

prof. dr hab. inż. Andrzej Obuchowicz

Zielona Góra, 12 września 2026

Uniwersytet Zielonogórski

Instytut Sterowania i Systemów Informatycznych

a.obuchowicz@issi.uz.zgora.pl

Recenzja rozprawy doktorskiej

Pana mgr inż. Jakuba Kwiatkowskiego

zatytułowanej:

Staged Self-Supervision for Solving Raven's Progressive Matrices

opracowana na zlecenie

Rady Dyscypliny Naukowej

Informatyka Techniczna i Telekomunikacja

Politechniki Poznańskiej

1. Problem badawczy i jego znaczenie

Wobec niezwykłego wzrostu, od początków obecnego wieku, dynamiki rozwoju technik sztucznej inteligencji, szczególnie opartej na tzw. uczeniu głębokim, istotnym problemem staje się weryfikacja, wiarygodność i wyjaśnialność na poziomie rozumienia człowieka proponowanych przez nią rozwiązań. Są to wyzwania, których, niestety, znakomita większość obecnie aplikowanych systemów AI, opartych na głębokich sieciach neuronowych, nie jest w stanie w pełni sprostać.

W kognitywistyce i metodyce edukacji pojęcie uczenia głębokiego nie wiąże się z liczbą warstw i parametrów systemu sztucznych neuronów, ale oznacza pozyskiwanie wiedzy i umiejętności, które prowadzi do trwałego zrozumienia ważnych idei i procesów, oraz uzdalnia do jej wykorzystania w rozwiązywaniu nowych proble-

mów w odmiennych kontekstach. Próby budowania koncepcji systemów AI, które podlegałyby temu praradygmatowi uczenia, stanowią obecnie szczególnie interesujący, ważny i aktualny obszar badawczy. W tym kontekście idea samonadzorowanego uczenia etapowego, zaproponowana w opiniowanej pracy doktorskiej pana mgr inż. Jakuba Kwiatkowskiego, jawi się jako interesująca odpowiedź na niezwykle ważne i aktualne zagadnienia uczenia systemów AI.

2. Wkład autora

Zasadniczym osiągnięciem doktoranta jest propozycja dwuetapowego uczenia modelu rozwiązującego zadanie progresywnych matryc Ravena (RPM). W pierwszej kolejności, w wyniku procesu nienadzorowanego, system buduje wiedzę o rozwiązywanym problemie w postaci rozpoznania i izolacji abstrakcyjnych cech i własności, oraz relacji pomiędzy nimi, prowadząc do zestawu przewidywanych własności rozwiązania. W drugiej, na bazie tak reprezentowanej wiedzy, podejmuje decyzję. Otrzymany model wykazał się znaczną odpornością na błędy zbioru danych i wysoką efektywnością, porównywalną z najlepszymi wynikami raportowanymi w literaturze przedmiotu. Realizacja powyższego procesu wymagała zaproponowania szeregu szczegółowych rozwiązań i przeprowadzenia licznych analiz. Skutkiem ich realizacji powstało wiele autorskich propozycji, których najistotniejszy zestaw autor rozprawy zamieścił w podrozdziale 1.3 i w pełni się z nim zgadzam. Na szczególne wyróżnienie zasługują propozycje podstawowych elementów budowanego modelu, czyli transformatora abstrakcyjnej kompozycji (*Abstract Compositional Transformer: ACT*), czy zestawu alternatywnych modułów podejmowania decyzji: *Direct Choice Maker (DCM)*, *Learned Choice Maker (LCM)*, and *Contrastive Choice Maker (CCM)*.

Trudnym do opisanie i docenienia jest wkład przeprowadzonych przez doktoranta badań i analiz budowanego systemu w celu zrozumienia specyfiki kreowanej przez model przestrzeni abstrakcyjnych cech, relacji i samego procesu nienadzorowanej ich kreacji. Wiedza może okazać się niezbędna w zagadnieniu wyjaśniania procesu podejmowania decyzji przez model AI.

3. Poprawność

Opiniowana praca, zawarta na 184 stronach, składa się z 7 rozdziałów uzupełnionych jednym dodatkiem, zawierającym szczegółowy opis przebiegu i wyników eksperymentów oraz dodatkowe analizy wspierające treści ujęte w zasadniczej części pracy, oraz bardzo obszerną bibliografią zawierającą wyczerpujący zbiór ponad 230 publikacji obejmujących stan badań prowadzonych w obszarze zainteresowania pracy.

Motywacja podjęcia tematu pracy, sformułowanie problemu i, na jego podstawie, postawienie trzech hipotez pracy, oraz zestawienie wkładu badań autora do dyscypliny, zostały zawarte w pierwszym wprowadzającym rozdziale. Rozdział ten uzupełnia opis zawartości następujących kolejnych rozdziałów, jak również przedstawiono profil zawodowy i badawczy kandydata wraz z listą jego publikacji i prezentacji naukowych. Zadaniem rozdziału drugiego jest przybliżenie czytelnikowi teoretycznych podstaw koncepcyjnych i metodyki związanej z rozumowaniem abstrakcyjnym w sztucznej inteligencji, wraz z analizą ograniczeń dotychczas publikowanych rozwiązań, szczególnie związanych z uczeniem głębokim. Tematem rozdziału trzeciego jest przedstawienie powszechnie stosowanych baz RAVEN, I-RAVEN oraz RAVEN-FAIR, reprezentatywnych dla zadań abstrakcyjnego rozumowania wizualnego. W szczególności omówiono ich obciążenia związane zestawami odpowiedzi i ich podobieństwa.

Zasadniczą częścią rozprawy, która opisuje wyniki badań doktoranta, stanowią rozdziały 4, 5 i 6. Pierwsze dwa dotyczą dwóch etapów realizacji zadania: ekstrakcji abstrakcyjnych cech i relacji z analizowanych obrazów, oraz wskazania ostatecznego rozwiązania. Rozdział 4 zawiera opis transformatora abstrakcyjnej kompozycji, jego strukturę, metodykę uczenia oraz szczegółową analizę jego własności. Rozdział ten opisuje strategie tokenizacji, proces uczenia samonadzorowanego oraz analizuje mechanizmy, dzięki którym model uczy się przewidywać ustrukturyzowane, symboliczne reprezentacje właściwości paneli. W rozdziale 5 przedstawiono i dokonano analizy skuteczności trzech modułów dokonywania wyboru: Direct Choice Maker (DCM), wykorzystujący stałe metryki podobieństwa do porównywania wektorów właściwości, Learnable Choice Maker (LCM), stosujący wytrenowane sieci neuronowe do uczenia się adaptacyjnych funkcji porównawczych, oraz Contrastive Choice Maker (CCM), wykorzystujący uczenie kontrastowe do udoskonalenia

przestrzeni reprezentacji przed dokonaniem porównań pod kątem podobieństwa. Wyniki badań nad zdolnościami uogólniającymi proponowanego modelu zawarto w rozdziale 6. Badania te przeprowadzono z wykorzystaniem macierzy generowanych proceduralnie (PGM), a nie dotychczas stosowanej bazy RAVEN i jej pochodnych. Rozdział ten uzupełniono o analizę wykorzystania dużych modeli językowych do wnioskowania abstrakcyjnego oraz omawia obiecujące kierunki przyszłych badań. Ostatni rozdział podsumowuje wyniki eksperymentów, analizuje stopień uzasadnienia poprawności wszystkich trzech hipotez badawczych postawionych na początku rozprawy, oraz omawia szersze konteksty wyników badań.

Obrona struktura pracy jest właściwa i poprawnie zrealizowana. Na wyróżnienie zasługują dobrze napisane wprowadzenia i podsumowania poszczególnych rozdziałów pracy, dzięki nim nie umyka czytelnikowi główny tok realizacji pracy i zasadności prowadzonych analiz i badań w kontekście postawionych hipotez. Jest to istotne wobec wielowątkowości informacji zawartych w poszczególnych rozdziałach. Na wyróżnienie zasługuje również wysoka jakość edytorska pracy. Interesującym rozwiązaniem zastosowanym przez autora jest wyprowadzenie szczegółowego opisu przebiegu badań do dodatku na końcu dysertacji. Dzięki temu zabiegowi utrzymał przejrzystość przedstawianych rozważań i analiz prowadząc czytelnika zasadniczym tokiem dowodu hipotez unikając jego zagubienia w licznych szczegółach procesu badań.

Mankamentem doskwierającym czytelnikowi jest brak spisu skrótów tak licznie stosowanych w tekście pracy, zwłaszcza, że znaczna ich część dotyczy terminów utworzonych na rzecz ocenianej rozprawy. Tekst pracy, bogaty w liczne skróty staje się fragmentami nieczytelny i wymaga od czytelnika częstego wertowania stron poprzedzających w celu poszukiwania ich wyjaśnień. Ponadto brak jest formalnych definicji, w postaci wyrażeń matematycznych bądź proceduralnych, wskaźników jakości, dzięki którym porównywane są poszczególne warianty rozwiązań - jest to element, którego oczekuje się od raportu badawczego z obszaru nauk inżyniersko-technicznych. Opisowa forma wprowadzenia tych pojęć, np. w punkcie 5.1.1, nie jest precyzyjna.

4. Wiedza kandydata

Realizując podjęte w rozprawie doktorskiej zadania doktorant wykazał się głęboką wiedzą z zakresu systemów uczących się, systemów uczenia głębokiego, czy systemów rozumowania abstrakcyjnego. Zawartość treści w rozdziałach 1 i 2, także w wielu późniejszych fragmentach, dogłębnie i wyczerpująco zapoznaje czytelnika z aktualnym stanem wiedzy dziedzinowej, a także opis przebiegu badań w rozdziałach następnych wskazuje na szeroką wiedzę autora rozprawy w powyższych obszarach i umiejętność z niej korzystanie. Ponadto doktorant wykazał się poprawnością terminologiczną i bardzo szeroką znajomością literatury przedmiotu.

5. Inne

Lektura rozprawy nasuwa pewne kwestie warte przedyskutowania.

- W niniejszej rozprawie idea nienadzorowanego uczenia etapowego w zadaniach abstrakcyjnego rozumowania zrealizowana została na bazie RAVEN i bazach pochodnych. W jakim stopniu ta idea byłaby możliwa do zastosowania w zadaniach bazujących na danych niewizualnych, szeregach czasowych itp. ?
- Czy zastosowane kodowanie cech i relacji pomiędzy nimi przez moduł ACT pozwala odczytać relacje pomiędzy własnościami w sposób czytelny dla człowieka? Byłby to istotny krok w kierunku wyjaśnialności decyzji systemu.

6. Podsumowanie

Biorąc pod uwagę opinie zaprezentowane w poprzednich punktach i wymagania zdefiniowane przez artykuł 187 ust. 1 i 2 Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (z późniejszymi zmianami – Dz. U. 2024 r., poz. 1571) moja cena rozprawy pod względem trzech podstawowych kryteriów jest następująca:

A. Czy rozprawa zawiera oryginalne rozwiązanie problemu naukowego?

Zdecydowanie TAK

- B. Czy, po przeczytaniu rozprawy, zgadzasz się, że kandydat posiada ogólną wiedzę teoretyczną w dyscyplinie informatyka techniczna i telekomunikacja?

Zdecydowanie TAK

- C. Czy kandydat posiada umiejętność samodzielnego prowadzenia pracy naukowej?

Zdecydowanie TAK

W związku z powyższym **wnoszę** o przyjęcie rozprawy i dopuszczenie jej do publicznej obrony.

