



UNIwersytet Medyczny im. Karola Marcinkowskiego w Poznaniu

Katedra i Zakład Chemii Farmaceutycznej i Farmacji Szpitalnej

ul. Rokietnicka 3
60-806 Poznań

Poznań, dnia 30 września 2026 r.

Recenzja pracy doktorskiej Pani mgr Andżeliki Lorenc

pt. *Zastosowanie algorytmów uczenia maszynowego w analizie predykcyjnej aktywności przeciwdrobnoustrojowej pochodnych imidazoliowych i wznowy raka płuca*

wykonanej pod kierunkiem Pana prof. dra hab. Adama Buczińskiego oraz

Pani dr hab. Anny Romaszko-Wojtowicz, prof. UWM

w Katedrze Biofarmacji Collegium Medicum im. Ludwika Rydygiera w Bydgoszczy

Uniwersytetu Mikołaja Kopernika w Toruniu

Celem przedstawionej do recenzji pracy była ocena przydatności sztucznych sieci neuronowych typu perceptron wielowarstwowy w modelowaniu (i) aktywności przeciwdrobnoustrojowej związków imidazoliowych oraz (ii) predykcji wznowy raka płuca na podstawie danych klinicznych. Temat poruszany przez Doktorantkę jest bardzo aktualny, a wykorzystanie danych, obliczeń i sztucznej inteligencji wkracza nieodwołalnie we wszystkie aspekty naszego życia. Sztuczne sieci neuronowe rozwijają się w bardzo szybkim tempie. Technologia ta daje z jednej strony olbrzymie możliwości, z drugiej stwarza poważne wyzwania społeczne, etyczne i techniczne, a jej przyszłość zależy w dużej mierze od wprowadzenia odpowiednich regulacji prawnych. Istotność zagadnień związanych z możliwościami wykorzystania sztucznych sieci neuronowych potwierdza także opublikowanie w Farmakopei Polskiej XIV rozdziału *Jakość danych*. Rozdział ten, publikowany w celach informacyjnych, dotyczy danych pochodzenia cyfrowego i zawiera przegląd koncepcji jakości danych oraz podsumowanie procedur zarządzania danymi, stosowanymi w nauce i w informatyce. Stanowi także uzupełnienie innych, wcześniej opublikowanych rozdziałów, takich jak np.: *Metody chemometryczne stosowane do danych analitycznych*. Dlatego w pełni uzasadnione jest podjęcie przez Doktorantkę badań związanych z możliwością wykorzystania algorytmów uczenia maszynowego.

Rozprawa doktorska została przedstawiona do recenzji w formie cyklu 3 prac, opublikowanych w latach 2023-2026, w czasopiśmie z listy filadelfijskiej, o łącznym

współczynnika oddziaływania $IF = 11,0$, co odpowiada 280 pkt MEiN. Publikacje mają charakter opracowań wieloautorskich. We wszystkich pracach Doktorantka jest pierwszym autorem. Jednocześnie, w opisie merytorycznym udziału w publikacjach, Kandydatka potwierdza, że Jej wkład w powstaniu prac wynosił od 45% do 65%. Publikacje wchodzące w skład rozprawy doktorskiej są następujące:

1. *Antimicrobial activity classification of imidazolium derivatives predicted by artificial neural networks*. Pharmaceutical Research, 2024, 41, s. 891-898.

W publikacji tej zostało przedstawione podejście oparte na analizie przed syntezą, co posłużyło do klasyfikacji 140 związków, pochodnych imidazoliowych z wykorzystaniem sztucznych sieci neuronowych, w połączeniu z innymi technikami uczenia maszynowego. Mechanizm działania badanych związków jest ściśle powiązany z ich strukturą chemiczną oraz właściwościami fizycznymi. Związki te jako substancje powierzchniowo czynne łatwo przenikają przez błony biologiczne o budowie białkowo-lipidowej, powodując zaburzenia ich spójności strukturalnej i funkcjonalnej, co czyniło je potencjalnie aktywnymi wobec szczepów *Klebsiella pneumoniae*. Podstawę podziału otrzymanych związków stanowiła wyznaczona wartość MIC, co umożliwiło ich zróżnicowanie na dwie klasy aktywności – niską i wysoką. Bardzo duża liczba deskryptorów molekularnych spowodowała zastosowanie wieloetapowego preprocessing danych. Aby zredukować liczbę zmiennych wykorzystano metodę drzew klasyfikacyjnych CART i wybrano 200 deskryptorów, które stanowiły dane wejściowe do dalszej redukcji wymiarowości z wykorzystaniem PCA. Odpowiednia walidacja krzyżowa miała na celu zwiększenie wiarygodności oceny zdolności generalizacyjnej modelu. W wyniku przeprowadzonych badań, Doktorantka wykazała, że integracja metod selekcji cech (CART), redukcji wymiarowości (PCA) oraz modelowania nieliniowego (MLP), stanowi skuteczne podejście w zadaniach klasyfikacji aktywności biologicznej związków chemicznych charakteryzowanych wielowymiarowymi zbiorami danych.

2. *Hybrid machine learning model for dual prediction of antimicrobial activity against Staphylococcus aureus and Candida albicans*. Journal of Pharmaceutical Innovation, 2026, 21, s. 1-9.

W publikacji numer dwa Autorka podjęła próbę rozwiązywania zadania bardziej złożonego – predykcji dokładnych wartości MIC. Analiza w tej pracy dotyczyła aktywności wobec bakterii *Staphylococcus aureus* oraz grzyba *Candida albicans*. Wybór został uzasadniony skłonnością do współwystępowania oraz wywoływania zakażeń mieszanych. Pierwszy etap badań polegał na wstępnym oczyszczeniu danych z użyciem języka Python i jego bibliotek.

W kolejnym etapie wykonano selekcję deskryptorów molekularnych z wykorzystaniem algorytmu drzew decyzyjnych CART. Badania prowadzono dla każdego z drobnoustrojów, po czym po porównaniu obu zbiorów, wyodrębniono predyktory wspólne dla *Staphylococcus aureus* oraz *Candida albicans*. W dalszym etapie badań zastosowano PCA jako metodę redukcji wymiarowości, co zredukowało wymiar przestrzeni oraz eliminowało współliniowość deskryptorów molekularnych. Ostatni etap badań stanowiło skonstruowanie sieci neuronowej typu MLP, której zadaniem była jednoczesna predykcja wartości MIC dla szczepów *Staphylococcus aureus* oraz *Candida albicans*.

Zastosowany przez Doktorantkę proces analityczny pozwolił skutecznie wyodrębnić informację predykcyjną z wielowymiarowej przestrzeni deskryptorów molekularnych oraz odwzorować zależności pomiędzy budową chemiczną badanych związków a wartościami MIC badanych mikroorganizmów.

3. *Exploring the efficacy of artificial neural networks in predicting lung cancer recurrence: a retrospective study based on patients records*. Translational Lung Cancer Research, Society for Translational Medicine, 2023, 10, s. 2083-2097.

W publikacji numer trzy przedmiotem analizy został problem kliniczny, czyli predykcja wznowy raka płuca na podstawie retrospektywnych danych medycznych pacjentów. W badaniach wykorzystano 2296 historii chorób pacjentów hospitalizowanych w latach 2012-2017 w Warmińsko-Mazurskim Centrum Chorób Płuc w Olsztynie. W badaniu tym dobór predyktorów wynikał z aktualnej wiedzy medycznej oraz z doświadczenia zespołu klinicznego, a o istotności zmiennych decydowały przesłanki kliniczne czynników potencjalnie związanych ze wznową. Największe znaczenie dla predykcji miały informacje dotyczące terapii. Wyniki uzyskane przez Doktorantkę potwierdzają, że sieci neuronowe typu MLP mogą stanowić efektywne narzędzie wspomagające ocenę ryzyka wznowy raka płuca i mogą być przykładem wykorzystania sztucznej inteligencji jako narzędzi wspomagania decyzji klinicznych. Badania tego rodzaju powinny być kontynuowane i należy w nich uwzględniać aktualne wytyczne leczenia tego typu nowotworu.

Układ pracy jest typowy dla publikacji będącej zbiorem prac własnych Autorki. Dysertacja składa się z pięciu podstawowych rozdziałów, w tym wstępu, celu pracy, materiałów i metod, wyników, podsumowania i wniosków. Dodatkowo, w opracowaniu zawarto literaturę, spis rycin, publikacje będące podstawą do ubiegania się o nadanie stopnia doktora oraz wymagane oświadczenia, życiorys naukowy Kandydatki i wykaz publikacji. Badania prowadzone były we współpracy z Katedrą Technologii Postaci Leku UMK, Katedrą

i Zakładem Biofarmacji i Farmakodynamiki GUMed, Wydziałem Technologii Chemicznej Politechniki Poznańskiej, a także Warmińsko-Mazurskim Centrum Chorób Płuc w Olsztynie.

W rozdziale *Wstęp* Autorka opisała typy uczenia maszynowego oraz wybrane algorytmy uczenia maszynowego, w których zawarła drzewa decyzyjne, CART, analizę składowych głównych, PCA oraz sztuczne sieci neuronowe, ANN. Zagadnienia powyższe zostały przedstawione w sposób naukowy, ale przystępny dla czytającego, który nie pracuje na co dzień z tego typu zagadnieniami.

W kolejnym rozdziale – *Cel pracy*, Autorka doprecyzowała cel swoich badań oraz przedstawiła, jakie badania eksperymentalne zostaną przeprowadzone. Zakres badań Doktorantka podzieliła na 4 cele szczegółowe (i) opracowanie i walidacja modeli sztucznych sieci neuronowych typu MLP do klasyfikacji aktywności przeciwdrobnoustrojowej pochodnych imidazoliowych wobec *Klebsiella pneumoniae*, na podstawie deskryptorów molekularnych, z oceną wpływu selekcji metodą CART oraz redukcji wymiarowości metodą PCA na jakość klasyfikacji i zdolność generalizacji modelu, (ii) opracowanie i walidacja regresyjnego modelu sztucznej sieci neuronowej typu MLP do ilości predykcji wartości lnMIC pochodnych imidazoliowych wobec *Staphylococcus aureus* oraz *Candida albicans*, z wykorzystaniem selekcji predyktorów metodą CART, redukcji wymiarowości metodą PCA oraz oceny stabilności i jakości predykcyjnej modelu, (iii) opracowanie i ocena modelu MLP do predykcji wznowy raka płuca na podstawie retrospektywnych danych klinicznych, z uwzględnieniem doboru zmiennych predykcyjnych o znaczeniu klinicznym, optymalizacji architektury sieci oraz algorytmów uczenia oraz analizy jakości klasyfikacyjnej, (iv) przeprowadzenie porównawczej analizy skuteczności zastosowanych metod uczenia maszynowego w modelowaniu danych o różnym charakterze (molekularne i kliniczne), ze szczególnym uwzględnieniem wpływu wysokowymiarowości danych na proces uczenia, roli selekcji zmiennych i redukcji wymiarowości w budowie stabilnych modeli oraz znaczenia procedur walidacyjnych w ocenie zdolności generalizacji.

Wszystkie etapy badań zostały prawidłowo zaplanowane, wykonane i szczegółowo opisane w załączonych publikacjach. W dalszej części Autorka przedstawiła metody badawcze oraz wykorzystane oprogramowanie, a także ocenę jakości modeli klasyfikacyjnych oraz ocenę jakości modeli regresyjnych.

W części *Wyniki*, Autorka przedstawiła w sposób zwięzły zakres badań poszczególnych prac oraz uzyskane najważniejsze wyniki.

Kolejny rozdział, najistotniejszy dla oceniającego to *Podsumowanie i wnioski*, które odnoszą się do 3 artykułów oryginalnych. W rozdziale tym Doktorantka przedstawiła wspólną

oś metodologiczną dla wszystkich prowadzonych badań, wskazując że skuteczność modeli MLP zależy w dużym stopniu od charakteru danych oraz sposobu ich przygotowania. Jednocześnie najważniejszym czynnikiem determinującym jakość predykcji jest świadome projektowanie całego procesu modelowania, uwzględniające charakter i strukturę analizowanych danych. Autorka wykazała, że metody uczenia maszynowego są bardzo przydatnym narzędziem w projektowaniu nowych substancji o potencjalnym działaniu biologicznym oraz w podejmowaniu decyzji klinicznych. W rozdziale tym nie znalazłam klasycznych wniosków, a jedynie podsumowanie badań.

Założenia pracy zostały przygotowane prawidłowo, a odpowiedzi na pytania zawarte w celu pracy, zostały przedstawione w postaci podsumowania a nie wniosków.

Dobór literatury w liczbie 59 pozycji, uwzględnia najnowsze piśmiennictwo i spełnia wymagania w tym zakresie. Przedstawiona dyskusja świadczy o znajomości i zrozumieniu poruszanych zagadnień. Wykonane przez Doktorantkę badania, ich omówienie i podsumowanie, potwierdzają umiejętność prowadzenia badań naukowych, stanowią oryginalne rozwiązanie problemu naukowego, zawierają elementy nowości naukowej i wnoszą znaczący wkład w rozwój dziedziny nauk medycznych i nauk o zdrowiu, w dyscyplinie nauki farmaceutyczne oraz są cenne dla przedstawicieli zawodów medycznych i pokrewnych.

Podsumowując, jest to bardzo dobrze przygotowana praca, która spełnia wymogi stawiane rozprawom doktorskim. Drobne uchybienia i nieścisłości wymienione poniżej nie mają wpływu na wysoką ocenę przedstawionej mi do recenzji pracy:

- Autorka często używa sformułowania związku syntetyzowane, powinno być syntezywane,
- na str. 28 w nazwie chlorków nie powinno być spacji,
- na str. 28 jest chlorków ...imidazol-1-ylo, powinno być imidazol-1-ilo.

Podczas obrony pracy chciałabym podyskutować na temat, jakie ograniczenia i bariery dostrzega Doktorantka w możliwości wprowadzenia algorytmów uczenia maszynowego do praktyki klinicznej? Czy rodzaj raka płuca był uwzględniony w badaniu? Czy losowy podział danych na zbiór uczący (50%), walidacyjny (25%) oraz testowy (25%) jest podejściem standardowym, czy bywa stosowany inny podział i jeżeli tak, to dlaczego?

W podsumowaniu pragnę podkreślić zasługującą na uznanie wysoką jakość przeprowadzonych badań. Doktorantka pracowała w bardzo doświadczonym zespole badawczym, pod kierunkiem Pana Profesora Adama Bucińskiego oraz Pani dr hab. Anny Romaszko-Wojtowicz, prof. UWM, jednakże bez Jej osobistego zaangażowania, realizacja

tak zaplanowanego zakresu badań byłaby niemożliwa. Godnym podkreślenia jest także znaczący dorobek naukowy Doktorantki, która jest współautorem w sumie 11 publikacji, z których 3 zostały włączone do cyklu będącego podstawą do ubiegania się o uzyskanie stopnia naukowego doktora. Sumaryczny współczynnik oddziaływania IF prac Doktorantki wynosi 83,503, co odpowiada 1190,00 punktom MEiN. Ponadto, Pani mgr Andżelika Lorenc realizowała 4 projekty badawcze, będąc kierownikiem jednego projektu. Na podkreślenie zasługuje także udział Doktorantki w licznych kursach oraz szkoleniach językowych związanych z podnoszeniem kwalifikacji zawodowych.

Wniosek końcowy

Rozprawa doktorska Pani mgr Andżeliki Lorenc pod tytułem *Zastosowanie algorytmów uczenia maszynowego w analizie predykcyjnej aktywności przeciwdrobnoustrojowej pochodnych imidazoliowych i wznowy raka płuca* zawiera elementy nowości naukowej oraz jest oryginalnym, bardzo dobrze zaplanowanym i przeprowadzonym badaniem naukowym. Autorka w pełni zrealizowała założone cele badawcze, a przedstawiona do oceny rozprawa jest pozycją wartościową z naukowego i aplikacyjnego punktu widzenia, spełniającą wymagania określone w art. 187 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce.

Oceniając pozytywnie, zarówno pod względem formalnym, jak i merytorycznym pracę doktorską mgr Andżeliki Lorenc pod tytułem *Zastosowanie algorytmów uczenia maszynowego w analizie predykcyjnej aktywności przeciwdrobnoustrojowej pochodnych imidazoliowych i wznowy raka płuca*, mam zaszczyt zwrócić się do Szanownego Pana Przewodniczącego Rady Dyscypliny Nauki Farmaceutyczne w Collegium Medicum im. Ludwika Rydygiera w Bydgoszczy Uniwersytetu Mikołaja Kopernika w Toruniu o przyjęcie pracy i dopuszczenie Pani mgr Andżeliki Lorenc do dalszych etapów przewodu doktorskiego. Jednocześnie, uwzględniając wysoki poziom przeprowadzonych badań, ich kompleksowe przedstawienie oraz opublikowanie wyników badań w czasopismach o wysokim współczynniku oddziaływania, zwracam się do Wysokiej Rady Dyscypliny Nauki Farmaceutyczne z wnioskiem o wyróżnienie rozprawy doktorskiej, co załączam także w oddzielnym wniosku.

A. Jeliński