

Recenzja rozprawy doktorskiej

Autor: Mgr inż. Jakub W. Wojciechowski

Tytuł: Computational studies of amyloids and their interactions

Promotorzy:

prof. dr hab. inż. Małgorzata Kotulska, Politechnika Wroclawska

dr Johannes Söding, Max Planck Institute for Multidisciplinary Sciences, Gottingen

Dziedzina: Nauki inżynieryjno-techniczne

Dyscyplina: Inżynieria biomedyczna

Niniejsza recenzja została przygotowana na zlecenie Rady Naukowej Dyscypliny Inżynieria Biomedyczna Politechniki Wroclawskiej, zgodnie z pismem z dnia 17.07.2023 roku.

Ogólna charakterystyka rozprawy i ocena wyboru tematyki rozprawy

Rozprawa doktorska mgr. inż. Jakuba Wojciechowskiego jest napisana w języku angielskim. Tekst rozprawy liczy 38 stron tekstu spełniającego rolę przewodnika po publikacjach oraz dołączonych do niego tekstów wybranych siedmiu publikacji. W przewodniku zamieszczono ogólne wprowadzenie do tematu badań, przedstawiono tezy rozprawy oraz podsumowania wyników opublikowanych we

wspomnianych publikacjach. Dołączono też streszczenie rozprawy w języku polskim. Całość zakończono krótkim podsumowaniem oraz dyskusją.

Tematyką rozprawy jest badanie i rozwijanie metod bioinformatycznych oraz technik uczenia maszynowego do przewidywania *in silico* występowania agregatów amyloidowych oraz badania ich interakcji. Amyloidy to białkowe struktury, które mogą gromadzić się w organizmach ludzkich i zwierzęcych, tworząc tzw. włókna amyloidowe. Wiele badań potwierdza obecność agregatów amyloidów w przebiegu wielu chorób. Jednymi z najbardziej znanych przykładów jest obecność beta-amyloidu w mózgach osób cierpiących na chorobę Alzheimera czy chorobę Parkinsona. Badania wykazują również powiązanie obecności amyloidów z chorobami serca czy cukrzycą typu 2. Pokazuje to jak niezwykle istotny jest ten temat badawczy, gdyż prowadzić może do lepszego zrozumienia procesów chorobowych, identyfikacji nowych celów terapeutycznych, opracowania skuteczniejszych metod diagnostycznych i wreszcie do rozwinięcia leków i interwencji terapeutycznych mających na celu zwalczanie chorób związanych z gromadzeniem się amyloidów. W aspekcie bardzo intensywnego rozwoju prac naukowych i klinicznych nad personalizowanymi terapiami tematyka recenzowanego doktoratu mieści się w bardzo aktywnym i szerokim polu badawczym.

Oczywiście uzyskiwane postępy są w bardzo dużym stopniu napędzane przez nowe techniki eksperymentalne biologii molekularnej, biochemii, chemii leków. Jednak w obliczu złożoności zjawisk i bardzo dużych wolumenów danych analitycznych istotną rolę we wspieraniu postępów w medycynie personalizowanej grają obliczenia komputerowe oraz nowe metody modelowania matematycznego. W tym dokładnie obszarze lokują się badania naukowe przeprowadzone przez Doktoranta i przedstawione w pracy. Doktorant zaproponował nowe metody badawcze i uzyskał oryginalne wyniki naukowe w zakresie zarówno biologii obliczeniowej jak i bioinformatyki. Oryginalność jego podejścia polega na odpowiednim doborze i zastosowaniu metod uczenia maszynowego i sztucznej inteligencji do przewidywania występowania agregatów amyloidowych i ich interakcji. Przeprowadzone badania naukowe są zrealizowane na bazie dostępnych internetowo danych eksperymentalnych.

Biorąc pod uwagę powyższe, tematykę badań podjętą w przedstawionej rozprawie doktorskiej uważam za niezwykle istotną i bez wątpienia wpisującą się w obszar inżynierii biomedycznej, ze szczególnym wskazaniem na jej zastosowania w obszarze biologii obliczeniowej.

Omówienie rozprawy

Rozprawa składa się z dwóch części – w pierwszej, poświęconej skrótownemu omówieniu wyników badań (ogółem 28 stron), umieszczono 4 rozdziały nazwane odpowiednio: „General Introduction”, „Thesis of this work”, „Results” i „Summary and Discussion”. Część druga zawiera teksty 7 publikacji, których współautorem jest Doktorant. Całość uzupełniono kilkustronicowym streszczeniem w języku polskim, krótkim jednostronicowym streszczeniem w języku angielskim, podziękowaniami oraz informacją o źródłach finansowania badań. W rozprawie zamieszczono również bogaty spis bibliograficzny zawierający 154 pozycje, głównie publikacje w renomowanych czasopismach naukowych z ostatnich lat. Taką konstrukcję pracy uważam za w pełni uzasadnioną, logiczną i przejrzystą.

W części pierwszej, w rozdziale „General Introduction”, Doktorant omówił podstawowe pojęcia związane ze strukturami amyloidnymi, przedstawił krótki zarys historyczny badań nad tymi cząsteczkami oraz umieścił szeroki przegląd literatury z zakresu metod obliczeniowych wykorzystywanych w tych badaniach. Skupił się tutaj przede wszystkim na narzędziach bioinformatycznych oraz zastosowaniu technik uczenia maszynowego do rozwiązywania postawionego zadania. Opracowane wprowadzenie do tematu, z natury swojej miejscami uproszczone i skrócone, pokazuje w mojej opinii szeroką wiedzę Doktoranta z zakresu biologii molekularnej, biochemii, bioinformatyki oraz uczenia maszynowego i potwierdza bardzo dobre przygotowanie do prowadzenia badań w wybranych przez siebie obszarach.

Znajomość i zrozumienie istniejących rozwiązań pozwala na określenie ich braków i niedociągnięć oraz postawienie odpowiednich pytań badawczych, na które poszukuje się

odpowiedzi. Przedstawione w części „Thesis of this work” cele i hipotezy uważam za oryginalne, poprawnie sformułowane i odpowiednio powiązane z prowadzonymi badaniami.

Kolejna część, nazwana przez Doktoranta „Results”, stanowi zwarte podsumowanie wyników badań zawartych w siedmiu publikacjach naukowych. Sześć z nich to publikacje w renomowanych czasopismach naukowych, jedna praca ma status tzw. preprint, i dostępna jest publicznie na platformie bioRxiv. W dwóch z sześciu wspomnianych artykułów, Doktorant jest pierwszym autorem, w jednej pełni rolę autora-seniora. Cała kolekcja stanowi spójny tematycznie cykl prac, pokazujący ewolucję opracowywanych narzędzi oraz ich potencjał zastosowań praktycznych. Zauważyć też można sięganie Doktoranta po coraz bardziej zaawansowane techniki uczenia maszynowego na kolejnych etapach rozwoju opracowywanych metod.

Pierwszy z artykułów, **Wojciechowski, J.W.**, Kotulska, M. (2020) *PATH - Prediction of Amyloidogenicity by Threading and Machine Learning*. Sci Rep 10, 7721, opublikowany został w czasopiśmie zaliczonym do grupy Top10 z zakresu informatyki i inżynierii biomedycznej. Praca prezentuje pierwszy kluczowy wynik Doktoranta, nowatorską metodę PATH (Prediction of Amyloidogenicity by THreading) do przewidywania amyloidogenności, w której dzięki wykorzystaniu w procesie decyzyjnym informacji o strukturach fragmentów amyloidogennych istotnie poprawia się skuteczność klasyfikacji. Opracowane narzędzi stanowi świetny przykład połączenia metod modelowania strukturalnego z technikami uczenia maszynowego. Dzięki takiemu podejściu autorzy zdołali również zidentyfikować najważniejsze cechy strukturalne i powiązane z nimi stabilne fragmenty białek. Zgodnie z oświadczeniami autorskim, Doktorant zaangażowany był we wszystkie etapy prowadzonych badań, począwszy od sformułowania problemu, poszukiwania rozwiązania, implementacji i walidacji numerycznej opracowanych algorytmów.

W drugim z artykułów, Kotulska, M., **Wojciechowski, J.W.** (2022). *Bioinformatics Methods in Predicting Amyloid Propensity of Peptides and Proteins*. In: Li, M.S., Kloczkowski, A., Cieplak, M., Kouza, M. (eds) Computer Simulations of Aggregation of Proteins and Peptides. Methods in Molecular Biology, vol 2340. Humana, New York, NY, umieszczono coś na kształt przeglądu istniejących metod

obecnie stosowanych do obliczeniowego przewidywania skłonności białka lub peptydu do tworzenia agregatów amyloidopodobnych. Ponieważ dokładność metod bioinformatycznych w dużym stopniu zależy od danych referencyjnych wykorzystywanych do opracowania i oceny predyktorów, w publikacji przedstawiono także dla kompletności główne bazy danych amyloidów, z których powszechnie korzystają autorzy narzędzi bioinformatycznych.

Trzecia ze wspomnianych prac, Szulc, N.; Gąsior-Głogowska, M.; **Wojciechowski, J.W.**; Szefczyk, M.; Żak, A.M.; Burdukiewicz, M.; Kotulska, M. (2021) *Variability of Amyloid Propensity in Imperfect Repeats of CsgA Protein of Salmonella enterica and Escherichia coli*. Int. J. Mol. Sci. 22, 5127, stanowi wymieniony przykład praktycznego zastosowania opracowanych narzędzi do analizy danych w konkretnym problemie badawczym z zakresu biologii. Wykazano w niej również, że predykcja *in silico* nie tylko jest zgodna z wynikami eksperymentów biofizycznych ale również pozwala na zidentyfikowanie nowych własności analizowanych białek.

W czwartej z zamieszczonych prac, Kamal, M., Tokmakjian, L., Knox, J., Mastrangelo, P., Ji, J., Cai, H., **Wojciechowski, J.W.**, Hughes, M.P., Takács, K., Chu, X., Pei, J., Grolmusz, V., Kotulska, M., Forman-Kay, J.D., Roy, P.J. (2022) *A spatiotemporal reconstruction of the C. elegans pharyngeal cuticle reveals a structure rich in phase-separating proteins* eLife 11:e79396, opublikowanej w czasopiśmie notowanym w Top10 w obszarze Biologii wg bazy ELSEVIER, kontynuowano i istotnie rozszerzono badania zawarte w poprzedniej publikacji. W pracy tej autorzy skupili się na powiązaniu aktywności enzymów katabolicznych oraz ekspresji genów kodujących syntazę chityny, czynniki sieciujące chitynę i homologi regulatorów amyloidu z procesem rozwoju nicienia *Caenorhabditis elegans*. Zgodnie z oświadczeniami współautorskimi, Doktorant odpowiedzialny był za przeprowadzenie formalnej analizy zebranych danych, doskonaląc swój warsztat bioinformatyczny.

Kontynuując swoją podróż w świecie amyloidów, w pracy umieszczonej na pozycji piątej listy, **Wojciechowski J.W.**, Tekoglu E, Gąsior-Głogowska M, Coustou V, Szulc N, Szefczyk M, et al. (2022) *Exploring a diverse world of effector domains and amyloid signaling motifs in fungal NLR proteins*. PLoS Comput Biol 18(12): e1010787, dzięki zastosowaniu zaawansowanych technik uczenia

maszynowego oraz opracowanego potoku przetwarzania danych, autorzy zdołali istotnie poprawić wrażliwość opracowywanych narzędzi i zidentyfikować znacznie więcej niż poprzednio motywów amyloidowych, co doprowadziło również do identyfikacji nowych rodzin. Wyniki analiz *in silico* zostały potwierdzone w eksperymentach biologicznych. Warty podkreślenia jest to, iż jest to kolejna publikacja w czasopiśmie o wysokiej renomie naukowej, notowane w grupie Top10 czasopism w swojej dziedzinie a Doktorant jest tutaj pierwszym autorem.

Z kolei szóstą z cytowanych prac, Burdukiewicz M., Rafacz D., Barbach A., Hubicka K., Bąkała L., Lassota A., Stecko J., Szymańska N., **Wojciechowski J.W.**, Kozakiewicz D., Szulc N., Chilimoniuk J., Jęskowiak I., Gąsior-Głogowska M., Kotulska M. (2023) *AmyloGraph: a comprehensive database of amyloid-amyloid interactions*, *Nucleic Acids Research*, 51(D1): D352–D357, stanowi naturalne podsumowanie prac zespołu prof. Kotulskiej, którego członkiem jest Doktorant. Utworzenie w pełni funkcjonalnej bazy danych jest dużym wsparciem dla środowiska naukowego, a związane z realizacją takiego projektu prace wymagają od twórców dużej wiedzy eksperckiej i olbrzymiego zaangażowania czego dowodem jest opublikowanie tego narzędzia w prestiżowym czasopiśmie (z grupy Top5 wg bazy SCOPUS). Częstokroć wiąże się to ze żmudnym i czasochłonnym procesem przygotowywania i weryfikacji danych. Docenić zatem należy udział Doktoranta w tego typu przedsięwzięciach.

Ostatnia z wybranych prac, **Wojciechowski J.W.**, Szczurek W., Szulc N., Szefczyk M., Kotulska M. (2022) *PACT - Prediction of Amyloid Cross-interaction by Threading*, bioRxiv 2022.07.07.499150, dotyczy kolejnego poziomu analiz bioinformatycznych a mianowicie wykrywania nie tylko regionów ale również interakcji pomiędzy nimi. Opracowany model predykcyjny oraz utworzony serwis internetowy należy uznać za uzupełnienie narzędzia przedstawionego w pierwszej z prac, gdzie skupiono się na poszukiwaniu jedynie motywów amyloidowych. Oba narzędzia pozwalają na stworzenie wieloaspektowego obrazu procesów komórkowych i z pewnością spotkają się z dużym zainteresowaniem ze strony innych zespołów badawczych. W mojej opinii opracowanie tego narzędzia jest drugim, istotnym osiągnięciem decydującym o wkładzie doktoranta w rozwój dyscypliny.

W trakcie lektury dysertacji nasunęło mi się kilka pytań natury ogólnej, które przedstawiam poniżej:

1. Czym kierował się Doktorant wybierając wykorzystane w badaniach algorytmy uczenia maszynowego oraz selekcji cech?
2. Jak Doktorant wyobraża sobie dalszy rozwój badań w zakresie zastosowania sztucznej inteligencji w diagnostyce chorób powiązanych z tworzeniem się złożeń amyloidowych i planowaniu ich terapii?
3. Czy Doktorant uważa, że rozwijane metody znajdą w przyszłości zastosowania kliniczne?

Ocena rozprawy

Moja ogólna ocena rozprawy jest **bardzo pozytywna**. Praca bazuje na oryginalnych, współautorskich publikacjach naukowych. Doktorant jest pierwszym lub ostatnim autorem w czterech z siedmiu prac. Ponadto, konstrukcja i kompozycja pracy, umieszczenie szeregu pomocniczych wyników i tłumaczeń, bardzo dobitnie wykazują bardzo dobrą orientację doktoranta w poruszanej problematyce oraz jego wiodącą rolę w badaniach. Należy podkreślić, że doktorat jest napisany ładnym językiem, jasnym i przejrzystym. Należy także docenić logikę konstrukcji wyводу w recenzowanej pracy.

Jak już wspomniałam na początku recenzji, tematyka doktoratu lokuje się w bardzo aktywnym i szeroko rozpracowywanym obszarze badawczym. Zastosowania sztucznej inteligencji i uczenia maszynowego obejmują obecnie wszystkie kierunki badań naukowych, w których udaje się zbudować odpowiednie formalizmy matematyczne. Biologia obliczeniowa jest także bardzo intensywnie rozwijana przez wyniki z zakresu uczenia maszynowego. Jednak w tym intensywnie eksploatowanym obszarze badawczym Doktorantowi udaje się sformułować nowe i oryginalne podejścia. Tworzy hybrydowe konstrukcje głębokich sieci neuronowych i systemów rekomendacyjnych, wyszukuje i opracowuje dla nich źródła danych treningowych i weryfikacyjnych. Przeprowadza wszechstronne badania obliczeniowe.

Doktorant wykazuje się kompetencjami i dojrzałością w swoim warsztacie naukowym w dyscyplinie inżynieria biomedyczna. Posiada także zdolność badań i współpracy o charakterze interdyscyplinarnym.

Syntetyczna o cena rozprawy:

- A. Czy rozprawa zawiera oryginalne rozwiązanie problemu naukowego? Zdecydowanie TAK
- B. Czy kandydatka posiada ogólna wiedzę teoretyczną w dyscyplinie? Zdecydowanie TAK
- C. Czy posiada umiejętność samodzielnego prowadzenia pracy naukowej? Zdecydowanie TAK

Konkluzja

Osiągnięcia i oryginalne elementy publikacji wchodzących w skład rozprawy, a także jakość całego tekstu rozprawy, są na pewno wystarczające do jej ogólnej pozytywnej oceny oraz spełniają z nadmiarem zwyczajowe wymagania stawiane rozprawom doktorskim.

Przedstawiony mi do oceny dorobek naukowy Jakuba W. Wojciechowskiego uważam za w pełni wystarczający do nadania mu stopnia doktora nauk technicznych w dyscyplinie inżynieria biomedyczna. W przedstawionych pracach Doktorant wykazał się głębokim zrozumieniem analizowanych zagadnień z zakresu biologii molekularnej, biofizyki, biochemii jak i uczenia maszynowego. Uważam, że zademonstrował dojrzały warsztat naukowy a otrzymane wyniki stanowią istotny i oryginalny wkład w rozwój dyscypliny inżynieria biomedyczna. Kandydat potrafi poprawnie sformułować problem badawczy, zaproponować jego rozwiązanie oraz zaplanować eksperymenty obliczeniowe weryfikujące słuszność/poprawność otrzymanych rozwiązań. Warta podkreślenia jest dodatkowa aktywność Kandydata w obszarze Open Science, postawa, którą uważam za godną polecenia.

Stwierdzam zatem z pełnym przekonaniem, że opiniowana rozprawa Pana mgr inż. Jakuba W. Wojciechowskiego pt. „Computational studies of amyloids and their interactions” zawiera samodzielne rozwiązanie ważnego i istotnego problemu naukowego, jednocześnie spełniając wszystkie wymagania przewidziane dla rozpraw doktorskich w aktualnie obowiązującej ustawie z dnia 20 lipca 2018 roku - Prawo o szkolnictwie wyższym i nauce.

W związku z tym stawiam wniosek o dopuszczenie rozprawy doktorskiej do publicznej obrony.

Ponadto, biorąc pod uwagę liczne nowatorskie elementy w pracy, opublikowanie jej wyników w bardzo prestiżowych czasopismach naukowych, dojrzałość naukową Doktoranta i jego bardzo dobry warsztat badawczy, wnioskuję o wyróżnienie rozprawy.

