

**RADA NAUKOWA DYSCYPLINY
INFORMATYKA TECHNICZNA I TELEKOMUNIKACJA POLITECHNIKI WARSZAWSKIEJ**

zaprasza na

OBRONĘ ROZPRAWY DOKTORSKIEJ

mgr. inż. Wojciecha Masarczyka

która odbędzie się w dniu **9 grudnia 2025 roku**, o godzinie **14:00** w trybie hybrydowym

Temat rozprawy:

„Data representations in non-stationary optimization”

Promotor: prof. dr hab. inż. Tomasz Trzciński – Politechnika Warszawska

Recenzenci: dr hab. Marek Cygan, prof. uczelni – Uniwersytet Warszawski

prof. dr hab. inż. Michał Woźniak – Politechnika Wrocławska

Joost van de Weijer, PhD – Computer Vision Center, Uniwersytet Autonomiczny w Barcelonie

Obrona odbędzie się zdalnie na platformie MS Teams oraz stacjonarnie w Sali 116 Wydziału Technik Informatycznych Politechniki Warszawskiej w Warszawie. Osoby zainteresowane uczestnictwem w obronie proszone są o zgłoszenie chęci uczestnictwa w formie elektronicznej na adres sekretarza komisji: dr hab. inż. Krzysztof Cabaj, email: krzysztof.cabaj@pw.edu.pl, do dnia 8 grudnia 2025r. do godz. 12:00.

Z rozprawą doktorską i recenzjami można zapoznać się w Czytelni Biblioteki Głównej Politechniki Warszawskiej, Warszawa, Plac Politechniki 1.

Streszczenie rozprawy doktorskiej i recenzje są zamieszczone na stronie internetowej: <https://www.bip.pw.edu.pl/Postepowania-w-sprawie-nadania-stopnia-naukowego/Doktoraty/Wszczete-po-30-kwietnia-2019-r/Rada-Naukowa-Dyscypliny-Informatyka-Techniczna-i-Telekomunikacja/mgr-inz.-Wojciech-Bogumil-Masarczyk>

Przewodniczący Rady Naukowej Dyscypliny
Informatyka Techniczna i Telekomunikacja
Politechniki Warszawskiej
prof. dr hab. inż. Jarosław Arabas

Głębokie uczenie się osiągnęło znaczące sukcesy w wielu dziedzinach, takich jak generowanie obrazów, przetwarzanie języka naturalnego, zwijanie białek czy projektowanie materiałów. Pomimo różnic w architekturach i strategiach optymalizacji, te przełomy opierają się na fundamentalnym założeniu: stałym zbiorze danych treningowych. Obecnie, złamanie tego założenia prowadzi do wielu problemów, takich jak katastrofalne zapominanie czy utrata plastyczności. Jednakże, wykroczenie poza paradygmat stałego zbioru danych umożliwiłoby ciągłą adaptację tych modeli do stale zmieniającego się środowiska — cechy charakterystycznej dla każdej inteligentnej jednostki. Aby to umożliwić, musimy odpowiedzieć na kilka fundamentalnych pytań badawczych. Pięć publikacji zawartych w tej pracy doktorskiej częściowo odpowiada na następujące pytanie:

Jaką rolę odgrywają reprezentacje danych w optymalizacji niestacjonarnej?

Praca [1] pokazuje, że modyfikacja reprezentacji danych przed treningiem może zmniejszyć katastrofalne zapominanie. W tym celu model-nauczyciel meta-uczy się syntetycznych próbek danych zaprojektowanych tak, aby złagodzić negatywne skutki niestacjonarnego treningu modeli-uczników. Praca podkreśla znaczenie reprezentacji danych w kontekście katastrofalnego zapominania. Jednakże, koszty meta-uczenia sprawiają, że to podejście jest niepraktyczne. Potencjalne rozwiązanie tego problemu przedstawiono w pracy [2], która modyfikuje reprezentacje danych poprzez uczenie się niezmienników danych przy użyciu zastępczej funkcji straty rekonstrukcji, co wiąże się z pomijalnym dodatkowym kosztem. Eksperymenty pokazują, że uczenie się tych niezmienników zwiększa odporność modelu na katastrofalne zapominanie.

Ograniczenia prac [1, 2] oraz wielu innych prac z zakresu ciągłego uczenia się wynikają z traktowania sieci neuronowych jako czarnych skrzynek, które przyjmują dane i zwracają przewidywania. Aby poszerzyć tę wąską perspektywę, prace [3, 4] koncentrują się na ukrytych reprezentacjach sieci neuronowych i badają ich wpływ na niestacjonarny trening. Praca [3] wprowadza Hipotezę Efektu Tunelu, która mówi, że warstwy wystarczająco głębokich sieci dzielą się na dwie odrębne części, które w różny sposób przyczyniają się do wykonania zadania. Początkowe warstwy generują reprezentacje liniowo separowalne, podczas gdy kolejne warstwy, zwane tunelem, kompresują te reprezentacje. Tunele odgrywają istotną rolę w generalizacji modelu i ciągłym uczeniu się, co podważa dotychczasowe poglądy na temat katastrofalnego zapominania. Kontynuując ten wątek, kolejna publikacja [4] teoretycznie analizuje dynamikę reprezentacji, aby zrozumieć źródła Efektu Tunelu, oraz dowodzi zaskakującej zdolności funkcji softmax do zwiększania rangi reprezentacji, co prowadzi do wielu konsekwencji, takich jak kompresja czy gorsza wydajność na danych spoza dystrybucji.

Ostatnia praca [5] podważa obecne rozumienie optymalizacji niestacjonarnej i empirycznie demonstruje, że ciągły trening na zadaniach z tym samym rozkładem reprezentacji danych może prowadzić do katastrofalnego zapominania i utraty plastyczności.

Deep learning has achieved remarkable success across a wide range of fields, including image generation, natural language processing, protein folding, and material design. Despite differences in architectures and optimization strategies, these breakthroughs share a fundamental assumption: fixed training dataset. Currently, breaking this assumption leads to many issues such as catastrophic forgetting or loss of plasticity. However, moving beyond the fixed-dataset paradigm would enable a continuous adaptation of these models to the ever-changing environment -- an ability of any intelligent entity. To make it possible we need to address several fundamental research questions. Five publications of the thesis provide partial answers to the following question:

What role do data representations play in non-stationary optimization?

Work [1] demonstrates that modifying data representations before the training can reduce catastrophic forgetting. To do so, the teacher-model meta-learns synthetic data samples designed to mitigate the negative effects of students' non-stationary training. The work highlights the importance of data representations in catastrophic forgetting. However, the costs of meta-learning make this approach infeasible. A potential way of addressing this shortcoming is presented in work [2], which modifies data representations by learning invariances of data using a surrogate reconstruction loss with a negligible additional cost. The experiments show that learning these invariances makes the model more robust to catastrophic forgetting.

Limitations of work [1, 2] and a large body of works in continual learning come from treating neural networks as black boxes taking data and returning predictions. To broaden this narrow perspective, works [3, 4] focus on hidden representations of neural networks and investigate their impact on non-stationary training. The work [3] introduces the Tunnel Effect Hypothesis saying that layers of sufficiently deep networks divide into two distinct parts that contribute differently to the task. The initial layers produce linearly separable representations, while the subsequent layers, the tunnel, compress these representations. The tunnel plays a non-trivial role in the model's generalization and continual learning, challenging existing perspectives on catastrophic forgetting. As a continuation, the next publication [4] theoretically analyzes representations' dynamics to understand the Tunnel Effect's origins and proves the softmax function's surprising ability to increase the representations' rank, showing it leads to multiple consequences such as compressed or and worse out-of-distribution performance.

The last work [5] challenges the current understanding of non-stationary optimization and demonstrates empirically that continual training on the tasks with the same distribution of data representations can lead to catastrophic forgetting and loss of plasticity.



UNIwersytet
Warszawski

Wydział Matematyki, Informatyki i Mechaniki
Instytut Informatyki

Recenzja rozprawy doktorskiej magistra Wojciecha Masarczyka pod tytułem *Data representations in non-stationary optimization*

Artykuły na których bazuje rozprawa

[P1] Wojciech Masarczyk and Ivona Tautkute. *Reducing catastrophic forgetting with learning on synthetic data*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, June 2020.

[P2] Wojciech Masarczyk, Kamil Deja, and Tomasz Trzcíński. *On robustness of generative representations against catastrophic forgetting*. International Conference on Neural Information Processing, 2021.

[P3] Wojciech Masarczyk, Tomasz Trzcíński, and Mateusz Ostaszewski. *On consequences of finetuning on data with highly discriminative features*. NeurIPS 2023: UniReps Workshop, 2023.

[P4] Wojciech Masarczyk, Mateusz Ostaszewski, Ehsan Imani, Razvan Pascanu, Piotr Miłoś, and Tomasz Trzcíński. *The tunnel effect: Building data representations in deep neural networks*. NeurIPS, 2023.

[P5] Wojciech Masarczyk, Mateusz Ostaszewski, Tin Sum Cheng, Tomasz Trzcíński, Aurelien Lucchi, and Razvan Pascanu. *Unpacking softmax: How temperature drives representation collapse, compression and generalization*. Under review at NeurIPS 2025, 2025.

Tematem rozprawy są badania na temat reprezentacji danych w modelach opartych o sieci neuronowe w przypadku zmieniającego się zbioru danych. Jednym z głównych ograniczeń trenowanych sieci neuronowych jest powszechnie obserwowane zjawisko katastrofalnego zapominania, gdzie optymalizacja modelu pod kątem nowego zadania powoduje załamanie skuteczności na wcześniej opanowanych zadaniach.

Zanim przejdę do omówienia i oceny wyników chciałbym zwrócić uwagę na podkreślenie przez autora stawianych hipotez, konkretnych pytań, które są weryfikowane za pomocą przedstawianych podejść. Jestem gorącym zwolennikiem takiego podejścia do prowadzenia i przedstawiania wyników prac naukowych.

W pierwszych dwóch pracach [P1], [P2] przedmiotem badań jest problem ciągłego uczenia, gdzie do rozwiązania jest kilka problemów klasyfikacyjnych o rozłącznych klasach. W artykule [P1] przedstawione zostało nowe podejście polegające na generowaniu sztucznych danych za pomocą generatora G w taki sposób, aby za pomocą meta-uczenia otrzymać model, który jest mniej narażony na zjawisko katastrofalnego zapominania. Przy okazji dyskusji wyników pracy [P1] zabrakło mi wyjaśnienia ile razy ostatecznie wytrenowany model jest ewaluowany na prawdziwych danych testowych i czy istnieje w tym wypadku ryzyko przeuczenia. Mam nadzieję, że autor odniesie się do tego zagadnienia w trakcie obrony.

Praca [P2] stanowi naturalne rozszerzenie pracy [P1] gdzie dokonywane jest porównanie reprezentacji w modelach dyskryminujących i generatywnych. Porównanie standardowych klasyfikatorów oraz modeli takich jak autokodery (w wersji wariacyjnej i standardowej) pozwala na wyciągnięcie bardzo interesujących wniosków na temat różnicy w możliwościach uogólniania przez reprezentacje uzyskane w te dwa różne sposoby. Przedstawione wyniki jednoznacznie wskazują na przewagę modeli generatywnych w tak zdefiniowanym zadaniu.

Praca [P3] skupia się na aspekcie wprowadzenia do zbioru danych bardzo dyskryminującej cechy w postaci pojedynczego piksela doskonale skorelowanego z docelową klasą. Autor najpierw trenuje model na pierwotnym zbiorze danych a następnie po wprowadzeniu owego specjalnego piksela trenuje model ponownie. Przedstawione wyniki wskazują, że reprezentacja modelu ulega załamaniu i model skupia się praktycznie wyłącznie na wartości pojedynczego piksela, kompletnie tracąc swoją plastyczność, czyli zdolność reprezentacji do generalizacji do nowych zadań.

W kolejnych dwóch pracach [P4], [P5] tematyka badań dotyczy tego co dzieje się z wewnętrzną reprezentacją sieci neuronowych na przestrzeni jej kolejnych warstw. W pracy [P4] główna obserwacja pokazuje bardzo ciekawą relację pomiędzy tym jak skutecznie cechy danego poziomu reprezentacji pozwalają rozróżniać docelowe klasy a tym jaki jest rząd reprezentacji. W pierwszych warstwach sieci bardzo szybko rośnie skuteczność reprezentacji, ale w momencie gdy zostanie ona wysycona następuje bardzo szybki spadek rzędu. To pozwoliło na wydzielenie części sieci nazwanej tunelem, gdzie reprezentacja ulega szybkiej

kompresji mierząc ją rzędem. Praca [P5] jest tematycznie powiązana, gdyż analizie poddana jest konkretna warstwa typu softmax a dokładnie temperatura użyta w tej warstwie. Warstwa typu softmax jest nieodzownym elementem mechanizmu uwagi, dlatego też zrozumienie jej działania oraz przełożenia na możliwości uogólniania i katastrofalnego zapominania są ciekawym aspektem badawczym.

W ostatnich latach obserwujemy trend coraz częściej występujących modeli fundamentalnych, które trenowane są na tak zróżnicowanych zbiorach danych, że zjawisko katastrofalnego zapominania zostało zminimalizowane. Przykładem takich modeli w dziedzinie przetwarzania języka naturalnego są praktycznie wszystkie modele opracowywane przez OpenAI, Anthropic, Gemini, etc. Podobnie w przypadku modeli do wizji maszynowej bardzo skuteczne są modele takie jak Segment Anything lub DINO. Jest to na tyle istotny trend, że w rozprawie zabrakło mi odniesienia się do tego zagadnienia.

Przy okazji przedstawiania swojej opinii na temat rozprawy, chciałbym zwrócić uwagę na potrzebę wypracowania przez środowisko naukowe formatu rozpraw w przypadku, w którym format zbioru publikacji (tzw. zszywka) nie jest formalnie możliwy. Wyjściem pośrednim jest formalny format monograficzny rozpoczynający się od wprowadzenia do zagadnień doktoratu a następnie przedstawieniu kolejnych publikacji jako osobnych rozdziałów, co ma miejsce w omawianej rozprawie. Moim zdaniem o jakości całego doktoratu świadczy zbiór artykułów, które się na niego składają, dlatego też uważam że jako środowisko nie powinniśmy kurczowo trzymać się sztywnych zwyczajowych reguł stawianych monografiom, co czasami ma miejsce w recenzjach prac doktorskich w dziedzinach gdzie nowe wyniki pojawiają się bardzo dużym tempie, co charakteryzuje dziedziny uczenia maszynowego i sztucznej inteligencji.

Ocena rozprawy

Podsumowując, uważam, że przedstawione wyniki dotyczą ważnych i aktualnych zagadnień a artykuły w których zostały one opublikowane zostały przedstawione na rozpoznawalnych międzynarodowych warsztatach i konferencjach. Na szczególną uwagę zasługuje spójność tematyczna przedstawionych prac, które autor ujął jedną spójną tematycznie narracją. W mojej ocenie praca spełnia wszystkie wymagania ustawowe i zwyczajowe i rekomenduję przyznanie stopnia doktora na podstawie przedstawionej rozprawy.

Marek Cygan

Thesis Review

Title: Data representations in non-stationary optimization

Author: Wojciech Masarczyk

Summary: The quality of the thesis is excellent and I recommend its acceptance. I also support awarding the degree Cum Laude.

The introduction clearly proposes the main research question of the thesis, and briefly outlines the main contributions for each of the chapters. In Chapter 2, the thesis explores whether synthetic data can be generated in a way that prevents forgetting when learned in sequence. They introduce a two-step optimization method using meta-gradients to create such data. The chapter is more exploratory in nature, and the preliminary results suggest that synthetic data can be used to mitigate forgetting. However, experiments on larger datasets and a deeper analysis are required for this to be a solid contribution.

In Chapter 3, the thesis compares the quality of generative and discriminative representations for the task of continual learning, considering both catastrophic forgetting (negative backward transfer) and generalization (positive forward transfer). Discriminative models forget considerably more than generative ones, highlighting the latter's resilience. The analysis includes both per task performance analysis and Centered Kernel Alignment analysis. The conclusions of this chapter are important because they also generalize to other non-discriminative representation learning strategies, such as self-supervised learning. This work has contributed to a better understanding of the strength of these representations within a continual learning context. The work has been published in the International Conference on Neural Information Processing.

Chapter 4 focuses on an interesting experiment, where a learner is first learning a task, and subsequently is provided with data with artificially simple features to solve the same classification problem. The results of this experiment provide intriguing insights into deep neural networks. This way of training leads to feature erosion, a dramatic drop in feature quality. The network focuses on the simple feature and forgets the previously learned features used for classification. In addition, experiments show that this hurts the plasticity of the network, permanently damaging the network.

Chapter 5 makes several fundamental observations on network training. Neural networks divide into two stages: early layers craft linearly separable features, and then a tunnel of later layers compresses them. This tunnel appears early in training and contributes little to final accuracy despite its depth. Its length grows when the network's capacity greatly exceeds the task's intrinsic complexity (the extractor remains constant, only becomes smaller for wider networks). Interestingly, the tunnel hurts out-of-distribution generalization and can aggravate continual learning forgetting. The conclusions of this chapter may have consequences for rethinking network depth, especially in the context of continual learning. The task-agnostic nature of the compression is particularly fascinating. This chapter has been published at NeurIPS 2023.

Chapter 6 studies the impact of the softmax function on representation collapse, learning dynamics, compression, and generalization to OOD data. The paper shows the surprising effect of the logit norm at initialization on the learning dynamics of deep neural networks. This can have profound effects on the generalization of the representation. The chapter then shows that adjusting the softmax temperature can steer the learning dynamics and lead to more generalizable representations that do not suffer from rank collapse. The chapter also establishes a fundamental trade-off between OOD generalization and OOD detection, highlighting the role of softmax temperature in balancing this trade-off.

Conclusion: This thesis is of very high quality. The thesis addresses several fundamental aspects of deep neural network training and identifies previously unobserved phenomena that deepen our understanding and can have a considerable impact on AI in general. The thesis is well written, and the graphics are of high quality. The empirical evaluation of the posed scientific hypotheses is extensive. The analysis is of high quality and scientific depth and will certainly lead to follow-up research within the community. The thesis has led to several publications, including one at NeurIPS, the most important and impactful AI conference in the field. **I consider the quality of the thesis to be excellent and recommend its acceptance. Furthermore, given the exceptional quality of the work, I would also support awarding the degree Cum Laude.**

Joost van de Weijer

Senior Scientist at the Computer Vision Center, Universitat Autònoma de Barcelona

Group Leader of the Learning and Machine Perception team

<http://www.cvc.uab.es/LAMP/joost/>



Recenzja rozprawy doktorskiej

Wojciecha Masarczyka

zatytułowanej:

Data representations in non-stationary optimization

1. Problem badawczy i jego znaczenie

W pracy rozważane są problemy reprezentacji danych w uczeniu modeli głębokich dedykowanych uczeniu ciągłemu. Praca ma charakter naukowy - doktorant proponuje w pracy autorskie algorytmy rozwiązujące powyższy problem oraz dokonuje ich ewaluacji na drodze eksperymentu komputerowego. Opracowane metody mogą znaleźć szerokie zastosowanie praktyczne.

2. Wkład autora

Na początku rozprawy autor sprecyzował cel pracy związany z rolą reprezentacji danych w optymalizacji niestacjonarnej, a następnie sformułował trzy główne pytania badawcze:

- (i) Jak reprezentacje danych wpływają na *catastrophic forgetting* i czy można nauczyć się reprezentacji danych odpornych na to zjawisko?
- (ii) Czy wszystkie reprezentacje danych są w równym stopniu podatne na wspomniane zjawisko?
- (iii) Jak dynamika uczenia wpływa na reprezentacje i ich zdolność do uogólniania poza rozkładem danych treningowych?

Postawione cele i pytania uważam za ciekawe, dobrze sformułowane i ważne dla dziedziny ciągłego uczenia maszynowego.

Do najważniejszych osiągnięć rozprawy zaliczam:

- (i) Zaproponowanie meta-uczenia się do generowania syntetycznych próbek danych wykorzystywanych w scenariuszach ciągłego uczenia się, a co za tym idzie wykazanie, że reprezentacje wyuczone przez modele generatywne są mniej podatne na *catastrophic forgetting* niż te wyuczone przez modele dyskryminacyjne.
- (ii) Zaproponowanie zastępczej straty rekonstrukcji (*surrogate reconstruction loss*) w celu mitygacji zjawiska *catastrophic forgetting*.
- (iii) Zaproponowanie i częściowe zbadanie hipotezy efektu tunelowego (*Tunnel Effect Hypothesis*).
- (iv) Wykazanie, że funkcja *softmax*, odgrywa kluczową rolę w kształtowaniu reprezentacji danych, w szczególności opisanie zjawiska *rank deficit bias*, gdzie modele głębokie oparte o *softmax* znajdują rozwiązania o randze znacznie niższej niż liczba klas.

Wyniki rozprawy są de facto zbiorem pięciu wieloautorskich artykułów opublikowanych na topowej konferencji NeurIPS (dwie prace) oraz na *workshopach* NeurIPS oraz CVPR, a jedna praca jest dostępna jako preprint (zgłoszona na konferencję NeurIPS). Dodatkowo, doktorant jest współautorem siedmiu

publikacji, w tym dwóch opublikowanych w materiałach konferencji o rankingu CORE A* (NeurIPS), dwóch w materiałach IJCNN (CORE rank A) oraz dwóch artykułach czasopismach indeksowanych w JCR (*Journal of Cultural Heritage, IEEE Access, Remote Sensing*). Na tym etapie kariery naukowej należy uznać dorobek doktoranta za wybitny. Warto także wspomnieć, że prace doktoranta są dostrzegane przez środowisko uczenia maszynowego, na co wskazuje liczba cytowań jego prac (268 wg Scopus i 440 wg Google Scholar¹).

3. Poprawność

Jak wspomniano praca jest zbiorem artykułów, które poprzedza krótkie wprowadzenie oraz omówienie najważniejszych wyników rozprawy. Praca jest zredagowana dobrze. Autor dużą wagę przykładą do dogłębnego wyjaśnienia zagadnień, którymi się zajmuje, a także określa precyzyjnie swój wkład w powstanie konkretnych publikacji.

Z punktu widzenia warsztatu badawczego, autor formułuje propozycje swoich metod na podstawie wnikliwej analizy literatury, dostrzegając możliwości poprawy zawartych tam metod. Jako narzędzi ewaluacji zaproponowanych metod przyjęto podejściu oparte głównie na eksperymencie komputerowym, jednakże część własności została także oceniona na drodze analitycznej.

Lektura rozprawy prowadzi do sformułowania kilku uwag.

1. Doktorant przyznaje, że opisany w pracy efekt tunelu nie jest dobrze poznany, szczególnie dlaczego i kiedy się pojawia oraz w jaki sposób można go wywołać. O ile opisanie tego zjawiska jest bardzo cenne i zostało dostrzeżone w środowisku związanym z uczeniem ciągłym, to należy zauważyć, że efekt ten (nie pod zaproponowaną przez doktoranta nazwą) był opisany we wcześniejszych pracach. Znane jest zjawisko, że reprezentacje wykorzystują wybrane warstwy. Stąd, intuicyjnie, długość tunelu jest związana z konkretną architekturą, a także trudnością zadania, która wpływa na długość „tunelu”, tj. im trudniejsze zadanie, tym tunel powinien być krótszy. Potwierdzają to także prace związane z pruningiem głębokich sieci neuronowych, tj. redukcji architektury (np. warstw sieci) bez wpływu na jej jakość.
2. W części opisowej pracy brakuje mi odniesienia się do dyskusji, którą wywołała praca o efekcie tunelowym. Co prawda, w pracy wspomniano pracę *What Variables Affect Out-of-Distribution Generalization in Pretrained Models?* (<https://arxiv.org/abs/2405.15018> - nr 215 wg spisu literatury), ale doktorant nie pokusił się o polemikę z zawartymi w tej pracy rezultatami. Stąd, jestem ciekawy, czy i doktorant zgadza się, czy odrzuca zredefiniowaną przez autorów hipotezę o efekcie tunelowym.
3. Jak wspomniano, praca jest de facto zbiorem artykułów opatrzonego krótkim wstępem. W mojej opinii, dla większej czytelności pracy, lepiej byłoby zamieścić artykuły w oryginalnej formie, zamiast sztucznie tworzyć na ich podstawie rozdziały.
4. Konsekwencją przyjętej formy pracy, jest także brak zebrania w jednym miejscu ew. planów badawczych doktoranta, które są częściowo umieszczone w rozdziałach 2 i 5.
5. O ile doktorant bardzo precyzyjnie określił swój wkład w powstanie publikacji składających się na rozprawę, to brakuje mi oświadczeń autorów tudzież określenie ich zaangażowania w powstanie wspomnianych prac, np. z wykorzystaniem taksonomii CRediT.

¹ Dane na 19 września 2025 r.

4. Wiedza kandydata

Doktorant posługuje się sprawnie zaawansowanym aparatem matematycznym, a także potrafi zaplanować eksperyment komputerowy w celu oceny jakości zaproponowanych metod. Przegląd literaturowy dotyczący zagadnień przedstawionych w rozprawie, zawarty głównie w rozdziale pierwszym i drugim pozwala stwierdzić, że doktorant posiada aktualną wiedzę z zakresu tematyki rozprawy, a także potrafi dokonać krytycznego przeglądu źródeł w celu wskazania ciekawych kierunków badań. Zawarty w dysertacji spis źródeł literaturowych, zawierających 225 pozycji, jest aktualny i kompletny. Należy zatem stwierdzić, że doktorant posiada wiedzę z zakresu informatyki technicznej i telekomunikacji, w szczególności w zakresie metod uczenia maszynowego.

5. Podsumowanie

Doktorant wykazał się w recenzowanej rozprawie właściwie stosowanym podejściem eksperymentalnym oraz dobrą znajomością aktualnej problematyki związanej z uczeniem modeli głębokich, szczególnie w scenariuszu uczenia ciągłego. Zostało to poparte dobrymi studiami literaturowymi, obejmującymi aktualne piśmiennictwo związane z problematyką rozprawy, co świadczy o bardzo dobrej wiedzy doktoranta z tego zakresu. Dla poruszanych problemów doktorant sformułował kilka ciekawych i użytecznych metod i obserwacji oraz wskazał możliwości dalszego rozwoju metod w obszarze związanym z rozprawą.

Recenzowana dysertacja przedstawia rozwiązanie ważnego i oryginalnego problemu, wzbogacając naszą wiedzę dotyczącą wykorzystania metod uczenia ciągłego. Zawarte w niej wyniki badań eksperymentalnych wskazują również na możliwość wykorzystania otrzymanych metod w praktyce. Przedstawione w punkcie 3 recenzji uwagi nie wpływają na jednoznacznie pozytywne wrażenie o przedłożonej rozprawie. Jestem przekonany, że doktorant zdaje sobie sprawę z możliwości kontynuowania rozpoczętej w pracy tematyki.

Reasumując, biorąc pod uwagę powyższe opinie i wymagania zawarte w art. 13 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym (z późniejszymi zmianami) stwierdzam, że rozprawa mgr inż. Wojciecha Masarczyka pt. „*Data representations in non-stationary optimization*” spełnia wymagania stawiane pracom doktorskim, w szczególności:

- Rozprawa zawiera oryginalne rozwiązanie problemu naukowego.
- Kandydat posiada ugruntowaną, głęboką wiedzę w dyscyplinie Informatyka techniczna i telekomunikacja
- Doktorant posiada umiejętność samodzielnego prowadzenia pracy naukowej.

Stąd, wnoszę o jej przyjęcie i dopuszczenie mgr inż. Wojciecha Masarczyka do publicznej obrony.

Ponadto, biorąc pod uwagę bardzo wysoki poziom rozprawy, a także ponadprzeciętny dorobek publikacyjny doktoranta, w skład którego wchodzi publikacje w materiałach konferencji o rankingu A* i publikacje w czasopiśmie indeksowanym w JCR, **wnioskuje o wyróżnienie rozprawy doktorskiej mgr inż. Wojciecha Masarczyka.**



Podpis