

Łódź, 4 kwietnia 2026 r.

dr hab. inż. Agnieszka Wosiak, prof. uczelni
Instytut Informatyki
Wydział Fizyki Technicznej, Informatyki i Matematyki Stosowanej
Politechnika Łódzka
Al. Politechniki 8
93-590 Łódź

RECENZJA ROZPRAWY DOKTORSKIEJ

TYTUŁ ROZPRAWY: Classification based on dispersed data with deep learning issues

AUTOR ROZPRAWY: mgr Kwabena Marfo

Podstawą niniejszej recenzji jest pismo Pana dr hab. inż. Rafała Doroza, prof. UŚ, Przewodniczącego Rady Naukowej Instytutu Informatyki z dnia 26 stycznia 2026 r.

CHARAKTERYSTYKA ROZPRAWY I OCENA UKŁADU ROZPRAWY (W TYM INFORMACJE O JEJ POSZCZEGÓLNYCH CZĘŚCIACH SKŁADOWYCH)

Przedstawiona do recenzji rozprawa doktorska składa się z cyklu jedenastu publikacji zebranych pod wspólnym tytułem „Classification based on dispersed data with deep learning issues”. Rozprawa została poświęcona problemowi klasyfikacji danych rozproszonych i heterogenicznych z wykorzystaniem metod głębokiego uczenia, przy czym szczególny nacisk położono na opracowanie rozwiązań umożliwiających budowę modeli globalnych bez konieczności centralizacji danych źródłowych. Już na poziomie ogólnej charakterystyki należy stwierdzić, że tematyka pracy jest dobrze osadzona w aktualnych problemach badawczych informatyki, zwłaszcza w obszarze uczenia maszynowego, analizy danych rozproszonych oraz metod ochrony prywatności w procesie uczenia modeli.

Rozprawa obejmuje część syntetyzującą dorobek naukowy Doktoranta oraz przedruki publikacji stanowiących podstawę doktoratu. Zgodnie ze spisem treści, praca zawiera: rozdział wprowadzający, część poświęconą charakterystyce danych rozproszonych, rozdział prezentujący autorskie podejścia, rozdział podsumowujący, następnie przedruki publikacji oraz oświadczenia potwierdzające informację o wkładzie Doktoranta w prace naukowe. Układ pracy jest właściwy dla rozprawy przygotowanej w formule cyklu publikacyjnego i umożliwia zarówno zapoznanie się z syntetycznym opisem problemu badawczego oraz głównych osiągnięć, jak i z pełną treścią prac stanowiących zasadniczy dorobek

naukowy Doktoranta. Całość liczy 262 strony, z czego część o charakterze syntetyzującym, stanowiąca swego rodzaju autoreferat, liczy 40 stron i poprzedza blok publikacji oraz część dokumentacyjną.

Część wprowadzająca rozprawy została zorganizowana w sposób przejrzysty. Autor rozpoczyna od nakreślenia problemu badawczego, wskazując na ograniczenia klasycznych metod uczenia w warunkach, gdy dane pozostają rozproszone pomiędzy wieloma niezależnymi źródłami, a ponadto charakteryzują się niejednorodnością przestrzeni cech. Następnie formułuje hipotezę rozprawy oraz porządkuje swój wkład naukowy poprzez wskazanie trzech głównych pytań badawczych i odpowiadających im linii metodycznych, logicznie rozwijających się w cykl badań: od metod fuzji lokalnych predykcji, przez agregację lokalnych sieci neuronowych, po ujednoczenie struktur danych, aż po podejścia oparte na ekstrakcji cech i odwzorowaniu danych do wspólnej przestrzeni reprezentacji. Całość jest czytelna i pozwala na analizę dorobku.

Rozprawa ma charakter metodyczny, teoretyczno-doświadczalny. Jej głównym celem jest opracowanie nowych podejść do klasyfikacji danych rozproszonych i heterogenicznych z wykorzystaniem sieci neuronowych oraz ich eksperymentalna weryfikacja. Rozważania teoretyczne, obejmujące analizę stanu wiedzy i identyfikację ograniczeń istniejących podejść, stanowią podstawę do zaproponowania autorskich rozwiązań, które następnie zostały poddane ocenie empirycznej w ramach cyklu publikacyjnego. W skład cyklu wchodzi niżej wymienione pozycje:

- [P1] Przybyła-Kasperek M., Marfo K.F. Neural network used for the fusion of predictions obtained by the k-nearest neighbors algorithm based on independent data sources. *Entropy*, 23(12), 2021. DOI:10.3390/e23121568.
- [P2] Przybyła-Kasperek M., Marfo K.F. Influence of noise and data characteristics on classification quality of dispersed data using neural networks on the fusion of predictions. *International Conference on Information Systems Development*, 2022. DOI: 10.62036/ISD.2022.
- [P3] Marfo K.F., Przybyła-Kasperek M. Radial basis function network for aggregating predictions of k-nearest neighbors local models generated based on independent data sets. *Procedia Computer Science*, 207:3234–3243, 2022, DOI: 10.1016/j.procs.2022.09.381.
- [P4] Marfo K.F., Przybyła-Kasperek M. Radial basis function neural network with a centers training stage for prediction based on dispersed image data. *International Conference on Computational Science*, 10476:89–103, 2023, DOI: 10.1007/978-3-031-36027-5_7.

- [P5] Marfo K.F., Przybyła-Kasperek M. Study on the Use of Artificially Generated Objects in the Process of Training MLP Neural Networks Based on Dispersed Data. *Entropy*, 25(5), 2023. DOI: 10.3390/e25050703.
- [P6] Marfo K.F., Przybyła-Kasperek M., Sulikowski P. Fragmented Image Classification Using Local and Global Neural Networks: Investigating the Impact of the Quantity of Artificial Objects on Model Performance. *International Conference on Computational Science*, 14838:280-194, 2024, DOI: 10.1007/978-3-031-63783-4_21
- [P7] Marfo K.F., Przybyła-Kasperek M. Exploring the Impact of Object Diversity on Classification Quality in Dispersed Data Environments. *ACIIDS*, 14796:250-262, 2024, DOI: 10.1007/978-981-97-4985-0_20.
- [P8] Marfo K.F., Przybyła-Kasperek M. Enhancing Dispersed Data Classification: A Hierarchical Model Based on Neural Networks. *Proceedings of the 5th Polish Conference on Artificial Intelligence*, 154-161, 2024, DOI: 10.17388/WUT.2024.0002.MiNI.
- [P9] Przybyła-Kasperek M., Marfo K.F. A multi-layer perceptron neural network for varied conditional attributes in tabular dispersed data. *PLoS ONE*, 19:1-56, 2024, DOI: 10.1371/journal.pone.0311041.
- [P10] Marfo K.F., Przybyła-Kasperek M. Objects Diversity and its Impact on Classification Quality in Dispersed Data Environments. *Vietnam Journal of Computer Science*, SN: 2196-8888 p253, 2025, DOI: 10.1142/s2196888824500180.
- [P11] Marfo K.F., Przybyła-Kasperek M., Decentralized Neural Network Modeling from Heterogeneous Data Sources: A Feature Mapping Approach *International Conference Information Systems Development*, 2025. DOI: 10.62036/ISD.2025.

Choć w żadnej ze wskazanych publikacji cyklu Kandydat nie jest jedynym autorem, to jego deklarowany udział w pracach można uznać za znaczący, zaś dodatkowo w ośmiu ze wskazanych 11 pozycji jest on pierwszym autorem.

OCENA PREZENTACJI WIEDZY TEORETYCZNEJ (W TYM LITERATURY ŚWIATOWEJ, AKTUALNEGO STANU WIEDZY ORAZ SFORMUŁOWANYCH WNIOSKÓW)

Podstawy teoretyczne rozprawy oraz analiza dotychczasowego stanu wiedzy zostały przedstawione przede wszystkim w rozdziale drugim pracy, zatytułowanym „Dispersed Data”, a także częściowo we wprowadzeniu oraz w omówieniu poszczególnych publikacji wchodzących w skład cyklu. Autor

poprawnie lokuje swoje badania w obszarze uczenia maszynowego dla danych rozproszonych i heterogenicznych, wskazując, że klasyczne podejścia do budowy modeli globalnych zakładają najczęściej centralizację danych lub jednorodność przestrzeni cech, co w wielu praktycznych zastosowaniach nie jest spełnione.

Autor nie ogranicza się do ogólnego opisu obszaru badawczego, lecz podejmuje próbę jego uporządkowania pojęciowego. W pracy zdefiniowano środowisko danych rozproszonych jako zbiór niezależnych tablic decyzyjnych, opisano jego podstawowe własności, a także sformułowano założenia dotyczące niejednorodności atrybutów, częściowej rozłączności obiektów oraz braku możliwości centralnego scalania danych.

Analiza literatury została skoncentrowana głównie wokół dwóch nurtów: „federated learning” oraz „distributed learning”. Autor omawia ich podstawowe założenia, a następnie wskazuje najważniejsze ograniczenia tych podejść w kontekście danych rozproszonych i heterogenicznych. W odniesieniu do federated learning Autor omawia między innymi problemy statystycznej heterogeniczności danych, przesunięcia rozkładów cech, zróżnicowania zasobów obliczeniowych klientów, kosztów komunikacyjnych oraz zagrożeń dla prywatności. Przy prezentacji kolejnych rozwiązań wskazuje ich zalety i ograniczenia, takie jak na przykład zależność od strojenia hiperparametrów, wzrost złożoności obliczeniowej, problemy ze sprawiedliwą reprezentacją klientów czy niepełną odporność na heterogeniczność danych. Analogicznie, w części dotyczącej „distributed learning” omawia rozwiązania związane z równoległym uczeniem modeli, redukcją kosztów komunikacji i odpornością na awarie, również uzupełniając opis o uwagi krytyczne, co świadczy o dojrzałym podejściu do analizy stanu wiedzy.

W mojej ocenie **przegląd literatury światowej jest wystarczająco szeroki i dobrze odpowiada zakresowi rozprawy. Poszczególne pozycje rozprawy potwierdzają dobrą znajomość stanu wiedzy Autora w odniesieniu do prowadzonych prac i bardzo dobrze wpisują się w dyscyplinę „informatyka”.** Wśród cytowanych prac znajdują się zarówno klasyczne pozycje fundamentujące analizowany obszar, jak i opracowania nowsze, odnoszące się do aktualnych problemów uczenia federacyjnego, rozproszonego oraz ochrony prywatności. Zestaw źródeł wskazuje, że Doktorant orientuje się zarówno w podstawach obszaru, jak i w jego współczesnych kierunkach rozwoju. Jednocześnie literatura została dobrana w sposób ukierunkowany, tj. nie nadmiernie szeroki, lecz podporządkowany problematyce rozprawy.

Analiza stanu wiedzy prowadzi do jasno sformułowanej luki badawczej. Autor wykazuje, że istniejące podejścia, zarówno z zakresu federated learning, jak i distributed learning, nie rozwiążą w satysfakcjonujący sposób problemu budowy globalnych modeli klasyfikacyjnych w sytuacji, gdy lokalne

zbiory danych różnią się zarówno obiektami, jak i przestrzeniami cech, a jednocześnie nie mogą być centralizowane. W konsekwencji rozprawa nie opiera się jedynie na zastosowaniu znanych metod do nowego zbioru danych, lecz na próbie zaproponowania nowych rozwiązań metodycznych odpowiadających na wyraźnie zidentyfikowane ograniczenia literatury. Autor dochodzi do dobrze uzasadnionego wniosku, że potrzebne są metody, które nie zakładają ani jednorodności przestrzeni cech, ani centralnego dostępu do danych. W tym sensie wnioski teoretyczne są spójne z hipotezą rozprawy oraz z trzema pytaniami badawczymi, które dotyczą odpowiednio: fuzji lokalnych predykcji, agregacji modeli po ujednoczeniu struktur danych oraz budowy modeli globalnych po transformacji lokalnych danych do wspólnej przestrzeni cech.

OCENA CELU PRACY, ZASTOSOWANYCH METOD BADAWCZYCH I UMIEJĘTNOŚCI PROWADZENIA BADAŃ

Cel rozprawy został sformułowany w sposób czytelny i odpowiada charakterowi pracy, która ma profil metodyczny, teoretyczno-doświadczalny. Autor przyjmuje jako punkt wyjścia problem budowy globalnych modeli klasyfikacyjnych dla danych rozproszonych, heterogenicznych i niemożliwych do centralizacji, a następnie stawia hipotezę, zgodnie z którą zastosowanie architektur opartych na sieciach neuronowych w trzech zaproponowanych nurtach badawczych może istotnie poprawić trafność i odporność klasyfikacji w takich środowiskach. Hipoteza ta została powiązana z trzema konkretnymi pytaniami badawczymi, dotyczącymi odpowiednio: wykorzystania sieci neuronowych do fuzji lokalnych predykcji, imputacji brakujących atrybutów i agregacji lokalnie wytrenowanych modeli oraz budowy modeli globalnych na podstawie transformacji lokalnych danych do wspólnej przestrzeni cech z użyciem metod takich jak PCA, SVD i UMAP.

W mojej ocenie cel pracy jest sformułowany adekwatnie do problematyki rozprawy i pozostaje spójny z przedstawionym wcześniej stanem wiedzy. Autor dąży do opracowania takich rozwiązań metodycznych, które będą mogły funkcjonować mimo niejednorodności przestrzeni cech i braku centralnego dostępu do danych. Jest to istotne, ponieważ właśnie te dwa ograniczenia stanowią zasadniczą trudność w rozważanym obszarze badań. Dobrze należy ocenić także to, że pytania badawcze zostały ułożone w porządku odpowiadającym rozwojowi kolejnych propozycji Autora: od prostszych schematów integrujących lokalne decyzje, poprzez bardziej zaawansowane podejścia agregujące wytrenowane modele, aż po rozwiązania wykorzystujące transformacje reprezentacji danych.

Zastosowane metody badawcze są dobrze dobrane do postawionego celu. Autor rozwija trzy zasadnicze klasy podejść. Pierwsza z nich opiera się na fuzji lokalnych predykcji: lokalne zbiory danych

służą do trenowania klasyfikatorów k-NN, a następnie ich wyniki są agregowane za pomocą sieci neuronowej podejmującej końcową decyzję klasyfikacyjną. Druga klasa metod polega na ujednoczeniu struktury lokalnych danych przez imputację brakujących atrybutów przy użyciu sztucznie generowanych obiektów, po czym następuje agregacja wag i stronniczości lokalnie wytrenowanych sieci neuronowych do postaci modelu globalnego. Trzecia linia badawcza zmierza do uniknięcia kosztów związanych z generowaniem sztucznych obiektów poprzez lokalne odwzorowanie danych do wspólnej, k-wymiarowej przestrzeni cech z wykorzystaniem metod ekstrakcji i redukcji wymiaru, a następnie łączenie wyników modeli trenowanych na tych reprezentacjach.

Umiejętność prowadzenia badań naukowych przez Doktoranta oceniam pozytywnie. Świadczy o niej konsekwencja w rozwijaniu jednej spójnej problematyki badawczej oraz umiejętność przechodzenia od identyfikacji ograniczeń znanych podejść do formułowania własnych rozwiązań. Autor poprawnie rozpoznaje, że w środowiskach danych rozproszonych podstawowym problemem nie jest wyłącznie sama decentralizacja danych, ale jednocześnie występowanie heterogeniczności atrybutowej i ograniczeń prywatności. Następnie projektuje takie procedury, które odpowiadają na te ograniczenia na różne sposoby: przez integrację decyzji lokalnych, przez sztuczne ujednoczanie przestrzeni atrybutów oraz przez budowę wspólnych reprezentacji cech. Wskazuje to na **dobrą orientację metodologiczną oraz zdolność do samodzielnego rozwijania programu badawczego.** Poszczególne publikacje stanowią realizację wskazanego programu badawczego. Prace [P1]–[P4] koncentrują się na problemie fuzji predykcji lokalnych i prowadzą od klasycznych architektur MLP do rozwiązań opartych na sieciach RBF, którym przypisano mniejszą złożoność architektoniczną, szybszą zbieżność oraz mniejsze ryzyko przeuczenia. Z kolei publikacje [P5], [P6] i [P9] rozwijają nurt związany z ujednoczaniem lokalnych zbiorów danych przez sztuczne obiekty i późniejszą agregację lokalnie uczonych sieci neuronowych, natomiast [P8] i [P11] przenoszą ciężar badań na harmonizację opartą na ekstrakcji cech i porównanie różnych architektur neuronowych we wspólnej przestrzeni reprezentacji. W publikacjach [P7] i [P10] Autor analizuje wpływ rozkładu i różnorodności obiektów na jakość klasyfikacji, wykazując, że proponowane podejścia pozostają relatywnie odporne nawet wtedy, gdy lokalne zbiory danych mają niewielką część wspólnych instancji.

Przeprowadzone przez Kandydata badania mają charakter porównawczy i rozwijają się etapowo. Autor proponuje różne architektury i strategie postępowania oraz analizuje ich skuteczność w zależności od przyjętej reprezentacji danych i parametrów modelu. W przytoczonych wynikach pokazuje między innymi zróżnicowanie skuteczności różnych architektur oraz przedstawia statystycznie istotne różnice pomiędzy nimi, co świadczy o poprawnym rozumieniu zasad prowadzenia badań eksperymentalnych, gdzie samo uzyskanie wyniku nie jest wystarczające, a konieczne jest również jego porównanie, interpretacja i ocena istotności. Zaproponowane metody zostały poddane weryfikacji

eksperymentalnej, a sformułowane odpowiedzi na kolejne pytania badawcze zostały przygotowane na podstawie wyników badań. Zatem **cel pracy sformułowany w jej wstępie został osiągnięty.**

Zastosowane metody badawcze są adekwatne do postawionych problemów, a ich dobór świadczy o dojrzałości metodologicznej Doktoranta. Autor wykazał się umiejętnością samodzielnego formułowania problemów badawczych, projektowania rozwiązań, przeprowadzania badań eksperymentalnych oraz wyciągania wniosków na podstawie uzyskanych rezultatów.

OCENA PREZENTACJI WYNIKÓW BADAŃ

Podstawową część rozprawy stanowią artykuły [P1] – [P11] odpowiednio opisujące zagadnienia podjęte w ramach pracy naukowej. Ich uzupełnieniem jest część wprowadzająca w badania, w ramach której zostały syntetycznie przedstawione podjęte w pracy problemy oraz zaproponowane przez Doktoranta ich rozwiązania. Poszczególne pozycje cyklu zostały odpowiednio przedstawione. Wyniki zostały zaprezentowane w sposób uporządkowany, powiązany z pytaniami badawczymi i uzupełniony o interpretację oraz, miejscami, o analizę statystyczną. Autor wykazał, że potrafi nie tylko przeprowadzić eksperymenty, ale również zaprezentować ich rezultaty w sposób merytorycznie uzasadniony. **Ogólna ocena prezentacji wyników badań jest bardzo pozytywna.**

OCENA ORYGINALNOŚCI ROZPRAWY

Rozprawa doktorska mgr. Kwabena Frimponga Marfo wpisuje się w aktualny i istotny nurt badań z zakresu uczenia maszynowego dla danych rozproszonych i heterogenicznych. Przedmiotem pracy jest problem budowy globalnych modeli klasyfikacyjnych w sytuacji, gdy dane pozostają zdecentralizowane, nie mogą zostać fizycznie scalone, a ponadto różnią się między sobą zarówno zakresem atrybutów, jak i rozkładem obiektów. Autor trafnie identyfikuje, że klasyczne podejścia, takie jak uczenie scentralizowane, standardowe metody zespołowe czy typowe warianty federated learning, nie rozwiązują w satysfakcjonujący sposób problemu pełnej heterogeniczności lokalnych zbiorów danych. Podjęcie tej problematyki należy ocenić wysoko, ponieważ dotyczy ona realnego ograniczenia współczesnych systemów analizy danych, w szczególności tam, gdzie występują jednocześnie wymagania dotyczące ochrony prywatności danych i strukturalna ich niejednorodność.

Jako najistotniejsze elementy dorobku Doktoranta uważam w szczególności:

- opracowanie metod fuzji lokalnych predykcji klasyfikacyjnych z wykorzystaniem sieci neuronowych jako narzędzia budowy decyzji globalnej dla danych rozproszonych,

- zaproponowanie podejścia do budowy globalnego modelu neuronowego dla danych o niejednorodnych przestrzeniach cech poprzez harmonizację lokalnych zbiorów danych z użyciem sztucznie generowanych obiektów,
- rozwinięcie metody budowy modeli globalnych opartej na odwzorowaniu lokalnych danych do wspólnej przestrzeni cech z wykorzystaniem metod PCA, SVD i UMAP,
- przeprowadzenie porównawczej oceny architektur MLP i RBF w zadaniach integracji lokalnych modeli oraz wykazanie użyteczności tych architektur w warunkach dyspersji i heterogeniczności danych,
- określenie wpływu kluczowych właściwości danych rozproszonych, w tym poziomu szumu, stopnia dyspersji i różnorodności obiektów, na skuteczność proponowanych metod klasyfikacyjnych,
- wykazanie, że skuteczna budowa modeli globalnych w środowiskach rozproszonych jest możliwa bez centralizacji danych źródłowych i bez przyjmowania założenia o jednorodności przestrzeni cech.

Podsumowując, **oceniana rozprawa doktorska mgr. Kwabena Frimponga Marfo zawiera oryginalne podejście i rozwiązanie analizowanych problemów.** Stanowi ona wkład w rozwój metod klasyfikacji danych rozproszonych i heterogenicznych z wykorzystaniem sieci neuronowych, a jej główne osiągnięcie polega na zaproponowaniu i eksperymentalnym zweryfikowaniu kilku komplementarnych strategii budowy modeli globalnych bez centralizacji danych.

UWAGI DYSKUSYJNE ORAZ ZAGADNIENIA, KTÓRE W MOJEJ OCENIE WARTO ROZWINĄĆ W TRAKCIE OBRONY

Pomimo ogólnie bardzo pozytywnego wrażenia na temat dorobku Doktoranta, poniżej przedstawiam te elementy rozprawy, które w moim odczuciu warto byłoby szerzej omówić. Nie podważają one wartości naukowej przedstawionego dorobku, lecz dotyczą kwestii, które mogą stać się przedmiotem pogłębionej dyskusji i lepiej uwidocznić znaczenie oraz ograniczenia zaproponowanych rozwiązań.

- 1) W pierwszej kolejności celowe wydaje się szersze wyjaśnienie relacji pomiędzy trzema głównymi nurtami badawczymi rozwijanymi w rozprawie, tj. podejściem opartym na fuzji lokalnych predykcji, podejściem wykorzystującym harmonizację danych przez sztucznie generowane obiekty oraz podejściem bazującym na odwzorowaniu danych do wspólnej przestrzeni cech. Choć z rozprawy wynika logiczny rozwój kolejnych metod, interesujące byłoby bardziej jednoznaczne wskazanie, czy należy je traktować jako konkurencyjne warianty rozwiązania tego samego problemu, czy raczej jako kolejne etapy dojrzewania jednej koncepcji metodologicznej. W szczególności warto

byłoby zapytać, jakie przesłanki skłoniły Autora do przechodzenia od jednego typu rozwiązania do następnego i które ograniczenia wcześniejszych metod były w ten sposób przewyciężane.

- 2) Drugim zagadnieniem, które warto rozwinąć podczas obrony, jest kwestia granic stosowalności proponowanych metod. Autor pokazuje, że uzyskane wyniki są obiecujące, jednak z przedstawionych analiz wynika również, że jakość klasyfikacji może obniżać się wraz ze wzrostem stopnia dyspersji danych lub poziomu zakłóceń. Interesujące byłoby zatem doprecyzowanie, jak Autor ocenia praktyczne granice skutecznego działania proponowanych podejść oraz które z nich uważa za najbardziej odporne w warunkach skrajnej heterogeniczności lokalnych zbiorów danych. Tego rodzaju komentarz byłby cenny zarówno z poznawczego, jak i aplikacyjnego punktu widzenia.
- 3) Za zasadne uważam również postawienie pytania o zakres przewagi proponowanych metod nad rozwiązaniami referencyjnymi. Rozprawa przekonująco uzasadnia, że klasyczne podejścia federacyjne i rozproszone nie rozwiązują w pełni problemu niejednorodności przestrzeni cech, jednak w trakcie obrony warto byłoby szerzej omówić, z jaką klasą metod Autor uważa swoje rozwiązania za najbardziej bezpośrednio porównywalne. Innymi słowy, interesujące byłoby doprecyzowanie, czy proponowane podejścia należy postrzegać przede wszystkim jako alternatywę dla klasycznych metod zespołowych, dla wybranych wariantów federated learning, czy też jako odrębną klasę metod przeznaczoną dla szczególnego rodzaju środowisk danych rozproszonych.
- 4) Kolejne pytanie, które warto rozwinąć, dotyczy skalowalności obliczeniowej proponowanych rozwiązań. W rozprawie wskazano, że podejście oparte na ekstrakcji cech stanowi między innymi odpowiedź na koszt związany z generowaniem sztucznych obiektów w metodzie harmonizacji danych. Warto byłoby jednak szerzej omówić, jak Autor ocenia praktyczną złożoność obliczeniową obu strategii oraz w jakich warunkach jedna z nich powinna być preferowana względem drugiej. Z punktu widzenia dalszego rozwoju tych badań szczególnie interesujące byłoby wskazanie, które rozwiązanie ma największy potencjał do zastosowań wielkoskalowych.
- 5) Warto także zwrócić uwagę na zagadnienie interpretacji wyboru architektur neuronowych. Autor analizuje między innymi modele MLP i RBF, a w późniejszych pracach również inne architektury, w tym modele sekwencyjne. W trakcie obrony interesujące byłoby syntetyczne wyjaśnienie, jakie własności tych architektur okazały się kluczowe z punktu widzenia danych rozproszonych i heterogenicznych oraz czy na podstawie przeprowadzonych badań można sformułować bardziej ogólną rekomendację co do doboru klasy modelu do określonego typu problemu.

Wszystkie powyższe kwestie traktuję jako naturalne pola dalszej dyskusji naukowej. Nie zmieniają one mojej pozytywnej oceny rozprawy, lecz przeciwnie — potwierdzają, że przedstawiona praca podejmuje problematykę na tyle istotną i wielowątkową, iż uzasadnia ona pogłębioną debatę podczas obrony.

WNIOSEK KOŃCOWY

Biorąc pod uwagę zaprezentowany dorobek Kandydata, **pozytywnie oceniam rozprawę**. Rozprawa stanowi oryginalne rozwiązanie problemu naukowego i potwierdza odpowiedni poziom wiedzy teoretycznej i praktycznej Kandydata do uzyskania stopnia doktora.

Praca spełnia wymagania stawiane rozprawom doktorskim w dziedzinie nauk ścisłych i technicznych w dyscyplinie informatyka, w tym określone w obowiązującej ustawie o stopniach naukowych i tytule naukowym. Wnoszę o dopuszczenie jej Autora do publicznej obrony.

KONKLUZJA – OCENA KOŃCOWA ROZPRAWY

- 1) **BARDZO DOBRA, ZASŁUGUJĄCA NA WYRÓŻNIENIE**
- 2) POZYTYWNA, SPEŁNIAJĄCA WYMAGANIA
- 3) WARUNKOWA, WYMAGAJĄCA WPROWADZENIA PRPRAWEK I PONOWNEGO RECENZOWANIA
- 4) NEGATYWNA, NIE SPEŁNIAJĄCA WYMAGAŃ STAWIANYCH ROZPRAWOM DOKTORSKIM PRZEZ OBOWIĄZUJĄCE PRZEPISY

Recenzję sporządziła:
dr hab. inż. Agnieszka Wosiak, prof. uczelni