

Recenzja rozprawy doktorskiej dla Rady Naukowej Dyscyplin Matematyka i Informatyka Uniwersytetu Warszawskiego

Tytuł rozprawy: Deep Learning for Planning and Reinforcement Learning

Autor rozprawy: Konrad Czechowski

1 Analiza strony merytorycznej rozprawy

Rozprawę doktorską stanowi zbiór pięciu opublikowanych powiązanych tematycznie artykułów naukowych.

1.1 Obszar problemowy

Praca doktorska Konrada Czechowskiego koncentruje się na problematyce integracji głębokiego uczenia (ang. deep learning, DL) z algorytmami planowania i uczenia ze wzmocnieniem (ang. reinforcement learning, RL). Głównym celem jest rozwój nowych metod, które łączą te dwa obszary w celu stworzenia bardziej efektywnych i efektywnie uczących się systemów decyzyjnych. Kluczowe obszary problemowe, które porusza praca są następujące:

1. uczenie ze wzmocnieniem bazującym na modelu, gdzie proponuje się rozwinięcie i zastosowanie algorytmów MBRL, szczególnie w kontekście gier Atari 2600, wykorzystując modele środowiska do symulacji i planowania, co ma na celu poprawę efektywności uczenia poprzez zmniejszenie potrzebnej liczby interakcji z rzeczywistym środowiskiem,
2. mitygacja błędów modelu środowiska, gdzie badane są strategie radzenia sobie z niedoskonałościami modeli środowiska używanymi w MBRL, w szczególności z użyciem metod, które mogą zminimalizować negatywny wpływ błędnych predykcji modelu na zdolności eksploracyjne agenta,
3. planowanie hierarchiczne z użyciem generowania podcelów, gdzie wprowadzane są nowatorskie podejścia do planowania hierarchicznego z wykorzystaniem sieci neuronowych do generowania podcelów w zadaniach o dużej złożoności kombinatorycznej, co umożliwi systemom decyzyjnym efektywniejsze radzenie sobie w trudnych przestrzeniach stanów,
4. algorytmy wyszukiwania związane z Monte Carlo Tree Search (MCTS), gdzie praca rozwija i integruje metody wyszukiwania oparte na MCTS z głębokim uczeniem, dostosowując je do skomplikowanych zadań decyzyjnych, takich jak gry i symulacje,

-
5. wykorzystanie abstrakcji czasowych w planowaniu hierarchicznym, gdzie przedstawia się nowe metody wykorzystujące różne horyzonty czasowe w hierarchicznym planowaniu.

Podsumowując, przedstawiony obszar problemowy ma na celu nie tylko teoretyczne rozszerzenie wiedzy na temat uczenia maszynowego, ale również praktyczne zastosowania w automatyzacji, grach komputerowych, a także w bardziej ogólnych zadaniach wymagających skutecznego planowania i podejmowania decyzji.

1.2 Organizacja pracy

Zawartość pracy została zorganizowana w czterech rozdziałach. W pierwszym z nich krótko zaprezentowano listę publikacji, które zostały uwzględnione w rozprawie doktorskiej. W szczególności, wymieniono pięć publikacji, w czterech z nich autor zadeklarował, że był jednym z głównych autorów, którzy w równym stopniu wnieśli swój wkład.

Rozdział drugi wprowadza podstawowe koncepcje związane z poruszaną tematyką, które w sposób skrótowy umiejscawiają zakres merytoryczny zbioru publikacji. Należą do niego (1) uczenie maszynowe i uczenie głębokie, (2) uczenie ze wzmocnieniem i samodoskonalenie, ze wskazaniem wpływu głębokiego uczenia na rozwój algorytmów uczenia ze wzmocnieniem i przykładów zastosowań algorytmów głębokiego uczenia ze wzmocnieniem, (3) planowanie i przeszukiwanie, z odwołaniem do ich roli w systemach decyzyjnych opartych na głębokim uczeniu i prezentacją dwóch algorytmów często wykorzystywanych w połączeniu z sieciami neuronowymi do planowania i przeszukiwania (Best-First Search, Monte Carlo Tree Search) oraz (4) uczenie modeli dynamiki środowiska, które umożliwia planowanie nawet w przypadku braku pełnego dostępu do stanu środowiska. Rozdział ten stanowi dość skrótowy, ale wewnętrznie spójny wypis fundamentów dla badań przedstawionych w rozprawie, zapewniając niezbędny kontekst i wprowadzenie do kluczowych pojęć i metod używanych w pracy.

Rozdział trzeci stanowi opis głównego wkładu naukowego autora. Doktorant szczegółowo opisuje swoje osiągnięcia zawarte w pięciu publikacjach. Wkład autora obejmuje propozycje zarówno teoretyczne, jak i praktyczne, które w czasie publikacji stanowiły novum w modelowaniu dynamiki środowiska, rozwijały nowe algorytmy planowania oraz poprawiały efektywność i stabilność metod uczenia ze wzmocnieniem.

Rozdział czwarty zawiera teksty opublikowanych prac, które stanowią zbiór publikacji, t.j.,

- P1** Kaiser, L., Babaeizadeh, M., Miłoś, P., Osiński, B., Campbell, R. H., Czechowski, K., ... & Michalewski, H. (2019). Model Based Reinforcement Learning for Atari. In International Conference on Learning Representations (ICLR).
- P2** Czechowski, K., Odrzygóźdź, T., Izworski, M., Zbysiński, M., Kuciński, E., & Miłoś, P. (2021). Trust, but Verify: Alleviating Pessimistic Errors in Model-Based Exploration. In 2021 International Joint Conference on Neural Networks (IJCNN). IEEE.
- P3** Czechowski, K., Januszewski, P., Kozakowski, P., Kuciński, E., & Miłoś, P. (2021). Structure and Randomness in Planning and Reinforcement Learning. In 2021 International Joint Conference on Neural Networks (IJCNN). IEEE.
- P4** Czechowski, K., Odrzygóźdź, T., Zbysiński, M., Zawalski, M., Olejnik, K., Wu, Y., ... & Miłoś, P. (2021). Subgoal search for complex reasoning tasks. *Advances in Neural Information Processing Systems*, 34, 624-638 (NeurIPS).
- P5** Zawalski, M., Tyrolski, M., Czechowski, K., Odrzygóźdź, T., Stachura, D., Piękos, P., ... & Miłoś, P. (2022). Fast and Precise: Adjusting Planning Horizon with Adaptive Subgoal Search. In The Eleventh International Conference on Learning Representations (ICLR).

W pracy nie został zawarty rozdział stanowiący zwięzłe podsumowanie oraz dyskusję na temat przyszłości badań nad prezentowanymi modelami w odniesieniu do całości zbioru publikacji. O przyszłych pracach badawczych można jednakże przeczytać w każdej z publikacji w zbiorze z osobna. Ciekawym uzupełnieniem rozprawy byłoby zbiorcze podsumowanie i dyskusja uaktualniająca do bieżącego stanu wiedzy, która wyczerpałaby interakcje i powiązania wniosków płynących z prezentowanych w zbiorze publikacji.

1.3 Ocena wyników oraz stopnia ich oryginalności

W pracy doktorskiej Konrada Czechowskiego zaprezentowano szereg oryginalnych osiągnięć badawczych, które znacząco przyczyniają się do rozwoju dziedzin uczenia ze wzmocnieniem i głębokiego uczenia w kontekście złożonych zadań decyzyjnych. Oryginalne osiągnięcia badawcze przedstawione w doktoracie dotyczą zarówno opracowania nowych metod, jak i ich zastosowaniu w różnorodnych scenariuszach decyzyjnych.

Rozwiązania zaproponowane w pracy opierają się na innowacyjnych koncepcjach lub istotnych modyfikacjach w odniesieniu do metod znanych z literatury naukowej. Warto podkreślić dążenie do tworzenia rozwiązań, które charakteryzują się zarówno wysoką jakością, jak i efektywnością. W rozprawie można znaleźć wiele propozycji nowych rozwiązań, usprawnień, dodatków technicznych i implementacyjnych, które są wynikiem zgłębionej analizy i konsekwentnej derywacji naukowej. Poniżej wymieniam uzyskane rezultaty.

P1 Praca: Model-based reinforcement learning for Atari. W pracy został podjęty problemem wysokiej złożoności interakcji wymaganych przez metody typu model-free reinforcement learning do nauki skutecznych polityk działania w grach Atari. Typowe algorytmy model-free RL, takie jak DQN czy Proximal Policy Optimization (PPO), aby osiągnąć dobre wyniki wymagają milionów interakcji, co znacznie przewyższa potrzeby człowieka do nauki tych samych gier. W pracy zaproponowano podejście model-based RL, które znacząco redukuje liczbę wymaganych interakcji przy zachowaniu porównywalnej jakości. Zaproponowana została metoda Simulated Policy Learning (SimPLe), która jest algorytmem opartym na modelach predykcji wideo przyszłych stanów gry i nagród, co pozwala na uczenie polityki w symulowanym środowisku przy mniejszej liczbie interakcji z rzeczywistym środowiskiem. W bogatych eksperymentach przeprowadzonych na różnych grach Atari, SimPLe wykazało znacznie wyższą efektywność w porównaniu z podejściami model-free RL. W szczególności, w większości gier SimPLe osiągało lepsze wyniki po zaledwie 100 tysiącach interakcji z rzeczywistym środowiskiem, co odpowiada około dwóm godzinom gry w czasie rzeczywistym. Dla niektórych gier, takich jak Freeway, SimPLe było ponad 10 razy bardziej efektywne w próbkowaniu niż algorytmy model-free, takie jak Rainbow i PPO. Co więcej SimPLe nie tylko przewyższyło podejścia model-free RL w zakresie efektywności, ale również osiągnęło wyniki porównywalne lub lepsze w kilku grach. Pokazuje to potencjał metody do nauki złożonych polityk działania przy znacznie mniejszej liczbie interakcji z rzeczywistym środowiskiem .

P2 Praca: Trust, but Verify: Alleviating Pessimistic Errors in Model-Based Exploration. Publikacja koncentruje się na problemie błędów modeli środowiskowych w podejściach model-based reinforcement learning (MBRL). Szczególnym wyzwaniem są tzw. "fałszywe pętle" (false loops), gdzie model błędnie przewiduje, że określone akcje nie zmieniają stanu środowiska. Tego typu błędy prowadzą do tego, że agent unika eksploracji pewnych stanów, co ogranicza jego zdolność do nauki optymalnych strategii. Aby rozwiązać ten problem zaproponowano mechanizm "Trust but Verify" (TBV), który wykorzystuje oszacowania niepewności modelu do kierowania eksploracją. Główne elementy TBV to oszacowanie niepewności modelu do identyfikacji stanów, w których przewidywania modelu mogą być błędne, mechanizm weryfikacji

priorytetyzujący wizyty w stanach niepewnych w celu oceny przewidywań modelu oraz mechanizm planujący np. MCTS lub BestFS. Można powiedzieć, że TBV wykorzystuje testowanie hipotez statystycznych, aby zdecydować, czy zaufać decyzjom planera, czy przeprowadzić dodatkową eksplorację. Dzięki temu TBV pomaga agentom efektywniej eksplorować przestrzeń stanów, zwłaszcza w obecności błędów modelu. Eksperymenty przeprowadzone w dwóch wymagających środowiskach z rzadkimi nagrodami, ToyMontezumaRevenge i Tower of Hanoi, wykazały, że TBV znacząco poprawia zdolności eksploracyjne agentów.

P3 Praca: Structure and Randomness in Planning and Reinforcement Learning. Publikacja ta skupia się na optymalizacji procesu planowania w uczeniu ze wzmocnieniem poprzez znalezienie równowagi między strukturą a losowością w algorytmach przeszukiwania. Tradycyjne algorytmy, takie jak Monte Carlo Tree Search (MCTS), mają ograniczenia wynikające z kompromisu między głębokością a szerokością przeszukiwania. Kluczowym wyzwaniem było stworzenie metody, która efektywnie balansuje prowadząc do bardziej efektywnego uczenia się i planowania. Zaproponowana metoda Shoot Tree Search (STS) integruje strukturalne aspekty MCTS z losowym przeszukiwaniem, aby lepiej zarządzać kompromisem między eksploracją a eksploatacją przestrzeni rozwiązań. W przeciwieństwie do MCTS, który rozwija pojedynczy liść w każdym kroku, STS rozwija wiele węzłów jednocześnie przez losowe rollouts, co pozwala na głębszą eksplorację przy mniejszej liczbie iteracji przeszukiwania. Wprowadzona losowość jest odpowiednio kontrolowana i zwiększa szanse na znalezienie optymalnych ścieżek. Eksperymenty przeprowadzone w domenach o wysokiej złożoności, takich jak łamigłówka Sokoban oraz zestaw zadań z Google Research Football (GRF), wykazały, że STS konsekwentnie przewyższa zarówno MCTS, jak i działanie losowe w prawie wszystkich testowanych zadaniach. Ponadto STS okazał się być bardziej skalowalnym i efektywnym podejściem do planowania w dużych i złożonych przestrzeniach stanów, oferując lepszą wydajność przy niższych kosztach obliczeniowych niż tradycyjne metody przeszukiwania.

P4 Praca: Subgoal Search for Complex Reasoning Tasks. W pracy skupiono się na opracowaniu metody przeszukującej złożone przestrzenie stanów, która miałaby zmniejszać koszty obliczeniowe znane z tradycyjnych metod, takich jak Best-First Search (BFS) czy Monte Carlo Tree Search (MCTS). Została przedstawiona metoda Subgoal Search (kSubS), której kluczowym elementem jest generator podcelów osiągalnych i zbliżających agenta do ostatecznego rozwiązania. Podejście to łączy modelowanie generatywne z klasycznymi algorytmami przeszukiwania, co pozwala na efektywne planowanie na wysokim poziomie. Metoda kSubS została zaimplementowana w dwóch wersjach: MCTS-kSubS (z użyciem Monte Carlo Tree Search) i BF-kSubS (z użyciem Best-First Search). Obie wersje korzystają z modelu transformatora do generowania podcelów oraz funkcji wartości do kierowania przeszukiwaniem. Weryfikację metody przeprowadzono na trzech domenach o wysokiej złożoności: grach logicznych Sokoban i Kostka Rubika oraz benchmarku dowodzenia nierówności (INT), a wyniki wykazały wysoką skuteczność metody kSubS.

P5 Praca P5: Fast and precise: adjusting planning horizon with adaptive subgoal search. W ostatniej z publikacji w zbiorze skoncentrowano się na problemie dostosowywania horyzontu planowania, w którym występują stany o różnej złożoności – wymagające różnych poziomów szczegółowości w planowaniu. Stały horyzont planowania może prowadzić do nieefektywnego przeszukiwania, zwłaszcza gdy środowisko jest zmienne pod względem trudności. W celu rozwiązania tego problemu został zaproponowany nowy algorytm o nazwie Adaptive Subgoal Search (AdaSubS), który dynamicznie dostosowuje horyzont planowania w zależności od lokalnej złożoności środowiska. Metoda realizuje generowanie podceli na różnych odległościach, co pozwala na elastyczne dostosowanie horyzontu planowania, a mechanizm weryfikujący, który ocenia, czy proponowany podcel jest osiągalny, pozwala na szybkie odrzucenie

nieprawidłowych podcelów. Wykorzystano również politykę (CLLP) umożliwiającą przejście między podcelami oraz mechanizm oceny odległości między stanem bieżącym a celem, co pomaga w priorytetyzacji rozszerzania węzłów. Algorytm AdaSubS adaptacyjnie wybiera podcele o różnych horyzontach, preferując dalsze podcele, jeśli są osiągalne, a w przeciwnym razie cofając się do bliższych, bardziej konserwatywnych podcelów. Dzięki temu możliwe jest szybkie przeszukiwanie łatwiejszych obszarów i dokładniejsze przeszukiwanie trudniejszych obszarów. Rozwiązanie zostało przetestowane podobnie jak w poprzedniej publikacji w trzech domenach: Sokoban, Kostka Rubika oraz dowodzenie nierówności (INT). Wyniki eksperymentów wykazały, że AdaSubS znacząco przewyższa algorytmy hierarchicznego planowania oraz ustala nowe standardy w benchmarku INT.

Rezultaty te zostały opublikowane na bardzo dobrych, wiodących, konferencjach o renomie międzynarodowej (ICLR, NeurIPS, IJCNN). Jest to niewątpliwie wielkie osiągnięcie. Należy zaznaczyć, że publikacje wchodzące w skład rozprawy zostały napisane przez wielu autorów (kolejne publikacje miały **P1** – 14, **P2** – 6, **P3** – 5, **P4** – 8, **P5** – 9 autorów). Stanowi to pewną trudność w ocenie samodzielności i zakresu merytorycznego wkładu naukowego Doktoranta w powstanie poszczególnych publikacji. Dlatego wparłem się opisem zaangażowania Doktoranta w powstanie poszczególnych publikacji, deklarowanego w sekcji "Significance and Author Contributions Summary" rozprawy:

- P1** zajmował się implementacją eksperymentów, refaktoryzacją, debugowaniem i rozwijaniem podstawowej implementacji algorytmu SimPLe; przeprowadził analizę danych i prezentację wyników, w tym przygotowanie tabel i przykładów opisanych w dodatku do artykułu; był także odpowiedzialny za tworzenie wewnętrznego narzędzia do badania jakości modelu środowiska SimPLe, używanego do debugowania,
- P2** pełnił rolę lidera projektu, koordynując wszystkie wysiłki badawcze, w tym implementację, analizę i projektowanie algorytmu Trust but Verify (TBV),
- P3** współprojektował główny algorytm Shoot Tree Search (STS), a także implementował i przeprowadzał eksperymenty w domenie gry Sokoban,
- P4** był głównym współautorem, kierując i koordynując badania; wkład obejmował strategiczne planowanie badań, przegląd literatury, projektowanie lub współprojektowanie wszystkich proponowanych ulepszeń algorytmicznych oraz codzienne kierowanie wszystkimi wątkami eksperymentalnymi,
- P5** deklarowana przez Doktoranta rola w publikacji jest taka sama jak w **P4**.

Tym samym, w tekście doktoratu została zawarta informacja na temat roli i udziału Doktoranta w procesie powstania rezultatów, ale w żadnym z oryginalnych manuskryptów publikacji nie odnalazłem tych informacji. Bazując na posiadanej wiedzy można jednakże podsumować, że wkład Doktoranta w przygotowanie i publikację pięciu artykułów **P1–P5** był znaczący, obejmując szeroki zakres działań od rozwijania podstawowych algorytmów, przez analizę danych, po kierowanie badaniami. Każdy z artykułów przyczynił się do zaawansowania wiedzy w dziedzinie sztucznej inteligencji, a Doktorant odegrał znaczącą rolę w osiągnięciu prezentowanych rezultatów. Praca ta była realizowana we współpracy z bardzo dobrymi naukowcami. Całość potwierdza, że rola Doktoranta jest wyraźna i stanowiła główną siłę napędową powstania prac.

1.4 Zagadnienia dyskusyjne

Zawarte poniżej uwagi i pytania nie mają na celu krytykowania prezentowanej rozprawy, a stanowią zaproszenie do dyskusji i dalszej refleksji nad przedstawionymi w pracy zagadnieniami.

-
- W pracy prezentującej metodę SimPLc (P1) została zgłoszona jej wada, która wiązała się z wysoką wariancją uzyskiwanych wyników w wielokrotnych próbach tych samych gier. Jakie konkretne czynniki mogą przyczyniać się do wysokiej wariancji tych wyników i jakie podejścia można zastosować, aby zmniejszyć tę wariancję, zwłaszcza w kontekście modeli dynamicznych i polityki zbierania danych?
 - Obecny stan nauki cechuje ciągły brak w pełni opisanych uniwersalnych podstaw teoretycznych wyjaśniających działanie metod uczenia i wnioskowania dla architektur głębokich. Stwarza to ograniczenie w analitycznym wykazywaniu własności i ograniczeń poszczególnych zaproponowanych w rozprawie podejść, w szczególności w kontekście złożoności obliczeniowej. Powstaje zatem pytanie jakie kroki możnaby podjąć w celu przeprowadzenia pogłębionych derywacji analitycznych dotyczących optymalności dla np. jednej z wybranych metod zaprezentowanych w rozprawie? Jakie konkretne własności i ograniczenia tej metody możnaby obliczyć? W jaki sposób Doktorant skonstruowałby wywód analityczny i formułował ograniczenia. Proszę o przedstawienie koncepcji przeprowadzenia takiego badania, biorąc pod uwagę współczesne wyzwania i oczekiwania w tej dziedzinie. Ciekawym byłoby odniesienie się tutaj do dyskusji na temat optymalności uzyskiwanych wyników np. w pracy P5 (epirycznie zostało to zweryfikowane w porównaniu z BFS).
 - Czy na etapie badań rozważany był wpływ różnorodności danych treningowych na efektywność MBRL? W jaki sposób różnorodność i jakość danych treningowych wpływa na skuteczność i niezawodność modeli środowiska w MBRL?

2 Analiza strony formalnej rozprawy

2.1 Ocena układu pracy i redakcji manuskryptu

Rozprawa doktorska napisana jest w języku angielskim i pod względem edytorskim jest opracowana starannie. Zawiera streszczenie w języku angielskim, spis treści, rozdział listujący publikacje należące do zbioru, rozdział wprowadzający do podstawowych zagadnień z bazowym przeglądem literatury oraz rozdział składający przedruki artykułów. Bibliografia zawiera 59 wymienionych i nieuporządkowanych alfabetycznie, wolnych od błędów pozycji bibliograficznych. Większość prac w bibliografii została opublikowana w ciągu ostatnich kilku lat, co świadczy o aktualności tematyki rozprawy na arenie międzynarodowej. Bibliografia jest również zamieszczona w każdej z publikacji należących do zbioru.

Układ rozprawy jest właściwy, aczkolwiek praca zyskałaby na walorze w przypadku uzupełnienia o rozdział stanowiący zwięzłe podsumowanie oraz dyskusję na temat przyszłości badań nad prezentowanymi modelami w odniesieniu do całości prezentowanego dorobku ze zbioru publikacji. Jest to jednakże nieistotny szczegół, w znikomym stopniu wpływający na przejrzystość pracy.

Użyta w pracy terminologia jest właściwa dla analizowanego obszaru badań, obejmującego uczenie głębokie, planowanie i uczenie ze wzmocnieniem. Autor szczegółowo opisuje zastosowane metody badawcze, co jest kluczowe dla oceny naukowej i możliwości reprodukcji wyników. Dokładny opis eksperymentów i algorytmów pozwala na głębsze zrozumienie pracy i podkreśla jej naukową wartość.

Praca posiada łącznie 127 stron, numerowanych osobno dla każdej z publikacji oraz dla części wprowadzającej.

2.2 Uwagi szczegółowe

Większość treści rozprawy jest poprawna pod względem językowym, stylistycznym i merytorycznym. W trakcie czytania zauważyłem jednak kilka drobnych błędów redakcyjnych i fragmentów, które w

toku dodatkowej weryfikacji możnaby poprawić. Kilukrotnie pominięte zostały przecinki, a inne pomniejsze błędy to:

- str. 6 "In 2013, [37] developed a DRL agent that performed at a human level in some of the Atari 2600 games." → "In 2013, a DRL agent was developed [37] that performed at a human level in some of the Atari 2600 games."
- str. 12 "if an agent undervalues certain types of behaviors, it may never reevaluate its knowledge and and remain stuck..." → "if an agent undervalues certain types of behaviors, it may never reevaluate its knowledge and remain stuck..."

3 Komentarz do oceny rozprawy

Ewaluacja rozprawy doktorskiej, która składa się z kilku wieloautorskich publikacji, stanowi wyzwanie dla recenzenta. Główną trudnością jest sprawiedliwe odczytanie i przypisanie wkładu Doktoranta w powstanie konkretnej koncepcji, metodologii czy wyników. Ten problem nabiera szczególnego znaczenia w kontekście rozprawy doktorskiej Konrada Czechowskiego, składającej się z pięciu wysokiej jakości publikacji naukowych, jednak wieloautorskich.

Jednym z najważniejszych aspektów, na które zwraca się w takim przypadku uwagę jest indywidualny wkład kandydata na doktora. W rozprawie Konrada Czechowskiego, mamy do czynienia z dokładnie opisanym wkładem w każdą z publikacji, co jest dobrą praktyką. Doktorant brał aktywny udział nie tylko w implementacji i analizie danych, ale również w projektowaniu eksperymentów i koncepcji badawczych, co jest kluczowe dla uznania jego autorskiego wkładu w osiągniętych rezultatach.

Podsumowując, ewaluacja rozprawy doktorskiej mgr. inż. Konrada Czechowskiego, składającej się ze zbioru publikacji wieloautorskich, podkreśla potrzebę szczegółowego i przejrzystego dokumentowania wkładu każdego z autorów. Dotyczy to również opisów wkładów już na etapie przygotowywania tekstów manuskryptów i publikowania ich. Ilustruje to również, jak ważne jest, aby procesy oceny były dostosowane do współczesnych realiów naukowych, które charakteryzują się współpracą, często między różnymi dziedzinami i kulturami badawczymi.

4 Konkluzja

Przedstawiona praca doktorska, autorstwa mgr. inż. Konrada Czechowskiego, nie tylko spełnia, ale wręcz przewyższa kryteria określone przez Ustawę z dnia 20 lipca 2018 r. - Prawo o szkolnictwie wyższym i nauce (Dz. U. 2018 poz. 1668). Praca doktorska ukazuje głęboką znajomość dziedziny uczenia ze wzmocnieniem oraz planowania i pozwala wywnioskować o wysokich umiejętnościach samodzielnego prowadzenia pracy naukowej przez Doktoranta. Praca koncentruje się na innowacyjnych odpowiedziach na kilka naukowych zagadnień, które łączy wspólny cel: rozwijanie i integracja metod głębokiego uczenia z algorytmami planowania i uczenia ze wzmocnieniem, w celu poprawy ich efektywności i zdolności do rozwiązywania złożonych problemów. Jest to praca mająca realny wpływ na dziedzinę, co można zaobserwować w licznych publikacjach Autora na uznanych międzynarodowych konferencjach z dziedziny informatyki. Przedstawione w pracy wyniki bardzo dobrze potwierdzają słuszność zgłoszonych przez Autora koncepcji oraz efektywność ich zastosowania w praktyce. Wszelkie uwagi krytyczne, które pojawiły się w recenzji, nie wpływają na moją jednoznaczną i bardzo pozytywną ocenę rozprawy, zarówno pod kątem oryginalności, jak i wartości merytorycznej. Określenie zadań, proponowane innowacyjne metody ich rozwiązania, realizacja badań i kompetencje w formułowaniu wniosków dowodzą, że Autor jest świetnie przygotowany do dalszej działalności naukowej. Bazując na tych przesłankach wnioskuję by mgr inż. Konrad Czechowski był dopuszczony do

kolejnych etapów postępowania. Dodatkowo, ze względu na wysoką jakość publikacji, potwierdzoną wysoką akceptacją przez liczne cytowania oraz prestiżowe miejsca ich zamieszczenia, proponuję wyróżnienie jego pracy doktorskiej.

Kejda

UNIwersytet Warszawski
BIURO RAD NAUKOWYCH

2024 -06- 05

WPLYNEŁO

L.dz. *122* Podpis. *Kejda*