



dr hab. inż. Michał Grochowski, prof. uczelni
Politechnika Gdańska,
Wydział Elektrotechniki i Automatyki
Katedra Inteligentnych Systemów Sterowania
i Wspomagania Decyzji

Gdańsk, 18.10.2024

**Recenzja rozprawy doktorskiej mgr inż. Weroniki Hryniewskiej-Guzik pt.
„A multi-level perspective on the deep learning models and human-oriented
explanations with applications to medical images”**

dla Rady Naukowej Dyscypliny Informatyka techniczna i telekomunikacja
Politechniki Warszawskiej, przygotowana na zlecenie prof. dra hab. inż. Jarosława
Arabasa, przewodniczącego rady dyscypliny.

Promotor rozprawy: **prof. dr hab. inż. Przemysław Biecek**
Dyscyplina: **Informatyka techniczna i telekomunikacja**

Niniejsza recenzja oparta jest o otrzymaną rozprawę doktorską, artykuły wymienione
w rozprawie, wkład Doktorantki w powstanie artykułów wieloautorskich oraz CV Doktorantki.

1. Problem badawczy analizowany w pracy

Recenzowana rozprawa doktorska Pani Weroniki Hryniewskiej-Guzik, zatytułowana:
*„A multi-level perspective on the deep learning models and human-oriented explanations
with applications to medical images”*, lokuje się merytorycznie w dziedzinie sztucznej
inteligencji, w szczególności wyjaśnialnej i wiarygodnej sztucznej inteligencji. Jakkolwiek
zaproponowane i opisywane w pracy podejścia i metody mają charakter generyczny,
przedstawione wyniki badań koncentrują się na aplikacjach medycznych, a dokładniej
diagnostyki na podstawie danych z obrazowania medycznego.

W pracy przyjęto tezę naukową: *„Multifaceted models and human-oriented explanations
have a positive impact on quality metrics”*, będącą podstawą do sformułowania
korespondujących pytań badawczych i zaplanowania odpowiednich badań.

Autorka sformułowała cztery pytania badawcze, tj.:

- Jakże są słabe strony wyjaśnień i modeli, które klasyfikują pacjentów z COVID-19?
- Jak łączyć modele, aby uzyskać lepszą skuteczność predykcyjną?
- Jak sprawić, by wyjaśnienia były bardziej przyjazne dla człowieka i zgodne z wiedzą dziedzinową?
- Jak połączyć wyjaśnienia, aby uzyskać lepszą wiarygodność i lokalizację?

Aby odpowiedzieć na postawione pytania badawcze, Doktorantka zaplanowała następujące
cele badawcze:

- Opracowanie standardów dla modeli klasyfikacji obrazów z COVID-19;
- Rozwój wielozadaniowego modelu głębokiego uczenia;
- Badanie skuteczności polireprezentacji;
- Zwiększenie objaśnialności modelu przy użyciu podejścia „human-in-the-loop”;
- Porównanie metod łączenia wyjaśnień;
- Ocena ilościowa przy użyciu łączenia wyjaśnień.



Zarówno pytania badawcze, jak i zadania badawcze zostały konkretnie powiązane z sześcioma artykułami naukowymi, które stanowią podstawę merytoryczną recenzowanej pracy doktorskiej.

W mojej ocenie, problem badawczy oraz towarzyszące mu pytania badawcze zostały dobrze zdefiniowane. Zaplanowane cele badawcze i przeprowadzone eksperymenty, są zgodne z podjętą tematyką i dobrze dobrane. Tematyka pracy, jaką jest opracowanie skutecznych i wiarygodnych systemów diagnostyki, wykorzystujących algorytmy głębokiego uczenia, jest aktualna i niezwykle istotna zarówno z naukowego, jak i praktycznego punktu widzenia. Zakres badań jest dobrze zdefiniowany, realistyczny i dobrze dobrany w stosunku do okresu czasu, w którym przeciętnie realizowana jest praca doktorska.

2. Znaczenie naukowe

Recenzowana rozprawa doktorska wnosi istotny wkład w dyscyplinę Informatyka techniczna i telekomunikacja, w dziedzinie sztucznej inteligencji, w obszarze głębokich sieci neuronowych, w szczególności wyjaśnialnej sztucznej inteligencji (XAI).

Praca ta ma duże znaczenie praktyczne, ponieważ odpowiada na realne problemy związane z odpowiedzialnym wykorzystaniem modeli i algorytmów sztucznej inteligencji w medycynie, gdzie skuteczność i precyzja systemów wspomaganie decyzji i zaufanie do nich są kluczowe dla absorpcji przez lekarzy tego typu systemów jak i dla zdrowia i życia pacjentów.

Dokonania pracy można analizować w następujących obszarach:

Opracowanie listy kontrolnej dla odpowiedzialnego stosowania AI w medycynie

Pierwszym, bardzo ważnym wkładem tej pracy w rozwój dziedziny jest dogłębna analiza 25 najbardziej znaczących artykułów dotyczących diagnostyki COVID-19. Do ich analizy opracowano listę kontrolną, która pozwoliła na ocenę stanu i próbę odnalezienia standardu tworzenia odpowiedzialnych rozwiązań sztucznej inteligencji w medycynie. Jakkolwiek analiza prac nie dostarcza pozytywnych wniosków co do istniejących podejść, opracowana lista jest cenna i może być dobrą podstawą dla dalszych badań nad etyką i odpowiedzialnością AI, gdzie bezpieczeństwo pacjentów jest priorytetem. Co więcej, praktyczne wskazówki i standardy wypracowane w ramach podjętych analiz mogą mieć szerokie zastosowanie w różnych obszarach nauki i przemysłu, gdzie wykorzystywana jest sztuczna inteligencja.

Rozwój algorytmów sztucznej inteligencji w medycynie

Praca przyczynia się do poprawy efektywności systemów opartych na modelach głębokiego uczenia, które są wykorzystywane w diagnostyce obrazowej, szczególnie w kontekście diagnostyki COVID-19. W czasach pandemii COVID-19, szybka i precyzyjna diagnoza była kluczowa, a modele sztucznej inteligencji miały za zadanie wspierać radiologów w analizie obrazów płuc. W ramach pracy rozwinięto nowatorskie modele wielozadaniowe, które są zdolne do jednoczesnego wykonywania zadań: rekonstrukcji, klasyfikacji, segmentacji i detekcji. Jest to niewątpliwie nowatorskie podejście w aspekcie diagnostyki, które jest również rozwiązaniem z generycznym, mogącym się sprawdzić w innych przypadkach chorobowych czy nawet innych dziedzinach.

Wprowadzenie pojęcia polireprezentacji

Jednym z kluczowych osiągnięć tej pracy jest wprowadzenie nowej koncepcji tzw. polireprezentacji. Według Autorki rozprawy i zespołu autorskiego, polireprezentacja jest zadaniem polegającym na łączeniu reprezentacji danych pochodzących z różnych źródeł, ale



posiadających tą samą modalność (np. różne techniki obrazowania płuc). Zaproponowane podejście pozwoliło na uzyskanie bardziej kompleksowych i dokładnych wyników, w szczególności w diagnostyce medycznej, gdzie różne metody obrazowania mogą dostarczać synergetycznych informacji. Propozycja nowej metodologii wykorzystania analizowanych danych poszerza pole badawcze i jest punktem wyjścia dla dalszych badań nad łączeniem danych z różnych źródeł w ramach jednego modelu diagnostycznego.

Rozwój metod wyjaśnialnej sztucznej inteligencji (XAI)

Wyjaśnialność modeli sztucznej inteligencji jest kluczowa w dziedzinach, w których AI może wpływać na decyzje o krytycznym znaczeniu, jak w przypadku zdrowia pacjentów, zastosowań militarnych czy praw człowieka. Modele głębokiego uczenia, są z punktu widzenia użytkownika „czarnymi skrzynkami”, których wewnętrzne działanie jest niejasne. Praca rozwija metody post-hoc wyjaśniania działania tych modeli, co ma na celu zwiększenie zaufania użytkowników końcowych, w tym przypadku lekarzy, do systemów AI. Jednym z głównych wkładów merytorycznych Autorki jest rozwinięcie interaktywnej metody LIMEcraft, która pozwala użytkownikom na wybieranie i modyfikowanie istotnych obszarów analizowanego obrazu. Zaproponowane w pracy interaktywne podejście angażujące użytkowników w trybie human-in-the-loop, daje możliwość ciągłego zwiększania dokładności i wiarygodności algorytmów XAI, a przede wszystkim zwiększenie stopnia ich użyteczności dla użytkowników końcowych.

Wprowadzenie nowych metod łączenia wyjaśnień

W pracy zaproponowano nowe, oryginalne sposoby na łączenie wyników różnych metod wyjaśniania głębokich modeli neuronowych. Autorka zaproponowała dwa nowe podejścia: NormEnsembleXAI oraz składanie (ensembling) wyjaśnień oparty o wykorzystanie sieci konwolucyjnych. Zaproponowane podejścia zostały kompleksowo przetestowane. Wykazano, że poprawiają one dokładność wyjaśnień, i co jest nawet bardziej istotne – zwiększają ich krzepkość. Obie cechy mają istotne znaczenie dla praktycznego zastosowania modeli AI.

3. Osiągnięte wyniki

Zaproponowane w pracy metody rozwiązania analizowanych problemów badawczych zostały dobrze dobrane. Autorka, aby osiągnąć cele wymienione we wstępie pracy, wykorzystwała zarówno znane metody, umiejętne i synergiczne połączenie istniejących metod, ale także zaproponowała własne, oryginalne rozwiązania.

Wybór zbiorów danych analizowanych w pracy jest trafny i dobrze uzasadniony. Autorka głównie wykorzystwała istniejące bazy danych, ale na potrzeby niektórych eksperymentów także przygotowała własne. Wszystkie wykorzystane w pracy zbiory danych są dobrze opisane i jest do nich wskazany dostęp.

Wyniki uzyskane w ramach badań zostały poprawnie zinterpretowane w kontekście postawionego problemu, założeń teoretycznych i dostępnej wiedzy. W przypadku każdego z zaproponowanych i wykorzystywanych algorytmów, zapewniono kody i dane umożliwiające reprodukcję i weryfikację wyników.

Podsumowując, stwierdzam, że przedstawione w pracy badania, analizy i wyciągnięte na ich podstawie wnioski mają istotne znaczenie dla rozwoju dyscypliny naukowej, w szczególności wielomodalnych i wieloreprezentacyjnych modeli głębokiego uczenia, jak



również technik objaśniania post-hoc modeli AI. Niezwykle ważne jest uwzględnienie człowieka (jako użytkownika końcowego) w procesie tworzenia i weryfikacji algorytmów decyzyjnych.

Na podkreślenie zasługuje kompleksowość i skrupulatność opisu w rozprawie poszczególnych badań. Ich elementami są zawsze: rzetelny przegląd literaturowy, rzeczowe porównanie wielu dedykowanych podejść i modeli, analiza wpływu różnych elementów systemu i hiperparametrów, wskazanie mocnych, ale i słabych stron i ograniczeń proponowanych podejść, a także wskazanie potencjału i kierunków rozwoju proponowanych metod.

4. Ocena indywidualnego wkładu Doktorantki w powstanie pracy doktorskiej

Oceniana praca doktorska jest oparta o 6 wieloautorskich artykułów naukowych.

Dwa z nich, tj. artykuły [P1] oraz [P4] zostały opublikowane w prestiżowych i cenionych czasopismach w dziedzinie sztucznej inteligencji, mianowicie w *Pattern Recognition* oraz *Machine Learning*, które posiadają wysokie IF=7.5 (na dzień publikacji artykułów). Artykuł [P6] został zaprezentowany na nowej, ale już prestiżowej konferencji *2nd World Conference on eXplainable Artificial Intelligence (2024)*, i umieszczony w zbiorowej monografii pokonferencyjnej. Artykuł [P6] ukazał się w polskim czasopiśmie *Progress in Polish Artificial Intelligence Research*, nieposiadającym współczynnika IF.

Artykuły [P3] oraz [P5], w momencie otrzymania do recenzji niniejszej rozprawy doktorskiej, nie zostały zaakceptowane ani zgłoszone do żadnego czasopisma ani na żadną konferencję, co jest dość rzadkim podejściem, jeżeli chodzi o przygotowywanie prac doktorskich opartych o cykl publikacji.

Doktorantka była pierwszą autorką we wszystkich publikacjach, jednak w rozprawie nie deklaruje swojego merytorycznego ani procentowego udziału w powstanie tych artykułów. Na podstawie korespondencji z Doktorantką, mogę stwierdzić następujący wkład Doktorantki w powstanie wskazanych artykułów:

- [P1] **Hryniewska, W., Bombiński, P., Szatkowski, P., Tomaszewska, P., Przelaskowski, A., Biecek, P.**, “Checklist for responsible deep learning modeling of medical images based on COVID-19 detection studies”. *Pattern Recognition* (2021).

Wkład Doktorantki polegał na: koordynacji prac zespołu autorów, przygotowaniu przeglądu literatury, wyciągnięciu wniosków w zakresie związanym z głębokim uczeniem, przygotowaniu schematów graficznych, napisaniu większości tekstu artykułu, przygotowaniu większości odpowiedzi na recenzje.

- [P2] **Hryniewska, W., Grudzień, A., Biecek, P.**, “LIMEcraft: handcrafted superpixel selection and inspection for Visual eXplanations”. *Machine Learning* (2022).

Wkład Doktorantki polegał na: sformułowaniu problemu badawczego, przygotowaniu przeglądu literatury, zaproponowaniu metody, przygotowanie i przeprowadzenie eksperymentów, przeprowadzeniu badań z użytkownikami, implementacji frontendu aplikacji, konsultacji dotyczące implementacji backendu aplikacji, przygotowaniu większości schematów graficznych, napisaniu większości tekstu artykułu, przygotowaniu większości odpowiedzi na recenzje.

- [P3] **Hryniewska-Guzik, W., Kędzierska, M., Biecek, P.**, “Multi-task learning for classification, segmentation, reconstruction, and detection on chest CT scans”, in: *Progress in Polish Artificial Intelligence Research 4* (2023).



Wkład Doktorantki polegał na: sformułowaniu problemu badawczego, zaproponowaniu metody, zaprojektowaniu architektury głębokiego uczenia, konsultacji sposobów implementacji, przygotowaniu części schematów graficznych, napisaniu większości tekstu artykułu.

- [P4] *Hryniewska-Guzik, W., Biecek, P., "X-ray transferable polyrepresentation learning". Preprint (2024).*

Wkład Doktorantki polegał na: propozycji i implementacji metody, przygotowaniu przeglądu literatury, przygotowaniu schematów graficznych, napisaniu całości tekstu artykułu.

- [P5] *Hryniewska-Guzik, W., Sawicki, B., Biecek, P., "NormEnsembleXAI: unveiling the strengths and weaknesses of XAI ensemble techniques". Preprint (2024).*

Wkład Doktorantki polegał na: sformułowaniu problemu badawczego, zaproponowaniu metody, zaproponowaniu struktury artykułu, przygotowaniu przeglądu literatury, napisaniu większości tekstu artykułu, przygotowaniu wszystkich schematów graficznych, przygotowaniu części analiz, wskazaniu kolejnych kroków możliwych do podjęcia.

- [P6] *Hryniewska-Guzik, W., Longo, L., Biecek, P., "CNN-based explanation ensembling for dataset, representation and explanations evaluation", in: 2nd World Conference on XAI (2024).*

Wkład Doktorantki polegał na: zaproponowaniu i implementacji metody, przygotowaniu przeglądu literatury, przeprowadzeniu eksperymentów, przygotowaniu schematów graficznych, napisaniu całości tekstu artykułu oraz większości odpowiedzi na recenzje.

Doktorantka jest również autorką trzech rozdziałów przedstawionej do oceny monografii. Na podstawie powyższego, nie mam wątpliwości, że powyższe prace stanowią istotny wkład naukowy w dyscyplinę. Ponadto należy podkreślić, że Doktorantka, poprzez swoje badania i powstałe w ich wyniku prace, wykazała się umiejętnością pracy zespołowej, co w dzisiejszych czasach jest cenną i niezbędną kompetencją.

5. Struktura i język pracy

Recenzowana praca doktorska składa się z wykazu skrótów, wykazu oryginalnych publikacji, 5 rozdziałów, w tym wstępu, podsumowania i literatury odwołującej się do 295 pozycji. Łącznie manuskrypt liczy 211 stron.

Praca jest napisana w formie monografii, jednak tylko część z jej rozdziałów (1, 2 oraz 5) ma tradycyjną strukturę monografii, natomiast rozdziały 3 oraz 4 składają się z artykułów naukowych których Doktorantka jest współautorką, a które zostały przedłożone do oceny jako immanentna część rozprawy doktorskiej. W rozdziale 1 Autorka przedstawia cel podjętych badań i listę sześciu publikacji w których je opisuje. W rozdziale 2 przedstawiono wiedzę teoretyczną ułatwiającą właściwe zrozumienie dalszej, głównej części pracy. Treści w rozdziale opisują podstawową wiedzę związaną z analizą obrazów medycznych, głębokimi sieciami neuronowymi oraz wyjaśnialną sztuczną inteligencją. Rozdziały 3 i 4 są głównymi rozdziałami, w których przedrukowano sześć artykułów przedstawiających wkład merytoryczny Doktorantki w rozwój dyscypliny i dziedziny nauki. W szczególności, w rozdziale 3: *Insights into Lung Representations* skupiono się na zastosowaniu modeli głębokiego uczenia do analizy obrazów medycznych, szczególnie w diagnostyce COVID-19. Autorka prezentuje w nim wyniki badań nad modelami wielozadaniowymi oraz koncepcję tzw. polireprezentacji. Omawia także standardy odpowiedzialnego stosowania AI w tej medycynie. W rozdziale 4: *New*



explanation methods, Autorka omawia zaproponowane nowe metody wyjaśniania modeli, takie jak interaktywna metoda LIMEcraft, a także nowe podejścia do łączenia wyjaśnień typu post-hoc. Rozdział koncentruje się na metodach zwiększających zaufanie użytkowników do modeli AI poprzez poprawę jakości wyjaśnień. W rozdziale 5, Autorka dokonuje podsumowania osiągnięć opisanych w pracy, umiejętnie odnosząc się przy tym do wskazanych w rozdziale 1 pytań i celów badawczych.

Przyjęta struktura recenzowanej pracy doktorskiej jest w mojej opinii prawidłowa i zarazem ciekawa. Pozwala oprzeć pracę na cyklu publikacji naukowych, które są uzupełnione o treści, pozwalające na płynne przyswajanie zawartej w nich wiedzy.

Praca jest dobrze zorganizowana, przejrzysta i łatwa w odbiorze. Napisana jest w języku angielskim, poprawnie gramatycznie i stylistycznie. Język jakim Autorka posługuje się w pracy spełnia wymagania jakie stawiane są przed monografiami naukowymi, m.in. pracami doktorskimi, Doktorantka w poprawny sposób używa terminologii naukowej właściwej dla dziedziny, w której lokowane są badania, w tym przypadku sztucznej inteligencji i medycynie. Wszystkie symbole użyte w pracy są zdefiniowane i wyjaśnione.

Przegląd i analiza cytowanej w pracy doktorskiej literatury jest bardzo silnym punktem recenzowanej rozprawy. W manuskrypcie odniesiono się do 296 pozycji literaturowych. Ich analiza pozwala stwierdzić, że w przeglądzie znalazły się praktycznie wszystkie najważniejsze pozycje literaturowe z dziedziny pracy. Tym samym, należy stwierdzić, że literatura jest trafnie dobrana, aktualna i dobrze koresponduje z opisywanymi w pracy zagadnieniami. Cytowane w pracy publikacje dopełniają prezentowane w monografii treści. Autorka pracy wybrała odpowiednie i aktualne podstawy teoretyczne i źródła literaturowe do analizy postawionego problemu badawczego, co świadczy o dobrej znajomości Autorki, aktualnego stanu wiedzy w tej tematyce.

6. Zagadnienia do dyskusji:

Natury ogólnej:

- W rozdziale 3 przedstawiono wielozadaniową strukturę neuronową. Proszę o doprecyzowanie informacji o sposobie uczenia tej złożonej architektury. W szczególności nie jest jasne czy owa struktura jest uczona jednocześnie w trybie end-to-end, czy sekwencyjnie (np. od klasyfikacji do segmentacji). W obu przypadkach proszę o przedstawienie diagramu przepływu informacji w trakcie treningu. Proszę również o wskazanie praktycznych korzyści wynikających z zaproponowanego podejścia wielozadaniowego.
- W sekcji 3.3.4 przedstawiono architekturę sieci syjamskiej oraz przedstawiono optymalizowaną w trakcie jej uczenia funkcję straty. Proszę o uzasadnienie doboru współczynników przyjętych w funkcji straty w proporcjach: 0.8/0.2.
- W sekcji 3.3.4 opisano sposób wyznaczania cech radiomicznych i dalszego ich wykorzystania dla celów uczenia w trybie polireprezentacji. Z opisu wynika, że wszystkie 126 cech radiomicznych zostało wykorzystane w dalszym procesie uczenia. Czy Autorka rozważała wykorzystanie metod regularyzacji (np. Lasso) lub dedykowanych do danych tabelarycznych sieci (np. TabNet) do wyboru optymalnego zestawu cech ?
- Proszę doprecyzować informacje odnośnie procesu redukcji danych przy wykorzystaniu algorytmu PCA (sekcja 3.3.5).



- W rozdziale 4 przedstawiono algorytm LIMCraft. Proszę przedyskutować kwestię wrażliwości algorytmu LIMCraft na inicjalizację algorytmu, m.in. z powodu wykorzystania niedeterministycznego algorytmu klasteryzacji k-means czy też manualnego wyboru obszarów przez użytkowników.
Czy efektem zaznaczania/wyboru superpixeli np. przez radiologów jest skuteczniejsze nauczanie modelu czy „jedynie” polepszenie wyjaśnień ?
- W rozdziale 4 kompleksowo porównano metody XAI przy wykorzystaniu wielu dedykowanych metryk. Trudno jest jednak wskazać metryki, które mówią o praktycznej „jakości”/”przydatności” danego wyjaśnienia. Czy zaproponowane modele wyjaśnień weryfikowane były (konsultowane) z potencjalnymi użytkownikami końcowymi, np. z radiologami.
- Tworzenie modeli będących złożeniem innych modeli ma swoje zalety i wady. Nie zawsze złożenie modeli daje synergetyczny skutek. Przykładowo, w przypadku w których jeden z modeli znacznie odróżnia się od pozostałych modeli (pomijając w tym momencie czy pozytywnie czy negatywnie). Co więcej, część z modeli XAI generuje dodatnie i ujemne wartości na mapach atrybucji a część jedynie pozytywne. Proszę przedyskutować w tym kontekście kwestię wyboru modeli XAI do procesu składania modeli XAI.
- Proszę wyjaśnić zasadność stosowania modeli wyjaśniających XAI w procesie segmentacji.
- Czy przedstawione w pracy modele XAI oraz towarzyszące im metryki pozwalają w jakimś stopniu na określenie stopnia niepewności decyzji systemu, pochodzącej z różnych źródeł, m.in. niedokładności modeli czy niepewności danych ?

Wybrane uwagi natury szczegółowej:

- Stosowanie formy „My” w przypadku pisania rozprawy doktorskiej jest uzasadnione w przypadku opisywania wspólnych publikacji, jednak jej użycie w sekcji 1.3 Research aims i Rozdziale 5 Conclusions and future work, jest niezręczne.
- Podział sekcji 2.2 Deep Learning na 2.2.1 Feedforward Neural Networks, 2.2.2 Convolutional Neural Networks oraz 2.2.3 Representation Learning jest dość dziwny.
- Na str. 38 Autorka pisze „A good representation should possess several desirable properties:” i następnie te własności wymienia. Nie jest dla mnie jasne czy są to wnioski Autorki czy informacje zaczerpnięte z literatury.
- W tabeli Table 2.1: Comparison of DL architectures for classification skupiono się na architekturach jedynie do 2020 r, podczas gdy do 2024 r powstało ich zdecydowanie więcej i skuteczniejszych, szczególnie w zagadnieniach medycznych i takich w których brakuje danych, lub gdy dane są oznaczone jedynie na poziomie obrazów.
- Brakuje podobnych tabel dla segmentacji, detekcji czy rekonstrukcji.
- Rysunek 2.2 jest słabej jakości.
- Podział metod wyjaśnialności zaprezentowany na Rys. 2.12 jest jednym z przyjętych, zatem warto w jego przypadku odwołać się do literatury źródłowej.
- W opisywanych w pracy eksperymentach brakuje informacji o sposobie doboru hiperparametrów związanych z procesem uczenia.

Wskazane uwagi i komentarze nie umniejszają wartości merytorycznej pracy, a są oczywistą konsekwencją prezentacji nowych i oryginalnych rozwiązań.



7. Podsumowanie

Recenzowana rozprawa doktorska Pani Weroniki Hryniewskiej-Guzik wnosi istotny wkład naukowy w dyscyplinę Informatyka techniczna i telekomunikacja, w szczególności w dziedzinie głębokiego uczenia i wyjaśnialnej sztucznej inteligencji, z naciskiem na praktyczne zastosowania analizowanych w pracy podejść i algorytmów w diagnostyce medycznej. Praca ma charakter interdyscyplinarny, łączy wiedzę z dziedziny informatyki, sztucznej inteligencji, analizy danych i medycyny.

Rozwiązania zaprezentowane w rozprawie są na poziomie międzynarodowym, odpowiadającym i częściowo wyprzedzającym aktualny stan wiedzy i podejścia prezentowane w literaturze światowej. Zaprezentowane badania niewątpliwie przyczyniły się do poprawy dokładności modeli, a także ich wyjaśnialności oraz wniosły swój wkład w opracowanie standardów dla odpowiedzialnego stosowania modeli AI w diagnostyce obrazowej, przyczyniając się tym samym do rozwoju bardziej bezpiecznych i wiarygodnych systemów AI.

Praktyczne wyniki pracy koncentrują się na zastosowaniu opracowanych metod do diagnostyki medycznej, jednak znaczenie tej pracy wykracza poza te zastosowania, mając duży potencjał wpływu na szerokie zastosowania sztucznej inteligencji w innych dziedzinach, gdzie transparentność, wyjaśnialność i wiarygodność opracowywanych systemów wspomagania decyzji są kluczowe.

Zarówno recenzowana rozprawa doktorska jak i pozostałe osiągnięcia naukowe Doktorantki m.in. praca w grantach (np. SONATA, INFOSTRATEG, granty IDUB), staż naukowy w National Taiwan University of Science and Technology czy pozostałe publikacje w których Doktorantka jest współautorką, niewątpliwie dowodzą umiejętności prowadzenie pracy naukowej przez Autorkę, zarówno samodzielnego jak i w zespole naukowym, prowadzonym przez prof. Przemysława Biecka.

Biorąc pod uwagę zakres, poziom merytoryczny i sposób prezentacji wyników przedstawiony w rozprawie, stwierdzam, że recenzowana rozprawa doktorska spełnia wymagania ustawowe stawiane rozprawom doktorskim. Stąd wnioskuję o dopuszczenie jej do publicznej obrony.

Ponadto, w związku z wysoką jakością części publikacji i rangi czasopism w których zostały opublikowane, w szczególności prac: [P1] Hryniewska, W., Bombiński, P., Szatkowski, P., Tomaszewska, P., Przelaskowski, A., Biecek, P., "Checklist for responsible deep learning modeling of medical images based on COVID-19 detection studies" opublikowanej w: *Pattern Recognition* oraz [P4] Hryniewska, W., Grudzień, A., Biecek, P., "LIMEcraft: handcrafted superpixel selection and inspection for Visual eXplanations". opublikowanej w: *Machine Learning*, a także dużą aktywność naukową Doktorantki przejawiającej się m.in. poprzez udział w czterech grantach badawczych, aktywny udział w siedmiu konferencjach naukowych, udział w pięciu szkołach letnich, staż naukowy National Taiwan University of Science and Technology czy recenzowanie publikacji dla prestiżowych czasopism i konferencji, a także członkostwo w Komitecie technicznym warsztatów Clinically-oriented and Responsible AI for Medical Data Analysis na MICCAI2023, wnioskuję o wyróżnienie rozprawy.

Michał Grochowski