



Politechnika Wrocławska

Katedra Sztucznej Inteligencji

Wrocław, 11 stycznia 2026 r.

ZNAK SPRAWY/ / 2025

Pan
dr hab. inż. Jacek Rak, prof. PG
Przewodniczący Rady Dyscypliny
Naukowej ITIT PG
Politechnika Gdańska
Ul. G. Narutowicza 11/12
80-233 Gdańsk

Szanowny Panie Profesorze,

W załączeniu przesyłam wykonaną recenzję rozprawy doktorskiej zatytułowanej
"Application of artificial intelligence algorithms for analysis of pure tone audiometry",
której autorem jest mgr Michał Kassjański.

Z wyrazami szacunku,

Prof. dr hab.inż. Urszula Markowska-Kaczmar

Załączniki

1. Recenzja rozprawy
2. Podpisana umowa o dzieło (2 egzemplarze)
3. Protokół zdawczo odbiorczy (2 egzemplarze)
4. Rachunek (2 egzemplarze)



HR EXCELLENCE IN RESEARCH

Politechnika Wrocławska

Wydział Informatyki
i Telekomunikacji

Katedra Sztucznej Inteligencji

Wybrzeże Wyspiańskiego 27
50-370 Wrocław

T: +48 71 320 24 54

www.pwr.edu.pl
ai.pwr.edu.pl
sekretariat.k46.wit@pwr.edu.pl

REGON: 00001614
NIP: 896-000-58-51

Nr konta:
37 1090 2402 0000 0006 1000 0434



RPW/1443/2026 N
Data: 2026-01-15



RECENZJA

pracy doktorskiej magistra Michała Kassjańskiego zatytułowanej
„*Application of artificial intelligence algorithms for analysis of pure tone audiometry*”

WSTĘP

Niniejsza recenzja została wykonana w odpowiedzi na prośbę wyrażoną w piśmie z dnia 5.11.2025, skierowaną przez Przewodniczącego Rady Dyscypliny Naukowej ITIT PG dr hab. inż. Jacka Raka, profesora PG. Przy jej realizacji kierowano się przepisami dotyczącymi recenzji określonymi w nowej Ustawie Prawo o szkolnictwie wyższym i nauce, regulującej proces awansu naukowego, a także w przepisami wykonawczymi i uchwałami Rady Doskonałości Naukowej.

Recenzowana rozprawa składa się z pięciu publikacji opatrzonych dość szczegółowym przewodnikiem. W skład cyklu publikacji wchodzi następujące prace:

1. Kassjański, M., Kulawiak, M., & Przewoźny, T. (2022). *Development of an AI-based audiogram classification method for patient referral*. W M. Ganzha, L. Maciaszek, M. Paprzycki & D. Ślęzak (Red.), *Proceedings of the 17th Conference on Computer Science and Intelligence Systems (FedCSIS 2022)* (ss. 163–168). Annals of Computer Science and Information Systems, Vol. 30. <https://doi.org/10.15439/2022F66> (5 cytowań)
2. Kassjański, M., Kulawiak, M., Przewoźny, T., Tretiakow, D., Kuryłowicz, J., Molisz, A., Koźmiński, K., Kwaśniewska, A., Mierzwińska-Dolny, P., & Grono, M. (2023). *Detecting type of hearing loss with different AI classification methods: A performance review*. In M. Ganzha, L. Maciaszek, M. Paprzycki, & D. Ślęzak (Eds.), *Proceedings of the 18th Conference on Computer Science and Intelligence Systems (ACSIS)* (Vol. 35, pp. 1017–1022). Annals of Computer Science and Information Systems. <https://doi.org/10.15439/2023F3083> (2 cytowania)
3. Kassjański, M., Kulawiak, M., Przewoźny, T., Tretiakow, D., Kuryłowicz, J., Molisz, A., Koźmiński, K., Kwaśniewska, A., Mierzwińska-Dolny, P., & Grono, M. (2024). *Efficiency of artificial intelligence methods for hearing loss type classification: An evaluation*. Journal of Automation, Mobile Robotics and Intelligent Systems, 18(3), Article 19. <https://doi.org/10.14313/JAMRIS/3-2024/19> (3 cytowania)

4. Kassjański, M., Kulawiak, M., Przewoźny, T., Tretiakow, D., Kuryłowicz, J., Molisz, A., Koźmiński, K., Kwaśniewska, A., Mierzwińska-Dolny, P., & Grono, M. (2024). *Automated hearing loss type classification based on pure tone audiometry data*. Scientific Reports, 14, 14203. <https://doi.org/10.1038/s41598-024-64310-2> Nature (6 cytowań)

5. Kassjański, M., Kulawiak, M., Przewoźny, T., Tretiakow, D., & Molisz, A. (2025). *Development and testing of an open source mobile application for audiometry test result analysis and diagnosis support*. Scientific Reports, 15, 14302. <https://doi.org/10.1038/s41598-025-99338-5>

II PROBLEM NAUKOWY ROZWIĄZywANY W ROZPRAWIE. OCENA WIEDZY TEORETYCZNEJ DOKTORANTA W KONTEKŚCIE NADANIA STOPNIA DOKTORA W DYSCYPLINIE INFORMATYKA TECHNICZNA I TELEKOMUNIKACJA

Rozprawa doktorska dotyczy problemu klasyfikacji rodzaju utraty słuchu na podstawie czystych danych audiometrycznych i została przedstawiona w formie spójnego cyklu publikacji. Poszczególne artykuły w sposób logiczny i stopniowy wprowadzają czytelnika w omawianą tematykę, a ich układ pozwala jednocześnie prześledzić rozwój zainteresowań naukowych Doktoranta oraz ewolucję podejmowanych przez niego problemów badawczych.

Cykl publikacji otwiera praca pt. „*Development of an AI-based audiogram classification method for patient referral*”, w której Autorzy porównują różne architektury sieci neuronowych, w tym wielowarstwową sieć perceptronową (MLP), sieć konwolucyjną (CNN) oraz modele rekurencyjne, analizując ich skuteczność na jednym zbiorze rzeczywistych danych klinicznych. Zaproponowane podejście ma charakter aplikacyjny i ilustruje praktyczne wykorzystanie metod uczenia maszynowego w analizie danych medycznych.

Druga praca w cyklu, „*Detecting type of hearing loss with different AI classification methods: a performance review*”, stanowi przegląd porównawczy metod sztucznej inteligencji stosowanych do klasyfikacji typu ubytku słuchu na podstawie danych z audiometrii tonalnej. Autorzy analizują skuteczność różnych algorytmów uczenia maszynowego i głębokiego, wskazując ich mocne i słabe strony w kontekście dokładności oraz stabilności klasyfikacji. Publikacja ta dostarcza uporządkowanego punktu odniesienia dla dalszych badań nad automatyczną analizą wyników badań słuchu.

Trzecia praca pt. „*Efficiency of artificial intelligence methods for hearing loss type classification: an evaluation*”, opublikowana w *Journal of Automation, Mobile Robotics and Intelligent Systems*, obejmuje porównanie szerokiego spektrum klasyfikatorów uczenia maszynowego — od modeli klasycznych, takich jak SVM i lasy losowe (*Random Forest*), po modele głębokie, w tym w pełni połączone sieci neuronowe (FNN), sieci konwolucyjne (CNN), klasyczne sieci rekurencyjne (RNN) oraz sieci z komórkami pamięci (LSTM i GRU) — w zadaniu rozpoznawania typu niedosłuchu. Klasyfikacja przeprowadzona została na podstawie 4007 audiogramów opisanych przez audiologów. Autorzy wykazują, że najlepsze wyniki uzyskuje sieć LSTM (dokładność rzędu 97,6%), a także analizują wpływ augmentacji danych z wykorzystaniem warunkowej sieci GAN oraz różnych procedur standaryzacji na skuteczność modeli. W pracy tej po raz pierwszy potraktowano audiogram jako sekwencję, w której kolejne elementy odpowiadają poszczególnym częstotliwościom.

Kontynuacją i rozwinięciem powyższych badań jest czwarty artykuł pt. „*Automated hearing loss type classification based on pure tone audiometry data*”, opublikowany w czasopiśmie *Scientific Reports*. Autorzy pokazują w nim, że wykorzystując wyłącznie wyniki audiometrii tonalnej można skutecznie klasyfikować typ niedosłuchu z zastosowaniem dwukierunkowej sieci rekurencyjnej BiLSTM, osiągając dokładność przekraczającą 99% na danych spoza zbioru treningowego. Należy jednak zaznaczyć, że badania te, podobnie jak wcześniejsze, przeprowadzono na jednym zbiorze rzeczywistych danych pochodzących z Kliniki Otolaryngologii Uniwersyteckiego Centrum Klinicznego w Gdańsku.

Ostatnia praca wchodząca w skład cyklu, opublikowana również w *Scientific Reports* i zatytułowana „*Development and testing of an open source mobile application for audiometry test result analysis and diagnosis support*”, opisuje rozwiązanie mające na celu wsparcie lekarzy w automatycznej klasyfikacji audiogramów. Dane wejściowe do aplikacji mobilnej stanowią obrazy audiogramów wprowadzane za pośrednictwem smartfona, co pozwala na realizację kompletnej ścieżki przetwarzania bezpośrednio na urządzeniu mobilnym.

Ocena wartości naukowej rozprawy

Rozprawa doktorska przedstawia szeroką i ugruntowaną wiedzę teoretyczną mgr. Michała Kassjańskiego, ubiegającego się o nadanie stopnia doktora w dyscyplinie informatyka techniczna i telekomunikacja. Autor wykazuje dobrą znajomość podstaw teoretycznych metod uczenia maszynowego i głębokiego, a także zagadnień z zakresu analizy danych, przetwarzania obrazów, modelowania systemów informatycznych oraz metod oceny i walidacji algorytmów. Istotnym elementem pracy jest umiejętne zastosowanie metod uczenia maszynowego i głębokiego w obszarze zastosowań medycznych, w szczególności w klasyfikacji niedosłuchu na podstawie danych audiometrycznych.

Na uwagę zasługuje nowatorskie podejście polegające na wykorzystaniu sieci rekurencyjnych do klasyfikacji audiogramów poprzez ich interpretację jako sekwencji, a także zastosowanie warunkowych sieci GAN do augmentacji wzorców danych. W kontekście niedoboru specjalistów medycznych, każde rozwiązanie zmierzające w kierunku przyspieszenia procesu diagnostycznego można postrzegać jako potencjalną korzyść dla pacjentów. Należy jednak mieć świadomość, że droga od wyników badań naukowych do rzeczywistego wdrożenia klinicznego jest długa i obciążona wieloma wymaganiami.

Z tego względu trudno w pełni zgodzić się z tezą postawioną przez Autora w hipotezie H1, zgodnie z którą „*the application of modern neural network architecture to classification of hearing loss types ... can ... deliver performance and accuracy viable for introduction in clinical practice*”. Sformułowanie to sugeruje bowiem, że zaproponowana metoda jest już na etapie wykonalności i gotowości do zastosowania w rutynowej praktyce klinicznej. Tymczasem w medycynie pojęcie „viable” odnosi się do rozwiązań, które przeszły odpowiednie badania kliniczne i wykazały potencjał poprawy wyników leczenia bez nadmiernych ryzyk czy kosztów. Jest to etap poprzedzający pełną adopcję, uwzględniający m.in. kwestie dostępności, zgodności z regulacjami prawnymi oraz akceptacji środowiska lekarskiego.

Ocena publikacji i ich znaczenia

Oceniając cykl publikacji wchodzący w skład ry Doktorskiej pod względem przypisanych im punktów ministerialnych, należy uznać dorobek Doktoranta za wysoko punktowany. Trzy prace zostały opublikowane w czasopismach naukowych. Jedna z nich ukazała się w czasopiśmie *Journal of Automation, Mobile Robotics and Intelligent Systems*, które nie posiada klasycznego współczynnika Impact Factor z listy JCR, jednak jest indeksowane w bazie Scopus i w obowiązującym wykazie czasopism otrzymało 70 punktów MEiN. Dwie kolejne prace zostały opublikowane w czasopiśmie *Scientific Reports*, które w 2024 roku posiadało współczynnik Impact Factor na poziomie 3,9 oraz 140 punktów MEiN.

Pozostałe dwie prace stanowią publikacje konferencyjne przedstawione na konferencji FedCSIS, która w latach ich publikacji miała przypisane 70 punktów MEiN.

Wady i słabe strony

W odniesieniu do omawianych prac składających się na rozprawę doktorską wnoszę jednak kilka zastrzeżeń.

1. Badania opisane w publikacjach stanowiących podstawę rozprawy doktorskiej opierają się na jednym, prywatnym zbiorze danych. Tymczasem zastosowane metody uczenia maszynowego mają charakter indukcyjny, co implikuje konieczność weryfikacji zdolności generalizacji na niezależnych zbiorach danych. Na obecnym etapie nie widzę możliwości bezpośredniego wdrożenia proponowanych rozwiązań do praktyki klinicznej, mimo pojawiających się w pracy sugestii o takim potencjale.
2. W nawiązaniu do uwagi zawartej w punkcie 1., sformułowanie hipotezy H1 należy uznać za niefortunne oraz zbyt ogólne, aby mogło zostać jednoznacznie zweryfikowane w ramach ocenianej rozprawy doktorskiej. Z całą pewnością można stwierdzić jedynie skuteczność zaproponowanej metody w odniesieniu do konkretnego zbioru danych, na którym przeprowadzono badania. W kontekście badań medycznych sformułowanie „*viable for introduction in clinical practice*” sugeruje możliwość integracji algorytmu z codzienną praktyką diagnostyczną, przy uwzględnieniu takich czynników jak skuteczność kliniczna, koszty, łatwość użycia oraz brak barier regulacyjnych. Pierwszym krokiem w kierunku weryfikacji takiej tezy powinno być przeprowadzenie badań na dodatkowych, niezależnych zbiorach danych. W przeciwnym razie jedynym wnioskiem, jaki można sformułować, pozostaje fakt osiągnięcia określonych wyników dla pojedynczego zbioru danych, którego charakterystyka jest ściśle związana z badaną populacją, użytym sprzętem oraz zastosowaną procedurą pomiarową.

Na usprawiedliwienie Doktoranta należy jednak dodać, że obecnie brak jest ogólnodostępnych, benchmarkowych zbiorów danych w tej dziedzinie. Również Autorce niniejszej recenzji nie udało się takich zasobów zidentyfikować. Niemniej jednak w poniższych publikacjach ich autorzy deklarują możliwość udostępnienia prywatnych zbiorów danych po złożeniu stosownego wniosku:

- Zhiyong Dou, Yingqiang Li, Dongzhou Deng, Yunxue Zhang, Anran Pang, Cong Fang, Xiang Bai, Dan Bing: "Pure tone audiogram classification using deep

AMU

learning techniques": Published in *Clinical Otolaryngology* in May 2024. <https://doi.org/10.1111/coa.14170> Badania przeprowadzono na 12 518 audiogramach od 6259 pacjentów.

- Lee, S.Y., Seo, H.W., Jung, S.M. *et al.* Assessing the accuracy and reliability of application-based audiometry for hearing evaluation. *Sci Rep* **14**, 7359 (2024). <https://doi.org/10.1038/s41598-024-57944-9>. Dane od 70 pacjentów. Zbudowali aplikację do oceny słuchu
- M. Cox and B. de Vries, "Bayesian Pure-Tone Audiometry Through Active Learning Under Informed Priors," *Frontiers in Digital Health*, vol. 3, Art. no. 723348, Aug. 2021, doi: 10.3389/fdgth.2021.723348. Dane od 88237 pacjentów z krajów nordyckich. Zbiór zawiera 176474 audiogramów dla 11 częstotliwości.

Zasadne więc byłoby zwiększenie liczby zbiorów danych, na których przeprowadzono eksperymenty, co pozwoliłoby na rzetelniejszą ocenę zdolności generalizacji proponowanych metod.

3. Nie podzielam optymizmu Autorów w odniesieniu do raportowanej wartości dokładności (99,33%) uzyskanej dla modelu BiLSTM. Choć testy przeprowadzono „outside of training” z wykorzystaniem metody stratyfikowanej walidacji krzyżowej (*Stratified K-Fold Cross-Validation, SKCV*), to dane te wciąż pochodzą z tego samego środowiska pomiarowego oraz tej samej populacji. W konsekwencji model może w pewnym stopniu wykorzystywać artefakty specyficzne dla danego ośrodka. Cała ocena jakości opiera się na SKCV przeprowadzonej na jednym zbiorze danych, przy braku zewnętrznego zbioru testowego pochodzącego z innego ośrodka, co sprawia, że kwestia generalizacji poza badaną populację pozostaje otwarta.
4. W żadnej z prac składających się na cykl publikacji stanowiących rozprawę doktorską nie przedstawiono opisu wydzielenia niezależnego zbioru walidacyjnego (*validation set*), wykorzystywanego np. do doboru hiperparametrów w sposób niezależny od walidacji krzyżowej. Ponadto w publikacjach nie podano szczegółowych wartości hiperparametrów (takich jak na przykład wartość współczynnika uczenia i metoda jego optymalizacji, metoda optymalizacji parametrów, rozmiar batcha itp.), co znacząco utrudnia odtworzenie przeprowadzonych badań. Podsumowując, nie jest jasne, czy i w jaki sposób hiperparametry były dostrajane oraz na jakim zbiorze danych. W celu uniknięcia potencjalnego przecieku informacji zasadne byłoby zastosowanie procedury *nested cross-validation* (w tym przypadku również w wariancie stratyfikowanym).
5. Z opisu (wzór 2 str. 20) wynika, że parametr w funkcji straty modelu L_{RNN} odpowiada długości sekwencji, tj. liczbie analizowanych częstotliwości w audiogramie, a każde rozwinięcie sieci rekurencyjnej odpowiada pojedynczemu progowi częstotliwości. Ponieważ etykieta przypisana jest do całego audiogramu, naturalnym rozwiązaniem wydaje się obliczanie funkcji straty wyłącznie dla ostatniego kroku sekwencji. W związku z tym nie jest jasne, dlaczego w zaproponowanej funkcji straty L_{RNN} wartości są sumowane po wszystkich krokach.
6. W rozprawie stosowane jest sformułowanie „SGD classifier”. Należy jednak zauważyć, że SGD jest metodą optymalizacji parametrów, a nie klasą modeli. Choć biblioteka *scikit-learn* wprowadza pojęcie *SGDClassifier*, w istocie odnosi się ono do liniowego klasyfikatora z funkcją progową, uczonego z wykorzystaniem stochastycznego spadku gradientu. Jest to pewien skrót terminologiczny przyjęty przez autorów biblioteki, który nie jest powszechnie spotykany w literaturze naukowej.

7. Przegląd literatury zawarty w przewodniku do publikacji, powinien zostać uzupełniony o najnowsze prace z omawianego obszaru, które z uwagi na czas publikacji nie mogły zostać uwzględnione w dołączonych artykułach.

III OCENA ORYGINALNOŚCI ROZWIĄZANIA PROBLEMU NAUKOWEGO

Rozwiązanie zaproponowane w rozprawie doktorskiej zawiera szereg elementów o charakterze oryginalnym. Do najważniejszych należy traktowanie audiogramu jako sekwencji oraz zastosowanie sieci rekurencyjnych do klasyfikacji audiogramów (klasyczne RNN, GRU, LSTM oraz BiLSTM). W odróżnieniu od wcześniejszych podejść, które opierały się głównie na klasycznych metodach uczenia maszynowego oraz sieciach konwolucyjnych, Autor proponuje spójne ujęcie problemu w modelu sekwencyjnym. Nowatorskim elementem pracy jest również zastosowanie warunkowych sieci generatywnych (CGAN) do augmentacji zbioru audiogramów.

Istotnym aspektem nowatorstwa pracy pt. *„Development and testing of an open source mobile application for audiometry test result analysis and diagnosis support”* jest zaproponowanie otwartoźródłowej aplikacji mobilnej integrującej automatyczną digitalizację audiogramów oraz ich analizę z wykorzystaniem algorytmów sztucznej inteligencji. Praca jako jedna z pierwszych prezentuje kompletną ścieżkę przetwarzania, realizowaną bezpośrednio na urządzeniu mobilnym. Zaproponowane rozwiązanie zostało zoptymalizowane pod kątem wydajności obliczeniowej oraz potencjalnych zastosowań klinicznych i zweryfikowane w realistycznym scenariuszu użycia.

Uzyskane wyniki posiadają potencjał wdrożeniowy i mogą znaleźć praktyczne zastosowanie po przeprowadzeniu dodatkowych badań na danych pochodzących z różnych ośrodków, pozyskanych w odmiennych warunkach pomiarowych i z wykorzystaniem różnych urządzeń, a także po spełnieniu obowiązujących wymagań regulacyjnych. Nadaje to pracy istotny wymiar społeczny, związany z możliwością wsparcia procesów diagnostycznych w ochronie zdrowia.

Praca wnosi istotny wkład w rozwój dyscypliny informatyka techniczna i telekomunikacja, w szczególności w obszarze zastosowań metod sztucznej inteligencji w analizie danych medycznych. Autor nie tylko rozwija i weryfikuje metody uczenia maszynowego w kontekście rzeczywistego problemu diagnostycznego, lecz także dostarcza nowych wyników empirycznych oraz demonstrowa możliwości praktycznego wykorzystania systemów informatycznych w ochronie zdrowia.

IV OCENA UMIEJĘTNOŚCI SAMODZIELNEGO PROWADZENIA PRACY NAUKOWEJ

Jak już wspomniano, rozprawa doktorska ma formę cyklu pięciu publikacji współautorskich. W każdej z nich deklarowany udział Doktoranta oszacowany został na poziomie 70%, co pozwala sformułować jednoznaczny wniosek, że przedstawiona rozprawa potwierdza zdolność Doktoranta do samodzielnego prowadzenia badań naukowych. Doktorant samodzielnie formułuje cele i problemy badawcze, dobiera oraz uzasadnia zastosowanie

metod sztucznej inteligencji adekwatnych do analizowanych danych medycznych, a następnie przeprowadza ich weryfikację eksperymentalną.

Przedstawiona analiza wyników oraz sformułowane wnioski świadczą o dojrzałości badawczej Autora, a także o jego umiejętności krytycznej oceny uzyskanych rezultatów.

V. KONKLUZJA KOŃCOWA

Na podstawie przeprowadzonej oceny stwierdzam, że przedstawiona rozprawa doktorska zadowalająco spełnia wymagania stawiane pracom doktorskim, w szczególności w zakresie oryginalności, istotnego wkładu w rozwój dyscypliny oraz samodzielności prowadzonych badań. Wnoszę zatem o dopuszczenie rozprawy do publicznej obrony oraz o podjęcie dalszych czynności w postępowaniu o nadanie Autorowi stopnia doktora w dyscyplinie informatyka techniczna i telekomunikacja.

U. Pawłowski - kaxmow