

Lublin, 25 maja 2022

dr hab. Przemysław Matuła, prof. UMCS
Katedra Matematyki Stosowanej
Instytut Matematyki
Uniwersytet Marii Curie-Skłodowskiej w Lublinie

Recenzja w postępowaniu habilitacyjnym doktora Wojciecha Rejchela

Poniżej przedstawię ocenę osiągnięcia naukowego, które wskazał doktor Wojciech Rejchel, jako podstawę nadania stopnia doktora habilitowanego, zgodnie z art. 219 ust. 1 pkt. 2 Ustawy z dnia 20 lipca 2018 "Prawo o szkolnictwie wyższym i nauce". Osiągnięcie to skrótowo i zamiennie nazywane będzie rozprawą habilitacyjną. Omówię również pozostały dorobek kandydata uzyskany po doktoracie, w tym dorobek naukowy, dydaktyczny i organizacyjny.

1. Sylwetka kandydata.

Pan Wojciech Rejchel urodził się 9 czerwca 1981 roku. Wiedzę matematyczną zdobywał pod kierunkiem wybitnych polskich specjalistów z dziedziny teorii prawdopodobieństwa i statystyki matematycznej. Kształcił się na Wydziale Matematyki i Informatyki w Uniwersytecie Mikołaja Kopernika w Toruniu. Pracę magisterską "*Estymacja monotonicznej gęstości*" napisał pod kierunkiem prof. dr hab. Tomasza Rychlika i obronił w roku 2005. Rozprawę doktorską "*Statystyczne modele regresji rangowej*" przygotował pod kierunkiem prof. dr hab. Wojciecha Niemirowicza i z wyróżnieniem obronił w roku 2011. Od roku 2010 do chwili obecnej pracuje na Wydziale Matematyki i Informatyki w Uniwersytecie Mikołaja Kopernika w Toruniu, początkowo przez rok na stanowisku asystenta, a później adiunkta.

2. Rozprawa habilitacyjna.

Rozprawę habilitacyjną stanowi cykl ośmiu artykułów naukowych pod wspólnym tytułem "*M – estymatory z karą w wyborze modelu*". Pięć prac ma współautorów, ale ze złożonych oświadczeń wynika, że wkład doktora Rejchela w ich powstanie był znaczący i kluczowy. Wszystkie prace zostały opublikowane w bardzo dobrych czasopismach naukowych, według aktualnie obowiązującej punktacji ministerialnej, cztery z nich w czasopismach, którym przypisano 140 punktów, jedna za 100 punktów, dwie za 70 punktów i jedna za 40 punktów. Cztery prace zostały opublikowane w czasopismach, do których nie została przypisana dyscyplina naukowa matematyka, są to – *Journal of Machine Learning Research*, *Entropy* i *Neurocomputing*. Czasopisma te przypisane są głównie do dyscypliny informatyka. Przyczyną jest oczywiście specyfika badań naukowych prowadzonych przez doktora Rejchela, są one nakierowane na zastosowania matematyki, statystyki oraz implementację metod i

algorytmów z użyciem technik informatycznych. Skutkiem takiego stanu rzeczy jest również fakt, że w matematycznej bazie Mathematical Reviews (MathSciNet) widnieje jedynie 11 spośród wszystkich prac Kandydata, a w bazie zbMATH Open (Zentralblatt MATH) – 10. Indeks cytowań jest w związku z tym, w tych bazach, bardzo skromny, warto więc posłużyć się innymi indeksami – Web of Sciences lub Scopus, gdzie ujęte są wszystkie prace autora. Oceniając rozprawę habilitacyjną należy, wobec powyższego, zwrócić uwagę na jej interdyscyplinarność oraz nakierowanie na zastosowania matematyki w różnych dziedzinach wiedzy.

Do dokumentacji dołączony został Autoreferat, w którym między innymi szczegółowo zostały opisane wyniki rozprawy habilitacyjnej. Dostrzegłem w nim nieliczne literówki – na przykład w tytule Sekcji 4 lub na stronie 4 linia +1. Kandydat nie był konsekwentny w używaniu formy bezosobowej lub pierwszej osoby liczby mnogiej i na przykład w paragrafie 4.1.3 zaczął używać pierwszej osoby liczby pojedynczej. Niemniej jednak Autoreferat jest bardzo dobrze napisany i w dużym stopniu pomaga w zorientowaniu się w dorobku Autora.

3. Ocena osiągnięcia.

W Autoreferacie rozprawy habilitacyjnej Pan doktor Rejchel szczegółowo omówił wyniki naukowe przedstawione jako osiągnięcie "*M-estymatory z karą w wyborze modelu*". W Autoreferacie dokonano pewnego ujednoczenia oznaczeń, które nie były spójne w różnych pracach. Przyjęta została konwencja pogrupowania prac stosownie do wspólnych zagadnień w nich poruszanych. Jednak, w niniejszej ocenie omówię prace włączone do habilitacji w takiej kolejności jak zostały wymienione.

W rozprawie rozważany jest problem doboru zmiennych objaśniających do modelu, gdy ich liczba p w stosunku do liczby obserwacji n jest duża (przypadek wysokowymiarowy, gdy p jest większe lub porównywalne z n) lub mniejsza (przypadek niskowymiarowy). Typowy schemat postępowania pojawiający się w rozprawie to estymacja parametrów modelu poprzez minimalizację pewnej funkcji straty powiększonej o karę za poziom skomplikowania modelu. Zwłaszcza przypadek wysokowymiarowy jest intensywnie badany w ostatnich latach. Szczególnie popularną metodą stała się metoda LASSO i jej modyfikacje. Temu zagadnieniu poświęcona jest obszerna praca [A1]. Zaproponowany został w niej algorytm SS (Screening-Selection), konkurencyjny względem używanych do tej pory metod TL, SCAD lub MCP. Zaproponowana metoda polega, w pierwszym kroku na wyznaczeniu estymatora LASSO i uporządkowaniu otrzymanych wartości, a następnie w drugim kroku na wyborze modelu, który minimalizuje GIC – uogólnione kryterium informacyjne. Zaletą tego algorytmu, w stosunku do konkurencyjnych, jest jego konstruktywność (czyli niezależność od nieznanymi parametrów) oraz większa efektywność obliczeniowa. W pracy [A1] zostały podane oszacowania eksponencjalne dla prawdopodobieństwa błędu selekcji. Prezentowane wyniki zostały uzyskane przy słabszych założeniach niż we wcześniejszych pracach innych autorów. Matematyka stająca za uzyskanymi

twierdzeniami nie jest zbyt wyrafinowana, sprowadza się do podstawowej wiedzy z zakresu teorii prawdopodobieństwa, statystyki i teorii optymalizacji. Znaczenie zaproponowanej metody podkreślają przeprowadzone symulacje i zastosowanie do danych rzeczywistych, ponadto algorytm SS został zaimplementowany w języku R, czyniąc go dostępnym dla wszystkich zainteresowanych praktyków.

W pracy [A2] rozpatrywany jest problem wyboru predyktorów, gdy zmienna odpowiedzi zależna jest od liniowej kombinacji zmiennych niezależnych poprzez funkcję, o której zakłada się jedynie monotoniczność, oraz szumu, który nie musi być nawet całkowalny. Zaproponowana została procedura nazwana RankLasso oraz jej modyfikacje – progowa oraz ważona. Wyniki teoretyczne zostały poparte symulacjami i analizą danych rzeczywistych wskazując na przewagę nowego podejścia nad procedurami stosowanymi dotychczas. Praca [A2] jest bardzo obszerna i niestety wprowadzenie nadmiaru dodatków (Appendixów) powoduje pewien chaos i nie ułatwia jej czytania. Nie ma to oczywiście wpływu na pozytywną ocenę uzyskanych rezultatów. W dowodach, zwłaszcza twierdzeń granicznych, wykorzystane zostały zaawansowane metody teorii U -statystyk.

Wśród rozpatrywanych przez Habilitanta modeli nie mogło zabraknąć przypadku, gdy zmienna odpowiedzi jest dwuwartościowa. Temu zagadnieniu poświęcony jest artykuł [A3]. Mamy do czynienia wtedy z klasyfikacją, w tym przypadku wielowymiarową, gdzie $p \gg n$. Rozpatrywany klasyfikator jest liniowy i zastosowane zostały techniki LASSO. Wyniki tej pracy wzbogacone zostały badaniami eksperymentalnymi i symulacjami.

Wysoko oceniam artykuł [A4], w którym przedstawiony został nowy algorytm, nazywany LASSO-SD. Stosowany jest on w przypadku wielowymiarowym, gdy liczba predyktorów jest dużo większa niż liczba danych. W opracowanej procedurze najpierw usuwana jest część najmniej istotnych predyktorów za pomocą algorytmu LASSO, a następnie wybierany jest model w drodze wielokrotnego testowania. Należy zaznaczyć, że jest to konstruktywny sposób postępowania, czyli nie ma zależności poszczególnych kroków od nieznanymi parametrów. W pracy dokonano symulacji, analizy danych rzeczywistych i porównania z konkurencyjnymi procedurami.

Badanie zależności między zmiennymi losowymi jest bliskie moim zainteresowaniom naukowym. Temu zagadnieniu poświęcona jest praca [A5], w której struktura zależności wektora losowego utożsamiana jest ze strukturą pewnego grafu, a ta dalej opisywana jest współczynnikami, które mogą być estymowane za pomocą technik LASSO i ich modyfikacji. W rozpatrywanym modelu Isinga podane są warunki na zgodną selekcję estymatora wyżej wymienionych parametrów, niestety procedura ta nie jest konstruktywna, ale w pracy przedstawione są sposoby obejścia tego i innych pojawiających się problemów. Wagę uzyskanych rezultatów podkreślają zastosowania, implementacje i numeryczne obliczenia.

W pracy [A6] empiryczna wersja ryzyka przyjmuje postać U -statystyki, a ostatecznie minimalizowana jest ta funkcja z dodanym członem kary. Artykuł ten można traktować jako udoskonalenie pracy Song i Ma z roku 2010. Główne ulepszenie polega na zastosowaniu wypukłej funkcji straty (bez dodatkowych założeń różniczkowalności lub monotoniczności), co pozwala zagwarantować istnienie minimum globalnego. Ponadto, osłabionych zostało szereg wcześniej wykorzystywanych założeń. Dla mnie, w tej pracy, szczególnie interesujące były dowody słabej zbieżności przy badaniu asymptotycznego zachowania ważonych estymatorów LASSO.

Teoria U -statystyk została również wykorzystana w pracy [A7], w której rozpatrywany jest przypadek wysokowymiarowy, a estymatory otrzymywane są poprzez minimalizację U -statystyk z karą LASSO. Otrzymane zostały nierówności pozwalające ocenić jakość estymatorów wykorzystywanych do predykcji oraz wyboru modeli opartych na rangowaniu. Wyniki teoretyczne ilustrowane są symulacjami.

Ostatnia praca [A8] wchodząca w skład osiągnięcia jest artykułem samodzielnym Habilitanta. Dotyczy ona przypadku niskowymiarowego – gdy liczba predyktorów p jest ustalona. Rozpatrywany jest estymator LASSO oraz ważona wersja tego estymatora. Najciekawszym wynikiem tej pracy są warunki – konieczny i dostateczny na selekcyjną zgodność estymatora LASSO. Udowodnione zostało również twierdzenie graniczne dla estymatora ważonego. Należy również odnotować konstruktywność tego estymatora, co umożliwia jego praktyczne zastosowanie. Uzyskane wyniki teoretyczne zostały zastosowane do analizy danych symulowanych oraz danych rzeczywistych.

Część artykułów włączonych do rozprawy habilitacyjnej została opublikowana w czasopiśmie nie przypisanym do dyscypliny matematyka, nie mam jednak wątpliwości, że są to prace matematyczne, zawierające rygorystyczne dowody. Chociaż w dowodach nie są stosowane wyrafinowane metody matematyczne, to rozprawa habilitacyjna wymagała szerokiej wiedzy z zakresu teorii prawdopodobieństwa, statystyki matematycznej, analizy wypukłej, teorii optymalizacji, a także algorytmiki i programowania. Opracowane lub udoskonalone zostały w niej procedury wyboru modeli, które mogą być bardzo szybko wykorzystane w praktyce. Podkreślić należy objętość prac składających się na rozprawę habilitacyjną, wykonaną pracę teoretyczną oraz symulacje lub analizy danych rzeczywistych. Wielkość i wszechstronność wykonanej pracy jest imponująca. Uzyskane wyniki należy zakwalifikować do zastosowań matematyki i stanowią one istotny wkład w rozwój dyscypliny naukowej matematyka.

4. Pozostały dorobek.

Na pozostały dorobek naukowy Habilitanta składają się dwie prace opublikowane przed uzyskaniem stopnia doktora i dziewięć prac opublikowanych po doktoracie. Spośród prac, które ukazały się drukiem po doktoracie – trzy są to prace samodzielne, natomiast sześć ma współautorów. Opublikowane zostały w bardzo

dobrych i dobrych czasopismach z przypisaną dyscypliną matematyka, jak choćby *Annales de l'Institut Henri Poincaré, Probability and Mathematical Statistics* oraz świetnych czasopismach przypisanych do dyscypliny informatyka – *Bulletin PAN: Technical Sciences, Neurocomputing, Journal of Machine Learning Research, Neural Computation*.

Pozostały dorobek naukowy koncentruje się wokół trzech zagadnień – estymatorów rangowych (tematyka zainicjowana w doktoracie), metod Monte-Carlo w zastosowaniu do estymatorów największej wiarygodności i zastosowaniu metod matematycznych w praktycznych problemach geodezji i kartografii. Szczególnie wysoko oceniam prace [B3], [B5] i [B7], w których otrzymywane i badane są estymatory największej wiarygodności metodami Monte-Carlo. W wielu przypadkach estymatory największej wiarygodności opisywane są trudnymi do rozwiązania układami równań, dlatego rozwój technik badania tych estymatorów jest szczególnie istotny. Prace [B6] i [B11] nie zawierają zaawansowanej matematyki, ale pokazują umiejętności Habilitanta do współpracy interdyscyplinarnej i wykorzystaniu matematyki w geoinformatyce.

Pan doktor Rejchel wielokrotnie prezentował swoje wyniki naukowe na konferencjach naukowych i seminariach. Po uzyskaniu stopnia doktora blisko trzydzieści razy wygłaszał odczyty zaproszone i zgłoszone na konferencjach krajowych i zagranicznych. Kilkukrotnie miałem okazję wysłuchać wystąpień Habilitanta. Był zawsze bardzo dobrze przygotowany, mówił ciekawie, w sposób przystępny i pewny. Dał się zapamiętać jako bardzo dobry mówca. Referaty Pana Rejchela były nagradzane na konferencjach w konkursach na najlepszy referat wygłoszony przez młodych naukowców.

Pozytywnie oceniam działalność Kandydata w zakresie realizacji projektów i pozyskiwania środków grantowych – kilkukrotnie uczestniczył w projektach OPUS z NCN jako wykonawca, a w projekcie FUGA z NCN pełnił funkcję kierownika.

Pan doktor Wojciech Rejchel współpracuje z wieloma naukowcami w kraju oraz za granicą (wizyta naukowa w Uniwersytecie Hubei w Wuhan w Chinach, której efektem była publikacja [B9]).

W zakresie działalności dydaktycznej Kandydata, należy podkreślić udział w kształceniu doktorantów – pełnienie funkcji promotora pomocniczego w dwóch przewodach oraz prowadzenie wykładów na studiach doktoranckich. Pan doktor Rejchel był również opiekunem naukowym 21 prac licencjackich. Ponadto prowadził zajęcia dla studentów różnych kierunków studiów, lat i stopni z kursowych przedmiotów, takich jak analiza matematyczna, rachunek prawdopodobieństwa, statystyka matematyczna oraz z przedmiotów specjalistycznych – na przykład: analiza danych lub uczenie maszynowe.

Autorytet i rozpoznawalność naukową Pana doktora Rejchela w kraju na świecie potwierdzają wykonane recenzje dla wiodących czasopism naukowych oraz dla AMS

Mathematical Reviews, pełnienie członka Jury Konkursu PTM na najlepszą pracę studencką z teorii prawdopodobieństwa i zastosowań matematyki oraz członkostwo w Komisji Statystyki Komitetu Matematyki PAN.

5. **Konkluzja.**

Zapoznanie się z dorobkiem naukowym, dydaktycznym i organizacyjnym Pana doktora Wojciecha Rejchela oraz możliwość wysłuchania kilku jego wystąpień konferencyjnych i seminaryjnych, pozwalają mi stwierdzić, że jest On dojrzałym matematykiem, posiadającym ciekawy, aktualny i perspektywiczny program badawczy. Potrafi pracować samodzielnie oraz w zespołach badawczych. Ma szerokie doświadczenie naukowe, dydaktyczne i organizacyjne. Przedstawione osiągnięcie naukowe oceniam jako znaczący wkład w rozwój dyscypliny matematyka. Pozytywnie oceniam również pozostałe obszary działalności akademickiej Habilitanta.

Uważam, że Kandydat spełnia wszystkie ustawowe i zwyczajowe wymogi stawiane kandydatom do stopnia naukowego doktora habilitowanego i wnoszę o dopuszczenie Pana doktora Wojciecha Rejchela do dalszych etapów przewodu habilitacyjnego.