

Kraków, dnia 04.09.2023 r.

dr hab. inż. Eliaż Kańtoch, prof. AGH
Wydział Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej
Akademii Górniczo-Hutniczej
im. Stanisława Staszica w Krakowie
al. Mickiewicza 30
30-059 Kraków

RECENZJA

Rozprawy doktorskiej mgra inż. Konrada Duraj pt.:
„Deep learning applications in biomedical engineering”
wykonanej na Wydziale Inżynierii Biomedycznej Politechniki Śląskiej

Recenzję rozprawy doktorskiej opracowano na zlecenie Rady Dyscypliny Inżynieria Biomedyczna Politechniki Śląskiej na posiedzeniu w dniu 22 czerwca 2023 roku (pismo nr RDIB.002.48.2023 z dn. 29.06.2023 roku) zgodnie z zasadami zawartymi w Ustawie z dnia 20 lipca 2018 roku Prawo o szkolnictwie wyższym i nauce: Dział V – stopnie i tytuł w systemie szkolnictwa wyższego i nauki.

Przedmiotem recenzji jest rozprawa doktorska w języku angielskim pt.: *„Deep learning applications in biomedical engineering”* przedstawiona do obrony w dyscyplinie Inżynieria Biomedyczna. Promotorem rozprawy jest profesor Politechniki Śląskiej dr hab. inż. Paweł Kostka.

Recenzowana rozprawa ma postać dzieła w języku angielskim zawierającego 101 stron i zbudowana jest z 5 rozdziałów zasadniczych, do których zaliczam: wprowadzenie do tematyki pracy oraz przegląd literaturowy dokonany w pierwszym rozdziale *„Introduction”*, kolejnym jest rozdział *„Aim of the doctoral thesis”*, następnie przedstawione są cztery publikacje wchodzące w skład rozprawy: *„Predicting Molecule Toxicity Using Deep Learning, Recognition of Drivers’ Activity Based on 1D Convolutional Neural Network”*, *„Heartbeat Detection in Seismocardiograms with Semantic Segmentation”*, *„Semantic Segmentation of 12-Lead ECG Using 1D Residual U-Net with Squeeze-Excitation Blocks”*. Kolejnym jest rozdział *„A new paradigm”* - uczenie reprezentacji rozmytej, a piątym *„Conclusions”*. Do rozdziałów uzupełniających pracę zaliczam: bibliografię, wykaz skrótów i oznaczeń oraz streszczenie w języku polskim i angielskim umieszczone na końcu rozprawy.

Pan mgr inż. Konrad Duraj w swojej rozprawie doktorskiej wykorzystuje algorytmy uczenia głębokiego do przetwarzania danych biomedycznych takich jak: przewidywanie toksyczności molekuł, rozpoznawanie akcji kierowców na drodze bazując na sygnałach EOG, detekcja uderzeń serca w sygnałach sejsmograficznych oraz segmentacja składowych EKG. Rozwiązania wymienionych wyzwań badawczych znajdują się w osobnych publikacjach stanowiących rozdziały rozprawy. Natomiast w rozdziale czwartym Autor zaproponował algorytm nazywany „uczeniem reprezentacji rozmytej”, który jest odpowiedzią na problem transparentności procesu decyzyjnego algorytmów uczenia głębokiego.

Z uwagi na charakter pracy i jej specyfikę ocenię poszczególne rozdziały podkreślając ich mocne i słabe strony.

Rozdział pierwszy pracy, zatytułowany „*Introduction*” jest szerokim wprowadzeniem do podejmowanych w rozprawie problemów badawczych. Doktorant wykazał się odpowiednią znajomością aktualnej literatury naukowej oraz trendów technologicznych. Bibliografia jest prawidłowo przygotowana i składa się z 65 zróżnicowanych źródeł (każda z opublikowanych prac zawiera osobny wykaz literatury), w tym z renomowanych czasopism naukowych, branżowych konferencji, źródeł internetowych i odpowiednich pozycji książkowych.

Autor poprawnie argumentuje swoje tezy z użyciem odpowiednich cytowań, co świadczy o jego umiejętności analize i oceny istniejącej wiedzy. W rozdziale pierwszym Autor przedstawia zarys historyczny rozwoju sztucznej inteligencji, odpowiednio definiuje systemy ekspertowe, struktury danych do reprezentowania wiedzy oraz logikę rozmytą. Omawia także wybrane algorytmy uczenia maszynowego, sieci neuronowe i dokonuje przeglądu najczęściej wykorzystywanych architektur ze szczególnym uwzględnieniem uczenia głębokiego. Doktorant w swojej dysertacji odnosi się do ważnych prac z obszaru sieci głębokich m.in. publikacji autorstwa Alex Krizhevsky, Ilya, Geoffrey E. Hinton pt. „*ImageNet Classification with Deep Convolutional Neural Networks*” i w sposób właściwy i klarowny prezentuje właściwości sieci głębokich w kontekście poruszanych wyzwań badawczych. Autor przytacza także dane sugerujące przewagę algorytmów uczenia głębokiego nad tradycyjnymi algorytmami uczenia maszynowego cytując m.in. dane pochodzące z *ImageNet Contest*. Autor stawia tezę, że uczenie głębokie doskonale znajduje zastosowania w obszarze inżynierii biomedycznej ze względu na wielowymiarowość danych medycznych i ich wysoce nieliniowe zależności.

W oparciu o analizę literatury Autor poprawnie zidentyfikował nierozwiązane problemy i zaproponował własne rozwiązania szczegółowo omówione w poszczególnych publikacjach.

Autor rozprawy w rozdziale „*Aim of the doctoral thesis*” przedstawia w sposób ogólny cel podjętych badań poprzez odwołanie się do opublikowanych prac jak również rozdziału czwartego, który jest odpowiedzią na zidentyfikowane niedoskonałości stosowanych podejść. W mojej ocenie rozdział jest zbyt mało precyzyjny i przydatny byłby bardziej szczegółowy opis zawierający zbiór celów szczegółowych wraz ze sposobem ich weryfikacji w oparciu o przyjęte metryki. Autor zdecydował o opisaniu szczegółów w kolejnych rozdziałach.

W rozdziale 3.1 zawierającym publikację [1], której mgr inż. Konrad Duraj jest pierwszym autorem podjęto próbę opracowania sieci neuronowej do przewidywania toksyczności molekuł. Podejmowany problem jest istotny, gdyż ocena toksyczności związku ma kluczowe znaczenie dla oceny jego szkodliwości dla organizmów żywych, zwłaszcza w kontekście tworzenia nowych leków. Natomiast obecnie stosowane chemiczne metody oceny toksyczności są długotrwałe, kosztowne i wymagają specjalistycznego sprzętu. Autorzy w sposób przejrzysty przedstawili materiał i metody badań. W prowadzonych eksperymentach wykorzystali bazę danych "SMILES (Simplified Molecular-Input Line-Entry System) Toxicity", składającą się z 6698 nietoksycznych molekuł i 964 toksycznych. Nierównowaga klas w zestawie danych stworzyła problem nadmiernego dopasowania (ang. „overfitting”), który mieli rozwiązać. Rozwiązanie problemu zostało oparte o hybrydową architekturę CNN-RNN pozwalającą na ekstrakcję cech zarówno w lokalnym polu recepcyjnym, jak i w sposób sekwencyjny. Rozwiązanie problemu nadmiernego dopasowania zostało zrealizowane poprzez zastosowanie następujących metod: zrównoważenie stosunku nietoksycznych i toksycznych molekuł, manipulowanie funkcją straty i procesem optymalizacji. Autorzy wykorzystali metody radzenia sobie z nadmiernym dopasowaniem w dwóch modelach o różnym poziomie dokładności w przewidywaniu toksyczności molekuł. Pierwszy model osiągnął dokładność na poziomie 96% na zbiorze treningowym, 89% na zbiorze walidacyjnym i 65% na zbiorze testowym. Drugi model osiągnął dokładność na poziomie 80% na zbiorze treningowym, 78% na zbiorze walidacyjnym i 77% na zbiorze testowym. Zastosowane wskaźniki oceny jakości są powszechnie wykorzystywane w podobnych analizach. Wyniki są satysfakcjonujące i świadczą o poprawności przyjętych założeń. Mgr inż. Konrad Duraj zadeklarował, że opracował koncepcję projektu, dokonał przeglądu literatury, opracował i przetestował metodologię wstępnego przetwarzania danych, uczenia sieci, ocenił model głębokiego uczenia oraz

przeanalizował wyniki jak i napisał część tekstu publikacji. Na tej podstawie Jego wkład w opracowanie publikacji oceniam jako znaczący.

W rozdziale 3.2 zawierającym publikację [2], której mgr inż. Konrad Duraj jest trzecim autorem przedstawiono badania skupiające się na opracowaniu klasyfikatora rozpoznającego różne scenariusze związane z nauką prowadzenia pojazdu na podstawie danych zebranych z inteligentnych okularów JINS MEME ES_R w rzeczywistych warunkach drogowych. Celem badań było zminimalizowanie liczby sygnałów wymaganych do rozpoznania aktywności. W badaniach wykorzystano dane elektrookulograficzne (EOG) i jednowymiarową splotową sieć neuronową (1D CNN). Znacznym osiągnięciem pracy jest także oryginalny zbiór nagrań składający się z danych zebranych od 20 kierowców poruszających się tą samą trasą w południowej Polsce, zarówno doświadczonych, jak i w trakcie nauki. Nagrania zostały podzielone na cztery kategorie czynności związane z prowadzeniem pojazdu: parkowanie, przejazd przez rondo, ruch miejski i przejazd przez skrzyżowanie. Wysoko oceniam utworzenie oryginalnej bazy danych na potrzeby prowadzonych badań. Uzyskane wyniki wynoszą 95,6% w przypadku zestawu walidacyjnego i 99,8% w przypadku zestawu treningowego. Zastosowane metody badawcze są właściwe. Na uwagę zasługuje przejrzysty i staranny sposób prezentacji i wizualizacji wyników badań. Przeprowadzone eksperymenty zostały dobrze zaprojektowane, a implikacje badania są istotne, gdyż w porównaniu z innymi badaniami opracowana metoda cechuje się wyższą dokładnością klasyfikacji niż istniejące rozwiązania wykorzystujące podejście oparte o BFS, k-NN czy SVM oraz wskazuje na możliwość rozpoznania wybranych czynności związanych z prowadzeniem pojazdu wyłącznie na podstawie danych EOG.

Wyniki pracy nie tylko posiadają znaczącą wartość naukową, ale także wysoki potencjał wdrożeniowy, gdyż opracowana metoda może być zastosowana szczególnie w pojazdach inteligentnych w celu poprawy bezpieczeństwa jazdy i umiejętności prowadzenia pojazdu. Mgr inż. Konrad Duraj zadeklarował, że był odpowiedzialny za przetwarzanie wstępne danych, opracował model sieci, dokonał jego wytrenowania i oceny jak również wizualizację wyników badań, napisał także część tekstu publikacji. Na tej podstawie Jego wkład merytoryczny w opracowanie publikacji szczególnie rozwoju modelu sieci głębokiej oceniam jako znaczący.

W rozdziale 3.3 zawierającym publikację [3], której mgr inż. Konrad Duraj jest pierwszym autorem zaprezentowano metodę wykrywania bicia serca z użyciem sejsmokardiografii (SCG) oraz konwolucyjnej sieci neuronowej. W badaniach wykorzystano publicznie dostępny zbiór danych CEBS (ang. Combined Measurement of ECG, Breathing and

Seismocardiogram) zawierający zsynchronizowane pomiary EKG, sygnałów oddechowych i sejsmokardiogramów. Metryki (m.in. Jaccard score – 97.1%) sugerują, że model poprawnie segmentuje uderzenia serca na sygnałach sejsmokardiogramu (SCG). Zastosowane metryki oceny jakości, wykorzystany materiał oraz metody badawcze są odpowiednie dla tego typu badań i zostały precyzyjnie zdefiniowane. Zaproponowane podejście wykazało wysoką czułość i precyzję, porównywalną z rozwiązaniami obecnymi w literaturze, ale różniło się wstępnym przetwarzaniem danych, architekturą sieci neuronowej, parametrami i reprezentacją sygnału wyjściowego. Do ograniczeń zaprezentowanego badania należy stosunkowo mała próba zdrowych osób w określonym przedziale wiekowym i akwizycja sygnałów w pozycji leżącej.

Uzyskane wyniki posiadają potencjał aplikacyjny np. mogą mieć zastosowanie w rozwoju systemów nieinwazyjnego monitorowania pracy serca. W przeprowadzonej obszernej dyskusji wyników pojawia się kilka tematów, które mogą być kierunkiem dalszego rozwoju np. wykorzystanie pomiarów od osób z wybranymi schorzeniami kardiologicznymi.

Mgr inż. Konrad Duraj zadeklarował, że opracował koncepcję projektu, dokonał przeglądu literatury, opracował i przetestował metodologię wstępnego przetwarzania danych, ocenił model głębokiego uczenia oraz przeanalizował wyniki jak i napisał część tekstu publikacji. Na tej podstawie Jego wkład w opracowanie publikacji oceniam jako znaczący.

W rozdziale 3.4 zawierającym publikację [4], której mgr inż. Konrad Duraj jest pierwszym autorem zaprezentowano system semantycznej segmentacji sygnałów elektrokardiograficznych (EKG). W badaniu wykorzystano architekturę U-Net zaadoptowaną do zadania jednowymiarowej semantycznej segmentacji. Zbiór danych pochodził z Uniwersytetu Lobachevsky'ego i składał się z 200 sygnałów 12-odprowadzeniowego EKG z adnotacjami. Zastosowane zostały adekwatne metody. Szczególnie starannie i w przejrzysty sposób przedstawiono wyniki korzystając z wykresów i tabel. Opracowany model osiągnął wysoką wydajność według przyjętych metryk (dokładność - 0.95, AUC - 0.99, swoistość – 0.95 i czułość – 0.99) w wyodrębnianiu kluczowych cech 12-odprowadzeniowego EKG, takich jak załamki P i T oraz zespoły QRS. Wskaźniki wydajności proponowanej metody są porównywalne lub lepsze niż te podawane w istniejącej literaturze, co wskazuje na jej odpowiednią skuteczność w segmentacji sygnału EKG. Wyniki sugerują, że opracowane rozwiązanie może być z powodzeniem zastosowane do wyodrębnienia ważnych składowych sygnału EKG równoległe ze wszystkich odprowadzeń sygnałowych, co czyni je oryginalnym, uniwersalnym

i efektywnym. Możliwe zastosowania uzyskanych wyników obejmują wsparcie wykrywania stanów nagłych i groźnych dla życia chorób w szczególności zaburzeń rytmu serca, niedokrwienia i martwicy mięśnia sercowego z użyciem urządzeń nasobnych umożliwiających rejestrację sygnału EKG szczególnie w zastosowaniach domowych.

Doktorant poprawnie zidentyfikował główne ograniczenia stosowanego podejścia - dostępność odpowiednio przygotowanych i zweryfikowanych baz danych zawierających sygnały fizjologiczne odpowiedniej jakości. Jako odpowiedź na to ograniczenie Doktorant sugeruje wykorzystanie generatorów sygnałów w celu utworzenia baz danych na potrzeby trenowania modeli.

Mgr inż. Konrad Duraj zadeklarował, że opracował koncepcję projektu, dokonał przeglądu literatury, opracował i przetestował metodologię wstępnego przetwarzania danych, uczenia i ocenił model głębokiego uczenia oraz przeanalizował wyniki jak i napisał część tekstu publikacji. Na tej podstawie Jego wkład w opracowanie publikacji oceniam jako znaczący.

W rozdziale nr 4 pt. „A new paradigm” Autor omawia ograniczenia związane ze stosowaniem technik deep learningu w zastosowaniach medycznych. Autor wskazuje na możliwą stronniczość lub słabość niektórych modeli i ograniczenia możliwości ich zastosowania w różnych scenariuszach klinicznych. Autor przytacza zidentyfikowane przez Garrego Marcusa ograniczenia głębokich sieci neuronowych takie jak m.in.: konieczność zapewniania odpowiednich zbiorów danych, brak transparentności, brak rozróżnienia związku przyczynowego od korelacji. Ponadto, wskazuje istotność optymalizacji algorytmów pod kątem zrozumienia, a nie tylko samej dokładności predykcji. Ma to związek także z bezpieczeństwem działania sieci neuronowej, której wynik działania może zostać zweryfikowany. Autor zaproponował nową metodę, która łączy głębokie sieci neuronowe i logikę rozmytą jako odpowiedź na zidentyfikowane ograniczenia. Podstawą opracowanej koncepcji jest stwierdzenie, że głęboka sieć neuronowa w oparciu o dane wejściowe powinna wygenerować system logiki rozmytej, który następnie może zostać wykorzystany do przewidywania wyniku. Autor dokonał weryfikacji opracowanego algorytmu poprzez adaptację i rozwiązanie nieliniowego problemu jakim jest bramka XOR. Zaproponowany eksperyment ukazuje znaczną wiedzę i dojrzałość naukową Doktoranta. Uzyskane wyniki świadczą o poprawności przyjętych założeń. Chociaż rozwiązanie nie jest doskonałe to może być bazą do rozwoju metody. Autor planuje udostępnienie opracowanej metody jako bibliotekę obliczeniową dla szerszego grona naukowców. Doktorant zdefiniował zastosowane podejście jako *fuzzy representation learning*

i zaklasyfikował je do kategorii *hybrid fuzzy inference neural network*. Przeprowadzona obszerna dyskusja wyników jest źródłem trafnych spostrzeżeń szczególnie w kontekście wyzwań związanymi z zastosowaniem algorytmów wykorzystujących głębokie sieci neuronowe w medycynie.

Poniżej przedstawiam zbiór uwag krytycznych i pytań, które w mojej ocenie mogłyby zostać szerzej omówione. Uwagi mają charakter dyskusji i nie wpływają na moją wysoką ocenę merytoryczną dysertacji.

Uwagi krytyczne i pytania:

- W pracy Doktorant nie dokonał oceny jakości baz danych zastosowanych do trenowania sieci. Jakie kryteria oceny jakości baz danych można by wprowadzić lub zastosować?
- W jakich konkretnych przypadkach algorytmy oparte o *deep learning* nie są odpowiednie dla zastosowań w dziedzinie inżynierii biomedycznej? Czy można podać przykłady scenariuszy, w których inne metody są bardziej efektywne lub bezpieczniejsze?
- Autor w swoich rozważaniach przedstawił tezę, odnośnie wyższości *deep learningu* nad tradycyjnym podejściem przetwarzania danych ze względu na automatyzację całego procesu decyzyjnego. Czy zdaniem Autora należy dążyć do pełnej automatyzacji procesu decyzyjnego, czy też systemy AI powinny być nadzorowane przez ludzi ?
- Autor wskazał w pracy potencjał wdrożeniowy opracowanych rozwiązań. Czy w związku z tym Autor rozważał wymagania związane z procesem certyfikacji oprogramowania zawierającego sztuczną inteligencję?
- W pracy nie omówiono w sposób wystarczający ryzyk związanych z wykorzystaniem sztucznej inteligencji do wspomaganie decyzji diagnostycznych. Jakie potencjalne zagrożenia i ryzyka Autor uważa za istotne w kontekście zapewnienia odpowiedniego bezpieczeństwa systemów opartych o sztuczną inteligencję.
- Jakie kroki lub strategie można podjąć, aby zapewnić bezpieczeństwo systemów opartych na sztucznej inteligencji w procesie diagnostyki?
- Jakie wyzwania związane z prywatnością danych medycznych Autor zidentyfikował podczas prowadzonych badań? Jakie środki można podjąć, aby zapewnić odpowiednią ochronę prywatności danych pacjentów?

- Praca jest opracowana w sposób staranny i występują tylko drobne uchybienia redakcyjne np. nie do wszystkich obrazków odnalazłem odnośniki w tekście - Figure 1.6.

Reasumując, uważam że przedstawiona do recenzji praca stanowi oryginalne rozwiązanie postawionego problemu naukowego w postaci wykorzystania algorytmów uczenia głębokiego do przetwarzania danych biomedycznych takich jak: przewidywanie toksyczności molekuł, rozpoznawanie akcji kierowców na drodze bazując na sygnałe EOG, detekcja uderzeń serca w sygnałach sejsmograficznych oraz segmentacja składowych EKG. Autor zaproponował algorytm nazywany „uczeniem reprezentacji rozmytej”, który jest odpowiedzią na zidentyfikowane wady istniejących rozwiązań. Doktorant dowiódł, że posiada szeroką wiedzę w dziedzinie inżynierii biomedycznej oraz w sposób właściwy potrafi zaplanować i zrealizować badania związane z poruszaną tematyką. Wysoko oceniam potencjał wdrożeniowy uzyskanych wyników. Praca nie budzi zastrzeżeń pod względem formalnym i merytorycznym. Wykorzystane w pracy metody badawcze, sposób prezentacji i analizy wyników jak i ich dyskusja wskazują na umiejętność samodzielnego prowadzenia pracy naukowej. Część opracowanych baz danych i wyników dostępnych jest w otwartych repozytoriach danych, co jest dobrą praktyką i ułatwia reprodukcję uzyskanych przez Autora wyników.

W związku z powyższym uważam, że **rozprawa doktorska zatytułowana „Deep learning applications in biomedical engineering”** spełnia wymagania stawiane pracom na stopień doktora w rozumieniu aktualnie obowiązującej Ustawy - **Ustawa z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce** i wnioskuję o dopuszczenie mgr inż. Konrada Duraj do dalszych etapów przewodu doktorskiego.

Równocześnie z uwagi na wysokie walory merytoryczne i aplikacyjne dające szansę na wdrożenie uzyskanych wyników w cyfrowych usługach medycznych jak i znaczące osiągnięcia naukowe w postaci publikacji w czasopismach o wysokim współczynniku wpływu (m.in. Applied Sciences IF 2.7, Electronics IF 2.9), udziału w prestiżowych konferencjach naukowych (m.in. 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society) oraz dużej ilości cytowań (opublikowane prace posiadają już łącznie 16 cytowań i ponad 4960 wyświetleń) wnoszę wniosek o wyróżnienie pracy.



Bibliografia:

1. Duraj, K.M., Piaseczna, N.J. (2022). Predicting Molecule Toxicity Using Deep Learning. In: Piaseczna, N., Gorczowska, M., Łach, A. (eds) Innovations and Developments of Technologies in Medicine, Biology and Healthcare. EMBS ICS 2020. Advances in Intelligent Systems and Computing, vol 1360. Springer, Cham. https://doi.org/10.1007/978-3-030-88976-0_3
2. Doniec RJ, Sieciński S, Duraj KM, Piaseczna NJ, Mocny-Pachońska K, Tkacz EJ. Recognition of Drivers' Activity Based on 1D Convolutional Neural Network. Electronics. 2020; 9(12):2002. <https://doi.org/10.3390/electronics9122002>
3. Duraj KM, Siecinski S, Doniec RJ, Piaseczna NJ, Kostka PS, Tkacz EJ. Heartbeat Detection in Seismocardiograms with Semantic Segmentation. Annu Int Conf IEEE Eng Med Biol Soc. 2022 Jul;2022:662-665. doi: 10.1109/EMBC48229.2022.9871477. PMID: 36086330.
4. Duraj K, Piaseczna N, Kostka P, Tkacz E. Semantic Segmentation of 12-Lead ECG Using 1D Residual U-Net with Squeeze-Excitation Blocks. Applied Sciences. 2022; 12(7):3332. <https://doi.org/10.3390/app12073332>