

Prof. dr hab. inż. Daoud Robert Iskander
Katedra Inżynierii Biomedycznej
Wydział Podstawowych Problemów Techniki
Politechnika Wrocławska
Wybrzeże St. Wyspiańskiego 27
50-370 Wrocław

RECENZJA

pracy doktorskiej mgr inż. Konrada Duraja
p.t. „Deep learning applications in biomedical engineering”

Recenzja została opracowana na podstawie uchwały Rady Dyscypliny Inżynieria Biomedyczna Politechniki Śląskiej z dnia 22 czerwca 2023r. oraz listu Przewodniczącego Rady, prof. dr. hab. inż. Ewy Piętki.

1. Układ i treść rozprawy

Praca doktorska, napisana w języku angielskim, posiada 101 stron i podzielona jest na pięć rozdziałów. Praca przedstawiona jest w postaci monografii i ma nietypowy układ, w którym cztery publikacje stanowią jeden z jej środkowych rozdziałów. Pomijając załączone artykuły, praca zawiera 28 rysunków, w większości zapożyczonych z literatury oraz Internetu. W rozdziale pierwszym autor rozprawy przedstawia wstęp do pracy, krótko opisujący historię sztucznej inteligencji i omawiający, kolejno, systemy eksperckie, logikę rozmytą, uczenie maszynowe, oraz głębokie uczenie. W drugim rozdziale autor przedstawia cel pracy oraz opisuje sposób przedstawienia dorobku w rozdziale 3. Ten kolejny rozdział obejmuje cztery opublikowane prace stanowiące dorobek naukowy autora oraz zwięzły opis wkładu własnego w każdą z nich. Rozdział czwarty poświęcony jest alternatywnemu podejściu, łączącego głębokie uczenie z logiką rozmytą. W ostatnim piątym rozdziale pracy autor przedstawia wnioski. Końcowa część pracy zawiera bibliografię (65 pozycji), listę symboli i skrótów, streszczenia pracy w języku angielskim i polskim oraz krótkie CV autora.

Akceptuję koncepcję autora pracy co do jej układu, ale uważam, że praca mogła być napisana spójniej, na przykład, poświęcając więcej miejsca w rozdziale trzecim opisowi dorobku i umieszczając opublikowane prace na końcu pracy w postaci załączników.

2. Uwagi merytoryczne

2.1. Uwagi do rozdziału pierwszego

Praca została napisana poprawną angielszczyzną z małą ilością literówek. Rozdział pierwszy, pomimo jego wstępnego charakteru, zawiera dużo niedociągnięć redakcyjnych (np., istnienie lub brak spacji przed cytowaniem prac, różne konwencje cudzysłowów, różne style literatury/bibliografii) oraz merytorycznych. Opisując historię sztucznej inteligencji autor oparł się na jednym dobrze znanym źródle, cytując je, na przykład na pierwszej stronie pracy aż sześć

razy. Podobnie, do opisu systemów eksperckich autor odnosi się do jednego niezbyt trafnie dobranego źródła z 1986 roku zamieszczonego IEEE Journal of Oceanic Engineering. Kolejno, do opisu logiki rozmytej, autor odnosi się do strony internetowej (codecruck.com), z której to niefrasobliwie kopiuje, wraz z błędami, wszystkie zamieszczone w tym podrozdziale rysunki. W szczególności, Rys. 1.2 nie odpowiada równaniu 1.2 (zamieniony miejscami jest parametr b z parametrem c ; brak oznaczenia osi odciętych), Rys 1.4 nie odpowiada równaniu 1.4 (parametr σ nie zawiera się pomiędzy $0.1m$ i $0.9m$; dodatkowo ten ostatni punkt nie ma sensu, gdyż $0.9m$ nie może być większe niż m , zakładając $m > 0$). Rys 1.5 nie odpowiada równaniu 1.5 (punkty przecięcia funkcji w punktach $c - a$ oraz $c + a$ powinny wynosić 0,5; oznaczenia x oraz y w środku wykresu nie mają sensu; z rysunku wynika, że $b = 2a$, co nie jest trafnym przedstawieniem funkcji, gdyż oba te parametry mogą być niezależne). Dalej, odniesienie się do pozycji literaturowej [5], która jest półstronicową recenzją książki, której tematyka jest spójna z omawianym tematem w danym miejscu pracy, uważam za niestosowne. Odchodząc w tym miejscu od opisu rozdziału pierwszego należy wskazać, że bibliografia przedstawiona na stronach 91–96 nie jest spójna i stanowi konglomerat informacji automatycznie zaciąganych, na przykład, przez stronę Google Scholar. W mojej opinii, „bogactwo” stylu przedstawiania literatury świadczy o tym, że autor pracy nie dokonał w pełni przeglądu tych prac ani weryfikacji zasadności ich cytowania. Pojawianie się w cytowaniu pracy jej ceny jak i adresu poczty elektronicznej do działu sprzedaży powinno (przy nawet pobieżnym przeglądzie pracy) być sygnałem ostrzegawczym dla autora, że jest to nieodpowiednio sformatowana pozycja literatury. Do oceny bibliografii wróć w dalszej części mojej recenzji.

W opisie uczenia maszynowego autor odnosi się do dwóch źródeł literaturowych: popularnej pozycji autorstwa Burkova (w pozycji [7] jest literówka, „hundred” zamiast „hundred”), podejmując się streszczenia na dwóch i pół stronach tego, co Burkov trafnie dokonał na stu, oraz bardzo dobrego podręcznika autorstwa Shalev-Shwartz i Ben-Davida (pozycja [9]). Opis słowny dotyczący tematu uczenia maszynowego jest akceptowalny, biorąc pod uwagę wstępny charakter rozdziału. Niestety, kiedy autor przechodzi do opisu matematycznego metod, pojawiają się problemy natury formalnej, takie jak brak opisu poszczególnych oznaczeń (np. y_i oraz j w $x^{(j)}$ na stronie 7, y_i w równaniu 1.9).

Lepiej opisaną częścią pierwszego rozdziału jest ta poświęcona głębokiemu uczeniu, chociaż tu też fragmenty opisu matematycznego nie są przedstawione w sposób rygorystyczny (brak formalizmu w równaniach 1.13–1.15; brak oznaczenia i w równaniu 1.29; brak oznaczenia θ w równaniach 1.20 oraz 1.21; $\sim$ vs $\backsim$ (w zapisie LaTeX) w równaniach 1.22 i 1.23; kompletny brak oznaczeń w równaniach 1.22, 1.23 oraz 1.31; to samo oznaczenie L dla funkcji start i jej logarytmu w równaniach 1.20 oraz 1.21; brak wyjaśnienia co do parametrów 10000 oraz d_{model} w równaniach 1.29 oraz 1.30). Lepiej opisana jest część dotycząca sieci konwolucyjnych, chociaż to też formalizm dla wartości dyskretnych wymaga aby t w równaniu 1.26 było zastąpione przez t_n i podane w nawiasie kwadratowym. Dodatkowo, na stronie 17 „asterik” powinno być „asterisk”.

Podsumowując, stronę redakcyjną rozdziału pierwszego oceniam nisko. Wszystkie rysunki w tym rozdziale są zapożyczone, część z nich załączona jest w słabej rozdzielczości. Opis metod jest miejscami zdawkowy bez uwzględnienia formalizmu matematycznego wymaganego od prac naukowych w dziedzinie nauk technicznych.

2.2. Uwagi do rozdziału drugiego

Rozdział drugi to zaledwie półstronicowy opis celu pracy, który w tłumaczeniu brzmi: „Głównym celem tej pracy jest ocena różnych zastosowań, w których głębokie uczenie może być pomocne w inżynierii biomedycznej”. W mojej opinii, nie jest to dobrze zredagowany cel pracy doktorskiej, gdyż pomimo tego, że autor, w przedstawionych w kolejnym rozdziale pracach stanowiących jego dorobek naukowy, rozważa różne zastosowania uczenia maszynowego, to jednak nie dokonuje w nich formalnej oceny samego zastosowania. Celem pracy była więc ocena metod głębokiego uczenia w zastosowaniach, a nie ocena zastosowań.

2.3. Uwagi do rozdziału trzeciego

Pierwsza praca, jaką autor dołączył do dysertacji dotyczy przewidywania toksyczności cząsteczek przy użyciu głębokiego uczenia się. Praca została opublikowana w materiałach międzynarodowej konferencji studenckiej pod egidą IEEE. Pracę uznaje jako pierwsze kroki jakie autorzy poczynili w swoich karierach naukowych, gdyż wśród autorów nie znalazłem starszych doświadczonych naukowców. Praca opisuje pewne manipulacje typowe dla technik uczenia maszynowego, a związek jaki ma z dyscypliną inżynierii biomedycznej jest fakt użycia w niej publicznie dostępnego zbioru danych pewnych molekuł (kaggle.com). Zgodnie z deklarowanym udziałem autora dysertacji, kierują do niego następujące pytania:

1. Abstrakt/wyniki. Dokładność modelu 1 i 2. Zakładając notację U, W, oraz T, odpowiednio dla zbioru uczącego, walidującego oraz testowego, jak wytłumaczyć ponad 10% lepszą dokładność modelu 2 względem model 1 dla T, przy ponad 10% gorszej dokładności tego modelu dla U i W?
2. Abstrakt. Ile było nietoksycznych molekuł, 6698 czy 6998?
3. Wstęp. Ostatni akapit. Do jakich obecnych metod odnosi się krytyka? Brak odniesień do źródeł literaturowych.
4. Materiały i Metody. Jak wyodrębniono 4500 próbek z dostępnych 6760? Dlaczego 4500? W jaki sposób powielono 937 próbek by otrzymać 3240? Dlaczego 3240? Nie znajduje w pracy informacji na temat uzasadnienia wybranych przez autora wielkości prób.
5. Czym różni się zaproponowany przez autora model 1 od modelu Fanconi’ego, umieszczonego na kaggle.com wraz z danymi o molekułach?
6. Jak dobrano wartości wag w model 2? Dotyczy wag 3 i 0,5 odpowiednio dla klasy toksycznych molekuł i tych nietoksycznych.
7. Skąd się wzięła liczba parametrów równa 329665?
8. Rys 2. Jak uzasadnić potrzebę dwukrotnego powielenia bloków (batch_normalization, conv1D oraz max_pooling1d_1)?
9. Wyniki. Rys. 3 oraz 4. Co jest oznaczone kolorem czerwonym, a co niebieskim? Rysunki nie mają oznaczonych osi. Rys 4c nie przedstawia krzywej ROC. Z związku z tym, co on przedstawia?
10. Dlaczego wyniki modeli 1 i 2 nie zostały skonstrastowane z wynikami metod, które były krytykowane na wstępie?

Podsumowując, autor kwestionuje zasadność użycia innych metod nauczania, lecz krytyka ta nie odnajduje uzasadnienia w wynikach. Autor wykorzystuje publicznie dostępny zbiór danych nie zrównoważonych, w których potrzeba jest zastosowania metod rozwiązujących problem

nadmiernego dopasowania. Poszczególne konfiguracje i parametry modelu zostały w tej pracy dobrane empirycznie bez uwzględnienia systematycznego podejścia do rozwiązywania tego problemu.

Drugi artykuł załączony w rozdziale 3 to wieloautorska praca dotycząca rozpoznawania aktywności kierowców za pomocą uczenia maszynowego w oparciu o dane pochodzące z sygnałów elektroencefalograficznych (EEG) i elektrookulograficznych (EOG). Praca została opublikowana w międzynarodowym czasopiśmie *Electronics* (MDPI). Zgodnie z deklarowanym udziałem autora dysertacji (m.in. przetwarzanie wstępne, opracowanie modelu uczącego, trening i testowanie), kierują do niego następujące pytania:

1. Przy tak dużej rozpiętości wielkości próbek (od 320 do 32357), jak uzasadnić procedurę konkatenacji danych? W jakim stopniu procedura ta obciąża (z ang. *bias*) uzyskane wyniki?
2. Czy istnieją dowody, że podział 20/80 % danych na uczenie/walidację daje najlepsze wyniki? Praca nie zawiera tej informacji.
3. Dlaczego w pracy nie ma zbioru testującego?
4. Rys. 5 i 6 nie pokrywają się z informacją wcześniej zamieszczoną w tekście. Czy jest na to wytłumaczenie? Uważam, że rysunki te mogły zostać zastąpione 2. lub 3. linijkami tekstu.
5. Rys. 7 (generowany przez Keras). Słaba jakość oraz brak oznaczeń. Czy „batch size” jest „None” (rysunek), czy też 20 (tekst)?
6. Skąd się wzięło 64 map cech?
7. Dlaczego ustawiono wskaźnik „dropout” na 0,5?
8. Czemu zdecydowano się na 100 epok?
9. Rys. 10. Brak oznaczeń wartości na osi odciętych.
10. Rys. 11. Co oznaczają etykiety 0, 1, 2, 3?
11. Strona 12 artykułu. Cztery różne postacie czcionek dla „F1 score”.
12. Dyskusja została przedstawiona na podstawie danych od 20 kierowców (10 doświadczonych i 10 uczących się). To bardzo małe grupy. Różnice pomiędzy grupami nie są omówione. Dlaczego więc zdecydowano się na dwie grupy?

Podsumowując, autor sprawnie używa narzędzi programistycznych w celu maksymalizacji oczekiwanych wyników nauczania maszynowego, ale tu też wyraźny jest brak stosowania systematycznych metod naukowych. Opracowany model został zwalidowany, ale nie został przetestowany na danych niezależnych.

Trzeci artykuł załączony w rozdziale 3 jest również wieloautorską pracą dotyczącą detekcji bicia serca na podstawie sejsmokardiogramu z segmentacją semantyczną. Tu autor dysertacji jest pierwszym autorem i deklaruje dominujący udział w powstanie pracy. Praca została opublikowana w materiałach flagowej międzynarodowej konferencji naukowej w dyscyplinie inżynieria biomedyczna pod patronatem IEEE. Jakość prezentacyjna tego artykułu jest dużo lepsza od tych w omawianych wyżej dwóch artykułach. Jakość to przejawia się nie tylko w postaci lepszych rysunków (choć wcióż, na przykład, brakuje opisu osi oraz jednostek na Rys. 2), ale również w formalizmie matematycznym i całej strukturze artykułu. Co do samego modelu, zrealizowanego na bazie sieci U-net, tu też w dużej mierze dobrano jego parametry

w sposób empiryczny. Ważne jest jednak, że ta empiryka jest tu wyartykułowana i skomentowana. Dobrze skonstruowane są też wyniki tej pracy do tych uzyskanych przez Suresh et al. (pozycja [28] literatury), którzy użyli podobnej architektury do rozwiązania problemu estymacji bicia serca. Podsumowując, pracę przedstawioną w artykule trzecim można uznać jako znaczący dorobek do doktoratu autora dysertacji.

Czwarty artykuł załączony w rozdziale 3 jest kontynuuje temat detekcji bicia serca na podstawie sygnałów elektrokardiograficznych oraz użycia segmentacji semantycznej. To również praca wieloautorska opublikowana w *Applied Sciences* (MDPI). Podobnie jak w artykule trzecim, autor dysertacji deklaruje dominujący udział w powstanie pracy. Do tej pracy mam następujące pytania:

1. Dane zostały podzielone w stosunku 80/20 % dla U/T (nomenklatura jak wyżej). Następnie ze zbioru T wyodrębniono 20% dla W. Skąd taki podział? Rys. 2 nie potwierdza tego, co napisane w tekście.
2. Skąd wzięła się liczba 34305156 parametrów i dlaczego 34816 z nich nie nadawały się do uczenia?
3. Zakładając, że na osiach odciętych w Rys. 4–8 są epoki (tu znów brak oznaczeń wszystkich osi), dlaczego wyniki są pokazane dla 22 epok, skoro w tekście jest mowa o wczesnym zakończeniu procesu po 20 epokach?
4. Dlaczego swoistość i czułość (Rysunek 5) spada po 10–14 epokach dla zbioru walidującego?
5. Dlaczego współczynnik strat (Rysunek 7) rośnie po 10–14 epokach dla zbioru walidującego?
6. Jak wyglądałyby te przebiegi dla dalszych epok?

Podsumowując, pracę przedstawioną w artykule czwartym można uznać jako znaczący dorobek do doktoratu autora dysertacji. Niemniej, tu znów nie ma rygoru w prezentacji wyników naukowych.

2.4. Uwagi do rozdziału czwartego

Rola rozdziału czwartego pracy zatytułowanego „Nowy paradygmat” jest dla mnie niezrozumiała, choć dopiero w tym miejscu, zrozumiałe się staje opisanie logiki rozmytej w rozdziale pierwszym. Włączenie logiki rozmytej do uczenia maszynowego można nazwać nowym podejściem, ale lepiej byłoby go nazwać *alternatywnym*, gdyż pierwsze prace na ten temat zaczęły pojawiać się już prawie 10 lat temu. Rozdział czwarty stanowi więc kolejny przegląd literatury, mający najprawdopodobniej na celu opisanie nowych kierunków rozwoju uczenia maszynowego zmniejszające ograniczania obecnie rutynowo stosowanych metod. Styl i prezentacja tego rozdziału są równie słabe jak rozdziału pierwszego. Większość rysunków jest zapożyczonych, brak rygoru w opisie równań matematycznych, brak opisu oznaczeń. Cel podrozdziału 4.4.3, pokazującego przykładowe rozwiązanie, jest dla mnie niezrozumiały. Autor deklaruje, że pojęcie uczenie się w oparciu o reprezentację rozmytą jest nowe. Tak może było, ale już nie jest (praca Zhang et al. z marca 2023 r. [Zhang CY, Lin YN, Chen CP, Yao HY, Cai HC, Fang WP. Fuzzy Representation Learning on Graph. *IEEE Transactions on Fuzzy Systems*. 2023 Mar 6.]). Podsumowując, rozdział czwarty mógłby być całkowicie pominięty,

gdyż przykłady w nim zaprezentowane są zbyt wczesne by stanowiłyby znaczący wkład do pracy naukowej autora.

2.5. Uwagi do wniosków

We wnioskach autor wpierw podsumowuje to co już zostało napisane wcześniej w pracy, a następnie wylicza kilka ogólnych i oczywistych refleksji na temat głębokiego uczenia maszynowego. W mojej opinii, rozdział ten mógłby być zatytułowany *Podsumowanie*.

2.6. Uwagi do bibliografii

Wyraźny jest brak spójności w pozycjach literatury:

1. imiona autorów są podawane czasami w całości, czasami tylko inicjały,
2. w wielu pozycjach brak źródła publikacji,
3. niespójny opis stron, numerów i woluminów,
4. Niespójny opis pozycji książkowych,
5. niespójność w używaniu pierwszych dużych liter w nazwie czasopism,
6. niespójność w hiperlinkach (aktywny/nieaktywny).

3. Uwagi końcowe

Rozprawa opisuje ciekawe badania związane z zastosowaniem narzędzi uczenia maszynowego do kilku zastosowań biomedycznych. W podsumowaniu uwag merytorycznych, pomimo wyraźnych niedociągnięć, jako recenzent doceniam umiejętności kandydata do samodzielnego prowadzenia pracy naukowej oraz przedstawienia w pracy oryginalnych rozwiązań problemów naukowych, stanowiących dorobek kandydata. Jednocześnie, zwracam uwagę na powierzchowny opis tego dorobku, który nie oddaje sprawiedliwości wykonanej pracy naukowej.

Podsumowując stwierdzam, że recenzowana przeze mnie praca spełnia formalne wymagania stawiane pracom doktorskim w ustawie: art. 187 ust 1 i 2 Ustawy z dnia 20 lipca 2018 r. *Prawo o szkolnictwie wyższym i nauce* i wnoszę o dopuszczenie mgr. inż. Konrada Duraja do dalszych etapów przewodu doktorskiego.

Wrocław, 4 września 2023 r.

Podpisał Robert Iskander (podpis odręczny)