

## *Prof. dr hab. inż. Ryszard Tadeusiewicz*

rtad@agh.edu.pl; www.tadeusiewicz.pl; 30-059 Kraków, al. Mickiewicza 30  
Katedra Biocybernetyki i Inżynierii Biomedycznej wydziału EAIIB AGH  
Doktor Honoris Causa piętnastu uczelni krajowych i zagranicznych  
Członek Polskiej Akademii Nauk; Członek Polskiej Akademii Umiejętności;  
Były Rektor AGH; Były Prezes Krakowskiego Oddziału PAN; Były członek CK  
Członek Akademii Inżynierskiej, член Российской Академии Естественных Наук  
Participle Pleno Jure Academiae Europensis Scientiarum Artium Litterarumque  
Fellow of World Academy of Art and Science; Euro-engineer FEANI  
Senior Member of IEEE; professional member of ACM; member of SPIE

Kraków, 02.12.2024

## Recenzja

Przedmiotem tej recenzji jest rozprawa doktorska mgr inż. Artura Chudzika zatytułowana „*Machine Learning Methods for Parkinson’s Disease Datasets*” (w tłumaczeniu „Metody uczenia maszynowego dla zbiorów danych o chorobie Parkinsona”). Promotorem rozprawy jest Dr hab. Andrzej Przybyszewski, prof. PJATK. Rozprawa jest przedstawiana Radzie Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja Polsko-Japońskiej Akademii Technik Komputerowych. Podstawą do sporządzenia niniejszej recenzji jest skierowane do mnie zlecenie Przewodniczącej wymienionej Rady Naukowej, prof. dr hab. Marii Elżbiety Ostrowskiej (pismo bez numeru datowane w Warszawie 7.11.2024) w zlecenie jest mi wykonanie niniejszej recenzji.

### Ocena celu i zakresu badań oraz kwestia tezy naukowej

Cel i zakres badań został wskazany w tytule ocenianej rozprawy. Wskazane zostało, że przedmiotem zainteresowania Doktoranta będą metody uczenia maszynowego wykorzystywane do różnych analiz zbiorów danych o chorobie Parkinsona. Ten obszerny cel był w rozprawie realizowany poprzez podejmowanie i rozwiązywanie szeregu problemów w publikacjach naukowych będących główną treścią rozprawy. **Natomiast cel ten nie został nigdzie skonkretyzowany w postaci wyrażonej jawnie tezy naukowej, dla której treść rozprawy mogłaby stanowić dowód prawdziwości.** To, co zapisano na przełomie stronic 9 i 10 pod nazwą Hipotezy tej szeroko przyjętej w Polsce formie nie odpowiada.

Nie zamierzam wykorzystywać tego stwierdzenia do formułowania krytycznych uwag pod adresem rozprawy i jej Autora. Jednak stwierdzam, że opiniowana praca **odbiega** w tym miejscu od schematu prac doktorskich przyjętego i powszechnie stosowanego w Polsce. Stwierdzam to na podstawie porównania ocenianej tu rozprawy z 357 rozprawami doktorskimi dla których do tej pory miałem okazję pisać recenzje. Ta jedna praca jest od wszystkich tamtych **zdecydowanie odmienna**. Zapewne Promotor, który jest – obok pracy w PJATK - wybitnym Uczonym o ogromnym prestiżu w USA pokierował Doktorantem zgodnie z normami przyjętymi w Stanach Zjednoczonych. Nie próbuję tych norm kwestionować, odnotowuję jednak, że przewód doktorski toczy się w Polsce zgodnie z

wymogami polskiej Ustawy „Prawo o szkolnictwie wyższym i nauce”, więc pewien usus przyjęty w Polsce dla rozpraw doktorskich powinien być brany pod uwagę. Dotyczy to wspomnianego braku wyartykułowania dowodzonej tezy naukowej, całej struktury i formy rozprawy, a także trzeciorzędnego w istocie szczegółu, dodatkowo odróżniającego opiniowaną tu rozprawę od zwyczajów przyjętych w Polsce. Otóż polskie prace naukowe z zasady pisane są w formie bezosobowej: „zbadano, stwierdzono, wykryto” itp. W języku angielskim taka forma bezosobowa rzadziej bywa wykorzystana, ale praktykuje się, że zamiast twardego „I” używa się często liczby mnogiej: „we”. Znana jest historia autora, który w nagłówku artykułu dopisywał imię swego kota, żeby uzasadnić owo „we”.

Natomiast mgr Chudzik wszędzie stosuje formę pierwszej osoby liczby pojedynczej: „Ja zbadałem, Ja stwierdziłem, Ja wykryłem”. Jest to oczywiście drobiazg, ale polski czytelnik może to odbierać jako przejaw arogancji. Nie zamierzonej, ale bijącej po oczach...

## **Komentarze i oceny dotyczące zawartości rozprawy**

Opiniowana rozprawa w całości napisana jest w języku angielskim, ale jako integralna część rozprawy występuje streszczenie w języku polskim. Wprawdzie streszczenie to nie odzwierciedla (według mojej oceny) całej zawartości rozprawy, to jednak pozwala zorientować się w tym, co Autor robił, do czego dążył i co uzyskał, zatem w tym zakresie spełnione są wymagania cytowanej wyżej Ustawy.

Rozprawa liczy (wraz ze stroną tytułową i różnymi wykazami oraz zestawieniami) 203 stronicę. Dysertacja oparta jest głównie na zestawie powiązanych tematycznie publikacji, w których Doktorant jest współautorem. Osiem publikacji zadeklarowano jako artykuły konstruujące rozprawę (oznaczone [A1] – [A8]), a ponadto przytoczono trzy publikacje jako dodatkowe składniki (oznaczone [C1] – [C3]) których teksty także w niej przywołano,

We Wstępie (numerowanym cyfrą rzymską I) wskazano, że badania podjęte w ramach rozprawy stanowią ważny krok w kierunku oceny metod obliczeniowych wspomagających lekarzy w diagnozowaniu chorób mózgu ze szczególnym uwzględnieniem choroby Parkinsona. Mgr Chudzik słusznie zauważa, że pomiędzy potrzebami a możliwościami jest tu spora luka (*gap*) i stawia hipotezę, że istnieje potrzeba oceny metod uczenia maszynowego wykorzystywanych do analizy zbiorów danych dotyczących choroby Parkinsona. Następnie Autor deklaruje, że przedkładana rozprawa właśnie takich ocen dostarcza.

Ponadto we Wstępie opisano strukturę pracy, co ułatwia orientację w jej dość skomplikowanej zawartości.

Rozdział II zawiera omówienie zbiorów danych dotyczących choroby Parkinsona (PD). Przedstawiony jest schematycznie przepływ sygnałów pobudzających i hamujących w mózgu ze wskazaniem zaburzenia w produkcji dopaminy pochodzącej z pojawiającego się procesu neurodegeneracyjnego w istocie czarnej które powoduje objawy kliniczne choroby. Przedyskutowano także nakładające się zestawy objawów w danych klinicznych i wpływ leczenia objawowego na dane diagnostyczne. Na tej podstawie opisano sposoby klasyfikacji postępów choroby i używane skale pomiarowe. Uwzględniono fenotypowanie cyfrowe i rolę obrazowania mózgu za pomocą magnetycznego rezonansu przy ustalaniu stopnia rozwoju choroby. Wspomniano także o zaburzeniach w ruchu gałek ocznych jako o metodzie oceny zaawansowania PD, a także o metodach opartych na testach generowanych komputerowo.

Rozdział III ocenianej rozprawy przedstawia modele, jakie są wykorzystywane do klasyfikacji choroby Parkinsona. Autor proponuje metodykę tworzenia i wykorzystania takich modeli, analizę statystyczną, jaka może być stosowana i metody uczenia maszynowego, będące głównym przedmiotem zainteresowania w ocenianej rozprawie. Zasygnalizowane jest stosowane podejście obliczeniowe, optymalizacja i weryfikacja. Przywołane są szeroko stosowane w odniesieniu do badań medycznych metody regresji logistycznej, zasygnalizowane jest użycie lasów losowych i obliczeń granularnych. Wspomniane są także prace dotyczące obliczeniowej teorii umysłu.

We wszystkich wymienionych wyżej omówieniach Doktorant nawiązuje do swoich prac (zaprezentowanych in extenso dalej) jako do źródeł prezentowanych spostrzeżeń i wniosków, ale owe własne wkłady szczegółowiej omawia mgr Chudzik w rozdziale IV. Rozdział ten uważam za główne jądro opiniowanej pracy i omówię go teraz bardziej szczegółowo.

Pierwszym zagadnieniem, które było przedmiotem badan mgra Chudzika, było analizowanie przydatności metod uczenia maszynowego do predykcji symptomów choroby Parkinsona na podstawie tensora dyfuzji DTI po dokonaniu zabiegu DBS (głębokiej stymulacji mózgu). Wyniki powadzonych badań były publikowane w 2019 roku w artykule zatytułowanym *“DTI Helps to Predict Parkinson’s Patient’s Symptoms Using Data Mining Techniques”*. W publikacji tej (oznaczonej w rozprawie jako **[A1]**) wykazano, że metody uczenia maszynowego pozwoliły przewidywać wyniki trzech rozważanych objawów z dokładnością odpowiednio 0,967, 0,824 i 0,878 - co trzeba uznać za wynik naprawdę bardzo dobry.

Drugim badanym zagadnieniem było wykorzystanie uczenia maszynowego do różnicowania stopnia zaawansowania choroby Parkinsona na podstawie szybkich ruchów gałek ocznych (sakkady, antysakkady i pogoń). Wyniki opublikowano w 2023 roku w pracy oznaczonej **[A2]**, zatytułowanej

*“Machine Learning and Eye Movements Give Insights into Neurodegenerative Disease Mechanisms”*

Stosowano wiele metod uczenia maszynowego i wykazano, że wielowymiarowa analiza danych na temat ruchów gałek ocznych pozwala na wyciągnięcie wielu wniosków na temat mechanizmów neurodegeneracyjnych w chorobach centralnego systemu nerwowego.

Trzecim zagadnieniem, które było przedmiotem badania mgr Chudzika, było porównanie różnych metod drążenia danych (Data Mining) w zastosowaniu do określania rozwoju choroby w odniesieniu do różnych grup pacjentów z chorobą Parkinsona. Wyniki opublikowano w 2020 roku w pracy zatytułowanej *“Comparison of Different Data Mining Methods to Determine Disease Progression in Dissimilar Groups of Parkinson’s Patients”* oznaczonej w tekście rozprawy symbolem [A3].

Zastosowane metody uczenia maszynowego miały posłużyć do tego, by pomagać w wyborze optymalnej metody terapii. Wyniki oceniam (po przestudiowaniu całego artykułu dołączonego dalej do tekstu rozprawy) jako bardzo ciekawe naukowo i obiecujące z punktu widzenia zastosowań praktycznych.

Kolejnym zagadnieniem poruszonym w rozprawie było ponowne zajęcie się sprawą parametrów ruchu gałek ocznych jako przesłanek do przewidywania symptomów choroby Parkinsona.

Odpowiedni artykuł, oznaczony jako [A4] ma tytuł *“Eye-Tracking and Machine Learning Significance in Parkinson’s Disease Symptoms Prediction”* i został opublikowany w 2020 roku. Tym razem zastosowano metody uczenia do przewidywania podklas na skali UPDRS dla pacjentów w zaawansowanych stadiach choroby Parkinsona. Wprawdzie Autor przestrzega, że praktyczne zastosowanie wniosków musi być ograniczone, ponieważ badania prowadzone były na dość ograniczonych liczebnie kohortach pacjentów, niemniej wyniki są zdecydowanie wartościowe.

Źródłem informacji w kolejnych dwóch publikacjach (oznaczonych odpowiednio [A5] oraz [A6] i mających tytuły *“Machine Learning and Digital Biomarkers Can Detect Early Stages of Neurodegenerative Diseases”* oraz *“Investigating the Impact of Parkinson’s Disease on Brain Computations: An Online Study of Healthy Controls and PD Patients”* były dane z internetowych testów kognitywno- behawioralnych. Zastosowano także metodę regresji logistycznej. Wyniki były zachęcające (dokładność ponad 91%), chociaż zdarzały się także pojedyncze klasyfikacje błędne.

W kolejnej publikacji zatytułowanej *“Classification of Parkinson’s Disease Using Machine Learning with MoCA Response Dynamics”*, opublikowanej w 2024 roku (oznaczenie [A7]), autorzy oparli się na wynikach testu MoCA (Montreal Cognitive Assessment) i zastosowali różne nowoczesne metody uczenia maszynowego uzyskując dobre wyniki klasyfikacji. Badanie było kontynuowane i stało się podstawą do stworzenia w 2024 roku kolejnej publikacji oznaczonej [A8], zatytułowanej *“Recognizing*

*Patterns of Parkinson's Disease using Online Trail Making Test and Response Dynamics – Preliminary Study*". Badano test w wersji TMT A oraz TMT B ponownie stosując całą gamę narzędzi uczenia maszynowego.

Dodatkowo opisano badania prowadzone w przestrzeni wirtualnej, także oryginalne i przynoszące ciekawe rezultaty.

Zamieszczone przez Doktoranta podsumowanie wkładu naukowego, jaki wniosły scharakteryzowane wyżej badania i publikacje oparte na wynikach owych badań pokazuje, jak wiele wartościowych wyników związanych z szeroko rozumianą diagnostyką choroby Parkinsona udało się uzyskać, dzięki mądrym i bardzo profesjonalnemu wykorzystaniu metod uczenia maszynowego – naprawdę jest imponujące. To wrażenie pogłębia zamieszczona w rozdziale V Dyskusja, a także trafne i w pełni uzasadnione wnioski, zawarte w bardzo krótkim rozdziale VI.

### **Ocena całościowa**

Muszę powiedzieć, że jestem pod wrażeniem ilości i jakości oryginalnych wyników naukowych składających się na ocenianą rozprawę. Jednak czytając (z uznaniem!) te artykuły widziałem w nich głównie osiągnięcia w obszarze dyscypliny naukowej, która nazywa się Inżynieria Biomedyczna. Tam owa praca ocierała by się wręcz nie o doktorat, ale o dobrą habilitację.

Natomiast przewód doktorski jest prowadzony w dyscyplinie Informatyka Techniczna i Telekomunikacja. Musiałem więc – odsuwając imponujące osiągnięcia Kandydata związane z Inżynierią Biomedyczną – szukać uzasadnienia dla nadania stopnia naukowego w Informatyce Technicznej.

Stwierdzam, że nie było to łatwe.

Nie ulega wątpliwości, że mgr Chudzik używa narzędzi informatycznych z ogromną wprawą, wręcz byłbym skłonny nazwać to wirtuozerią. Ale nie zdobywa stopni naukowych w informatyce każdy, kto trafnie użyje komputera. Musiałem sobie zadać pytanie, w jakim stopniu publikacje Doktoranta wzbogaciły informatykę jako taką. Przyjąłem, że prace Kandydata wnoszą wkład do informatyki, ponieważ realizując swoje cele (medyczne!) projektował on, budował, a potem także opisywał nowe platformy cyfrowe. Co więcej, prezentował on sposób „przepływu pracy” w zakresie stosowania metod uczenia maszynowego, a także opracowywał i publikował metody uzyskiwania potrzebnych wyników naukowych w sytuacjach, kiedy zbiory danych były bardzo złożone oraz mało liczne. Porównując to do wielu prac doktorskich z dziedziny czystej informatyki, jakie miałem okazję

recenzować – stwierdzam, że stopień naukowy doktora w dyscyplinie Informatyka Techniczna i Telekomunikacja może być magistrowi Arturowi Chudzikowi nadany *lege artis*.

### **Ocena dodatkowych elementów rozprawy**

Ocenianą rozprawę doktorską, której merytoryczny trzon oceniłem wyżej, uzupełniają elementy dodatkowe, które także chciałbym ocenić.

Pierwszym takim dodatkowym elementem jest słownik ważniejszych terminów i skrótów. Bardzo pochwalam jego zamieszczenie w rozprawie, bo podczas studiowania zasadniczej części dysertacji wielokrotnie sięgałem do tego słownika, żeby upewnić się, że dobrze rozumiem wywody Autora. Z przyjemnością odnotowuję, że ani razu się nie zawiodłem!

Kolejnym elementem nieodzownym w każdej pracy naukowej jest wykaz literatury. Nie jest on ponumerowany, więc nie mogę stwierdzić z ilu materiałów źródłowych mgr Chudzik korzystał, ale z całą pewnością jest to liczba pokaźna. Nie byłem także w stanie sprawdzić, czy wszystkie pozycje podane w wykazie literatury zostały w tekście rozprawy przywołane i wykorzystane. Generalnie jednak stwierdzam, że warunek (zapisany w polskiej Ustawie!) że rozprawa doktorska ma dowodzić, iż Doktorant ma odpowiednią wiedzę i orientację w najnowszych trendach uprawianej dyscypliny naukowej – jest spełniony.

Dosyć nietypowym elementem ocenianej pracy jest wykaz streszczeń artykułów (tych głównych i tych dodatkowych), których mgr Chudzik jest współautorem i które zostały wykorzystane w rozprawie. W tym wykazie streszczeń podano dokładnie dla każdej pracy, gdzie była ona publikowana, ile punktów ministerialnych przyniosła, a także Doktorant określił, na czym polegała jego rola w jej tworzeniu. Jest to bardzo rzetelne opracowanie.

Poczynając od 65 strony dysertacji aż do jej końca (na stronie 203) są w niej przytoczone in extenso teksty wykorzystanych w rozprawie publikacji, których mgr Artur Chudzik był współautorem. To bardzo ważny element, bowiem ilekroć czytałem w głównym tekście pracy jakieś elementy opisujące, co w tej czy innej publikacji zawarto – mogłem sięgnąć do źródła i rozwiązać ewentualne wątpliwości. Samych publikacji nie oceniam, gdyż przeszły one stosowną ewaluację przez recenzentów powołanych przez organizatora konferencji lub wydawcę czasopisma – i zostały przez nich ocenione pozytywnie, skoro ukazały się drukiem.

## Podsumowanie i wniosek końcowy

Przechodząc do finalnej oceny rozprawy stwierdzam, że opiniowana rozprawa doktorska mgr inż. Artura Chudzika zatytułowana „*Machine Learning Methods for Parkinson’s Disease Datasets*” zdecydowanie spełnia warunki stawiane rozprawom doktorskim.

Biorąc pod uwagę powyższe konstatacje wnioskuję do wysokiej Rady Naukowej Dyscypliny Informatyka Techniczna i Telekomunikacja PJATK o jej **przyjęcie** oraz o dopuszczenie jej autora, mgra Artura Chudzika do publicznej **obrony**, a po jej pozytywnym zakończeniu będę z całym przekonaniem głosował za **nadaniem mu stopnia naukowego doktora** w dyscyplinie Informatyka Techniczna i Telekomunikacja.

Ponadto biorąc pod uwagę zdecydowanie ponadprzeciętną ilość i wartość przedstawionych w rozprawie opracowań, koncepcji i wyników naukowych, a także uwzględniając fakt, że Doktorant może się dodatkowo wylegitymować wieloma wartościowymi publikacjami – wnioskuję o **wyróżnienie** tej rozprawy.

