjaśniono skrót IVOWA podając angielski odpowiednik, a wcześniej przy akronimie OVA nie uwzględniono wyjaśnienia.
- Brak odwołań do literatury gdy wspomina się np. o systemach sterowania metrem, strona 17 czy w zdaniu “*W 1983 roku F. Allen zaproponował...*” na stronie 21.
- Angielskie odpowiedniki przywoływanych pojęć z reguły są zapisane z użyciem kursywy podczas gdy przy objaśnieniu “*ang. necessary similarity measure*”, na stronie 48, nie zastosowano takiego stylu formatowania.
- Brak konsekwencji przy numerowaniu metryk. Na stronie 58: “*poniżej zamieszczone przykłady metryk dla danych numerycznych*” lecz tylko jedna metryka jest numerowana pozostałe już nie. Na stronie 64 nie mamy numerowania dla metryk $P(C_i|X)$ oraz C .
- Błędy literowe/stylistyczne:
 - brak znaku kropki “.” przed nowym zdaniem “*W części aplikacyjnej...*” w streszczeniu pracy,
 - zdanie bez orzeczenia – tj. “*Wyniki przeprowadzonych doświadczeń skuteczność opracowanych metod.*”, ostatnie zdanie przed rozdziałem 1.2 na stronie 9,
 - “*najefektywniejsze rezultaty...*” -> “*najbardziej efektywne*” na stronie 12,
 - “*Rozprawa składa się z sześciu rozdziałów*” a tak naprawdę składa się z 7 rozdziałów - strona 14, rozdział 1.3.,
 - “*małe zarobki*” -> “*niskie zarobki*”, strona 17,
 - błędy literowe typu “*Zadech*” zamiast “*Zadeh*”, strona 36, podrozdział 2.2.4, czy “*operatorem W*” zamiast “*operator W*”, w ostatnim akapicie strony 37
 - niefortunna kombinacja słów, np. na stronie 59: “*prosty proces użycia*” zamiast “*prostotę użycia*”,
 - “*...składa się z jednego kryterium opartym na atrybucie dochód*” powinno być “*opartego*”, strona 61,

- odnośnie drzew decyzyjnych w zdaniu “...w którym z każdego z atrybutów wychodzą najwyżej dwie krawędzie...” -> zamiast atrybutów powinno być węzłów, strona 62,
- “zbór” zamiast “zbiór” na stronie 75,
- brak nawiasu zamykającego przy pierwszym akapicie “parametru k (rozważano $k \in \{1,3,5\}$ ” na stronie 76,
- “Model 2. Model oparty o miary możliwości!”, na stronie 76 powinno być “możliwości” i bez znaku !,
- niefortunna kombinacja słów w pod tabelą 5.1. “wyników uzyskanych w wyniku”,
- w Algorytmie 2 mamy nieprecyzyjne określenie “jakąś dużą wartość”, na stronie 82,
- należy zweryfikować stylizację zdania w podrozdziale 5.3.1. “wybór algorytmu Relief jako bazy nowo tworzonego algorytmu służącego do selekcji cech reprezentujących dane z niepewnością był podyktowany jego wyjaśnialnością i zastosowania miar uwzględniających niepewność”,
- niepotrzebne powtórzenie frazy “za pomocą algorytmu IV-kNN” na stronie 88,
- konieczne przeredagowanie treści akapitów dotyczących liczności próby testu senDD dla 750 studentów gdyż ta sama treść ma trzy powtórzenia, strona 94,
- “uzyskane rezultaty pokazały, że stosując operatory i miary możliwości poprawiają efektywność diagnostyki depresji” powinno być “że zastosowanie operatorów i miar możliwości poprawia efektywność...”, strona 98.

6. Wniosek końcowy

W podsumowaniu stwierdzam, że rozprawa doktorska mgr. inż. Dawida Kosiora, mimo pewnych uwag krytycznych, spełnia wymagania stawiane kandydatom do stopnia naukowego doktora, określone w art. 187 ust. 1 i 2 Ustawy z dnia 20 lipca 2018r. Prawo o szkolnictwie wyższym i nauce (Dz.U. z 2020r. poz. 85, z późn. zm.). W mojej ocenie zawiera oryginalne rozwiązanie problemu naukowego. Doktorant osiągnął postawione cele, wykazując się niezbędną wiedzą teoretyczną w dyscyplinie „Informatyka techniczna i telekomunikacja” i umiejętnościami w zakresie prowadzenia badań naukowych z wykorzystaniem metod informatycznych. Wnoszę zatem o dopuszczenie jej do publicznej obrony.



dr hab. Agnieszka Nowak – Brzezińska, Prof. UŚ

Uniwersytet Śląski, Wydział Nauk Ścisłych i Technicznych, Instytut Informatyki