dr hab. inż. Tomasz Kajdanowicz, prof. uczelni
Katedra Sztucznej Inteligencji
Wydział Informatyki i Telekomunikacji
Politechnika Wrocławska

Wrocław, 09.06.2026

Recenzja rozprawy doktorskiej

dla Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Uniwersytetu Jagiellońskiego w Krakowie

Tytuł rozprawy: *Computationally Efficient Neural Networks (Obliczeniowo efektywne sieci neuronowe)*

Autor rozprawy: mgr inż. Bartosz Wójcik

Promotor: dr hab. Marek Śmieja

Dyscyplina naukowa: informatyka techniczna i telekomunikacja

Charakter rozprawy: cykl powiązanych tematycznie publikacji naukowych (pięć artykułów)

1 Analiza strony merytorycznej rozprawy

1.1 Obszar problemowy i temat rozprawy

Rozprawę doktorską mgr. inż. Bartosza Wójcika stanowi cykl pięciu powiązanych tematycznie publikacji naukowych, opatrzonych spójnym tekstem ramowym o objętości 55 stron, integrującym poszczególne artykuły w jednolitą narrację badawczą. Wszystkie pięć prac zostało opublikowanych w czołowych światowych konferencjach i czasopismach z zakresu uczenia maszynowego (NeurIPS, AAAI, ECML PKDD oraz Neural Networks), co już na wstępie sytuuje dorobek w najwyższym przedziale jakościowym dyscypliny.

Rozprawa koncentruje się na problemie efektywności obliczeniowej głębokich sieci neuronowych, a ściślej na redukcji kosztu obliczeniowego modeli przy zachowaniu ich skuteczności predykcyjnej. Doktorant trafnie formułuje problem jako zadanie optymalizacji wielokryterialnej polegające na przesuwaniu frontu Pareto w przestrzeni kompromisu „koszt obliczeniowy vs. jakość", a jako wiodący paradygmat przyjmuje obliczenia warunkowe (conditional computation), w których graf obliczeniowy modelu dostosowuje się dynamicznie do trudności konkretnego wejścia.

Temat jest istotny naukowo i wybitnie aktualny. Rosnące koszty energetyczne, sprzętowe i środowiskowe trenowania oraz wdrażania dużych modeli stanowią obecnie jedno z centralnych wyzwań sztucznej inteligencji, o bezpośrednich implikacjach dla dostępności technologii i jej skalowalności. Znaczenie naukowe pracy nie wynika jednak z samej popularności tematu, lecz z konkretnego kierunku obranego przez Doktoranta: zamiast statycznych technik kompresji (przycinanie, kwantyzacja, destylacja), które stosują tę samą liczbę operacji do każdego wejścia, rozprawa systematycznie identyfikuje i eliminuje formy redundancji wykraczające

poza nadmiarowość parametrów, redundancję wzdłuż głębokości oraz wzdłuż szerokości sieci. To rozróżnienie organizuje całą rozprawę i nadaje jej wyraźną oś koncepcyjną.

1.2 Organizacja pracy i przegląd literatury

Tekst ramowy jest zorganizowany w pięć rozdziałów. Rozdział 1 (Wprowadzenie) przedstawia motywację, statyczne metody kompresji, paradygmat obliczeń warunkowych oraz formalnie definiuje cel badawczy jako minimalizację pary (koszt, -jakość). Rozdział 2 omawia redundancje wzdłuż głębokości modelu i odpowiadające im prace [I–III]. Rozdział 3 dotyczy dynamicznie regulowanej szerokości modelu i prac [IV–V]. Rozdział 4 prezentuje pozostały dorobek naukowy oraz aktywności Doktoranta, a rozdział 5 zawiera wnioski, ograniczenia i kierunki przyszłych badań. Pełne teksty pięciu publikacji zamieszczono w Załączniku A.

Struktura ta jest przejrzysta, logiczna i konsekwentna. Podział na dwie części - głębokość i szerokość sieci - jest naturalny, dobrze umotywowany i pozwala czytelnikowi zrozumieć każdy artykuł jako odpowiedź na zidentyfikowaną wcześniej formę redundancji. Każdy podrozdział ma powtarzalną, czytelną kompozycję (Motywacja → Metody → Wyniki → Podsumowanie → Wkład), co znacząco ułatwia lekturę i ocenę. W przeciwieństwie do wielu rozpraw opartych na cyklu publikacji, narracja metodologiczna nie sprawia wrażenia sztucznie nałożonej ex post, a dwuczęściowy podział wynika organicznie z natury rozważanych problemów.

Przegląd literatury jest prowadzony kompetentnie, choć co typowe i uzasadnione dla formatu cyklu, jest rozproszony po częściach poszczególnych rozdziałów oraz wewnątrz załączonych artykułów. Doktorant wykazuje dobrą orientację w aktualnym stanie badań nad obliczeniami warunkowymi, wczesnymi wyjściami, modelami Mixture-of-Experts oraz rzadkością aktywacji. Bibliografia tekstu ramowego jest aktualna i adekwatna. Pewnym mankamentem jest to, że tekst ramowy nie zawiera osobnego, przekrojowego rozdziału przeglądowego, który usytuowałby cały cykl na wspólnej mapie metod efektywności obliczeniowej i wyraźniej rozgraniczył wkład własny od dorobku innych zespołów; jest to jednak uwagą drugorzędną, niewpływająca na ocenę wartości merytorycznej.

1.3 Ocena publikacji, wykazania wymogów ustawowych i wkładu Doktoranta

Poniżej oceniam kolejno każdą z pięciu publikacji wchodzących w skład cyklu. Dla każdej pracy analizuję jej oryginalność i znaczenie naukowe, deklarowany wkład Doktoranta oraz to, czy praca przyczynia się do poszerzenia ogólnej wiedzy teoretycznej w dyscyplinie, potwierdza umiejętności samodzielnego prowadzenia pracy naukowej oraz jest wynikiem oryginalnego rozwiązania problemu naukowego.

Praca [I] Finding the Optimal Network Depth in Classification Tasks (ECML PKDD 2020; CORE: A; 140 pkt MEiN)

Współautorzy: Wójcik, Wołczyk, Bałazy, Tabor.

Praca proponuje metodę wyznaczania optymalnej głębokości sieci w trakcie treningu, bez potrzeby kosztownego przeszukiwania architektur (NAS). Do warstw ukrytych dołączane są głowice klasyfikujące, a ich predykcje są agregowane jako ważona suma logarytmów prawdopodobieństw z wagami w_k podlegającymi uczeniu. Autorzy wykazują z odwołaniem do

nierówności Jensena, że zaproponowana funkcja straty jest minimalizowana, gdy dokładnie jedna waga dąży do jedności, co pozwala jednorazowo „odciąć" zbędne warstwy. Dodatkowy człon regularyzacyjny steruje kompromisem koszt-jakość. Oryginalność polega na eleganckim, analitycznie uzasadnionym mechanizmie selekcji głębokości, tańszym i prostszym w implementacji niż NAS.

Wkład Doktoranta: Doktorant jest pierwszym autorem i deklaruje opracowanie schematu agregacji logarytmów prawdopodobieństw i członu regularyzacyjnego, implementację metody, przeprowadzenie wszystkich eksperymentów oraz analiz (podobieństwo kosinusowe gradientów, zbieżność wag w_k) i istotny udział w redakcji manuskryptu. Opis jest konkretny i wskazuje na wiodącą rolę intelektualną i wykonawczą.

Praca [II] Zero Time Waste: Recycling Predictions in Early Exit Neural Networks (NeurIPS 2021; CORE: A*; 200 pkt MEiN)

Współautorzy: Wołczyk†, Wójcik†, Bałazy, Podolak, Tabor, Śmieja, Trzciński (†równy wkład).

Praca wprowadza metodę wczesnych wyjść Zero Time Waste (ZTW), w której predykcje wcześniejszych klasyfikatorów wewnętrznych (IC) nie są odrzucane, lecz ponownie wykorzystywane przez kolejne klasyfikatory poprzez połączenia kaskadowe oraz ważone uśrednianie geometryczne z balansowaniem klas. Autorzy identyfikują dwa źródła nieefektywności (1) marnotrawstwo predykcji oraz (2) zjawisko „przemyślenia" (overthinking) i wprowadzają oryginalną metrykę Hindsight Improvability mierzącą utracone informacje. Metoda osiąga wyniki state-of-the-art, a praca jest pierwszą stosującą wczesne wyjścia w uczeniu ze wzmocnieniem. Praca odnotowała znaczący oddźwięk (85 cytowań wg Google Scholar), a wprowadzona metryka jest adoptowana przez kolejnych badaczy.

Wkład Doktoranta: Doktorant deklaruje współautorstwo wiodące (równy wkład z M. Wołczykiem), a jego rola i wkład to konceptualizacja, przegląd literatury, implementacja i testy metod bazowych, wprowadzenie i implementacja techniki połączeń kaskadowych, badania eksperymentalne i analiza decyzji projektowych oraz konceptualizacja i implementacja eksperymentów z uczeniem ze wzmocnieniem.

Praca opublikowana na konferencji A* o najwyższej randze, prezentująca oryginalne rozwiązanie problemu naukowego (recycling predykcji). Oznaczenie równego wkładu dwóch pierwszych autorów oznacza, że dla pełnej przejrzystości postępowania celowe byłoby precyzyjne rozgraniczenie, które z elementów (kaskady, ensembling, metryka HI, eksperymenty RL) są autorskim wkładem Doktoranta. Z opisu wynika, że połączenia kaskadowe oraz eksperymenty RL są przypisane wprost Doktorantowi, co stanowi wyraźny, identyfikowalny wkład.

Praca [III] Zero Time Waste in Pre-trained Early Exit Neural Networks (Neural Networks, 2023; 200 pkt MEiN)

Współautorzy: Wójcik, Przewięźlikowski, Szatkowski, Wołczyk, Bałazy, Krzepkowski, Podolak, Tabor, Śmieja, Trzciński.

Praca jest rozszerzeniem [II] na scenariusze uczenia transferowego oraz modele pre-trenowane, dotychczas pomijane w badaniach nad sieciami wielowyjściowymi. Wnosi pogłębioną analizę umieszczenia i architektury klasyfikatorów wewnętrznych, eksperyment z destylacją IC, analizę miskalibracji oraz znacznie poszerzoną ewaluację na różnych modalnościach danych. Stanowi dojrzałe, czasopismowe domknięcie wątku wczesnych wyjść.

Wkład Doktoranta: Doktorant jest pierwszym autorem i poza wkładem przeniesionym z wersji konferencyjnej deklaruje przegląd prac pokrewnych, dobór metod bazowych, implementację metody L2W, rozszerzoną ewaluację w wizji komputerowej, autorskie eksperymenty (destylacja IC, umiejscowienie i architektura IC, analiza miskalibracji), przygotowanie większości manuskryptu oraz prowadzenie całego procesu zgłoszeniowego i korespondencji z recenzentami.

Jest to praca, w której rola pierwszoautorska i samodzielność Doktoranta są udokumentowane przekonująco od koncepcji, przez implementację i eksperymenty, po zarządzanie procesem publikacyjnym. Doskonale dokumentuje umiejętność samodzielnego prowadzenia pracy naukowej.

Praca [IV] Adaptive Computation Modules: Granular Conditional Computation for Efficient Inference (AAAI 2025; CORE: A*; 200 pkt MEiN)

Współautorzy: Wójcik, Devoto, Pustelnik, Minervini, Scardapane.

Praca wprowadza Adaptive Computation Modules (ACM), generyczny moduł dostosowujący nakład obliczeń do pojedynczego tokena wejściowego, co skutkuje przestrzenną adaptacyjnością inferencji. Zaproponowano schemat inicjalizacji wag umożliwiający „ACM-izację" istniejących modeli statycznych, dwie funkcje pomocnicze (L_e, L_d) zapobiegające kolapsowi do trybu statycznego oraz wydajną implementację w języku Triton. Metoda osiąga front Pareto względem trzech silnych metod bazowych (A-ViT, MoEfication, ZTW) na zadaniach wizji i mowy.

Wkład Doktoranta: Doktorant deklaruje, że wyznaczył kierunek badawczy projektu, zaproponował architekturę ACM, schemat inicjalizacji wag oraz funkcje straty L_e i L_d, stworzył wspólną bazę kodu, przeprowadził eksperymenty w wizji komputerowej, zaimplementował wydajny wariant w Tritonie oraz wykonał większość analiz, odgrywając wiodącą rolę w redakcji artykułu. Jest to opis jednoznacznie wskazujący na dominującą rolę intelektualną i wykonawczą.

Praca na konferencji A*, w której Doktorant pełni rolę lidera naukowego i pierwszego autora stanowi oryginalne rozwiązanie problemu (granularna, przestrzennie adaptacyjna inferencja) wraz z weryfikacją sprzętową i silnie dokumentuje zarówno oryginalność, jak i samodzielność.

Praca [V] Exploiting Activation Sparsity with Dense to Dynamic-k Mixture-of-Experts Conversion (NeurIPS 2024; CORE: A*; 200 pkt MEiN)

Współautorzy: Szatkowski†, Wójcik†, Piórczyński†, Scardapane (†równy wkład trzech pierwszych autorów).

Praca proponuje D2DMoE, metodę konwersji statycznych podsieci feedforward na warstwy Mixture-of-Experts z dynamicznym doborem liczby aktywowanych ekspertów (dynamic-*k*), opartą na wymuszaniu rzadkości aktywacji (regularyzator Hoyera) oraz routingu regresyjnym przewidującym normę wyjścia eksperta. Praca wyjaśnia, dlaczego Mixture-of-Experts działa (rola rzadkości aktywacji) i wykazuje, że sztywna reguła top-k jest nieoptymalna dla modeli konwertowanych z architektur gęstych. W pracy dostarczono dedykowane jądra GPU (Triton) z istotnym przyspieszeniem czasu rzeczywistego oraz wykazano komplementarność z przycinaniem i kwantyzacją.

Wkład Doktoranta: Doktorant deklaruje wyznaczenie kierunku badawczego oraz autorstwo kluczowych elementów, alternatywnego schematu routingu, selekcji dynamic-*k*, fazy wymuszania rzadkości aktywacji dla ReLU i GELU oraz zastąpienia warstw projekcji MHA; napisanie wspólnej bazy kodu, przeprowadzenie eksperymentów ViT-B oraz implementację własnych jąder Triton, a także wiodącą rolę w redakcji i dyskusji z recenzentami.

Na szczególne podkreślenie zasługuje, że dołączony do rozprawy pełny tekst artykułu zawiera szczegółową sekcję wkładu autorskiego, która rzeczowo potwierdza deklaracje z tekstu ramowego. Ta zgodność deklaracji autoreferatu z udokumentowanym opisem w publikacji jest wzorcowa i znacząco ułatwia ocenę samodzielności. Mimo oznaczenia równego wkładu trzech pierwszych autorów, autorski zakres Doktoranta jest tu wyjątkowo dobrze rozgraniczony i identyfikowalny.

1.4 Całościowa ocena wkładu naukowego i samodzielności Doktoranta

Dorobek publikacyjny stanowiący rozprawę jest bardzo silny i znacząco przewyższa typowe wymagania stawiane rozprawom doktorskim w dyscyplinie informatyka. Cztery z pięciu prac ukazały się na konferencjach i w czasopismach o najwyższej randze (A*/200 pkt), a piąta na renomowanej konferencji A* (ECML PKDD, 140 pkt). Łączna punktacja ministerialna pięciu prac cyklu wynosi 940 punktów, do czego dochodzi obszerny dorobek towarzyszący (m.in. publikacje w ICML, WACV, Applied Soft Computing) opisany w rozdziale 4.

Wszystkie publikacje są wieloautorskie, co jest normą w uczeniu maszynowym. Doktorant w każdym przypadku opisuje swój wkład w wydzielonych akapitach „Contribution”. Analiza tych opisów pozwala na następującą ocenę samodzielności: w pracach [I], [III] oraz [IV] Doktorant pełnił rolę jednoznacznie wiodącą jako pierwszy autor i twórca głównych idei oraz wykonawca eksperymentów; w pracach [II] i [V] był jednym z autorów o równym, lecz wyraźnie wyodrębnionym i dominującym wkładzie koncepcyjnym i implementacyjnym. Wiarygodność tych deklaracji wzmacnia fakt, że co najmniej w przypadku pracy [V] dołączony pełny tekst publikacji zawiera niezależną, szczegółową sekcję wkładu autorskiego potwierdzającą deklaracje Doktoranta, jest to praktyka wzorcowa, której rozszerzenie na pozostałe prace dodatkowo umocniłoby przejrzystość postępowania.

Całościowy wkład Doktoranta w rozwój dyscypliny jest spójny i wielowymiarowy. Obejmuje on: (1) nowe metody redukcji redundancji wzdłuż głębokości sieci i wyznaczanie optymalnej głębokości w trakcie treningu [I] oraz recycling predykcji we wczesnych wyjściach [II–III] wraz z oryginalną metryką diagnostyczną Hindsight Improvability; (2) nowe metody

dynamicznej regulacji szerokości, granularne, przestrzennie adaptacyjne moduły ACM [IV] oraz konwersję D2DMoE wiążącą rzadkość aktywacji z architekturą Mixture-of-Experts [V]; (3) wymiar inżynieryjno-systemowy, czyli dedykowane implementacje GPU w języku Triton dla metod [IV] i [V], dokumentujące rzeczywiste przyspieszenia czasu rzeczywistego, a nie jedynie redukcję FLOPs. Ten ostatni aspekt jest szczególnie cenny: Doktorant konsekwentnie odróżnia teoretyczną redukcję operacji od faktycznego zysku sprzętowego, co świadczy o dojrzałości badawczej i zrozumieniu praktycznych ograniczeń dziedziny.

1.5 Zagadnienia dyskusyjne i pytania do Doktoranta

Poniższe uwagi i pytania nie mają charakteru zarzutów wpływających na ocenę rozprawy, lecz stanowią zaproszenie do dyskusji w trakcie obrony i refleksji nad poruszonymi zagadnieniami.

1. Tekst ramowy nie zawiera jednej, spójnej sekcji ilościowo zestawiającej wyniki wszystkich pięciu prac na wspólnej osi kompromisu koszt–jakość. Czy Doktorant mógłby pokusić się o takie zestawienie i wskazać, która z form redundancji (głębokość vs. szerokość) daje większy zysk obliczeniowy w typowych zastosowaniach?
2. Metody [IV] i [V] były rozwijane głównie na modelach pre-trenowanych. Doktorant sam wskazuje w ograniczeniach, że modele pre-trenowane są optymalizowane pod kątem pełnego wykorzystania pojemności. Na ile zaproponowane podejścia zyskałyby, gdyby modele były trenowane od początku jako dynamiczne, i jakie są tego koszty?
3. ACM [IV] podejmuje decyzję o liczbie „learnerów" na podstawie pojedynczego tokena, co jak wskazano prowadzi do alokowania maksymalnych zasobów na nieinformatywne regiony wysokoczęstotliwościowe. Czy routing uwzględniający szerszy kontekst (np. sąsiedztwo tokenów) usunąłby ten artefakt?
4. Od czasu publikacji prac pojawiły się liczne alternatywy efektywnej inferencji (np. natywnie rzadkie MoE w dużych modelach językowych, spekulacyjne dekodowanie). Jak Doktorant pozycjonuje swoje metody, w szczególności wczesne wyjścia, względem dekodowania spekulatywnego, które adresuje pokrewny problem marnotrawstwa obliczeń?

2 Analiza strony formalnej rozprawy

2.1 Ocena układu pracy i redakcji manuskryptu

Rozprawa jest napisana w języku angielskim, poprawnym i klarownym, z dbałością o terminologię specjalistyczną. Zawiera streszczenia w językach angielskim i polskim, spis treści, listę publikacji, tekst ramowy złożony z pięciu rozdziałów oraz załącznik z pełnymi tekstami artykułów. Liczba stron tekstu ramowego jest adekwatna dla rozprawy opartej na cyklu publikacji, w której tekst ramowy pełni funkcję integrującą.

Aparat pojęciowy jest właściwy, notacja matematyczna wprowadzana w miejscach jej użycia jest na ogół konsekwentna, a ryciny (zwłaszcza mapy obciążenia obliczeniowego ilustrujące przestrzenną adaptacyjność oraz wykresy frontu Pareto) są czytelne i dobrze wspierają wywód. Kompozycja każdego rozdziału według powtarzalnego schematu jest zaletą redakcyjną. Na

uznanie zasługuje udostępnienie kodu źródłowego wszystkich prac na licencjach otwartych, co czyni wyniki weryfikowalnymi i odtwarzalnymi, jest to standard, którego rozprawa wzorcowo dochowuje.

2.2 Uwagi szczegółowe

Większość treści jest poprawna pod względem językowym i merytorycznym. W trakcie lektury odnotowałem drobne usterki edytorskie, niewpływające na odbiór pracy:

- pojedyncze literówki w tekście ramowym (np. „folloing" zamiast „following" w sformułowaniu problemu optymalizacji w rozdziale 1);
- tekst ramowy miejscami referuje wyniki w sposób skrótowy („selected subset of the results"), co jest uzasadnione formatem, lecz w kilku miejscach utrudnia samodzielną ocenę bez sięgania do załączników;
- brak jednej zbiorczej tabeli zestawiającej wszystkie prace cyklu (rangę, punktację, rolę autorską i główny wynik), taka tabela podniosłaby przejrzystość autoreferatu.

3 Komentarz do oceny rozprawy

Ocena rozprawy opartej na cyklu wieloautorskich publikacji zawsze wymaga od recenzenta rozważenia kwestii indywidualnego wkładu kandydata. W przypadku rozprawy mgr. inż. Bartosza Wójcika kwestia ta została rozwiązana w sposób satysfakcjonujący: Doktorant jest pierwszym lub współpierwszym autorem wszystkich pięciu prac cyklu, w każdej z nich zamieszcza rzeczowy opis swojego wkładu, a przynajmniej w jednym przypadku ([V]) deklaracje te są niezależnie potwierdzone szczegółową sekcją wkładu autorskiego w samej publikacji. Spójność tematyczna cyklu, od głębokości, przez wczesne wyjścia, po dynamiczną szerokość i Mixture-of-Experts, świadczy o przemyślanym, konsekwentnie realizowanym programie badawczym, a nie o przypadkowym zbiorze współprac.

Na pozytywny odbiór wpływa również dojrzałość metodologiczna i inżynieryjna Doktoranta. Konsekwentne dostarczanie wydajnych implementacji GPU oraz krytyczne odróżnianie redukcji FLOPs od rzeczywistych zysków czasu rzeczywistego pokazują, że autor rozumie zarówno teoretyczną, jak i praktyczną stronę efektywności obliczeniowej. O wysokiej jakości pracy Doktoranta świadczą także samodzielne osiągnięcia organizacyjne: pozyskanie i realizacja grantu NCN PRELUDIUM, współorganizacja warsztatu na ICML 2022, staż badawczy oraz recenzowanie dla ICLR i JMLR. To profil naukowy w pełni dojrzałego, samodzielnego badacza.

4 Ocena spełnienia wymogów ustawowych

Rozprawa prezentuje szeroką i ugruntowaną wiedzę Autora w zakresie głębokiego uczenia, ze szczególnym uwzględnieniem obliczeń warunkowych, dynamicznych sieci neuronowych, architektur Transformer i Mixture-of-Experts oraz rzadkości aktywacji. Doktorant swobodnie porusza się między formalizmem matematycznym, projektowaniem architektur i implementacją sprzętową.

Umiejętność samodzielnego prowadzenia pracy naukowej przez Doktoranta jest zademonstrowana w sposób niebudzący wątpliwości. Doktorant jest pierwszym lub współpierwszym autorem oryginalnych idei algorytmicznych, projektantem i wykonawcą eksperymentów, autorem implementacji (w tym niskopoziomowych jąder GPU) oraz osobą prowadzącą proces publikacyjny. Potwierdzeniem samodzielności jest również kierowanie własnym grantem badawczym NCN.

Rozprawa stanowi oryginalne rozwiązanie istotnego problemu naukowego redukcji kosztu obliczeniowego sieci neuronowych przy zachowaniu skuteczności poprzez systematyczne wprowadzenie i weryfikację nowych metod obliczeń warunkowych wzdłuż głębokości (ZTW, wyznaczanie głębokości w treningu) oraz szerokości (ACM, D2DMoE) modelu. Praktyczna wartość rozwiązań jest potwierdzona udostępnionym, otwartym kodem oraz zmierzonymi przyspieszeniami na sprzęcie GPU.

5 Konkluzja

Stwierdzam, że rozprawa doktorska mgr. inż. Bartosza Wójcika pt. „Computationally Efficient Neural Networks" spełnia wymagania określone w art. 187 Ustawy z dnia 20 lipca 2018 r. – Prawo o szkolnictwie wyższym i nauce (Dz. U. 2018 poz. 1668 z późn. zm.). Praca dowodzi głębokiej znajomości dziedziny głębokiego uczenia i efektywności obliczeniowej, a także wysokich umiejętności samodzielnego prowadzenia badań naukowych przez Doktoranta.

Rozprawa koncentruje się na spójnym, przemyślanym programie redukcji redundancji obliczeniowej sieci neuronowych wzdłuż dwóch komplementarnych osi głębokości i szerokości modelu i wnosi oryginalny, wielowymiarowy wkład obejmujący nowe metody algorytmiczne, narzędzia diagnostyczne oraz wydajne implementacje sprzętowe. Dorobek publikacyjny, na który składają się prace w NeurIPS, AAAI, ECML PKDD oraz czasopiśmie Neural Networks, jest wybitny pod względem rangi i jakości, a indywidualny wkład Doktoranta jako pierwszego lub współpierwszego autora jest wyraźnie udokumentowany.

Wszelkie uwagi krytyczne zawarte w niniejszej recenzji mają charakter dyskusyjny i nie wpływają na moją jednoznacznie pozytywną ocenę rozprawy pod względem oryginalności i wartości merytorycznej. Klarowne sformułowanie celu badawczego, nowatorskie i poprawnie zweryfikowane metody jego realizacji oraz dojrzałość w łączeniu perspektywy teoretycznej, eksperymentalnej i inżynieryjnej dowodzą, że Autor jest w pełni przygotowany do samodzielnej działalności naukowej.

Na tej podstawie wnoszę o dopuszczenie mgr. inż. Bartosza Wójcika do dalszych etapów postępowania w sprawie nadania stopnia doktora w dyscyplinie informatyka. Jednocześnie, ze względu na wybitną jakość dorobku publikacyjnego (cztery prace na konferencjach i w czasopismach rangi A*/200 pkt) oraz dojrzałość samodzielnego warsztatu badawczego, rekomenduję rozważenie przez Radę Dyscypliny wyróżnienia rozprawy.

Tomasz Kajdanowicz

Digitally signed by Tomasz Kajdanowicz
Date: 2026.06.09 22:20:46 +02'00'