

Łódź, 31.08.2026 r.

dr hab. inż. Lidia Jackowska-Strumiłło, prof. PŁ
Instytut Informatyki Stosowanej
Wydział Elektrotechniki, Elektroniki, Informatyki i Automatyki
Politechnika Łódzka
ul. Bohdana Stefanowskiego 18,
90-537 Łódź

RECENZJA
Rozprawy doktorskiej
mgr Nishadha Himanshi Hikkaduwa Liyanage

pt. „Advancing Explainability and Predictive Accuracy in AI-Driven Cardiovascular Risk Assessment (A Hybrid Statistical and Neural Network Approach)”,

w języku polskim: „Poprawa wyjaśnialności i dokładności prognoz w opartej na sztucznej inteligencji ocenie ryzyka sercowo-naczyniowego (podejście łączące metody statystyczne i sieci neuronowe)”

dla Komisji Uniwersytetu Łódzkiego ds. stopni naukowych w dyscyplinie **Informatyka**.

1. Temat, cel i zakres rozprawy oraz rozpatrywany w niej problem naukowy

Tematyka przedłożonej rozprawy doktorskiej dotyczy prognozowania ryzyka chorób sercowo-naczyniowych, które są obecnie jednym z poważniejszych problemów zdrowotnych współczesnych społeczeństw. Skuteczne przewidywanie tego ryzyka dla indywidualnego pacjenta ma kluczowe znaczenie dla właściwego procesu leczenia.

Tradycyjne metody statystyczne jak metoda regresji liniowej, czy model proporcjonalnego hazardu Coxa CPHM (ang. *Cox Proportional Hazards Model*) oraz wskaźniki kliniczne stosowane w medycynie jak np. Skala Ryzyka Framinghama pozwalają na oszacowanie prawdopodobieństwa rozwoju chorób układu krążenia dla grupy pacjentów określonej zdefiniowanymi czynnikami ryzyka, takimi jak np.: wiek, płeć, poziom cholesterolu, ciśnienie tętnicze, palenie tytoniu, cukrzyca, itp. Metody te dostarczają precyzyjne i proste do zinterpretowania wyniki, ale ze względu na uproszczenia modelowe nie uwzględniają złożonych nieliniowych zależności pomiędzy danymi klinicznymi. Sztuczne sieci neuronowe (SSN) mają zdolność modelowania systemów nieliniowych i zapewniają większą dokładność predykcji, ale ze względu na to, że ich struktura nie odzwierciedla rzeczywistych fizycznych zależności pomiędzy danymi, ich wyniki są trudne do interpretacji w praktyce klinicznej.

Celem recenzowanej rozprawy było opracowanie modeli obliczeniowych, które zapewniłyby poprawę dokładności prognozowania oceny ryzyka sercowo-naczyniowego z zachowaniem zdolności do interpretacji klinicznej wyników.

Autorka rozprawy zaproponowała nowatorskie rozwiązanie wyżej wymienionego problemu, w którym zastosowała podejście hybrydowe łączące konkurencyjne uczenie maszynowe typu WTA (*Winner Takes All*) do grupowania danych z modelami proporcjonalnego hazardu Coxa budowanymi niezależnie dla każdej z wyodrębnionych grup pacjentów.

Jest to oryginalne rozwiązanie złożonego problemu, które uważam za główne osiągnięcie pracy. Zaproponowana przez Doktorantkę integracja metod statystycznych i uczenia konkurencyjnego spełniła określone w celu pracy wymagania.

Temat rozprawy dotyczy złożonego i wciąż aktualnego problemu naukowego, ma bardzo ważne znaczenie praktyczne i mieści się w zakresie dyscypliny naukowej Informatyka.

Autorka rozprawy sprecyzowała następujące szczegółowe problemy badawcze:

- Jak wypada porównanie skuteczności predykcyjnej sieci FFNN z proponowaną strukturą hybrydową?
- Czy modelowanie oparte na podgrupach może poprawić interpretowalność bez obniżania dokładności prognozowania?
- Jakie kompromisy pojawiają się między globalnymi modelami FFNN a podejściem hybrydowym w zadaniach prognozowania klinicznego?
- W jaki sposób można uporządkować wyniki modeli uzyskanych w obu podejściach, aby lepiej wspierały one podejmowanie decyzji klinicznych?

Doktorantka nie sformułowała tezy naukowej, ale **zagadnienie naukowe zostało jasno określone w rozprawie**. Jest ono trafne i wskazuje na **złożony i trudny problem badawczy** polegający na opracowaniu metody obliczeniowej zapewniającej większą dokładność prognozowania ryzyka chorób sercowo-naczyniowych przy jednoczesnym zachowaniu przejrzystości struktury modelu umożliwiającej interpretację kliniczną danych.

2. Zawartość i układ rozprawy, analiza źródeł

Rozprawa jest napisana w języku angielskim; liczy 81 stron oraz dodatkowych 7 stron początkowych i 33 strony końcowe ponumerowane za pomocą cyfr rzymskich zawierające streszczenie, podziękowania, spis treści, spisy rysunków i tabel oraz bibliografię i załączniki. Praca jest podzielona na 4 rozdziały; jest napisana w sposób zwięzły i jasny. Układ pracy jest prawidłowy. Praca jest napisana bardzo starannie pod względem edycyjnym.

Rozdział pierwszy to wprowadzenie. Autorka opisała w nim przegląd stanu wiedzy na temat stosowanych obecnie metod prognozowania ryzyka chorób sercowo-naczyniowych, uzasadniła ważność podjętego tematu, sformułowała problem badawczy i określiła jaki oryginalny wkład do nauki wnosi proponowane przez Nią rozwiązanie. Scharakteryzowała też bazę danych CVD SL (*The Cardiovascular Disease SL Dataset*) wykorzystaną w badaniach przeprowadzonych w pracy doktorskiej. Baza ta zawiera zbiór danych dotyczących chorób

sercowo-naczyniowych zebranych w trzech instytucjach ochrony zdrowia w Sri Lance dla ponad 66 tysięcy osób. Zestawiono w niej następujące dane pacjentów: wiek, wskaźnik masy ciała (BMI), płeć, ciśnienie rozkurczowe (DBP), poziom cholesterolu, palenie tytoniu, spożywanie alkoholu, aktywność fizyczna, poziom glukozy we krwi na czczo (FBS), które stanowią zmienne niezależne oraz występowanie chorób układu krążenia (CVD), które jest zmienną docelową (zależną). Autorka przeprowadziła wnikliwą analizę eksploracyjną w/w danych (EDA) w celu oceny i wyboru głównych czynników ryzyka CVD najistotniejszych z punktu widzenia budowy modeli obliczeniowych i wyciągnęła konstruktywne wnioski.

W rozdziale drugim rozprawy po przeglądzie stanu wiedzy na temat sztucznych sieci neuronowych typu FFNN Autorka przedstawiła autorski model obliczeniowy oparty na sieciach FFNN do oceny ryzyka występowania chorób układu krążenia. Opisała architekturę zbudowanego modelu FFNN, proces uczenia oraz ocenę jego dokładności i skuteczności prognostycznej z wykorzystaniem standardowych wskaźników klasyfikacyjnych.

W rozdziale trzecim Doktorantka przedstawiła opracowany nowatorski hybrydowy model obliczeniowy łączący grupowanie pacjentów metodą WTA z modelami proporcjonalnego ryzyka Coxa zbudowanymi dla każdego z klastrów. Do oceny skuteczności przewidywania ryzyka długoterminowego wystąpienia chorób sercowo-naczyniowych zastosowała opracowany na potrzeby oceny modelu wskaźnik skumulowanej częstości występowania (CPR).

Rozdział czwarty zawiera porównanie obu rozważanych metod oraz podsumowanie i wnioski.

Rozdziały 2-4 zawierają opis głównych osiągnięć naukowo-badawczych Doktorantki, które zostały szerzej omówione w sekcji 3. niniejszej recenzji.

Wykaz literatury zamieszczony w pracy liczy 67 pozycji, w tym 1 publikację naukową z 2026 roku, w której Doktorantka jest współautorką. Literatura jest związana z tematyką rozprawy i jest aktualna. Należy podkreślić, że w wykazie literatury są zarówno kluczowe pozycje klasyków takich jak Cox, McCulloch i Pitts, Minsky i Papert, Rosenblatt czy Rumelhart, jak i pozycje najnowsze; 16 pozycji z w/w wykazu stanowią publikacje, które ukazały się od 2020 r. do chwili obecnej, a 6 z nich opublikowano w latach 2024-2026.

3. Ocena metod i wyników naukowych, oryginalność rozprawy

Aby osiągnąć założony cel pracy Doktorantka przeanalizowała czynniki ryzyka istotne dla predykcji i zbadała skuteczność predykcji ryzyka sercowo-naczyniowego przy wykorzystaniu jednokierunkowych sztucznych sieci neuronowych FFNN (ang. *Feedforward Neural Networks*). Zaproponowała wspomniane wcześniej podejście hybrydowe i nową miarę skumulowanego ryzyka CPR (ang. *Cumulative Prevalence Ratio*), która umożliwia porównanie wyników dla różnych grup pacjentów i ułatwia ich interpretację w czasie. Opracowała hybrydową platformę obliczeniową, która łączy modelowanie oparte na danych z interpretowalnym szacowaniem ryzyka. Autorskie rozwiązania i osiągnięcia naukowo-badawcze Doktorantki zostały opisane w rozdziałach 2-4

Na etapie projektowania modelu obliczeniowego opartego na sieci FFNN do oceny ryzyka występowania chorób układu krążenia Kandydatka przeprowadziła wnikliwą analizę stanu wiedzy oraz eksploracyjną analizę danych z bazy CVD SL opisaną w rozdziale 1.

Architektura zbudowanego modelu przedstawiona jest na Rysunku 2.3. Zaproponowana sieć FFNN ma dziewięć wejść, trzy warstwy ukryte z funkcją aktywacji ReLU i jedno wyjście z sigmoidalną funkcją aktywacji. Do wejść FFNN wprowadzane jest dziewięć zmiennych niezależnych, w tym cztery ciągłe i pięć dyskretnych, które zostały wybrane do budowy modelu jako główne czynniki ryzyka CVD. Na wyjściu modelu estymowane jest prawdopodobieństwo wystąpienia chorób układu krążenia.

Doktorantka zastosowała do uczenia FFNN metodę wstecznej propagacji błędów z funkcją straty binarnej entropii krzyżowej i algorytmem optymalizacyjnym Adam (Adaptive Moment Estimation). Wyboru najlepszej architektury SSN dokonała metodą eksperymentalną w procedurze testowania zbioru modeli FFNN w oparciu o metodę walidacji skróśnej i standardowe wskaźniki klasyfikacji, takie jak: dokładność, precyzja, czułość, wskaźnik F1, pole powierzchni pod krzywą ROC (AUC). Wybór zastosowanych metod jest trafny w odniesieniu do rozważanego problemu i zastosowanej architektury SSN. Zaproponowany model FFNN do predykcji ryzyka sercowo-naczyniowego jest oryginalnym autorskim rozwiązaniem, które zapewniło dokładność klasyfikacji na poziomie 74.5%. W rozpatrywanym zastosowaniu jest to bardzo dobry wynik. Autorka przeprowadziła szczegółową wielostronną analizę uzyskanych wyników, m.in. wyznaczyła macierz pomyłek, krzywą ROC i przeprowadziła dyskusję na temat wyników i ograniczeń modelu. Opracowany model FFNN został wykorzystany przez Autorkę jako baza porównawcza dla dalszych prac.

Głównym osiągnięciem Autorki jest opracowanie nowatorskiego hybrydowego podejścia obliczeniowego do przewidywania ryzyka wystąpienia chorób sercowo-naczyniowych integrujące grupowanie pacjentów metodą „Winner-Takes-All” z modelami proporcjonalnego ryzyka Coxa wyznaczonymi dla każdej z podgrup pacjentów z osobna. Podejście to umożliwia estymację ryzyka dla każdej z podgrup o określonym stopniu zagrożenia chorobą, dzięki czemu uwzględniona jest różnorodność populacji pacjentów, a zbudowana struktura hybrydowego modelu może być wykorzystana do interpretacji klinicznej wyników. Na potrzeby oceny skuteczności przewidywania ryzyka długoterminowego wystąpienia chorób sercowo-naczyniowych Doktorantka zaproponowała oryginalny autorski wskaźnik skumulowanej częstości występowania (CPR), który nie tylko umożliwia porównanie w/w ryzyka dla różnych grup pacjentów, ale także ocenę zmian tego ryzyka w czasie.

Autorka opracowała także *framework* programistyczny do stratyfikacji ryzyka wystąpienia chorób sercowo-naczyniowych zbudowany w oparciu o proponowane hybrydowe podejście modelowe rozszerzone o analizę przeżywalności klastrow Kaplana–Meiera. Określiła schemat postępowania w analizie danych pacjentów, których celem było:

- modelowanie wyników dotyczących czasu do wystąpienia zdarzenia przy użyciu analizy przeżywalności,
- uwzględnienie heterogeniczności populacji poprzez grupowanie,
- oszacowanie ryzyka specyficznego dla poszczególnych podgrup przy użyciu modeli proporcjonalnych zagrożeń Coxa,
- dostarczenie zarówno chwilowych, jak i skumulowanych miar ryzyka sercowo-naczyniowego.

Koncepcję procesu prognozowania dla nowych pacjentów i stratyfikacji ryzyka sercowo-naczyniowego zaimplementowaną w w/w *frameworku* Autorka przedstawiła na rysunku 3.5.

Przeprowadziła także dogłębne badania eksperymentalne zaproponowanej metody na dużym zbiorze danych i jej systematyczną walidację na wielu etapach procesu modelowania, w tym analizę jakości klastrowania, ocenę przeżywalności klastrów, ocenę kalibracji, analizę dyskryminacyjną oraz ocenę spójności prognoz, a także badanie adaptacyjnego zachowanie mechanizmu uczenia konkurencyjnego WTA, w tym redukcję nieaktywnych klastrów podczas uczenia oraz stabilność tworzenia podgrup, i inne. Należy tutaj podkreślić bardzo wnikliwe i szerokie podejście do tematu.

Doktorantka dokonała oceny dokładności i skuteczności proponowanych rozwiązań stosując jako miary skuteczności indeks zgodności (C-index) i pole powierzchni pod krzywą (AUC) ROC wskazując na znaczącą poprawę w/w miar po wprowadzeniu do modelu hybrydowego analizy przeżywalności klastrów.

W rozdziale 4 Autorka przedstawiła podsumowanie wyników i analizę porównawczą modelu bazowego w postaci FFNN oraz proponowanego modelowania hybrydowego. Główne wyniki porównała w tabeli 4.1. Wyniki przeprowadzonych badań potwierdziły, że model FFNN i podejście hybrydowe zapewniają porównywalną dokładność rzędu 74,5% i 73%, natomiast analiza miary AUC-ROC (ok. 0.75 dla FFNN i 0,87 dla struktury hybrydowej) potwierdziła większą skuteczność proponowanego podejścia hybrydowego w identyfikacji pacjentów o podwyższonym ryzyku sercowo-naczyniowym i grupowaniu pacjentów z jednoczesnym zachowaniem zdolności do interpretacji klinicznej wyników.

Zaproponowane przez Doktorantkę modele obliczeniowe, przeprowadzone badania, oraz analiza ich wyników stanowią oryginalny i twórczy wkład Autorki rozprawy w dziedzinie predykcji ryzyka wystąpienia choroby sercowo-naczyniowej u pacjenta.

4. Uwagi dyskusyjne i krytyczne

Recenzentka nie ma uwag krytycznych. Praca jest napisana bardzo starannie zarówno pod względem merytorycznym jak i edytorskim. Zauważyłam tylko nieliczne niedociągnięcia edytorskie, jak np. brak numeru sekcji (see Section ??) na stronie 62, które nie wpływają na ocenę pracy.

Prosiłabym Doktorantkę o odpowiedź na dwa pytania dyskusyjne.

1. Jaka strategia została przyjęta przy wyborze i redukcji liczby klastrów w metodzie uczenia konkurencyjnego WTA i jaki jest jej wpływ na jakość predykcji ryzyka sercowo-naczyniowego?
2. W jaki sposób oceniana była liczba nierozpoznanych przypadków wystąpienia chorób sercowo-naczyniowych (false negative) w podejściu hybrydowym? Czy któraś z zastosowanych miar jakościowych to uwzględnia?

5. Ocena dorobku naukowego i osiągnięć doktorantki

Mgr Nishadha Himanshi Hikkaduwa Liyanage jest współautorką 7 publikacji i opracowań naukowych w dyscyplinie Informatyka, w tym między innymi 1 artykułu w czasopiśmie z listy JCR, jednego artykułu w czasopiśmie w bazie Scopus i 1 publikacji w materiałach cyklicznej konferencyjnej międzynarodowej w bazie Scopus i Web of Science. Dwie publikacje kandydatki z lat 2025-2026 są związane tematycznie z ocenianą rozprawą doktorską, a Doktorantka jest w nich pierwszą współautorką. Za najważniejszy uważam

artykuł w czasopiśmie *Archives of Control Sciences* (IF = 0,8; 100 punktów na liście MNiSW), w którym opisano główne wyniki przeprowadzonych w rozprawie badań.

Dorobek publikacyjny doktorantki spełnia wymagania stawiane pracom doktorskim. Jej prace są cytowane 14 razy bez autocytowań w bazie *Web of Science* i 26 razy w bazie *Scopus*, a indeks Hirsha wynosi $H=2$ w bazie *Scopus*.

Ocena dorobku publikacyjnego Doktorantki była utrudniona ze względu na fakt, iż pozycja nr 1 z załączonej listy Jej dorobku naukowego wpisana została do baz danych *Scopus* i *WoS* pod nazwiskiem Liyanage, a pozycje 3 i 4 pod nazwiskiem Himanshi.

6. Podsumowanie i wniosek końcowy

Przedłożona do recenzji rozprawa dokumentuje samodzielne i oryginalne rozwiązanie złożonego problemu naukowego i zawiera opis oryginalnych osiągnięć badawczych stanowiących istotny wkład do dyscypliny naukowej Informatyka.

Autorka rozprawy wykazała się rozległą i dogłębną wiedzą w w/w. dyscyplinie naukowej, ale także bardzo dobrą znajomością metod statystycznych oraz umiejętnością planowania i prowadzenia badań eksperymentalnych.

W związku z tym stwierdzam, że rozprawa doktorska Pani mgr Nishadha Himanshi Hikkaduwa Liyanage spełnia wymagania Ustawy z dnia 20 lipca 2018 roku Prawo o szkolnictwie wyższym i nauce.

Wnioskuje do Komisji Uniwersytetu Łódzkiego ds. stopni naukowych w dyscyplinie Informatyka o dopuszczenie rozprawy doktorskiej mgr Nishadha Himanshi Hikkaduwa Liyanage do publicznej obrony.

Biorąc pod uwagę wysoki poziom merytoryczny rozprawy, oryginalność zaproponowanego rozwiązania, wnikliwe i szerokie podejście do tematu oraz całokształt dorobku publikacyjnego Kandydatki wnioskuję o wyróżnienie recenzowanej rozprawy doktorskiej.